Competition for Imperfect Consumers

John Vickers*

April 2021

Abstract

This lecture provides a selective discussion of some issues in the economics of competition for imperfect consumers. Limited consumer awareness – about what deals are on offer, or about the consequences of purchasing decisions – is a central feature of many markets. Consumer imperfection can give rise to market power, but can also co-exist with strong rivalry between firms, and then is more for 'consumer policy' than competition policy. Two recent court cases illustrate some of the issues involved. The lecture then reviews a line of theoretical work on competition for imperfect consumers. The analysis highlights some policy tradeoffs, and a key theme is how patterns of consumer awareness matter for patterns of competition among firms.

JEL codes: D11, D18, D43, D83, L13, L51

Keywords: Price dispersion, price discrimination, consumer search, consideration sets, consumer policy.

1 Introduction

Much of the theory of industrial organization concerns imperfect competition for perfectly informed customers. In this lecture, however, I want to talk about competition for consumers that are imperfect, for example in their awareness of what deals are on offer, or about the consequences of deals they enter into. Given costs of information acquisition, there is nothing irrational in being imperfect in that sense. Anyway consumer imperfection is a primary fact about how many markets work. It is manifested in the widespread consequence that some consumers get much worse deals than others for essentially the same product in the same market.

*Department of Economics and All Souls College, University of Oxford. This paper was for the presidential address at the annual conference of the European Association for Research in Industrial Economics on 29 August 2020. It largely discusses joint work with Mark Armstrong, to whom I am especially grateful. Thanks also to Simon Cowan, Amelia Fletcher, Leslie Marx, Jidong Zhou, and a referee.

This can happen in at least three ways. The first is *price dispersion* – different firms charging different (uniform) prices for the same product. Second, there is *price discrimination* – the same firm charging different prices to different consumers, a kind of intra-firm price dispersion. Third, there is *add-on pricing*, whereby naive or unlucky consumers incur high follow-on charges that the savvy or lucky avoid. In short, and despite the internet, we are a long way from the law of one price in many consumer markets.

Consumer imperfection can give rise to market power, and hence questions for competition policy. For example, weak customer response was a central feature of the findings of the Competition and Markets Authority (2016a) report on UK energy markets. Although the supply side was moderately, rather than highly, concentrated – with six large firms plus challengers – the CMA concluded that "some suppliers have a position of unilateral market power, arising from the extent of customer lack of engagement in the market, and that these suppliers have the ability to exploit such a position, for example through price discrimination". The disengaged retail customers on poor deals tended to be lower-income households, which appears to be a widespread phenomenon as documented by Byrne and Martin (2021).[1] Following the CMA report, price caps were introduced for consumers on pre-payment meters and for those on default tariffs with 'standard variable' rates, which disengaged consumers typically paid for their energy supplies. This example directly exhibits the issues of price dispersion and price discrimination mentioned above. And it relates to the issue of add-on pricing if 'loyal' customers on default tariffs get substantially worse deals than those who shop around.

Consumer imperfection can co-exist with otherwise effective competition. For example, in markets with add-on prices or charges that arise only in certain contingencies, which some consumers might not have foreseen, a pattern of bargains (for savvy consumers) and ripoffs (for the naive ones) can emerge quite consistently with vigorous rivalry. We are then not within the realm of competition policy, at least as normally understood, but consumer policy. Some especially important examples of these issues come from consumer financial markets, the subject of John Campbell's (2016) Ely lecture to the American Economic Association – "an area where competitive markets may deliver substandard outcomes that can be improved through intervention".

[1] Byrne and Martin (2021) review evidence from a broad set of markets that shows an inverse U-shaped relationship between search intensity and income. They discuss reasons why low-income households fail to search, and the distributionally regressive price dispersion that may result.

When I was at the UK's Office of Fair Trading[2] 15-20 years ago, with responsibilities spanning both competition and consumer policy – or rather their union – it was striking how economists were ubiquitous in competition policy but rarely to be seen in consumer policy, despite both dealing with markets, and sometimes the very same market. That contrast was a theme of a lecture published as Vickers (2004), but a lot has happened since – in our economic understanding, in market developments (especially online), and to some extent in policy.[3] This is a large topic and a survey of those developments would go far beyond the scope of this paper.[4] Instead, the main aim of this paper is to give a selective account of a line of theoretical work with Mark Armstrong, some of it currently in progress, on competition for imperfect consumers. That work mostly addresses the first two issues above – price dispersion and price discrimination – in moderately competitive markets.

But I will begin with the issue of add-on charges, and with law more than economics, by drawing attention to two recent cases in consumer and contract law decided by the UK's Supreme Court. I do this to suggest that consumer and contract law cases, which are much less familiar to industrial economists than antitrust and merger cases, perhaps deserve more attention from us. They can have considerable economic significance for how markets work, and they may be of interest to economists in their own right.

2 Two add-on pricing cases: overdrafts and parking charges

Unauthorised overdraft charges

The overdraft charges case that came to the UK's Supreme Court in 2009 concerned a classic instance of add-on pricing in the form of contingent charges.[5] When a consumer made a purchase that made their bank balance negative – specifically, more negative than any pre-agreed overdraft limit – then the banks would impose insufficient fund charges, which were widely regarded as unfairly high. They were typically £20-£25 per item, and

[2] In 2014 the newly-established CMA took over the functions of the OFT and also those of the Competition Commission.

[3] On complementarities and tensions between competition and consumer policies, see Armstrong (2008).

[4] The book by Spiegler (2011) offers a broad synthesis of theoretical work on markets where profit-maximizing firms interact with boundedly rational consumers. The limited consumer awareness in the models reviewed below can be viewed as a type of bounded rationality.

[5] *Office of Fair Trading (Respondents) v Abbey National plc and Others (Appellants)* [2009] UKSC 6.

in 2006 such charges were about 30 per cent of the banks' revenue from current account customers. Under the UK's free-if-in-credit model of retail banking, a large cross-subsidy operated in favour of the majority of consumers (42 million) who never or rarely incurred the charges and the minority (12 million) who regularly did so, sometimes to the tune of hundreds of pounds a year.

The OFT wanted to take action against the charges under the Unfair Terms in Consumer Contracts Regulations 1999, which implemented an EU Directive.[6] In 2006 the OFT had intervened with some success against the similar issue of default charges for late payments on credit cards and for exceeding credit limits on cards. The OFT set a threshold of £12 per item as a guide for when it would, and would not, take action. Default fees had often been twice that level. The OFT's view, and the basis for its £12 guide, was that unfairness arose if a charge exceeded the level of common law damages based on a reasonable pre-estimate of cost to the card issuer. (This common law damages point is central to the parking charges case discussed below.) On credit card charges the OFT's statement changed market practice without formal action being taken. On unauthorised overdraft charges, however, the banks resisted the OFT's approach, arguing that the OFT did not have the power to challenge the fairness of the charges under the Regulations. The OFT won in the High Court and Court of Appeal but then the matter came to the Supreme Court.

The fairness of the charges was not the question for the Court. The issue was whether the OFT had the power under the Regulations to challenge the charges as unfair. The Regulations specified that: "In so far as it is in plain intelligible language, the assessment of fairness of a term shall not relate – (a) to the definition of the main subject matter of the contract, or (b) to the adequacy of price or remuneration, as against the goods or services supplied in exchange".

The Court ruled that the charges were part of the price or remuneration for the package of services that a current account provides. Their being contingent, and not incurred by most consumers, was held to be irrelevant. As the contract terms were largely in plain intelligible language, they were therefore outside the scope of the Regulations. So the level of charges could not be assessed for fairness under the Regulations.

Events have since moved on, especially following the CMA (2016b) investigation of the

[6] The Regulations have since been superseded by the Consumer Rights Act of 2015.

UK retail banking market, of which overdrafts were one of many elements. Unsurprisingly, users of unarranged overdrafts were found to have poor engagement with the market, and indeed often to be unaware that they had become overdrawn beyond pre-agreed limits. They were found to suffer particularly from competition failures in the market, and it was unrealistic to expect competition to solve their problems. A package of remedies, enshrined in the regulatory order adopted as a result of the investigation, was therefore aimed at them. It included a requirement on banks to auto-enrol customers in overdraft alerts, usually by text, and a maximum monthly charge cap.[7] In 2019 the UK Financial Conduct Authority set out new rules for overdraft pricing, which require banks to express overdraft charges as an annual percentage interest rate, and with unarranged overdrafts no more expensive than arranged ones. Most banks have set an overdraft interest rate near 40% – often an increase on prior percentage rates but with fixed fees banned. The FCA estimated that 7 out of 10 overdraft users would be better off, or see no change in cost, as a result of the new rules.

Following the court judgment, but before the CMA investigation, we set out a very simple model of some of the issues in the case in Armstrong and Vickers (2012).[8] Banks in perfect competition supply two tied services – the basic service (current accounts) with price P and a contingent add-on service (unauthorised overdrafts) at price p, which a consumer that does not take care pays with probability τ. The two services have respective unit costs C and c. Exogenous proportion λ of consumers are savvy in that they take care not to incur the add-on price p if it is high. The other consumers don't pay attention to p, perhaps because of undue optimism, and are attracted to the bank with the lowest basic service price. Suppose that $\bar{p}$ is some high limit on the add-on price. Then there is an equilibrium with high overdraft charges ($p = \bar{p}$) and *subsidised* current account charges:

$$P = C - \tau(1 - \lambda)(\bar{p} - c) \ .$$

Banks break even, and savvy consumers get a subsidy equal to the expected profit from the contingent charges paid by the non-savvy.

A bank that deviated by lowering its p would make less from non-savvy consumers, and savvy ones are loss-making, so the market could get stuck in this sort of equilibrium.

[7] Adam et al (2018) report results from a field experiment on text alerts. They appear significantly to reduce overdraft fees paid.

[8] The model was based on Gabaix and Laibson (2006) but without 'shrouding' by firms.

It is bad in distributional terms if the non-savvy are low-income or otherwise vulnerable and have a higher welfare weight. It is also inefficient if (as in our paper) effort incentives matter. In particular, high add-on charges may cause the savvy to make socially excessive efforts avoid them, when it might be efficient for them to incur them occasionally.

An important feature of this bad equilibrium is that raising λ, the proportion of savvy consumers, *worsens* the outcome for both consumer types because the subsidy falls. (But if λ is increased enough, equilibrium flips to being efficient.) Armstrong (2015) calls this a ripoff externality, in contrast to the positive externality of having more savvy consumer that arises in other settings, as will be discussed below. However, constraining $\bar{p}$ and reducing τ (the probability of incurring the charges) are welfare-improving in this model, especially if distributional concerns have weight.

The OFT consumer law case was in these terms an attempt to constrain $\bar{p}$. It failed because the Court decided that the charges were part of the price for the package of current account services, and so beyond the scope of the relevant regulations. Following the review by the CMA, regulation now requires text alerts to improve consumer awareness (which relates to τ), relative price regulation as between arranged and unarranged overdrafts, and regulation of the form of pricing in terms of annual percentage rates and the ban on fixed fees (which relate to $\bar{p}$).

Parking charges

The second case that I want to mention concerns an £85 parking charge, which might seem trivial by comparison with overdraft charges, but in the case in question the UK Supreme Court fundamentally re-stated the law on penalty clauses for breach of contract.[9] It therefore has widespread economic importance.

In April 2013 Mr Beavis parked in a car park in a shopping centre near Chelmsford railway station. Parking for up to two hours was free but prominent, indeed orange, notices warned that staying longer would trigger an £85 charge. Mr Beavis's car was left there for almost an hour beyond the two-hour limit, and he received a demand for the £85. His challenge to the charge concerned a profound point in contract law in the UK and a range of other jurisdictions – the ancient rule that penalty clauses for breach of contract are unenforceable.[10] For a century before the 2015 judgment of the Supreme Court in the

[9] *ParkingEye Limited (Respondent) v Beavis (Appellant)* [2015] UKSC 67.

[10] He also (unsuccessfully) challenged the charge as unfair under the consumer law regulations discussed above in connection with the overdrafts case.

Beavis case the leading case on the point was the *Dunlop* case of 1914. Ironically this concerned the consequences of breaching a resale price maintenance agreement for tyres, which in the modern era would run afoul of competition law. A test formulated in that case was that a clause was penal if "the sum stipulated for is extravagant and unconscionable in amount in comparison with the greatest loss that could conceivably be proved to have followed from the breach". It followed that damages for breach are not penal if they are a genuine pre-estimate of the loss to the innocent party of the breach by the other. But are they necessarily penal if they exceed such a pre-estimate?

If so, it appears that Mr Beavis would have won, because the hour's over-stay of his car would have cost the operator of the car park nothing like £85. In fact Mr Beavis lost his case. The Court ruled that "the fact that the clause is not a pre-estimate of loss does not ... mean that it is penal". The true test, it declared: "is whether the impugned provision is a secondary obligation which imposes a detriment on the contract-breaker out of all proportion to any legitimate interest of the innocent party in the enforcement of the primary obligation". While punishment was not a legitimate interest, mere compensation for breach was not the only kind of legitimate interest. In the context of the car park, the £85 charge promoted the efficient use of the parking space for the nearby shops and their customers, and it yielded revenue for the operator. It aimed to attract customers to the shops while deterring long-stayers not shopping, such as commuters by rail going to the station nearby, but not to punish them. The sum of £85 was not considered extravagant or unconscionable for those purposes.

Although I was sorry that the OFT lost the case about unauthorised overdraft charges, the outcome of the parking charge case makes economic sense in my view. There is an argument for aligning the payment for breach with a pre-estimate of loss in terms of *ex post* incentives for efficient breach, but there are other aspects of efficiency too, notably *ex ante* relationship-specific investment incentives.[11] In the parking case, however, a key issue was the mix of consumer types – screening in short – and it seems natural and efficient for the landowner, via the car park operator, to give an incentive to some consumer types and a disincentive to others. If charges above the level of pre-estimated loss were unenforceable, such incentives would be unduly constrained. The prominence of the notices, as distinct from the situation of a vulnerable consumer going overdrawn, is surely important too.

[11] Spier and Whinston's (1995) analysis of damages provisions covers investment incentives, renegotiation, and entry deterrence.

Whereas one can easily go overdrawn without realising, one knows when one has entered a car park, and with prominent signage it is clear what the deal is.

The reason I draw attention to this contract law case, however, is not to give any economic analysis of it. It is simply to highlight that interesting and important questions about markets and pricing, beyond the scope of competition policy, are being decided in the courts, and to suggest that economists could usefully engage with them more.

3 Some early economics of price dispersion with imperfect consumers

The economic theory literature of competition for imperfect consumers has a long pedigree. A remarkable symposium on the economics of information published in the *Review of Economic Studies* 1977 contains three seminal papers on equilibrium price dispersion arising from limited consumer awareness of prices. In the bargains and ripoffs paper by Salop and Stiglitz (1977) there is free entry and no lack of competition on the supply side but, depending on the configuration of search costs faced by heterogeneous consumers, a two-price equilibrium can exist in which well-informed consumers buy at the perfectly competitive price while many ignorant consumers, who shop at random, pay a high price – the monopoly price or, if it is lower, the highest price that deters price search by less-informed consumers. (This is consistent with zero-profit equilibrium because firms are assumed to have the same u-shaped average cost curve.) Salop (1977) examines monopolistic price discrimination and resulting *intra*-firm price dispersion.

In the same symposium Butters (1977) analyses equilibrium price dispersion in a model with advertising. There are many consumers and firms, which send ads that randomly inform consumers about the price and availability of products. Each consumer that receives one or more ads buys one unit of the product from the firm advertising the lowest price (provided that's no greater than its reservation price). Butters solves for the equilibrium distribution of prices advertised and transactions prices.

In Varian's (1980) model of sales, which is the other main point of departure for much of what follows, firms are thought of as varying their prices over time – sometimes pricing high, sometimes having a 'sale' and pricing low. So again there is price dispersion, but (unlike in Salop and Stiglitz) it is not the case that firms persistently price high or low. Again there are well-informed and ignorant consumers, with the former buying at the

lowest price in the market and the others buying at random. Thus a consumer effectively sees 'all-or-one' price offer(s). Again there is free entry and so zero profits prevail.[12] Varian solves for the symmetric equilibrium in mixed pricing strategies. Firms sometimes price low to fight for the custom of the informed consumers, but often keep price high and exploit the uninformed who come to them randomly and who are therefore akin to being captive customers. As with Salop and Stiglitz, many consumers get a bad deal despite there being no competition problem on the supply side.

A considerable literature has developed from the Varian model. An important extension is Burdett and Judd's (1983) model of 'noisy' search, in which each consumer that pays a search cost becomes aware of k prices with probability ϕ_k. Thus consumers are *ex post* heterogeneous, some being better informed than others (but not 'all-or-one' as in Varian), while firms are symmetric with respect to each other. At equilibrium, which often involves mixed strategies, the distribution of prices depends on the ϕ_k. If all consumers see at least two prices, so that $\phi_1 = 0$, then we have standard Bertrand perfect competition. If all see just one price, so that $\phi_1 = 1$, there is monopoly pricing despite there being a multiplicity of firms. Otherwise, if $0 < \phi_1 < 1$, there is a dispersed price equilibrium involving mixed strategies. As in the models of Salop and Stiglitz, and Varian, the intensity of competition depends on the structure of consumer awareness rather than on supply-side features of the market.

A straightforward policy implication of this kind of analysis, which appears to be quite robust, is that better consumer awareness is pro-competitive and beneficial. But what about other policy interventions, for example to cap prices or, where firms could potentially set different prices to different consumers, to stop price discrimination? When consumer behaviour, as well as firm behaviour, is endogenous, the effects of such interventions are not immediately obvious. As I will now discuss, models of the kind just sketched, from forty years ago, can be used to illustrate some of the risks and tradeoffs involved. And one of their main limitations – an assumption of symmetry between firms – is in need of relaxation. When that happens, new features of the models, including aspects relevant to policy, can come to light.

[12] The model applies straightforwardly when the number n of firms is fixed.

4 Consumer protection and search incentives

In markets with bargains and ripoffs, two common policy prescriptions are (i) caps to limit the *absolute* level of prices, and (ii) anti-discrimination rules to constrain *relative* prices. In this section I will talk about (i) with reference to the papers by Fershtman and Fishman (1994) and Armstrong, Vickers and Zhou (2009). The subsequent section picks up issue (ii).

While price caps may seem a natural remedy to ripoff problems, there are ways in which they could be anti-competitive and hence counter-productive. First, they could blunt incentives for entry. But although new entry is often positive for imperfect consumers, that is not always so. For example, if it tends to be unaware consumers who are prone to ripoffs, they might well be hard for new entrants to reach. It often seems more plausible that the savvy, aware consumers will be more responsive to offers from new competitors. In that case entry might not be an effective cure to the ripoff problem. Indeed, as we will see in section 6, entry that only the savvy see can even lead to worse consumer outcomes overall. Second, price caps might lessen competition by facilitating tacit collusion, as found by Knittel and Stango (2003) using data from US credit card markets in the 1980s.

Third, and the focus here, is via the effect on incentives for consumer search. Specifically, suppose that each consumer has demand for one unit of the product up to reservation value 1 and that by paying upfront search cost s she gets to see more price offers, so a better chance of a good deal. An informed consumer sees k prices with probability ϕ_k^I, and never sees just one, whereas a less-informed consumer does so with probability ϕ_k^U, where $\phi_1^U > 0$ so that some are captive in the sense of seeing just one price. Let λ denote the proportion of consumers who choose to become informed. In equilibrium with mixed strategies the average price paid by a random consumer – and the expected profit per consumer with production cost normalised to zero – is

$$P = (1-\lambda)\phi_1^U \ . \tag{1}$$

Let the average prices paid by informed and uninformed consumers be P_I and P_U respectively. Because the informed also pay the search cost s, their average total outlay is $P_I + s$, which equals P_U in equilibrium with positive search because

$$s = P_U - P_I \tag{2}$$

As more consumers become informed – i.e. as λ rises – P_I and P_U both fall.

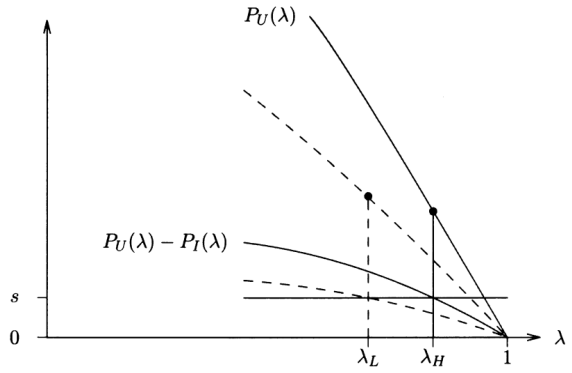

Figure 1: How expected prices paid vary with the proportion of informed consumers

As illustrated in Figure 1, λ_H is the equilibrium proportion of consumers that become informed. In contrast to the economics of the overdraft example, here the uninformed gain from there being more informed consumers: informed consumers generate positive externalities for all consumers. However, and crucially, the average price paid by the informed consumers falls proportionately more than that paid by others as more become informed: the ratio P_I/P_U decreases with λ.[13]

Now consider what a price cap does to average prices. Its direct effect is obviously positive for consumers. Holding λ fixed (or if no-one is searching to begin with) a cap lowers average prices paid. Indeed with price cap $\bar{p} < 1$, condition (1) becomes $P = \bar{p}(1-\lambda)\phi_1^U$, and P_I and P_U reduce by the same proportion, so the ratio P_I/P_U is invariant to the level of the cap for given λ. (In the Figure a price cap pushes the solid curves downward to become the dashed curves.) But a price cap blunts search incentives, so reduces λ (from λ_H to λ_L in the Figure), and the indirect effect goes the other way, as firms have more captive customers and less elastic demand. What is the net effect? If there is still some search once the cap is imposed, the average expenditure of each consumer, including on search costs, is

$$P_U = \frac{s}{1-(P_I/P_U)}$$

given the equilibrium search condition (2). Lower λ means higher (P_I/P_U) and so higher P_U. A price cap makes consumers *worse off* on average. For example, suppose that the uninformed are captive to the single firm that they see, so $\phi_1^U = 1$, and that the informed see two price offers. With $s = \frac{1}{20}$ about 95% of consumers become informed and $P_U \approx \frac{1}{10}$

[13]The proof of this (non-obvious) point used the Cauchy-Schwarz inequality.

under *laissez-faire*. With a price cap of $\frac{1}{2}$, i.e. half the monopoly price, fewer than 75% of consumers become informed, P_U rises by about 70%, and profit doubles. Much further tightening of the cap in this example kills off search altogether, the regime flips from competition to monopoly regulation, and consumers lose unless the cap is much tighter.

This example illustrates the counter-intuitive possibility of price caps raising average prices for reasons that have nothing to do with tacit collusion. Rather, the cap reduces competitive pressure by weakening consumer engagement. In the model as stated, and if distributional issues count for nothing, this is not necessarily a bad thing. Indeed, given inelastic demand, 'total welfare' increases as search is discouraged despite the detriment to consumers, provided that all consumers get served. New effects enter the analysis if, as seems wholly realistic, search costs vary across the consumer population, and the counter-intuitive result may then be overturned. It is still a cautionary tale about consumer policy possibly having effects contrary to pro-consumer intentions.

5 Discrimination against captive consumers

Another way for policy to attempt to protect consumers vulnerable to exploitation is to ban discrimination against captive consumers. Such a policy would not bite on Salop-Stiglitz ripoffs because in their model different firms charged different prices, but there seem to be plenty of examples, including aspects of retail energy pricing as mentioned earlier, where the same firm charges different prices to different consumer groups, and partly on the basis of how 'captive' they are. For example, the recent investigation by the UK CMA (2018) into the practice of 'loyalty penalty pricing' – whereby consumers who stay with their pre-existing suppliers are charged more than new customers or those who renegotiate their deals – found such practices in markets for mobile phone contracts, broadband, household insurance, cash savings, and mortgages, none of which is monopolistic. Such dynamic pricing strategies are an example of behaviour-based price discrimination – see, for example, the analysis of duopoly with advertising by De Nijs (2013).

Policy towards discrimination against captive customers is not straightforward. While bargains-and-ripoffs are bad, ripoffs-and-ripoffs are worse, and as Corts (1998) and Thisse and Vives (1988) showed, price discrimination by imperfectly competitive firms can intensify competition so as to lower prices for *all* consumers. While that cannot happen in models of the limited awareness type with a homogeneous product, there is still a trade-

off: banning discrimination against captive customers is good for them but bad for other customers, and the effect on profits is not immediately clear. In Armstrong and Vickers (2019) we investigate this tradeoff in a simple static duopoly model where the representative consumer has demand $q(p) = -v'(p)$ for the homogeneous product, where $v(p)$ is consumer surplus, and the common, constant unit cost is normalised to zero. Consumers see one firm or both. Specifically, fraction σ_i see firm i (so σ_i is i's 'reach') and fraction $\alpha_i = 1 - \sigma_j$ for $j \neq i$ see only firm i. Label firms so that $\sigma_1 \geq \sigma_2$. We call $\rho_i \equiv \alpha_i/\sigma_i$ firm i's captive-to-reach ratio, and we have that $\rho_1 \geq \rho_2$.

If firms can identify who are their captives, and can discriminate against them, then captives are charged the monopoly price p^*, yielding profit π^*, while there is Bertrand competition for other consumers, who pay the competitive price $p = 0$. If discrimination is banned, or otherwise impossible, there is equilibrium in mixed strategies.[14] Firm 1 must make profit of at least $\alpha_1\pi^*$ in that equilibrium with uniform pricing because it could get that by monopolising its captive consumers. So it will never price such as to get profit per consumer below $\rho_1\pi^*$ because that would yield less profit even if it served all its potential consumers. So firm 2 must also get profit per consumer of at least $\rho_1\pi^*$, and must therefore make profit of at least $\sigma_2\rho_1\pi^* \geq \alpha_2\pi^*$, with strict inequality if $\sigma_1 > \sigma_2$. So total profit is higher – and strictly so unless firms are symmetric – if discrimination is banned.[15] That blunts competition, and effectively allows the smaller firm to shelter under the price umbrella of the larger one.

But what about consumers overall? To address this question a good move is to think in terms of consumer surplus v rather than price p.[16] With discrimination allowed, the $(\alpha_1 + \alpha_2)$ captive consumers get surplus $v^* \equiv v(p^*)$ while the rest get $v_0 \equiv v(0)$. With uniform pricing each consumer faces some probability distribution over v. Take the case where firms are symmetric. Then each gets the same expected profit with uniform pricing as when discrimination is allowed. But profit variation across consumers is greater, indeed maximised, in the latter case. Thus profit is riskier in the mean-preserving spread sense. It follows that if consumer surplus $V(\pi)$ as a function of profit is concave, then consumer overall do worse with discrimination if expected industry profit stays the same. When is $V(\pi)$ concave? The answer is when demand elasticity is increasing in p, which holds

[14] Narasimhan (1988) analysed mixed strategy equilibrium in asymmetric duopoly.

[15] Specifically, the larger firm makes the same expected equilibrium profit, and the smaller firm gains $\sigma_2(\rho_1 - \rho_2)\pi^*$ from a ban on price discrimination.

[16] This follows the competition-in-utility-space approach in Armstrong and Vickers (2001).

widely, especially remembering that cost has been normalised to zero, so p is price net of cost. So surplus variation across consumers is not just an issue if you care about inequality; it relates to consumers on average doing worse.

But with asymmetric firms there is a force in the other direction because expected profit is higher with uniform pricing than with discrimination. So profit is lower, but riskier, with discrimination. So long as the deadweight loss from monopoly is no greater than monopoly profit – i.e. if $(v_0 - v^*) \leq 2\pi^*$, which log-concave demand ensures – then consumers overall do better with discrimination despite the risk effect if firms are asymmetric enough. The comparison of overall welfare $w(\pi) \equiv v(\pi) + \pi$ is similar. Welfare is higher on average with uniform pricing when firms are sufficiently symmetric, but the opposite holds with enough asymmetry.

Supposing that firms can observe which of their customers is captive is stark, but the analysis sketched above can be the basis for analysis of the more general case where firms receive other kinds of information about consumer (un)awareness of competing offers.[17] One can think of an information structure as a partition of consumers into segments, each with its respective captive-to-reach ratios. If the aggregate market is symmetric, then any partition into segments will increase profit if discrimination by segment is allowed, so will be doubly bad for consumers overall. But if the aggregate market is very asymmetric, as when firm 2 has no captive customers at all, industry profit is highest with uniform pricing. An important live research question is how market outcomes relate to information structures – see, for example, Armstrong and Zhou (2019) and Bergemann, Brooks and Morris (2020).

6 Patterns of competitive interaction

Several possible patterns of consumer awareness have been mentioned in the discussion so far, including the 'all-or-one' price awareness in Varian (1980), the more general symmetric analysis of Burdett and Judd (1983), and the asymmetric duopoly model in section 5. But what can be said more broadly about how patterns of consumer awareness influence patterns of competition? Do the features of the special cases just mentioned apply generally,

[17] Fabra and Reguant (2020) analyze a model of search with price discrimination based on the size of customer demand. Customers with high demand have more to gain from search but are therefore willing to engage in search with higher search costs than low-demand customers. Because of these countervailing forces, the relationship between discriminatory prices and customer size need not be monotonic.

or might new and different patterns of price competition emerge? What do entry, exit, and mergers mean for profits and consumers in models that allow for more general patterns of consumer awareness? Such questions are not only of theoretical interest, because, as reported in Armstrong and Vickers (2020), the empirical evidence on patterns of consumer awareness often does not accord with neat symmetric models. And the ordered manner of presentation of internet search results, for example, immediately suggests that in some important contexts with multiple sellers, asymmetry rather than symmetry is the norm.

In our current project, Mark Armstrong and I seek to address these questions, and I will give an overview of it now. We analyse a general model of consideration sets, as follows. There are n price-setting firms supplying a homogeneous product with cost normalised to zero. Consumers are each willing to pay up to 1 for one unit of the good. (So, unlike in the previous section, demand is inelastic.[18]) Each consumer buys from the lowest-price firm that she sees provided that its price does not exceed 1. Consumers differ in their awareness of firms' price offers. Specifically, the fraction that consider offers from subset S of firms is α_S.

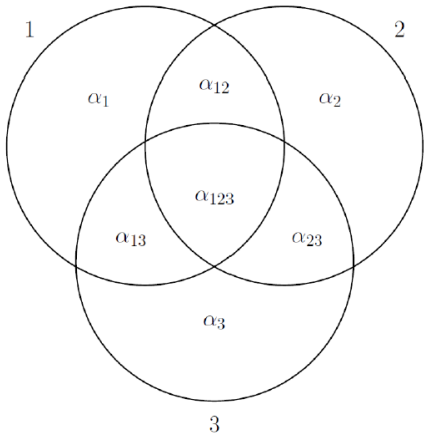

Figure 2: Consideration sets with three firms

The Venn diagram in Figure 2 illustrates a possible pattern of awareness when there are three firms. Thus, for example, α_i are captive to firm i, and as before its captive-to-reach ratio is $\rho_i = \alpha_i/\sigma_i$. We assume that some consumers consider at least two firms (so there is some competition) and that in every subset S of firms at least one has some consumers that consider no other firm in S (so competition is imperfect). It follows that there is no

[18] Except for welfare propositions, results generalise to the case with elastic demand.

equilibrium in pure strategies, but equilibrium exists, so it involves mixed strategies for at least some firms. What can we say about equilibria? How are pricing patterns shaped by patterns of consumer awareness?

Symmetric interactions

Progress in theoretical work often depends on looking at the problem the right way. In our experience with the model just outlined, a good way to proceed is in terms of 'competitive interactions'. First, define σ_S as the fraction of consumers who consider all firms in subset S including along with other firms too. Then define the interaction term

$$\gamma_S \equiv \frac{\sigma_S}{\Pi_{i \in S} \sigma_i}$$

as the fraction of consumers who consider all the firms in S divided by the product of the individual reaches of those firms. The γ_S capture correlations in the reaches of subsets of firms. We always have $\gamma_i = 1$ for a single firm i, and if firms i and j have positively correlated reach, then $\gamma_{ij} > 1$.

We say that interactions are *symmetric* if γ_S is the same for all sets S with the same number of firms. This notion of symmetry allows the γ_S to vary by size of S; for example, there can be more or less interaction between pairs of firms than triples. It also allows firms to differ widely in terms of how many consumers they reach: firms that differ in size can still *interact* symmetrically. All the special cases mentioned above, including asymmetric duopoly, have symmetric interactions. So does the case of 'independent reach' used in some of the literature, where $\sigma_S = \Pi_{i \in S} \sigma_i$ and so all $\gamma_S = 1$. The class of symmetric interactions therefore generalises well beyond those cases: it is a broad class.

With symmetric interactions, equilibrium takes a simple and economically intuitive form given the mild assumption that some consumers consider each pair of firms in isolation (i.e. $\alpha_{ij} > 0$). Label firms in order of increasing reach, so $\sigma_1 \leq \sigma_2 \leq ... \leq \sigma_n$. Then firm n has the highest captive-to-reach ratio ρ_n. In the unique equilibrium firm i uses a mixed strategy with price support $[p_0, p_i]$, where the base price $p_0 = \rho_n$ is the same for all firms, and the highest price p_i used by i is increasing in i, so price supports are nested. Firms with larger reach set stochastically higher prices in that $F_i(p)$ decreases with i for all $p \in [p_0, 1]$. Firm i gets expected profit $\sigma_i \rho_n$. So every firm is a direct competitor of every other firm in the sense that all use prices in a range above p_0. But if reaches differ, smaller firms don't price high. Indeed close to the monopoly price $p = 1$, only the largest two firms operate.

Comparative statics are simple and intuitive too. If a firm exits, the minimum price p_0 rises (e.g. firm n gets more captives if another firm exits) and consumer surplus falls. The same is true if, starting from symmetric interactions, two firms merge – in the sense that the reach of the merged entity is the union of the reaches of the constituent firms – and the merger is profitable.[19]

It turns out, however, that very different things can happen in the less-charted territory of asymmetric interactions.

Asymmetric interactions

A stark example of asymmetric interactions is the pattern of 'nested reach' illustrated in Figure 3. Here, all consumers that consider a given firm also consider all larger firms, so σ_S is the reach of the smallest firm in S. Only the largest firm has any captive consumers. This seems a natural configuration in which to think about situations where consumers consider firms in an ordered pattern (that is common to all consumers), which seems a reasonable approximation to internet search.

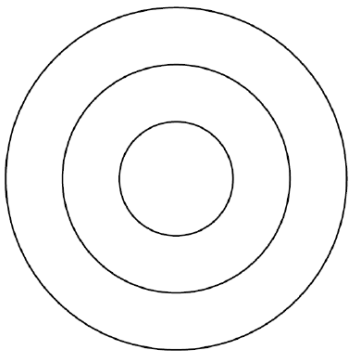

Figure 3: Nested reach

Suppose that firms have distinct reaches, and label firms so that $\sigma_1 < \sigma_2 < ... < \sigma_n$. When there are few firms with sufficiently similar reaches, then equilibrium is as with symmetric interactions: base price $p_0 = \rho_n = (\sigma_n - \sigma_{n-1})/\sigma_n$ is the lowest price in the support of all firms. In general, however, not all firms use the same base price. While there is always direct price competition between neighbouring firms, in the sense that there is a price range used by both i and $(i+1)$, it is quite possible that the *only* direct price competition is between neighbouring firms. Specifically, if there are 'increasing differences'

[19] Assuming that the merged firm sets a single price rather than separate prices for each 'brand'.

in reaches so that $(\sigma_2 - \sigma_1) \leq ... \leq (\sigma_n - \sigma_{n-1})$, then there is an equilibrium of 'overlapping duopoly' form, as follows. There is an increasing sequence of price thresholds $p_0 = p_1 < p_2 < ... < p_{n-1} = p_n = 1$ such that firm i uses price interval $[p_{i-1}, p_{i+1}]$. So prices in the range between p_i and p_{i+1} are used only by firms i and $(i+1)$. Smaller firms price low and larger firms price high; they never compete head-to-head. This is entirely different from the pricing pattern with symmetric interactions.

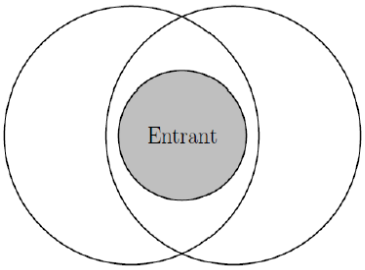

Figure 4: Entry into the contested segment

Whereas the comparative statics of exit and merger were intuitive with symmetric interactions, that is not necessarily so otherwise. Consider entry by a new firm into a market segment that is already contested, as illustrated in Figure 4. This seems quite a natural type of entry to consider if a new firm is most likely to be considered by consumers already aware of competing alternatives. The Figure depicts a pre-entry situation of symmetric duopoly, when each firm's expected profit is equal to it captive portion α_i. Entry into the already-contested area neither expands the total number of consumers that can be served nor reduces the number of those captive to incumbents. So the profit of the incumbents remains undiminished at α_i, while total welfare stays the same. It would be a dominated strategy for an incumbent to price below its captive-to-reach ratio ρ_i and so the entrant can make at least as much profit as that times its own reach. Consumer surplus reduces by an amount equal to the entrant's profit. So, in this example, entry simply shifts surplus from consumers to the entrant.

Here the effect of entry is to blunt competition between the incumbents, who retreat toward their captive bases rather than fighting each other. This leaves the entrant free to set high prices too. The implication is that, in models of this type, the nature of entry matters. Entry that eats into the captive bases of incumbents and which expands the market can be strongly positive for consumers, but not all entry has those characteristics.

This example is but one illustration of the wider issue being explored in this line of work – how does the pattern of consumer consideration of firms determine the equilibrium pricing pattern? Understanding that will provide a basis for further work on search by consumers and advertising by firms to influence patterns of consumer awareness.

7 Concluding comments

Let me conclude with two general comments on public policy towards markets with imperfect consumers in which some get much worse deals than others.

The first is to beware the risks of counter-productive policy intervention. Price caps, bans on discrimination, and promotion of entry can all seem appealing policies, and they will be effective in many circumstances, especially if one puts a higher welfare weight on vulnerable consumers than others. But as the models outlined above illustrate, they can have counter-productive effects when account is taken of endogenous responses, e.g. in consumer search behaviour or on the pattern of competitive interaction between firms. This certainly does not amount to an argument against policy intervention, but it is a real issue for policy design. A task for further work, both theoretical and empirical, is to identify which types of intervention are most likely to be robust – i.e. not prone to unintended consequences. It would be premature to draw strong policy conclusions from simple models of the kinds surveyed here, but they serve to highlight questions that policy makers should consider, such as whether and how restraints on pricing might blunt incentives for consumer search and rivalry between firms. They also underline the importance of diagnosing the root of the problem – the source of consumer, hence market, imperfection – that policy seeks to address.

The second comment is about the appropriate policy tools. The core elements of competition law – aimed at abuse of dominance, anti-competitive agreements, and mergers – are the policy instruments most familiar to industrial economists. But they do not seem well-adapted to issues of consumer imperfection in many markets that appear at least moderately competitive by traditional standards. With sufficient ingenuity, I suppose, it could be argued that those issues are fundamentally competition cases after all. They raise issues about excessive pricing, and even if the suppliers did not have dominant positions in the market as a whole, arguably they did in relation to some consumer groups – e.g. unauthorised overdraft users or consumers that do not consider alternative suppliers. But

I think it would be unduly contrived to seek to extend antitrust law against abuse of dominance to cover such issues, and that doing so would dangerously dilute the meaning of dominance in competition law. Moreover, excessive pricing cases in competition law are rare, and the tests that must be met for an abuse case to succeed are demanding (I would say rightly so).[20] It is important to note that the UK CMA cases mentioned earlier were not abuse of dominance cases but were market investigations under broader competition law powers that led to regulatory remedies.

Finally, when is consumer law, such as the regulations against unfair contract terms, the right approach? The answer to that must depend on the issue being addressed. As I tried to illustrate by way of the two court cases in section 2, questions of major economic importance are being decided under consumer and contract law, but for the most part we as economists have little knowledge of them, still less involvement with them. I hope that some of us will get more engaged. The problems are interesting and important, and we have things to say about them. Of course it is for the courts to say what the law is, but, as antitrust has shown, economics can have a major role in shaping how law and regulation develops, and there is a role to play in the sphere of consumer policy too.

References

Adam, P., M. Grubb, D. Kelly, J.Nieboer and M. Osborne (2018), 'Time to act: a field experiment on overdraft alerts', FCA occasional paper 40.

Armstrong, M. (2008), 'Interactions between competition and consumer policy', *Competition Policy International*, 4, 97-147.

Armstrong, M. (2015), 'Search and ripoff externalities', *Review of Industrial Organization*, 47, 273-302.

Armstrong, M. and J. Vickers (2001), 'Competitive price discrimination', *RAND Journal of Economics*, 32, 1-27.

Armstrong, M. and J. Vickers (2012), 'Consumer protection and contingent charges', *Journal of Economic Literature*, 50, 477-493.

[20] For a recent legal analysis of those tests, see the recent judgment of the UK Court of Appeal in *The Competition and Markets Authority v Flynn Pharma Limited and Pfizer Inc.* [2020] EWCA Civ 339.

Armstrong, M. and J. Vickers (2019), 'Discriminating against captive customers', *American Economic Review Insights*, 1, 257-272.

Armstrong, M. and J. Vickers (2020), 'Patterns of competitive interaction', working paper.

Armstrong, M., J. Vickers, and J. Zhou (2009), 'Consumer protection and the incentive to become informed', *Journal of the European Economic Association*, 7, 399-410.

Armstrong, M., and J. Zhou (2019), 'Information structure and price competition', working paper.

Bergemann, D, B. Brooks and S. Morris (2020), 'Search, information and prices', Cowles Foundation working paper 2224R.

Burdett, K. and K. Judd (1983),'Equilibrium price dispersion', *Econometrica*, 51, 955-969.

Butters, G.(1977), 'Equilibrium distributions of sales and advertising prices', *Review of Economic Studies*, 44, 465-491.

Byrne, D. and L. Martin (2021), 'Consumer search and income inequality', *International Journal of Industrial Organization*, [this issue].

Campbell, J. (2016), 'Restoring rational choice: The challenge of consumer financial regulation', *American Economic Review*, 106(5), 1-30.

Competition and Markets Authority (2016a), *Energy market investigation: Final report*, available via www.gov.uk/cma.

Competition and Markets Authority (2016b), *Retail banking market investigation: Final report*, available viawww.gov.uk/cma.

Competition and Markets Authority (2018), *Tackling the loyalty penalty: Response to a super-complaint made by Citizens Advice on 28 September 2018*, available via www.gov.uk/cma.

Corts, K. (1998), 'Third-degree price discrimination in oligopoly: all-out competition and strategic commitment', *RAND Journal of Economics*, 29, 306-323.

De Nijs, R. (2013). 'Information provision and behaviour-based price discrimination', *Information Economics and Policy*, 25, 32-40.

Fabra, N. and M. Reguant (2020). 'A model of search with price discrimination', *European Economic Review*, 129, 103571.

Fershtman, C. and A. Fishman (1994), 'The 'perverse' effects of wage and price controls in search markets', *European Economic Review*, 38, 1099-1112.

Gabaix, X. and D. Laibson (2006), 'Shrouded attributes, consumer myopia, and information suppression in competitive markets', *Quarterly Journal of Economics*, 121, 505-540.

Knittel, C. and V. Stango (2003), 'Price ceilings as focal points for tacit collusion: evidence from credit cards', *American Economic Review*, 93, 1703-1729.

Narasimhan, C. (1988), 'Competitive promotional strategies', *Journal of Business*, 61, 427-449.

Salop, S. (1977), 'The noisy monopolist: imperfect information, price dispersion and price discrimination', *Review of Economic Studies*, 44, 393-406.

Salop, S. and J. Stiglitz (1977), 'Bargains and ripoffs: a model of monopolistically competitive price dispersion', *Review of Economic Studies*, 44, 493-510.

Spiegler, R. (2011), *Bounded rationality and industrial organization*, Oxford University Press.

Spier, K. and M. Whinston (1995), 'On the efficiency of privately stipulated damages for breach of contract: entry barriers, reliance, and renegotiation', *RAND Journal of Economics*, 26, 180-202.

Thisse, J., and X. Vives (1988), 'On the strategic choice of spatial pricing policy', *American Economic Review*, 78, 122-137.

Varian, H. (1980), 'A model of sales', *American Economic Review*, 70, 651-659.

Vickers, J. (2004), 'Economics for consumer policy', *Proceedings of the British Academy*, 125, 287-310.