Competition and Consumer Confusion

Xavier Gabaix
MIT and NBER

David Laibson
Harvard University and NBER

Current Draft: April 30, 2004*

Abstract

In many markets consumer biases do not affect prices, since competition forces firms to price their products close to marginal cost; competition protects the consumer. We show that noisy consumer product evaluations undermine the force of competition, enabling firms to charge high mark-ups in equilibrium, even in highly competitive environments. We analyze markets in which rational firms sell goods to consumers who evaluate products with noise. Using results from extreme value theory, we show that competition generally has a remarkably weak impact on markups. For normally distributed evaluation noise, we show that markups are proportional to the inverse of $\sqrt{\ln n}$, where n is the number of competitors. In this setting, a highly competitive industry with $n = 1,000,000$ firms will retain 1/3 of the markup of a highly concentrated industry with only $n = 10$ competitors. When we make noise an endogenous variable, we find that firms choose excess noise by making their products inefficiently confusing. Moreover, competition exacerbates this effect: a higher degree of competition causes firms to choose even more excess complexity. Firms with lower intrinsic quality and higher production costs choose the most excess complexity. Educating consumers to reduce their evaluation noise would generate

*For useful suggestions we thank Simon Anderson, Roland Bénabou, Douglas Berheim, Andrew Caplin, Victor Chernozhukov, Casper de Vries, Avinash Dixit, Edward Glaeser, Penny Goldberg, Robert Hall, Sergei Izmalkov, Julie Mortimer, Barry Nalebuff, Aviv Nevo, Nancy Rose, José Scheinkman, Andrei Shleifer, Wei Xiong and seminar participants at Berkeley, Columbia, Harvard, MIT, NBER, New York University, Princeton, Virginia, the 2003 European Econometric Society meeting, the 2003 SITE meeting, and the 2004 AEA meeting. We acknowledge financial support from the NSF (SES-0099025). Gabaix thanks the Russell Sage Foundation for their hospitality during the year 2002-3. Xavier Gabaix: MIT, 50 Memorial Drive, Cambridge, MA 02142, xgabaix@mit.edu. David Laibson: Harvard University, Department of Economics, Cambridge, MA, 02138, dlaibson@arrow.fas.harvard.edu.

large welfare gains. But the gains accrue mostly to the consumer, so firms can't profitably educate consumers and steal them away from competitors. Finally, we introduce an econometric framework that measures bounded rationality and confusion in the marketplace.

JEL classification: D00, D80, L00.

Keywords: bounded rationality, complexity, confusion, extreme value theory, discrete choice, profit, behavioral economics, behavioral industrial organization, mutual fund industry, consumer protection.

1 Introduction

In standard markets competition protects consumers from their own cognitive biases. For example, even if consumers overweight small probability events[1] and overestimate the value of life insurance, the equilibrium price of life insurance will still equal marginal cost. Cognitive errors do not affect the price since competing life insurance companies will undercut each other until price equals marginal cost.

We study a small perturbation to the traditional economic approach and find that it makes competition lose most of this price-cutting force. When consumers have noisy product evaluations, firms have market power that barely decreases as competition rises. We represent a consumer's ex-ante estimate of the value of a good as the sum of the (true) expected consumption value plus evaluation noise. Assuming that consumers have noisy beliefs doesn't seem like a very strong assumption. Mutual fund investors, for instance, don't know the expense ratios of the funds they buy (Alexander et al 1998, Barber et al. 2002). Desktop printer buyers don't know the cost of ink per page (Hall 1997). Wine store customers — like the second author of this paper — don't know the difference between Gamay and Grenache. Wine may be an example of a good with extremely noisy in-store evaluations, but almost every good gets sized up with at least a little noise.

We analyze markets in which there are many perfectly rational firms selling goods to consumers that have noisy product evaluations. The paper can be divided into an analysis of five questions about those markets.

First, we ask whether firms will exploit the noisy consumer evaluations. Following Perloff and Salop (1985), we find that equilibrium markups are proportional to the amount of noise. Higher levels of noise increases the chance that a consumer will either overestimate or underestimate the surplus associated with the firm's good. Firms take advantage of this noise by *raising* their prices. Such price increases reflect the fact that noise reduces the sensitivity of consumers to small differences in product attributes. This in turn reduces the elasticity of each firm's demand curve, leading firms to raise equilibrium prices.

Second, we ask how increased competition affects markups. Using results from extreme value theory, we find that competition typically has remarkably little impact on markups. Our leading

[1] See Kahneman and Tversky (1979).

example is the case of normally distributed noise. For this case, we show that markups are proportional to $\left(\sqrt{\ln n}\right)^{-1}$, where n is the number of competitors.[2] A highly competitive industry with $n = 1,000,000$ firms will have 1/3 the markup of an industry with only $n = 10$ competitors. When consumers have typical (thin-tailed) noise distributions, competition — even extreme competition — barely reduces markups. Moreover, we show that competition actually *increases* markups when consumers have fat-tailed distributions.

Third, we ask how firms will manipulate the amount of noise when they are able to do so. In this analysis, complexity is itself an endogenous variable chosen by each firm. For example, a firm can create an unnecessarily complex fee schedule, which makes it harder for a consumer to determine the true cost of the good. We show that firms will generally prefer such excess complexity. A small amount of excess complexity has only a second-order negative impact on the intrinsic quality of the good, but generates a first order increase in the (confusion-driven) demand for the good. So firms choose inefficiently high levels of complexity.

Fourth, we ask what determines a firm's choice of excess complexity. We show that higher levels of competition *increase* the equilibrium amount of excess complexity. Firms in highly competitive markets have small market shares and have more to gain from excess complexity. We also show that firms with higher intrinsic quality and lower production costs choose less excess complexity. Intuitively, high quality firms maximize profits by making their competitive advantage relatively transparent (i.e., they reduce noise). By contrast, average or high cost firms will pick a high degree of complexity, maximizing profits by taking advantage of the fact that their over-priced product will be misevaluated by some fraction of consumers.

Fifth, we ask whether firms have an incentive to educate consumers and thereby turn naive consumers (with noisy product evaluations) into sophisticated consumers (with less noisy evaluations). We show that such incentives are quite weak, since sophisticated consumers are much less profitable to firms than naive consumers. The large gains from education disproportionately accrue to the consumer, generating a wedge that makes it impossible for firms to profitably educate consumers and thereby steal them away from other firms.

We also introduce an econometric framework that can be used to measure bounded rationality

[2] Hence, mark-ups converge very slowly to zero with n. This is by contrast with the Cournot model in which markets are proportional to $1/n$.

and confusion in the marketplace. The model exploits the fact that populations of consumers with *identical* underlying objective functions should consume similar bundles of goods. When otherwise identical sophisticated and naive consumers buy different bundles of goods, then the naive consumers suffer from some confusion about the goods they are buying. We introduce an econometric framework that can measure such effects by exploiting randomized educational interventions.

The rest of this paper formalizes these claims. Section 2 describes the general model where consumers have noisy signals about product value and firms can control the level of noise. Section 3 shows that the existence of noise increases firms' market power and that competition does remarkably little to offset these effects. Section 4 shows that firms have an incentive to generate excess complexity, and that excess complexity increases with the amount of competition and decreases with product quality. Section 5 explains why firms won't educate consumers. Section 6 describes a practical econometric framework for measuring the magnitude of confusion. Section 7 concludes.

Before proceeding, we start with a review of the literature. We show how the market power coming from bounded rationality differs from market power coming from search costs and heterogeneous tastes, both for positive and normative analysis. Of course, all three (and more) sources of market power coexist in most markets.

Literature Review Since its inception in the late 1970's, psychological principles have been applied in every traditional field of economics. Our paper contributes to the emergent field of behavioral industrial organization[3] (DellaVigna and Malmendier 2003 and Oster and Morton 2004 apply hyperbolic discounting; Heidhues and Koszegi 2004 and Koszegi and Rabin 2004 apply loss aversion; Gabaix and Laibson 2004 and Spiegler 2003 apply boundedly rational heuristics).

Our paper is partially motivated by empirical work that suggests that consumers are sometimes confused about the decisions that they make. Woodward (2003) documents confusion in the mortgage market, which decreases with the level of household education. Madrian and Shea (2002) and Choi et al (2003a, 2003b) show that workers are extraordinarily sensitive to the defaults in their 401(k) plan (e.g., automatic enrollment), suggesting that consumers do not have a clear model of

[3] The first papers in "behavioral IO" may be Hausman (1979) and Hausman and Joskow (1982), who find that for durables consumers seem to care more about upfront costs that future flow costs.

how to invest for retirement. Benartzi and Thaler (2001, 2002) show that investors allocate their financial assets noisily. For example, Benartzi and Thaler (2002) find that the 401(k) investors choose a mix of stocks and bonds that is driven by the proportion of stock funds available in the investor's 401(k) plan. Asset allocation appears to derive at least partially from randomization over the set of available funds.

Our model is motivated by the standard Luce (1959)-McFadden (1981) random utility framework, (see Anderson, de Palma and Thisse 1992 for a review of this literature and Sheshinski 2003 for a recent implementation). We extend the markup calculations of Perloff and Salop (1985). Like the previous literature, our analysis uses extreme value theory. Analytic markup calculations for Gumbel noise (Anderson et al 1992) and exponential noise (Perloff and Salop 1985) were already known. We derive asymptotic approximations for a much wider class of distributions, including Gaussian, exponential, log-normal, and power-law densities. Our results for endogenous noise are also original. We extend previous analyses interpreting this noise as a form of bounded rationality (Anderson et al. 1998, McFadden 1981, De Palma et al. 1994, Sheshinksi 2003). A related IO literature analyzes rational (Bayesian) consumers who have noisy product evaluations (Judd and Riordan 1994 and Anantham and Ben-Shoham 2004).

Our work is related to the literature on advertising (e.g. Becker and Murphy 1990, Dixit and Norman 1978; see Bagwell 2002 for a remarkable review of the advertising literature). We believe that advertising and marketing play a key role in generating the noise in consumer product evaluations. Consumers have a hard time filtering misleading marketing signals about goods. Imperfect (but unbiased) filters will create noisy impressions about product value.

Our paper is also related to the literature on search (e.g., Stigler 1961, Diamond 1971, Salop and Stiglitz 1977). Like search models, our model predicts that consumers will not always purchase the most competitively priced good. However, our framework has little else in common with the search framework. Our model has a different microfoundation and explores different phenomena: e.g., endogenous market power arising from endogenous noise.

2 General Model with Complexity

2.1 Consumers

Our model is motivated by the standard Luce (1959)-McFadden (1981) random utility framework.[4] Each consumer must pick one good from a set of n goods. For consumer a, good i has complexity σ_i, value v_i, and price p_i. Consumers do not directly observe either σ_i, v_i, or p_i. Instead, consumer (agent) a observes only a noisy signal of good i's net value, where the noise component scales with the complexity of the good.

$$U_{ia} = \underbrace{v_i - p_i}_{\text{true value}} + \underbrace{\sigma_i \varepsilon_{ia}}_{\text{noise}}.$$

The scaled noise term, $\sigma_i \varepsilon_{ia}$, captures consumer a's noise in evaluating product i. We assume that ε_{ia} is zero mean[5] and i.i.d. across consumers and goods. Let ε_{ia} have unit variance, density f, cumulative distribution F, and 'complexity' scaling factor σ_i.

These assumptions imply that before consumers purchase a good, the consumers do not know the true expected value or the true expected cost of the good. This imperfect observability may arise for many possible reasons. Marketing campaigns create noise. Mental simulations of the future use value create noise. Indeed, all complex mental calculations are associated with error. Costs are also perceived with noise, since the sticker price is often not the end of the cost story. Many goods have complex repair costs or other add-on costs. Durables have complex financing arrangements.

We summarize the consumer's noisy evaluations of both value and price by assuming that consumer a only observes a utility signal U_{ia} for each good and does not observe its constituent components. We assume that the consumer uses a very simple and sensible decision-rule: pick the good with the highest signal value. So consumer a chooses the firm i with the highest value of U_{ia}.[6] In our baseline model, this sensible heuristic rule will also be an optimal policy.[7]

[4] See Anderson, de Palma and Thisse (1992) for an excellent review of this literature.

[5] In the leading cases, the mean of the noise does not matter for the equilibrium. In particular, the mean does not matter when firms have identical noise intensity, σ_i.

[6] We assume that the consumer must have (and will buy) exactly one good, even if the largest U_{ia} is negative for consumer a.

[7] Appendix C offers a class of situations where this is the optimal rule, and discusses how the conclusions of our model change in other cases.

To simplify notation, we suppress the consumer specific subscript a for the rest of the paper, so $U_i \equiv U_{ia}$ and $\varepsilon_i \equiv \varepsilon_{ia}$ unless otherwise noted.

2.2 Firms

Each firm needs to pick an endogenous price, p_i, and a level of product complexity, σ_i. We assume that changes in complexity, σ_i, have two effects. Product complexity influences the underlying intrinsic valuation of the product, so $v_i = v(\sigma_i)$. Second, product complexity influences the standard deviation of the noise that will be perceived by the consumers (cf. subsection 2.1).

If a social planner designed goods and assigned them to consumers, such efficient products would typically have a level of complexity $\sigma > 0$. For example, an efficient computer will trade off the costs of complexity — e.g. "It's so complex that I can't figure out how to use it." — with the benefits of complexity — e.g. "The computer can be used to do many different things." If a product is too complex it will be inefficiently too hard to use, but if a product is too simple it will have inefficiently too few features.

We capture these trade-offs by assuming that complexity σ gives rise to a hump-shaped valuation function $v(\sigma)$. Figure 1 presents an example of such a function. There is a maximum at $\sigma^* \geq 0$, the "bliss point" for complexity.

We study the Bertrand equilibrium with endogenous complexity, where firms maximize profit, π_i, by choosing (p_i, σ_i):

$$\max_{p_i, \sigma_i} \pi_i \equiv (p_i - c)\, \mathcal{D}(p_i,\, \sigma_i)\,, \tag{1}$$

where c is the marginal cost of production and $\mathcal{D}$ is the firm's demand function. In a symmetric equilibrium, the demand function of firm i is equal to the probability that a consumer receives the best noisy signal from firm i, so

$$\begin{aligned}
\mathcal{D}(p_i, \sigma_i) &= P\left(v(\sigma_i) - p_i + \sigma_i \varepsilon_i > \max_{j \neq i} v(\sigma) - p + \sigma \varepsilon_j \right) \\
&= P\left(v(\sigma_i) - p_i - [v(\sigma) - p] + \sigma_i \varepsilon_i > \max_{j \neq i} \sigma \varepsilon_j \right).
\end{aligned}$$

It is convenient to rewrite the firm's maximization problem by introducing a change of variables. Define a new demand function D that takes as its first argument the intrinsic surplus x of firm i

relative to its competitors,

$$D(x,\sigma_i) \equiv P\left(x + \sigma_i \varepsilon_i > \max_{j \neq i} \sigma \varepsilon_j\right). \tag{2}$$

Define

$$x_i \equiv v(\sigma_i) - p_i - [v(\sigma) - p]. \tag{3}$$

Now the firm's optimization problem may be rewritten

$$\max_{x_i, \sigma_i} (p_i(x_i, \sigma_i) - c)\, D(x_i,\, \sigma_i),$$

where $p_i(x_i, \sigma_i)$ is defined by rearranging equation (3),

$$p_i(x_i, \sigma_i) \equiv v(\sigma_i) - x_i - [v(\sigma) - p]. \tag{4}$$

Call

$$M_{n-1} = \max_{j \in \{1,\ldots,n\}, j \neq i} \varepsilon_i, \tag{5}$$

so M_{n-1} is the highest of $n-1$ noise realizations. Then,

$$\begin{aligned} D(x,\sigma_i) &= P\left(\varepsilon_i > \frac{-x + \sigma M_{n-1}}{\sigma_i}\right) \\ D(x,\sigma_i) &= E\left[\overline{F}\left(\frac{-x + \sigma M_{n-1}}{\sigma_i}\right)\right], \end{aligned} \tag{6}$$

where $\overline{F}(x) = \int_x^\infty f(y)\, dy$ is the countercumulative distribution function. This formulation emphasizes the property that the demand for good i is driven by the right-hand tail properties of the countercumulative distribution function, $\overline{F}$.

The properties of the symmetric equilibrium[8] can be derived from the behavior of $D(x, \sigma_i)$ at $(x, \sigma_i) = (0, \sigma)$. Specifically, (6) gives:

[8] Section 9.9 treats the existence of the symmetrical equilibrium.

$$
\begin{aligned}
\frac{\partial}{\partial x} D(0,\sigma) &= \frac{1}{\sigma} E\left[f\left(M_{n-1}\right)\right] && (7) \\
\frac{\partial}{\partial \sigma_i} D(0,\sigma) &= \frac{1}{\sigma} E\left[f\left(M_{n-1}\right) M_{n-1}\right] && (8) \\
D(0,\sigma) &= \frac{1}{n}. && (9)
\end{aligned}
$$

Proposition 1 *In a symmetric Bertrand equilibrium,*

$$
\begin{aligned}
p - c &= \frac{1}{nE\left[f\left(M_{n-1}\right)\right]} \sigma && (10) \\
v'(\sigma) &= -\frac{E\left[f\left(M_{n-1}\right) M_{n-1}\right]}{E\left[f\left(M_{n-1}\right)\right]}, && (11)
\end{aligned}
$$

where M_{n-1} is a random variable with cumulative density function $P\left(M_{n-1} \leq x\right) = F(x)^{n-1}$.

These results generalize the findings in Perloff and Salop (1985), who consider the case in which σ_i is not a choice variable. It can be shown that our markup equation is equivalent to their markup equation (for fixed σ). We write the markup equation differently than they do to anticipate the application of some asymptotic approximations from extreme value theory. These approximations give the model a wide scope of applicability, by yielding analytic results for the leading classes of distributions.

Before proceeding with a formal proposition, we first explain the intuition for our result. We begin by characterizing the right-hand tail of the noise distribution. Recall that M_{n-1} is the maximum value of $n-1$ draws. First, we observe that $E\left[\overline{F}\left(M_{n-1}\right)\right] = 1/n$. On average there is a $1/n$ chance of drawing a noise realization that dominates the largest element in a random set of $n-1$ noise realizations. This suggest that if we define

$$
A_n \equiv \overline{F}^{-1}(1/n) \tag{12}
$$

M_{n-1} will be close to A_n.

Given the mean of $\overline{F}\left(M_{n-1}\right)$ is $1/n$, one can usefully rewrite $\overline{F}\left(M_{n-1}\right) = u/n$ for u a random variable near 1. We therefore expand $M_{n-1} = \overline{F}^{-1}\left(\frac{u}{n}\right)$ around $u = 1$.

$$\begin{aligned} M_{n-1} &= \overline{F}^{-1}\left(\frac{u}{n}\right) \simeq \overline{F}^{-1}\left(\frac{1}{n}\right) + \left(\overline{F}^{-1}\right)'\left(\frac{1}{n}\right)\frac{u-1}{n} \\ &= A_n - \frac{1}{f(A_n)}\frac{u-1}{n}. \end{aligned}$$

So the dispersion in M_{n-1} is the dispersion of $\frac{1}{f(A_n)}\frac{u-1}{n}$, which implies that variation in M_{n-1} is proportional to $1/[nf(A_n)]$. We conclude that the typical difference between the best draw and the second best draw is proportional to $1/[nf(A_n)]$. Using standard optimization arguments, a firm sets its markup proportional to this dispersion, so that $p - c \sim 1/[nf(A_n)]$.

The following Proposition shows that this heuristic argument generates the right approximation for the Gaussian, exponential, Gumbel and lognormal distributions. The Proposition also shows that the approximation remains accurate up to a corrective constant $\Gamma(2+\xi)$ in other cases.

Proposition 2 *In a symmetric Bertrand equilibrium:*

$$p - c \sim \frac{1}{nf(A_n)\Gamma(2+\xi)}\sigma, \tag{13}$$

$$v'(\sigma) \sim -\frac{A_n}{\Gamma(2+\min(\xi,0))}. \tag{14}$$

where A_n satisfies $P(\varepsilon \geq A_n) = 1/n$, $\Gamma(z) = \int_0^\infty t^{z-1}e^{-t}dt$ is the Gamma function, and ξ is the characteristic index of the distribution (see Appendix A for a definition). Table 1 presents values of A_n and ξ for many distributions.

Proof. See Appendix B. ■

This final proposition yields very useful formulae, since the key mathematical objects, A_n, $f(A_n)$, and ξ are easy to calculate for most distributions of interest. It is useful to remember $\Gamma(1) = \Gamma(2) = 1$.

2.3 Distributions

To analyze the impact of competition on markups, we examine the equilibrium markup for various noise distributions. It is useful to consider seven well-studied analytically tractable distributions.

First, we consider the case in which ε is uniformly distributed between -1 and 1,

$$f_{\text{Uniform}}(\varepsilon) = \frac{1}{2} 1_{|\varepsilon|<1}. \tag{15}$$

which generalizes to a density in $[-1, 1]$ that is power law around $\varepsilon = 1^-$,

$$f_{\text{Bounded power law}}(\varepsilon) \sim \alpha k^{-\alpha} (1-\varepsilon)^{\alpha-1}, \tag{16}$$

with $\alpha > 0$. For a large number of firms, only the right tail matters. So it is enough to characterize the behavior of the density near the right boundary.

We also consider the Gaussian density,

$$f_{\text{Gaussian}}(\varepsilon) = \frac{1}{\sqrt{2\pi}} e^{-\varepsilon^2/2}, \tag{17}$$

the Gumbel density (where $\theta \simeq 0.577216$ is Euler's constant),

$$f_{\text{Gumbel}}(\varepsilon) = \exp\left(-e^{-\varepsilon-\theta} - \varepsilon - \theta\right), \tag{18}$$

the exponential density,

$$f_{\text{Exponential}}(\varepsilon) = e^{-(\varepsilon+1)} 1_{\varepsilon>-1}, \tag{19}$$

the log-normal density,

$$f_{\text{Lognormal}}(\varepsilon) = \frac{1}{(\varepsilon + \sqrt{e})\sqrt{2\pi}} e^{-\ln\left(\varepsilon+\sqrt{e}\right)^2/2} 1_{\varepsilon>-\sqrt{e}}, \tag{20}$$

and the power law density for large ε,

$$f_{\text{Power law } \zeta}(\varepsilon) \sim \zeta k^{\zeta} \varepsilon^{-\zeta-1}. \tag{21}$$

The shift factors θ, 1, and $\sqrt{e}$ ensure that the mean of ε is 0 in each distribution. The densities are ranked from thinnest to fattest tails.[9]

The reader may be uncomfortable with the use of unbounded noise distributions. But we

[9] A density g has fatter tails than a density f if there is a positive constant D such that for all x above a certain threshold $f(x) \leq Dg(x)$.

use unbounded distributions only for analytical convenience, not because one needs to assume that distributions are truly unbounded.[10] Indeed, the results that follow will still hold if the distribution of the noise is truncated on the right by some upper bound.[11] We only fundamentally need to evaluate the behavior of the noise density in the part of the right-hand-tail with cumulative probability $1/n$.

We calculate the Bertrand outcome for the seven distributions discussed above. Some of our calculations are asymptotic expansions, which hold for large n and small positive t. Table 1 reports values for the key ingredients in our calculations.[12] In this table, f is the density, $\overline{F}(x) \equiv \int_x^\infty f(y)\,dy$ is the countercumulative function, $A_n \equiv \overline{F}^{-1}(1/n)$, $h(t) \equiv f\left(\overline{F}^{-1}(t)\right)$, and ξ is the characteristic index of F (i.e., an index of the fatness of the distribution, see Appendix A). For application of Proposition (2), note that $f(A_n) = h(1/n)$.

Table 1: Distributions and Associated Functions.

	$A_n \equiv \overline{F}^{-1}(1/n)$	$h(t) \equiv f\left(\overline{F}^{-1}(t)\right)$	ξ
Uniform	$1 - 2/n$	$1/2$	-1
Bounded power law	$1 - kn^{-1/\alpha} + o\left(n^{-1/\alpha}\right)$	$\sim \alpha k^{-1} t^{1-1/\alpha}$	$-1/\alpha$
Gaussian	$\sim \sqrt{2\ln n}$	$\sim t\sqrt{2\ln\frac{1}{t}}$	0
Gumbel	$\sim \ln n$	$\sim t$	0
Exponential	$\ln n - 1$	t	0
Lognormal	$\sim e^{\sqrt{2\ln n}}$	$\sim te^{-\sqrt{2\ln\frac{1}{t}+\frac{1}{2}\ln\left(2\ln\frac{1}{t}\right)}}$	0
Power law	$\sim kn^{1/\zeta}$	$\sim \zeta k^{-1} t^{1+1/\zeta}$	$1/\zeta$

Key quantities for Proposition 2.

[10] The same issues arise when economists model GDP growth as a Gaussian variable.
[11] See Appendix B.
[12] The proof is a consequence of e.g. Embrechts et al. (1997, p.155-7) and simple calculations.

3 Toothless competition

3.1 Will Competition Protect Consumers?

We now answer our second question: When consumers are confused, how do markups respond to intensified competition? Naturally, the answer to this question depends on the distribution of the noise in consumer evaluations. We assume that the standard deviation of noise, σ, is fixed and focus analysis on endogenous markups and eventually endogenous entry. Proposition 3 provides closed form expressions for the markups in different distributional cases for fixed σ and a fixed number of competitors, n.

Proposition 3 *The Bertrand equilibrium generates the following markups. For uniform noise (15),*

$$p - c = \frac{2}{n}\sigma. \tag{22}$$

For bounded power law noise (16) with $\alpha > 1/2$,

$$p - c \sim \frac{k}{\alpha\Gamma\left(2 - 1/\alpha\right)} n^{-1/\alpha}\sigma. \tag{23}$$

For Gaussian noise (17),

$$p - c \sim \frac{1}{\sqrt{2\ln n}}\sigma. \tag{24}$$

For Gumbel noise (18),

$$p - c = \frac{n}{n-1}\sigma. \tag{25}$$

For exponential noise (19),

$$p - c = \sigma. \tag{26}$$

For log-normal noise (20),

$$p - c \sim e^{\sqrt{2\ln n} - \frac{1}{2}\ln(2\ln n)}\sigma. \tag{27}$$

For power-law noise (21) with exponent $\zeta > 1$,

$$p - c \sim \frac{k}{\zeta\Gamma\left(2 + 1/\zeta\right)} n^{1/\zeta}\sigma. \tag{28}$$

Proof. To obtain exact results we use equation (10) in Proposition 1. For approximate results we use equation (13) in Proposition 2. We also exploit the distributional statistics in Table 1. ■

The distributions in Proposition 3 are presented in increasing order of fatness of the tails[13]. For the uniform distribution, which has the thinnest tail, the markup falls relatively rapidly as the number of competitors, n, increases: the markup is proportional to $1/n$.

For the distributions with the fattest tails, the markups paradoxically[14] *rise* as the number of competitors *increases*.[15] Markups rise since the price elasticity *falls* as n gets large. Intuitively, for fat tailed noise, as n increases, the difference between the best draw and the second best draw, which is proportional to $1/\left[nf\left(A_n\right)\right]$, increases with n. However, even though markups rise with n, profits per firm go to zero since firm prices increase with $n^{1/\zeta}$ but sales per firm are proportional to $1/n$. We do not yet know whether the fat-tailed case is empirically relevant. We speculate that it might apply in markets with fat tailed distribution of sales – for instance, the book market.[16] We describe a way to test for distributional form in section 6.2.

Thin-tailed distributions (e.g., uniform) and fat-tailed distributions (e.g., power-laws) are the extreme cases in Proposition 3. Most of the distributional cases imply that competition typically has remarkably *little* impact on markups. For instance with Gaussian noise, the markup, $p-c$, is proportional to $1/\sqrt{\ln n}$. So $p-c$ converges to 0, but this convergence proceeds at a glacial pace. To illustrate this fact, we normalize the markup at $n=10$ to be 1 and calculate the markup as the number of competitors expands by multiple factors of 10. Table 2 shows that a highly competitive industry with $n=1,000,000$ firms will retain 1/3 of the markup of a highly concentrated industry with only $n=10$ competitors.

[13] Additionally, in the case $F(\varepsilon) = \left(1-\varepsilon^{-3}\right) 1_{\varepsilon>1}$, one gets the closed form $p-c = \Gamma(n+2)\left[3n(n-1)\Gamma(7/3)\Gamma(n-1/3)\right]^{-1}\sigma$.

[14] See Bénabou and Gertner (1993), Rosenthal (1980), Spector (2002) for paradoxes along lines very different from ours.

[15] In this rather perverse scenario, consumer surplus goes to negative infinity as $n \to \infty$, a limit result that only arises because we made the simplifying assumption that consumers must buy one good. Three immediate fixes would eliminate this perverse result. First, one could assume that consumers do not buy any good at all if their best signal, U_i, is not sufficiently positive. Second, one could assume that consumers only sample a finite number of goods, which effectively bounds n. Third, once one endogenizes entry, n won't go to infinity because any positive fixed entry cost will eventually swamp firm profits for large enough n.

[16] See Chevalier and Goolsbee (2004) and Sornette et al. (2003). Movies (De Vany 2004) also have power law distributions. This is a general property for markets where word of mouth creates snowballing effects and power laws (Simon 1955, Gabaix 1999, and the survey in Gabaix and Ioannides 2003). If consumers base their book choice on the popularity of a book, then the noise may be power law distributed.

Table 2: Mark-ups with Gaussian noise as a function of the number of competitors, n.

n	Markup
10	1.00
100	0.61
$1,000$	0.48
$10,000$	0.40
$100,000$	0.35
$1,000,000$	0.32

We normalize the markup for $n = 10$. We integrate numerically Eq. (10). The asymptotic result (24) provides a good approximation for these exact results.

In cases with moderate fatness, such as the Gumbel, exponential, and log-normal densities, the markup again shows little (or no) response to changes in n. Finally, we do not need an infinite support to generate such results. In the case of *bounded* power law noise (23), the decay is slow when α is large: the markup is proportional to $1/n^{1/\alpha}$.

In practical terms, these results imply that in markets with noise, we should not expect increased competition to dramatically reduce markups. The mutual fund industry exemplifies such stickiness. Currently 10,000 mutual funds are currently available in the U.S., and many of these funds offer very similar portfolios. Even in a narrow class of fairly homogenous products, such as medium capitalization value stocks or S&P 500 index funds (Hortacsu and Syverson 2003), it is normal to find 100 or more competing funds. Despite the large number of competitors in such sub-markets, mutual funds still charge high annual fees, often more than 1% of assets under management. Most interestingly, these fees have not fallen as the number of competing funds has increased by a factor of 100 over the past several decades.

In a classical industrial organization model with n firms, the Cournot model, the markup is proportional to $1/n$. This suggests that competition lowers markups very fast. Proposition 3 implies, in contrast, that when consumers have noisy evaluations the markup is likely to decrease much more slowly with n — for instance in $1/\sqrt{\ln n}$ in the Gaussian case. In our interpretation, most consumers are not financially savvy and find the choice of mutual funds confusing. Noisy evaluations make it difficult to choose among mutual funds. Because of this noise, fees remain stubbornly high, even with 100 mutual funds in a homogeneous market. This is consistent with equation (26), which implies that in the case of exponential noise markups do not change with the number of firms. More generally equations (24)-(27) imply that markups will be only weakly sensitive to the number of firms. Even when the evaluation noise is bounded, the noise can generate a markup that decreases very slowly with n, as illustrated by the case with large α in equation (23).

The next two subsections extend analysis of the equilibrium with exogenous noise. First we endogenize the number of firms by considering entry (subsection 3.2) and then we consider the impact of multiple noise sources (subsection 3.3). Readers who wish to skip these extensions can proceed without loss of continuity to Section 4.

3.2 Extension: Endogenous Entry

We endogenize the number of firms (i.e., goods) in the industry, by assuming that firms pay a fixed cost C of competing in an industry (in addition to marginal cost c for producing incremental units). The marginal profit per good is $p - c$. Entry will endogenously determine the largest number of firms n so that profit per firm is non-negative.

$$\Pi_n = \frac{p-c}{n} - C \geq 0.$$

Ignoring integer constraints, the free entry condition implies $\Pi_n = 0$.

The following Proposition describes the impact of confusion σ and entry costs C on the number of firms and consumer welfare. We will use the notation

$$b_n = \frac{1}{nE\left[f\left(M_{n-1}\right)\right]}, \tag{29}$$

so that by Proposition 1 the markup is $p - c = b_n \sigma$.

Proposition 4 *Suppose that b_n/n is decreasing in n, and that firms enter the market until the zero profit condition binds. Then the number n of firms is decreasing in C/σ. Consumer welfare* *decreases in σ. Consumer welfare increases in the entry cost C iff b_n is a decreasing function.*

Proof. The zero profit condition $\Pi_n = 0$ implies

$$b_n/n = C/\sigma, \tag{30}$$

which proves that n decreases in C/σ. Consumer welfare decreases in $p-c$, and $p-c = b_n\sigma = nC$. So markups rise with σ (since n rises with σ) and welfare falls with σ. Finally, $p-c = b_n\sigma$, so consumer welfare decreases in C iff b_n is a decreasing function. ■

Intuitively, as complexity, σ, increases, markups increase, consumer surplus falls, producer rents increase, and n increases. As C increases, n falls and producer rents increase. But the effect of C on consumer welfare depends on how b_n varies with n. If $b_n' < 0$, then markups increase and consumer surplus falls with C. If $b_n' > 0$, then markups will fall and consumer surplus will rise with C. The sensitivity of the effect of C on markups and welfare is explored in Proposition 5.

Proposition 5 *For our noise distributions (15)–(21), the limiting equilibrium markups ($C/\sigma \to 0$) are given by,*

$$\begin{aligned}
(p-c)_{Uniform} &= \sqrt{2\sigma C} \\
(p-c)_{Bounded\ power\ law} &= \left(\tfrac{k}{\alpha\Gamma(2-1/\alpha)}\right)^{\alpha/(1+\alpha)} C^{1/(1+\alpha)}\sigma^{\alpha/(1+\alpha)} \\
(p-c)_{Gaussian} &\sim \tfrac{\sigma}{\sqrt{\ln 2\sigma/C}} \\
(p-c)_{Gumbel} &= \sigma + C \\
(p-c)_{Exponential} &= \sigma \\
(p-c)_{Lognormal} &\sim \sigma e^{\sqrt{2\ln\frac{\sigma}{C}} - \frac{1}{2}\ln\left(2\ln\frac{\sigma}{C}\right)+\frac{1}{2}} \\
(p-c)_{Power\ law\ \zeta} &\sim \left(\tfrac{k}{\zeta\Gamma(2+1/\zeta)}\right)^{\zeta/(\zeta-1)} C^{-1/(\zeta-1)}\sigma^{\zeta/(\zeta-1)}
\end{aligned}$$

and the number of firms is $n = (p-c)/C$.

Only for densities with thinner tails than the exponential does a decrease in entry costs reduce markups and therefore increase consumer welfare (cf. Proposition 4). For the exponential distribution, a decrease in entry costs has no effect on markups. For distributions fatter than the exponential (i.e., log normal and power law), a decrease in entry costs *raises* markups and *reduces* consumer welfare for the same reason that greater competition (n) raises markups in Proposition 4.

For densities at neither extreme of thin tails or fat tails (i.e. Gaussian, Gumbel, exponential, and lognormal), welfare and markups change very little with entry costs, C. The elasticity of $p-c$ with respect to C is asymptotically zero for these distributions. By contrast, the number of firms, $n = (p-c)/C$, is very sensitive to entry cost, with asymptotic elasticity of -1. So the bulk of our analysis implies that while low entry costs generate a great deal of competition, low entry costs do not necessarily generate low markups or high consumer surplus.

3.3 Extension: Principle of Maximum Fatness

In general, we expect the noise term $\sigma\varepsilon$ to be the sum of many components. For instance, one might have $\sigma\varepsilon = \sigma_u u + \sigma_v v$ with u and v independent, mean zero, unit variance random variables. One might guess that the equilibrium markup should depend on both σ_u and σ_v and that $p-c$ should be proportional to $\sqrt{\sigma_u^2 + \sigma_v^2}$. Appendix B shows that this is not the case. For instance, if u and v are exponentially distributed and σ_u and σ_v are non-negative, then the markup is: $p - c = \max(\sigma_u, \sigma_v)$. Intuitively, only the fatter-tailed noise component matters. This principle of maximum fatness simplifies analysis.[17] If there are several sources of noise, one only needs to track the noise with the fattest tails, a point that we make more generally in Appendix B.

4 The Supply of Confusion

We return to the general case in which firms choose both the markup, $p-c$, and the confusion variable, σ. For simplicity, we continue to consider the case of symmetric technologies and symmetric equilibria. These symmetry assumptions will be in force until we relax them in subsection 4.2. We begin by showing that firms will choose to make their products excessively complex in equilibrium.

Recall that Proposition 1 characterizes the equilibrium complexity of the good

$$v'(\sigma) = -\kappa_n,$$

[17] This property that the fattest variable domains is the key simplifying fact in theories of extreme movements. For more rules of that type, see Gabaix et al. (2004), Appendix A.

where

$$\kappa_n \equiv \frac{E\left[f\left(M_{n-1}\right)M_{n-1}\right]}{E\left[f\left(M_{n-1}\right)\right]}. \tag{31}$$

Proposition 6 *If the distribution of noise is symmetric and $n > 2$, then $\kappa_n > 0$.*

Proposition 7 *If $\kappa_n > 0$, then in the symmetric Nash equilibrium σ is strictly greater than σ^*, the bliss point for complexity.*

Proof. Since $v'(\sigma) = -\kappa_n$, it follows that $\kappa_n > 0$ iff $v_i'(\sigma) < 0$. Hence, $\sigma > \sigma^*$ since $v(\sigma)$ is hump-shaped. ■

To gain intuition for this result, consider a counterfactual equilibrium at which all producers set $\sigma = \sigma^*$. Let firm i increase complexity. This increase in complexity has a first-order positive effect on firm $i's$ market share, since more symmetric noise leads more and more consumers (at least half of the population as $\sigma_i \to \infty$) to buy good i.[18] The increase in complexity decreases the value $v(\sigma_i)$ of good i, partially offsetting the gain from complexity. However, this decrease in quality is a second-order effect local to $\sigma_i = \sigma^*$. Hence, in equilibrium firms will set $\sigma > \sigma^*$, so their products are excessively complex.

4.1 Competition Increases the Supply of Confusion

We next show that competition tends to *exacerbate* the production of excess complexity. We work out the analysis for three leading distributions that yield closed form expressions for κ_n.

Proposition 8 *The equilibrium amount of complexity σ is characterized by the following expressions for $\kappa_n = -v'(\sigma)$. For uniformly distributed noise (15),*

$$\kappa_n = 1 - \frac{2}{n}. \tag{32}$$

[18]This is true as long as $\kappa_n > 0$. However, $\kappa_n < 0$ if the distribution of f is sufficiently skewed to the right. For this case, a *decrease* in complexity *increases* market share at the counterfactual equilibrium $\sigma = \sigma^*$. The right-skewed distribution for ε has more mass below zero than above zero. Think about a limiting case in which all of the mass lies below zero, except a small amount of mass far above zero. The mass above zero helps the producer relatively little, since the producer does not charge a price conditional on the realization of the noise. Hence, on net the noise hurts the producer, leading the producer to reduce the noise by opting for an inefficiently low level of complexity. Though this case is mathematically possible, we believe that it is empirically uncommon. For instance, Proposition 2 shows that for all our distributions, for n high enough, $\kappa_n > 0$.

For Gumbel noise (18),

$$\kappa_n = \ln n - 1. \tag{33}$$

For exponential noise (19),

$$\kappa_n = -1 + \sum_{j=2}^{n} \frac{1}{j} = \ln n + \theta - 2 + O\left(\frac{1}{n}\right), \tag{34}$$

where θ is Euler's constant.

Proof. By direct calculations of the right hand side of (31). ■

Corollary 9 *There is excess complexity for $n \geq 3$, $n \geq 3$ and $n \geq 4$ for respectively uniform, Gumbel and exponential noise. The amount of excess complexity increases in n.*

Proof. Excess complexity arises iff $\kappa_n > 0$. Formulae for κ_n are given in the previous Proposition. ■

For instance, with the functional form for $v(\sigma) = v_0 - (\sigma - \sigma^*)^2 / \chi$, we get the equilibrium values of noise:

$$\sigma = \sigma^* + 2\chi\kappa_n.$$

In these three cases the equilibrium has excess noise, $\sigma > \sigma^*$.[19] Moreover, the level of excess noise *increases* with the intensity of competition, n. This contradicts the standard economic intuition that competition increases consumer welfare. In the current model, competition exacerbates the incentives for excess (inefficient) complexity. Though we illustrate that here for only three distributions, Proposition 2 shows that this is true for all distributions, as A_n increases with n.

To gain intuition for this result, again consider a counterfactual equilibrium at which all producers set $\sigma = \sigma^*$ and all products have market share $1/n < 1/2$. Let firm i increase complexity. Holding product quality fixed, this increase in complexity has a first-order positive effect on firm $i's$ market share, since more noise leads more and more consumers (exactly half of the population as $\sigma_i \to \infty$) to buy good i. As n increases, the benefits of adding noise get stronger, since the

[19] We refer to the case of "large" n's. For extremely low n's the effects reverse for right-skewed distributions ($n = 2$ or 3 for the exponential, $n = 2$ for the Gumbel). In these cases, the equilibrium is characterized by excess simplicity. Intuitively, there is a high chance that the best draw is negative if n is small, so it's optimal to have a low σ.

starting market share, $1/n$, is lower. So, the incentive for higher complexity increases as n gets large.

In summary, for the case of uniform noise, competition eliminates price inefficiencies ($\lim_{n\to\infty} p - c = 0$), but competition *exacerbates* the incentive for excess complexity. So the net effect on consumer welfare is ambiguous. In the Gumbel and exponential cases, greater competition does not eliminate pricing inefficiencies ($\lim_{n\to\infty} p - c = \sigma$) and competition drives complexity to infinity, so ever greater increases in competition unambiguously reduce consumer welfare. This effect is general for unbounded distributions. Eq. (14) shows that $\kappa_n \sim A_n/\gamma$, and $A_n \to \infty$ for unbounded distributions (see Table 1).

4.2 The Worst Firms Supply The Most Confusion

The discussion above showed that when firms are symmetric they will choose excess complexity of products, $\sigma > \sigma^*$ where σ^* is the socially efficient product complexity. By continuity, $\sigma > \sigma^*$ will still apply, even when firms are heterogeneous, as long as the heterogeneity is minor.

In this subsection, we show that the firms with the highest extrinsic quality choose to produce the least excess confusion. Also, we will see that when a firm is *much* better than the others, it will actually choose to make its product excessively simple: $\sigma < \sigma^*$.

We formalize this argument with an illustrative example rather than by treating the general case. We consider a single firm that makes endogenous decisions. It provides a good with value $v(\sigma)$, from which consumers receive the signal $v(\sigma) - p + \sigma\varepsilon$. The value provided by exogenous competitor firms is assumed to be a fixed number v_0.[20] The endogenous firm chooses σ and p to maximize profits

$$\pi = (p - c)\, P\left(v(\sigma) - p + \sigma\varepsilon > v_0\right). \tag{35}$$

To simplify analysis we suppose that ε has an exponential distribution (19), and that the value of the product has the functional form,

$$v(\sigma) = a - 2\sigma \ln \sigma. \tag{36}$$

[20] This benchmark is also the equilibrium in which the competitors have zero noise and value minus marginal cost equal to v_0.

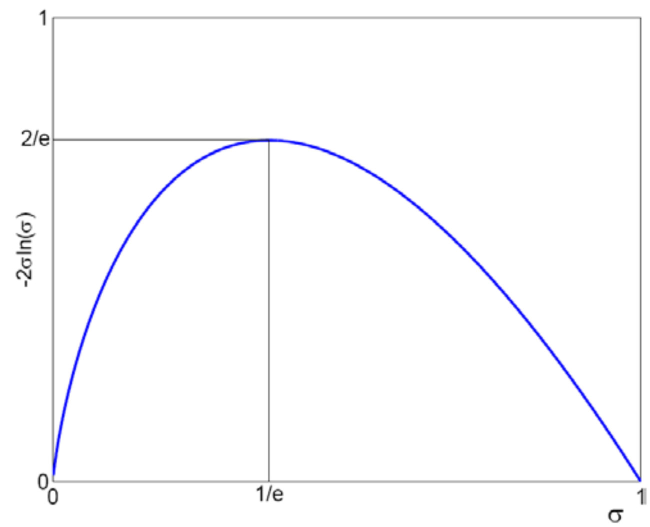

Figure 1: The product's quality $v(\sigma) = a - 2\sigma \ln \sigma$ as a function of its complexity σ. The value is maximized at the "bliss" level $\sigma^* = e^{-1}$. The function is plotted with $a = 0$.

Figure 1 plots the shape of the product value function $v(\sigma)$. Product value has a maximum at $\sigma^* = e^{-1}$. The next Proposition describes the equilibrium choice of price p and complexity σ.

Proposition 10 *When the firm maximizes profit (35) over price p and complexity σ, it sets complexity equal to*

$$\sigma = \max\left(e^{-3/2}, c + v_0 - a\right). \tag{37}$$

In particular, if $c + v_0 - a > e^{-3/2}$, then complexity decreases with the product's quality a, increases with the product's production cost c, and increases with the value v_0 offered by the competition. If $c + v_0 - a > e^{-1}$, then the firm chooses to make the product excessively complex, while if $c + v_0 - a < e^{-1}$, the firm chooses to make the product excessively simple.

The proof is in Appendix B. Figure 2 plots the equilibrium[21].

By assumption, firms have quality a and marginal cost c. If a firm has low quality or high production costs ($c + v_0 - a > e^{-1}$) then it will choose excessively complex products, setting

[21] In the more general case with $v(\sigma) = a - \chi\sigma \ln \sigma$ for $\chi > 1$, the same yields $\sigma = \max\left(e^{-1-1/\chi}, (c + v_0 - a) / (\chi - 1)\right)$. As is intuitive, if the value of the product does not change much with its complexity (χ is low), then the complexity σ^{**} will be very sensitive to the advantages of cost or quality.

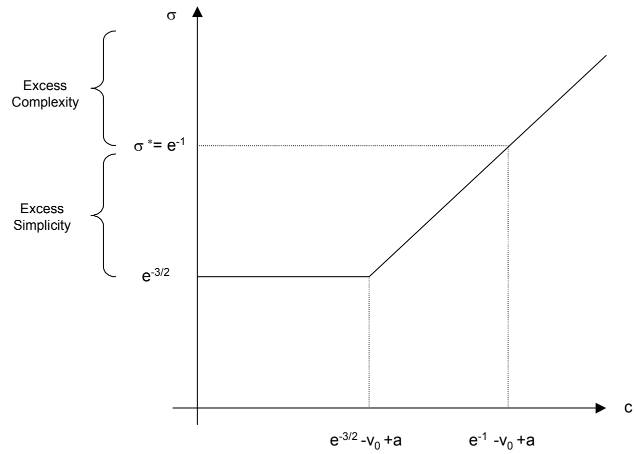

Figure 2: This Figure plots the value of the complexity σ chosen by the firm as a function of its marginal cost c. See Eq. (37). Better firms (lower marginal cost) choose a lower level of complexity. There is excess complexity for $c > e^{-1} + a - v_0$, and excess simplicity for $c < e^{-1} + a - v_0$. $\sigma^* = e^{-1}$ is the bliss level of complexity (see Figure 1), a is the quality of the product for zero complexity, and v_0 is the consumer surplus offered by the firm' competitors.

$\sigma = c + v_0 - a > e^{-1} = \sigma^*$. Intuitively, such "bad" firms generate noise with the hope that the noise will mislead the consumers who happen to draw a noise realization in the right-hand tail. For these consumers the noise will mask the firm's low quality and high prices.

The better the firm, the lower the excess complexity, and the best firms choose excess simplicity of their product. If $c + v_0 - a < e^{-1}$, then the equilibrium complexity will be less than σ^*. For intuition, fix prices and consider a firm that has a particularly high quality, $a >> 0$. If there is no noise in the consumers' decisions, the superior firm will have a market share of 1. If there is some noise, then its market share will decrease. So, a superior firm has an incentive to have very low noise, or excess simplicity: $\sigma < \sigma^*$.

4.3 A Framework For Testing Endogenous Complexity

It would be easy to test these predictions about endogenous noise. For instance, Proposition 10 predicts that a mutual fund offering very low management fees would have a clear prospectus, while funds with high fees would have an opaque prospectus. Likewise, in the cellular phone market, plans with low prices would be simple to understand, while plans with high prices would be more complicated.

The following example illustrates how such a test could be practically implemented. First, identify mutual funds with low fees (e.g., Vanguard S&P 500 fund) and mutual funds with high fees (e.g., Morgan Stanley S&P 500 fund). Then objectively measure the complexity of the respective fee descriptions.

There are many sensible ways to measure complexity. For example, one could measure the quantity of fee numbers in the fee structure. One could also count the lines of footnotes.

For example, Morgan Stanley's S&P 500 fund prospectus contains 54 fee numbers on the 1.3 pages that summarize their fee structure. Vanguard's S&P 500 prospectus contains 15 fee numbers on the 0.8 pages that summarize their fee structure. Similarly, Morgan Stanley's fee summary contains 13 lines of footnotes, whereas Vanguard's fee summary contains 4 lines of footnotes.[22]

5 The Curse of Education

So far we have assumed that firm i picks the standard deviation of noise that applies to good i (i.e., σ_i). We have shown that in most cases, firms will want to make their own products excessively complex, $\sigma_i > \sigma^*$.

In this section, we consider another manipulation of noise. Now we assume that firm i can educate a consumer a, and that this educated consumer will consequently have a low value of σ_{aj} at *all* firms $j = 1, ..., I$. If consumer a is educated, the consumer will become a better judge of all products in the marketplace.

Firms have an incentive to educate consumers — i.e., to reduce evaluation noise in the market in general — and then to win the business of the educated consumers by offering them a low markup. In this section, we study these incentives and show they are actually quite weak. The reasoning is simple. After being educated, consumer a becomes a relatively low margin consumer, since consumer a can better pick out the best deals. Hence education greatly benefits consumer a, but only moderately benefits the firm who invests in the education. Firms that educate the public may attract new customers, but the firms will have a small profit margin on those customers.

[22] Even with all of Morgan Stanley's footnotes one still can not estimate fees without reading another seven page section of the prospectus that discusses "share class arrangements."

The wedge between the consumer's benefits and the firm's benefits creates a potential inefficiency. For large education costs, the firm does not have an incentive to undertake the costly education, even though the social benefits are positive.

Economists will naturally wonder why the consumer does not purchase education services from a third party. In fact, we do observe such third party education. However, though the financial education industry exists, it does not offer reliable advice. In the U.S., publications like *Money Magazine*, *Worth*, and many others, try to market their advice every month, and hence offer a variety of complicated strategies. Such publications have little incentive to reveal the benefits of passive investment strategies like index funds. A person listening to such advice wouldn't need to hear it repeated every month. So even the financial advice industry has perverse incentives to distort information.

We now present a formal model of the curse of education. Assume there are two types of consumers, Sophisticates and Naives. Fraction ϕ_S of the population is sophisticated and they have confusion parameter σ_S, while fraction ϕ_N is Naive, with $\sigma_N \gg \sigma_S$. Assume that the total population of firms can be decomposed into n_S firms that market to Sophisticates and n_N firms that market to Naives. In the mutual fund industry, S firms will choose to charge low fees, while N firms will choose to charge high fees. Profits must be the same in both markets, which leads to[23]

$$\pi = \phi_S b_{n_S} \sigma_S / n_S = \phi_N b_{n_N} \sigma_N / n_N, \tag{38}$$

where $b_n \sigma = p - c$ by Proposition 1 and Eq. (29).

We will say that if a firm educates a Naive consumer, it changes that consumer into a Sophisticate for a cost K. Such education reduces the consumer's confusion from σ_N to σ_S, and consequently saves the consumer an amount $\sigma_N b_{n_N} - \sigma_S b_{n_S}$. However, the firm will be able to make little ex-post profit on that consumer. The consumer is now a low-margin Sophisticated consumer, so the firm will earn an expected profit of only $\sigma_S b_{n_N} / n_S$ on the newly educated consumer. We summarize this conclusion with the following Proposition.

Proposition 11 *A Sophisticated firm will educate a naive consumer only if the cost of education*

[23] Indeed, $\sigma_S b_{n_S}$ is the profit per consumer in the S market. The total profit is thus $\phi_S \sigma_S b_{n_S}$, and is divided among the n_S firms serving this market.

K satisfies $K < b_{n_S}\sigma_S/n_S$. But the consumer's benefit from his education is $b_{n_N}\sigma_N - b_{n_S}\sigma_S$. Hence there is an undersupply of education by firms when $\sigma_N \gg \sigma_S$.

6 Measuring Confusion

In principle, economists should measure the amount of confusion in the marketplace. In practice, such measurement is difficult, because it is hard to separate intrinsic sources of demand and noise-based motives for demand. Does an investor hold a mutual fund with a high management fee for a good reason (e.g., he believes that the mutual fund employs great managers), or because he just doesn't understand the fee structure?

To cleanly measure the effects of confusion, economists should use randomized education treatments (Duflo and Saez 2003, Choi et al 2004). For example, Choi et al ask subjects to allocate a wealth windfall among a set of n mutual funds. The subjects are randomly divided into a treatment group and a control group. The treatment group reads an educational background statement before making the wealth allocation decision (e.g., "mutual funds charge management fees, which are reported in the prospectus, etc..."). The control group makes the allocation without any educational intervention. Randomized assignment to the treatment and control groups creates two sets of subjects who necessarily have the same ex-ante utility functions but who have different ex-post levels of financial sophistication.

To simplify exposition, assume that the subjects in the treatment (i.e., education) group become perfect Sophisticates as a result of the educational intervention. By contrast, the Naive subjects in the control condition have a normal level of noise in their evaluations. Formally, the Sophisticates — i.e., the treatment group — have no noise in their evaluations $\sigma^S = 0$.[24] The Naives — i.e., the control group — have $\sigma^N > \sigma^S = 0$.

In this example, Sophisticates and Naives have the *same* underlying objective function. They differ only in their ability to maximize their preferences. Sophisticates choose optimally. Naives would like to replicate those sophisticated choices, but end up choosing with some additional noise.

We will measure that additional noise by comparing a product's market share of Sophisticates to the product's market share of Naives. When a product has a high level of confusing complexity,

[24] The analysis can easily be generalized to handle the case $\sigma^N > \sigma^S > 0$.

it will attract a lower share of Sophisticates than of Naives.

To formalize these ideas we will compare the demand vectors of Sophisticates and Naives: $\left(D_i^S\right)_{i=1\ldots n}$ and $\left(D_i^N\right)_{i=1\ldots n}$. In the mutual fund example, D_i^S represents the fraction of assets invested by the Sophisticates in mutual fund i. We will focus on a particularly important summary statistic, ρ, the correlation between $\left(D_i^S\right)_{i=1\ldots n}$ and $\left(D_i^N\right)_{i=1\ldots n}$. If there is perfect correlation ($\rho = 1$) between the Sophisticates' and Naives' demand vectors, then there is no confusion on the part of the Naives, implying $\sigma^N = 0$.[25] If there is no correlation ($\rho = 0$), then there is full confusion ($\sigma^N = \infty$). More generally, the correlation ρ between the two demand vectors allows us to infer the precise amount of confusion.

6.1 A Tractable Econometric Framework for Measuring the Relative Importance of Confusion and True Taste Differences

We propose a tractable implementation of the above idea. We begin by assuming that a Sophisticate has *true utility* $U_{ia}^S = v_i + H^S\left(T_{ia}\right)$ where we index consumers by a and goods by i. This utility function depends on the idiosyncratic (true) tastes of the Sophisticate, T_{ia}. By contrast, a Naif will not observe her true utility, but instead only observes utility signals

$$U_{ia}^N = v_i + H^N\left(T_{ia}, N_{ia}\right)$$

where N_{ia} is noise, i.e. confusion. If the noise is always 0, i.e. $N_{ia} \equiv 0$, then we will have

$$H^S\left(T_{ia}\right) = H^N\left(T_{ia}, N_{ia} \equiv 0\right). \tag{39}$$

In general, the respective market shares for product i will be given by, $D_i^\tau = P\left(U_{ia}^\tau > \max_{j\neq i} U_{ja}^\tau\right)$ for $\tau = S, N$.

We assume that idiosyncratic tastes are captured by a unidimensional source of variation, $T_{ia} = \sigma_T u_{ia}$, and that confusion noise (for Naives) is captured by a two-dimensional source of variation $N_{ia} = \left(\sigma_N \varepsilon_{ia}, \sigma_i \eta_{ia}\right)$, where u_{ia}, ε_{ia} and η_{ia} are independent standard normals. The

[25] Another possible interpretation is that the education intervention was too weak to remove the confusion. In this sense, the experiment provides a lower bound on the amount of confusion σ^N.

following parametrization proves to be useful:

$$H^N\left(T_{ia}, N_{ia}\right) = \sqrt{\sigma_T^2 + \sigma_N^2}\, G\left(\frac{\sigma_T u_{ia} + \sigma_N \varepsilon_{ia}}{\sqrt{\sigma_T^2 + \sigma_N^2}}\right) + \sigma_i \eta_{ia}. \tag{40}$$

For reasons that will be clear soon, we choose the increasing transformation function

$$G\left(s\right) = -\ln \ln 1/\Phi\left(s\right), \tag{41}$$

where Φ is the cumulative distribution function of the standard Gaussian distribution. To interpret (40), let us start with the Sophisticates. Because of (39), their tastes shocks are $H^S\left(T_{ia}\right) = \sigma_T G\left(u_{ia}\right)$. We show later that $G\left(u_{ia}\right)$ has a Gumbel distribution, which is convenient for demand analysis. So, our model nests the basic model of product choice, which has been used so productively in the industrial organization literature (e.g. McFadden (1981), Berry, Levinsohn and Pakes (1995), Goldberg (1995), Nevo (2001)).

For economic and econometric generality, we allow two types of noise for the Naives: general confusion, scaled by σ_N, and good-specific noise, scaled by σ_i. Among the naifs, we expect a high value of σ_N to smooth market shares (since goods will effectively be chosen randomly), and a high value of σ_i to increase the market share of good i (when $n > 2$). Those intuitions are formalized in the next Proposition.

Proposition 12 *With the above formulation, the demand of sophisticated consumers is*

$$D_i^S = \frac{\exp\left(\lambda^S v_i\right)}{\sum_{j=1}^n \exp\left(\lambda^S v_j\right)} \tag{42}$$

$$\lambda^S = 1/\sigma_T \tag{43}$$

and the demand of the naives is

$$D_i^N = E\left[\frac{\exp\left(\lambda^N v_i + \lambda^N \sigma_i \eta_{ia'}\right)}{\sum_{j=1}^n \exp\left(\lambda^N v_j + \lambda^N \sigma_j \eta_{ja'}\right)}\right] \tag{44}$$

$$\lambda^N = 1/\sqrt{\sigma_T^2 + \sigma_N^2}. \tag{45}$$

In the limit where all products have small market shares (formally, $n \to \infty$, and there is an M s.t. for all i, var $\left(e^{\lambda^N v_i}\right) < M < \infty$), we have the following equivalence:

$$D_i^N \sim \frac{\exp\left(\lambda^N v_i + \lambda^N \sigma_i^2/2\right)}{\sum_{j=1}^n \exp\left(\lambda^N v_j + \left(\lambda^N\right)^2 \sigma_j^2/2\right)}. \tag{46}$$

If $\sigma_i = 0$, as $\lambda^N = 1/\sqrt{\sigma_T^2 + \sigma_N^2}$ is lower than $\lambda^S = 1/\sigma_T$, the naive's market shares are more dispersed. Now consider a vector of market shares within a product category $(D_i^S, D_i^N)_{i=1}^n$. Consider running the following regression of the Naives' demand on the Sophisticates' demand:

$$\ln D_i^N = \alpha + \beta \ln D_i^S + r_i. \tag{47}$$

If σ_i has zero (or small) covariance with v_i, we will have,

$$\beta = \sigma_T/\sqrt{\sigma_T^2 + \sigma_N^2}$$

Hence we expect a slope $\beta < 1$. The amount of general noise in the market is given by

$$\frac{\sigma_N}{\sigma_T} = \sqrt{\beta^{-2} - 1}.$$

We can also infer the amount of good-specific noise for good i

$$\frac{\sigma_i^2 - \overline{\sigma_j^2}}{\sigma_T^2} = 2\frac{r_i}{\beta^2}.$$

Hence the strategy for measuring confusion is quite simple: run regression (47) and impute the value of σ_N/σ_T and $\frac{\sigma_i^2 - \overline{\sigma_j^2}}{\sigma_T^2}$ from the resulting parameter estimates.

To implement this framework a researcher will need to identify a marketplace in which Sophisticated and Naive agents can be distinguished ex-ante and in which the Sophisticated and Naive agents have the same underlying distributions of objective functions. These are not trivial requirements, but they can be surmounted with randomized field experiments: randomly select a subpopulation of consumers, educate them about an industry (e.g., mutual funds), and then compare their demand vectors to a control group.

6.2 Measuring the Distribution of the Noise

Until now, the analysis in this section has assumed a particular functional form for the distribution of noise. This subsection shows how to estimate the distribution when the functional form is unknown.

Suppose that a decision-maker evaluates a set of n goods (e.g., mutual funds). The decision-maker then ranks his two favorite goods and reports the price difference, Δp_n, that makes him indifferent between them.[26] It turns out that the relationship between Δp_n and n, the number of goods, can be used to impute the shape of the noise distribution.[27] To see how this works, first observe that $\Delta p = \sigma\left(\varepsilon_{(1)} - \varepsilon_{(2)}\right)$, where $\varepsilon_{(1)}$ and $\varepsilon_{(2)}$ are the highest and second highest draws of n i.i.d. signals. We will want to exploit the following Proposition.

Proposition 13 *The expected value of Δp_n is:*

$$\Delta p_n \sim b_n \Gamma(2+\xi)\,\Gamma(1-\xi) \tag{48}$$

where $b_n = \frac{p-c}{\sigma}$, and $p-c$ is calculated in Proposition 10. In particular, for the Gaussian, Gumbel, exponential, and log-normal distributions $\Delta p_n \sim b_n$.

The proposition implies that Δp_n is a good approximation of b_n, up to a constant, $\Gamma(2+\xi)\,\Gamma(1-\xi)$, which does not depend on n and which will be approximately equal to one in the Gaussian, Gumbel, exponential and lognormal cases.

We now have all of the results that we need to infer the underlying noise distribution. The imputation relies on the following data: measures of the gap Δp_n for a variety of market sizes n, for instance $n = 10$, 50 and 200. To impute the distribution of noise, graph $\ln \Delta p_n$ as a function $\ln n$. Its slope should be the characteristic index ξ of the distribution.

[26] We assume that all preferences are measured with incentive compatible mechanisms.

[27] We have exact results in two cases. For the uniform distribution, $\Delta p_n = 2\sigma/(n+1)$, and the the exponential distibution $\Delta p_n = \sigma$. The derivations are by simple calculations from e.g. Reiss (1989, Chapter 1).

7 Conclusion

We have analyzed an environment in which consumers have noisy evaluations of products. The noise effectively increases the market power of firms, raising equilibrium markups. Competition barely reduces this market power. If consumers have noisy product evaluations, high levels of competition have little effect on markups.

We also analyze markets in which firms can influence the level of noise, or product complexity. In equilibrium, firms choose to make products excessively complex (i.e., excessively noisy) and competition only exacerbates this problem. Higher levels of competition *increase* the equilibrium level of excess complexity. Here again, competition does not protect consumers from the consequences of noisy evaluations. We also show that firms with the highest unit costs and the lowest product quality will choose the greatest excess complexity.

We present an econometric framework that measures confusion in the marketplace. Randomly divide a population of consumers into two groups: control and treatment. Observe the control group's market purchases (e.g., mutual fund selection). Educate the treatment group and then observe their market purchases (e.g., teach the treatment group how to read a mutual fund prospectus and then observe their mutual fund selection). If the control group chooses differently than the treatment group, then the choices of the control group are driven partially by confusion. We show how to use market demand vectors to econometrically impute the quantity of noise in the marketplace.

8 Appendix A: Elements of Extreme Value Theory

Coefficients of Regular Variation We recommend Embrechts *et al.* (1997) and Resnick (1987) for excellent expositions of extreme value theory. The following concept will be important in the proofs.

Definition 14 *A function g defined in a right neighborhood of 0 has regular variation with exponent ρ if*

$$\forall u > 0, \lim_{t \to 0^+} g(ut)/g(t) = u^\rho. \tag{49}$$

This means that, for small $t > 0$, $g(t)$ behaves like t^ρ, perhaps up to a constant or slowly varying function. For instance, (49) holds if $g(t) = t^\rho$ and $g(t) = t^\rho [\ln(1/t)]^\alpha$ for some α. If the following limit exists, it provides a convenient characterization of ρ (see Resnick 1987, p.21):

$$\rho = \lim_{t \to 0^+} \frac{tg'(t)}{g(t)}. \tag{50}$$

Three Types of Distributions In extreme value theory, there are three classes of distributions. They are ordered by increased fatness of the right tail. A useful indicator of their fatness is a the characteristic index ξ. ξ is the coefficient of regular variation[28] of $f\left(\overline{F}^{-1}(t)\right)/t$ for $t \to 0$. There are alternative, similar, definitions (Embrechts et al. 1997, p.152-158). As Table 1 indicates, ξ is an index of the fatness of the distribution. Distributions with fatter tails have a (weakly) larger ξ.

The first type is the "domain of attraction of the Weibull". It comprises very "thin tailed" distributions of the type (16), such as the uniform distribution. Their support has an upper bound. For them $\xi = -1/\alpha$.

The second type is the "domain of attraction of the Gumbel". It comprises distributions of medium thinness, such as the Gaussian, Gumbel, Exponential, and Gamma distributions. Their support is unbounded on the right, and $\xi = 0$.

The third type is the "domain of attraction of the Fréchet." It comprises power law tail distributions of the type (21). Here $\xi = 1/\zeta$.

[28] So $\xi \equiv -1 + \lim_{t\to 0^+} t \frac{d}{dt} \ln\left[f\left(\overline{F}^{-1}(t)\right)\right]$.

9 Appendix B: Longer Derivations

9.1 Two Basic Propositions

The following result will characterize the large n behavior of $E[J(M_{n-1})]$, where J is a function, and $M_{n-1} = \max_{i=1...n-1} \varepsilon_i$ is the maximum of $n-1$ random variables with CDF F. Define $\overline{F}(x) \equiv 1 - F(x)$. The following Proposition allows us to replace the complicated integral $E[J(M_{n-1})]$ by the simpler deterministic expression $J(A_n)\Gamma(1+\rho)$.

Proposition 15 *Consider a function J such that $J\left(\overline{F}^{-1}(t)\right)$ has coefficient of regular variation ρ, and M_{n-1} is the maximum of $n-1$ i.i.d. random variables with CDF F. Set $(A_n) = \overline{F}^{-1}(1/n)$. Then:*

$$\lim_{n\to\infty} \frac{E[J(M_{n-1})]}{J(A_n)\Gamma(1+\rho)} = 1 \tag{51}$$

Proof. We call $m_{n-1} = \min_{i=1...n-1} y_i$ the minimum of $n-1$ i.i.d. uniform $[0,1]$ variables y_i. Its cumulative is:

$$P(m_{n-1} > t) = P(\forall i = 1...n-1, y_i > t) = \prod_{i=1}^{n-1} P(y_i > t) = (1-t)^{n-1} \tag{52}$$

M_{n-1} has the distribution: $P(M_{n-1} < X) = F(x)^{n-1}$. So M_{n-1} has the same distribution as $\overline{F}^{-1}(m_{n-1})$. Indeed:

$$\begin{aligned} P\left(\overline{F}^{-1}(m_{n-1}) < X\right) &= P\left(m_{n-1} > \overline{F}(x)\right) = \left(1 - \overline{F}(x)\right)^{n-1} \text{ by (52)} \\ &= F(x)^{n-1} = P(M_{n-1} < X). \end{aligned}$$

Defining $j(t) = J\left(\overline{F}^{-1}(t)\right)$, this yields the representation

$$I_n \equiv E[J(M_{n-1})] = E\left[J\left(\overline{F}^{-1}(m_{n-1})\right)\right] = E[j(m_{n-1})].$$

Before proceeding, we give some intuition for why the result should be true. Given that m_{n-1} is concentrated around its mean, $1/n$, the last equation makes plausible that $I_n \simeq j(1/n)$. This reasoning gives the leading order term for I_n. Taking into consideration the fluctuations of m_n

around its mean, we get a subleading correction.

$$\begin{aligned} I_n &= E\left[j\left(m_{n-1}\right)\right] = \int_0^1 j\left(x\right)\left(n-1\right)\left(1-x\right)^{n-2} dx \\ &= \int_0^n j\left(\frac{u}{n}\right)\frac{n-1}{n}\left(1-\frac{u}{n}\right)^{n-2} du. \end{aligned}$$

As $\left(1-\frac{u}{n}\right)^{n-1} \to e^{-u}$, we have

$$\begin{aligned} I_n &\sim \int_0^n j\left(\frac{u}{n}\right)e^{-u}du = j\left(\frac{1}{n}\right)\int_0^n \frac{j\left(u/n\right)}{j\left(1/n\right)}e^{-u}du \\ &\sim j\left(\frac{1}{n}\right)\int_0^n u^\rho e^{-u}du \sim j\left(\frac{1}{n}\right)\int_0^\infty u^\rho e^{-u}du = j\left(\frac{1}{n}\right)\Gamma\left(1+\rho\right). \end{aligned}$$

■

The assumption of unbounded support is not crucial. What matters here is the behavior of the density around A_n. So the above results hold if the distribution of the noise is truncated on the right by some X, if $\overline{F}\left(X\right) \ll 1/n$. One replaces the index ρ by its local equivalent, $\rho = j'\left(1/n\right)/\left[nj\left(1/n\right)\right]$. Numerical simulations show that this gives very good results.

Characterizing the Exponents The following two Propositions will be useful.

Proposition 16 *For a regular distribution with characteristic index ξ, the coefficient of regular variation of $f\left(\overline{F}^{-1}\left(t\right)\right)$ is $1+\xi$.*

Proposition 17 *For a regular distribution with characteristic index ξ, if the upper bound of F's support is ∞ or a non-zero real number, then the coefficient of regular variation of $f\left(\overline{F}^{-1}\left(t\right)\right)\overline{F}^{-1}\left(t\right)$ is $1+\min\left(0,\xi\right)$.*

The proofs are typically easy for power law distributions, which have $\xi \neq 0$. The results for $\xi = 0$ can be guessed by continuity, but require slightly more careful arguments.

Proof of Proposition 16. To reduce the algebraic clutter, in the derivations we omit the corrections due to the "slowly varying functions", and set the scaling constants k at 1. If F is in the domain of attraction of the Fréchet (as in Eq. 21), write $f\left(x\right) = \zeta x^{-\zeta-1}$, and $\overline{F} = x^{-\zeta}$. Then $\overline{F}^{-1}\left(t\right) = t^{-1/\zeta}$, and $f\left(\overline{F}^{-1}\right)\left(t\right) = \zeta t^{1+1/\zeta}$. Because in this case $\xi = 1/\zeta$, this means $\rho = 1+\xi$.

If F is in the domain of attraction of the Weibull (as in Eq. 16), the same proof works. We say that the upper bound is 1, and have $f(x) = \alpha(1-x)^{\alpha-1}$, and $\overline{F} = (1-x)^{\alpha}$. Then $\overline{F}^{-1}(t) = 1 - t^{1/\alpha}$, and $f\left(\overline{F}^{-1}\right)(t) = \alpha t^{1-1/\alpha}$. Because in this case $\xi = -1/a$, this means $\rho = 1 + \xi$. If F is in the domain of attraction of the Gumbel, Resnick (1987, p. 66) shows: $t\frac{d}{dt}\ln\left[f\left(\overline{F}^{-1}(t)\right)\right] = -\overline{F}(x) f'(x)/f(x)^2 \to 1$, which implies $\rho = 1$. □

Proof of Proposition 17. We call ρ_g the coefficient of regular variations of a function g. If $\xi > 0$ and $\overline{F}^{-1}(t) = t^{-1/\zeta} = t^{-\xi}$, then $\rho_{\overline{F}^{-1}} = -\xi$. See Resnick (1987, pp. 63-67), along with the case $\xi = 0$. For $\xi < 0$, $\overline{F}^{-1}(t)$ converges to a non-zero upper bound of the distribution, so $\rho_{\overline{F}^{-1}(t)} = 0$. We conclude that $\rho_{\overline{F}^{-1}(t)} = \min(-\xi, 0)$. To finish the proof, we observe that, for two functions g and h, (49) implies $\rho_{g \cdot h} = \rho_g + \rho_h$. We apply this to find

$$\begin{aligned} \rho_{f\left(\overline{F}^{-1}(t)\right)\overline{F}^{-1}(t)} &= \rho_{f\left(\overline{F}^{-1}(t)\right)} + \rho_{\overline{F}^{-1}(t)} \\ &= 1 + \xi + \min(-\xi, 0) = 1 + \min(0, \xi). \ \square \end{aligned}$$

9.2 Proof of Proposition 1

The firm's maximization problem is given by $\max_{x_i,\sigma_i}(p_i(x_i,\sigma_i) - c)\, D(x_i,\, \sigma_i)$. Using (4), the associated first order conditions are

$$\begin{aligned} -D(x_i,\, \sigma_i) + (p_i - c)\frac{\partial}{\partial x_i} D(x_i,\, \sigma_i) &= 0, \\ v_i'(\sigma_i)\, D(x_i,\, \sigma_i) + (p_i - c)\frac{\partial}{\partial \sigma_i} D(x_i,\, \sigma_i) &= 0. \end{aligned}$$

At a symmetric equilibrium ($p_i = p$, $x_i = 0$, $\sigma_i = \sigma$), substitution of equations (7), (8), and (9) yields

$$\begin{aligned} -\frac{1}{n} + (p - c)\frac{1}{\sigma} E\left[f(M_{n-1})\right] &= 0, \\ v'(\sigma)\frac{1}{n} + (p - c)\frac{1}{\sigma} E\left[f(M_{n-1})\, M_{n-1}\right] &= 0, \end{aligned}$$

which can be rearranged to produce the required results.

9.3 Proof of Proposition 2

For $p-c$, Proposition 16 says that the coefficient of regular variations of $f\left(\overline{F}^{-1}(t)\right)$ is $\rho = 1+\xi$. So applying Proposition 15 to $J(x) = f(x)$ gives: $E[f(M_{n-1})] \sim f(A_n)\Gamma(2+\xi)$. Substituting this into (10) we get (13).

For $v'(\sigma)$, Proposition 17 says that the coefficient of regular variations of $f\left(\overline{F}^{-1}(t)\right)\overline{F}^{-1}(t)$ is $\rho = 1+\min(0,\xi)$. Proposition 15 applied to $J(x) = f(x)x$ gives: $E[f(M_{n-1})M_{n-1}] \sim f(A_n)A_n\Gamma(2+\min(0,\xi))$. Eq. (11) gives:

$$-v'(\sigma) = \frac{E[f(M_{n-1})M_{n-1}]}{E[f(M_{n-1})]} \sim \frac{f(A_n)A_n\Gamma(2+\min(0,\xi))}{f(A_n)\Gamma(2+\xi)} = \frac{A_n}{\Gamma(2+\max(0,\xi))}.$$

9.4 Proof of Proposition 5

The markup is $p-c = b_n\sigma = (b_n\sigma/n)\,n = Cn$ and n solves (30). By homogeneity, it is enough to consider the case $C=1$ and the limit $\sigma \to \infty$. In the case where n cannot be solved exactly, we use the method of iterated approximations.[29] This calculation is straightforward, except in the Gaussian and lognormal cases. In the Gaussian case, we have $b_n/n \sim 1/\left[n\sqrt{\ln 2n}\right] \sim 1/\sigma$, so

$$\begin{aligned}
\ln n &= \ln\sigma - \frac{1}{2}\ln(\ln 2n) + o(1) = \ln\sigma - \frac{1}{2}\ln\left(\ln 2 + \ln\sigma - \frac{1}{2}\ln(\ln 2n)\right) + o(1) \\
&= \ln\sigma - \frac{1}{2}\ln(\ln 2\sigma + o(\ln\sigma)) + o(1) \\
&= \ln\sigma - \frac{1}{2}\ln(\ln 2\sigma) + \frac{1}{\ln 2\sigma}o(\ln\sigma) + o(1) \text{ by Taylor expansion} \\
&= \ln\sigma - \frac{1}{2}\ln(\ln 2\sigma) + o(1)
\end{aligned}$$

which proves the result for $(p-c)_{\text{Gaussian}}$. The lognormal case is similar.

[29] See e.g. Resnick (1987, p.67-74).

9.5 Proof of Proposition 6

The denominator of (31) is positive, and the numerator is $E\left[f\left(M_{n-1}\right)M_{n-1}\right]=\int\left(n-1\right)xf\left(x\right)^{2}F\left(x\right)^{n-1}dx$. If f is symmetric around 0, then

$$E\left[f\left(M_{n-1}\right)M_{n-1}\right]=\int_{x>0}\left(n-1\right)xf\left(x\right)^{2}\left[F\left(x\right)^{n-1}-F\left(-x\right)^{n-1}\right]dx>0.$$

9.6 Proof of Proposition 10

By shifting a into $a'=a-c-v_0$ and p into $p'=p-c$ we can assume $c=v_0=0$ in the proof. We are looking for the optimal complexity level, which we will call $\widehat{\sigma}$ in this proof. The profit function is:

$$\pi=pP\left(v\left(\sigma\right)-p+\sigma\varepsilon>0\right)=pP\left(\varepsilon>\frac{p-v\left(\sigma\right)}{\sigma}\right)=p\max\left(e^{-\frac{p-v(\sigma)}{\sigma}-1},1\right).$$

We consider first the case where we have an interior solution ,

$$e^{-\frac{p-v(\sigma)}{\sigma}-1}<1. \tag{53}$$

Then the price p that maximizes the profit π is $p=\sigma$, and the profit is $\pi=\sigma e^{v(\sigma)/\sigma-2}$. Using the function form (36) we get: $\ln\pi=\frac{a}{\sigma}-\ln\sigma-2$. If $a\geq0$, the maximum is at $\sigma=0$, which violates (53). If $a<0$, the maximum is at $\widehat{\sigma}=-a$. Condition (53) holds iff

$$0>-\frac{\widehat{\sigma}-v\left(\widehat{\sigma}\right)}{\widehat{\sigma}}-1=\frac{v\left(\widehat{\sigma}\right)}{\widehat{\sigma}}-2=\frac{a}{\widehat{\sigma}}-2\ln\widehat{\sigma}-2=-2\ln\widehat{\sigma}-3,$$

i.e. $-a=\widehat{\sigma}>e^{-3/2}$.

If this last condition does not hold, then (53) binds, and $\pi=p=v\left(\sigma\right)-\sigma$, and whose maximization gives $\sigma=\arg\max_{\sigma}v\left(\sigma\right)-\sigma=e^{-3/2}$. This proves (37).

9.7 Proof of Proposition 12

First we observe that for s a standard normal, $G\left(s\right)$ has a Gumbel distribution,

$$P\left(G\left(s\right)<x\right)=P\left(s<G^{-1}\left(x\right)\right)=\Phi\left(G^{-1}\left(x\right)\right)=e^{-e^{-x}}.$$

So $G\left(\frac{\sigma_T u_{ia}+\sigma_N \varepsilon_{ia}}{\sqrt{\sigma_T^2+\sigma_N^2}}\right)$ is a Gumbel variable, and we have, by the standard result on logit demands[30],

$$D_i = D_i^N = E\left[\frac{\exp\left(\lambda^N v_i + \lambda^N v_i \eta_{ia'}\right)}{\sum_{j=1}^n \exp\left(\lambda^N v_j + \lambda^N v_j \eta_{ja'}\right)} \mid (v_j)_{1\leq j\leq n}\right] \tag{54}$$

with λ^N as in (45). Replacing σ_n and v_i in the above formula yields the formula (42) for the sophisticates.

The derivation of asymptotic formula (46) is straightforward.

9.8 Proof of Proposition 13

We use the results in Reiss (1989, p.161), and direct calculations. For instance, in the unbounded power law case with $k=1$, the asymptotic CDFs of $n^{-1/\zeta}\varepsilon_{(1)}$ and $n^{-1/\zeta}\varepsilon_{(2)}$ are $F_{(1)} = e^{-x^{-\zeta}}$ and $F_{(2)} = e^{-x^{-\zeta}}\left(1+x^{-\zeta}\right)$, so:

$$\begin{aligned} n^{-1/\zeta}\Delta p_n &= E\left[n^{-1/\zeta}\varepsilon_{(1)} - n^{-1/\zeta}\varepsilon_{(2)}\right] \to \int_0^\infty x\left(F'_{(1)} - F'_{(2)}\right)dx \\ &= \left[x\left(F_{(1)} - F_{(2)}\right)\right]_0^\infty - \int_0^\infty \left(F_{(1)} - F_{(2)}\right)dx = \int_0^\infty e^{-x^{-\zeta}} x^{-\zeta} dx = \frac{1}{\zeta}\Gamma\left(1-\frac{1}{\zeta}\right) \end{aligned}$$

The calculation is similar for the other types of distributions.

9.9 Existence of a Symmetric Equilibrium

This section discusses conditions under which a symmetric equilibrium exists. By shifting all the variables by $-c$, it is sufficient to study the case $c=0$. By homogeneity, it is sufficient to study the case $\sigma = 1$. We define $D(x) = D(x, \sigma = 1)$.

By Proposition 1, a symmetric equilibrium can only exist at $p = b_n$. A global symmetric equilibrium exists iff the profit function $pD(b_n - p)$ has a global maximum for $p = b_n$. A local equilibrium exists iff the profit function $pD(b_n - p)$ has a local maximum at $p = b_n$. We next review what is known about the existence of equilibrium in such cases.

It can be shown that $p = b_n$ is a global equilibrium if $\ln f$ is concave. However, this sufficient condition is not necessary. A global equilibrium also exists in other natural cases, e.g., for the

[30] See e.g. Anderson et al. (1992), p. 39.

log-normal distribution. We can prove the existence of a local equilibrium in the log-normal case, and can verify that the equilibrium is global using numerical methods. Complete analytic proofs of the existence of a symmetric equilibrium for non-log concave distributions (such as the lognormal distribution) is an open problem in mathematical economics (cf. Caplin and Nalebuff 1991).

While it would be useful to have analytic results for the global equilibrium, for many purposes this is not necessary. For a given distribution and value of n, the check can be done numerically. Hence as a check for our Proposition that the markup increases in n for a lognormal distribution, one can check that, for $n = 2...100$, that $p - c = b_n$ is indeed a global equilibrium. This confirms our statement that the markup increases in n for lognormal distributions, as least for $n = 2...100$.

We now detail our arguments. First, we recall standard conditions for the existence and uniqueness of an equilibrium.

Proposition 18 *If* $\ln f$ *is concave, an equilibrium exists and is unique.*

Proof. We apply Caplin and Nalebuff (1991), Theorem 2 and Proposition 7. ■

The above proposition is useful for distributions less fat than the exponential (such as the distributions (15)– (19)). But the proposition is not powerful enough for fatter distributions, such as the lognormal. We verified numerically that the equilibrium exists for distributions (20)- (21), which have finite variance. Also, we can derive the following mathematical criterion to show that the equilibrium is at least local, i.e. that $pD(b_n - p)$ has a local maximum at $p = b_n$. We then derive its limit behavior.

Proposition 19 *The profit function* $\pi(p) = pD(b_n - p)$ *satisfies* $\pi'(b_n) = 0$*, and we have* $\pi''(b_n) < 0$*, i.e. a strict local maximum, iff* $\rho_n < 2$*, where the normalized curvature* ρ_n *is:*

$$\rho_n \equiv \frac{-E\left[f'\left(M_{n-1}\right)\right]}{nE\left[f\left(M_{n-1}\right)\right]^2}. \tag{55}$$

Proof. Profit $\pi_i(p_i) = p_i D(b_n - p_i)$. At $p = b_n$, we have $p = 1/(nD')$ and $\pi''(b_n) = pD'' - 2D'$, implying that $\pi''(b_n) < 0$ iff $\frac{D''}{nD'} - 2D' < 0$, i.e., $\rho_n = D_i'' \left[nD_i'^2\right]^{-1} < 2$. We conclude using $D''(0) = E\left[-f'\left(M_{n-1}\right)\right]$ and $D'(0) = E\left[f\left(M_{n-1}\right)\right]$. ■

Proposition 20 *We have*

$$\lim_{n\to\infty} \rho_n = \rho_\infty = \frac{\Gamma(2+2\xi)}{\Gamma(1+\xi)\,\Gamma(2+\xi)}, \tag{56}$$

$\rho_\infty < 2$, *for all the distributions in Table 1 (including power law distributions provided they have finite variance).*

Proof. Define $\overline{h}(t) = -f'\left(\overline{F}^{-1}(t)\right)$. Proposition 15 implies $\rho_n \sim \rho(1/n)$ where

$$\rho(t) \equiv \frac{t\overline{h}(t)\,\Gamma\left(1+\rho_{\overline{h}}\right)}{h(t)^2\,\Gamma(2+\xi)^2}.$$

First, we consider the case $\xi = 0$. Embrechts et al. (1997, p.140) show that as $x \to \infty$, $-f'(x)/f(x) \sim f(x)/\overline{F}(x)$. This implies that for $t \to 0$, $\overline{h}(t) \sim h(t)^2/t$. So $\rho_{\overline{h}} = 1$ and $\rho(t) = t\frac{h(t)^2/t}{h(t)^2} \sim 1$. Hence $\lim_{n\to\infty}\rho_n = 1$, and the Proposition is verified as $\xi = 0$.

Next, we consider $\xi > 0$. (The proof is the same if $\xi < 0$.) For $\overline{F}(x) \sim k^\zeta x^{-\zeta}$, $\overline{F}^{-1}(t) \sim kt^{-1/\zeta}$, and

$$\begin{aligned} h(t) &= f\left(\overline{F}^{-1}(t)\right) \sim \frac{\zeta}{k}t^{1+1/\zeta} \\ \overline{h}(t) &= -f'\left(\overline{F}^{-1}(t)\right) \sim \frac{\zeta(\zeta+1)}{k^2}t^{1+2/\zeta}, \end{aligned}$$

so that $\rho_h = 1 + 1/\zeta$, $\rho_{\overline{h}} = 1 + 2/\zeta = 2\rho_h - 1$ and

$$\rho(t) = \frac{t\overline{h}(t)\,\Gamma\left(1+\rho_{\overline{h}}\right)}{h(t)^2\,\Gamma(1+\rho_h)^2} \sim \frac{\frac{\zeta(\zeta+1)}{k^2}t^{2+2/\zeta}\Gamma(2\rho_h)}{\left(\frac{\zeta}{k}t^{1+1/\zeta}\right)^2\Gamma(1+\rho_h)^2} = \frac{\Gamma(2+2\xi)}{\Gamma(1+\xi)\,\Gamma(2+\xi)}.$$

■

For instance, for the Gaussian, exponential and lognormal distributions (which have $\beta_h = 1$), $\rho_\infty = 1$, which implies that for large n the equilibrium is indeed locally stable.

Also, for power law distributions with exponent ζ (Eq. 21), for large n, the equilibrium is locally stable if $\zeta \geq 2$, which implies $\rho_n < 2$. Extensive numerical work leads us to conjecture that it is globally stable, but we could not prove this formally. Interestingly, the equilibrium is not globally stable for the power laws that have $\zeta < 2$, i.e. that have infinite variance.

In summary, using analytic and numerical results we have shown that the symmetric equilibrium exists for our distributions with finite variance, but a completely analytic proof is beyond the scope of this paper. We offer a simple numerical way to check that $p = b_n$ is a global equilibrium, so that in any particular application one can check that our assumption of a symmetric equilibrium is verified.

9.10 The Maximum Fatness Principle

In general, we expect the noise term $\sigma\varepsilon_i$ to be the sum of several influences. For instance, one can have $\sigma\varepsilon = \sigma_u u + \sigma_v v$ with u and v independent. One might expect that the markup will depend on both noise intensities, implying that $p - c$ would be proportional to $\sigma_u + \sigma_v$ for instance. The following Proposition shows that this is not the case.

Proposition 21 *If u and v are exponentially distributed with mean 0 and variance 1 and the total noise is $\sigma\varepsilon = \sigma_u u + \sigma_v v$, then the markup is $p - c = \max(|\sigma_u|, |\sigma_v|)$.*

Proof. By homogeneity it is enough to consider the case $\sigma_u = 1 > \sigma_v > 0$. The density of u and v is $e^{-x-1}1_{x>-1}$, hence the tail distribution of $\sigma_u u + \sigma_b v$ is:

$$
\begin{aligned}
G(x) &= P(u + \sigma_v v > x) = \int\int_{u>-1,v>-1} e^{-u-v-2}1_{u+\sigma_b v>x}dudv \\
&= \frac{e^{-\sigma_v-1}}{1-\sigma_v}e^{-x} + \frac{\sigma_v e^{-/\sigma_v-1}}{1-\sigma_v}e^{-x/\sigma_v} \sim \frac{e^{-\sigma_v-1}}{1-\sigma_v}e^{-x}
\end{aligned}
$$

hence $G(x)$ is tail equivalent (Resnick 1987, p.67) to an exponential distribution. We use Proposition 2 to conclude. ■

Intuitively, the fatter tail dominates the sum. This is likely to be quite general. We conjecture that if u has fatter tails than v, in the sense that $\lim_{t\to\infty}\left[\overline{F}_u(t)/f_u(t)\right]/\left[\overline{F}_v(t)/f_v(t)\right] = \infty$ then $\sigma_u u + \sigma_b v$ is tail-equivalent to $\sigma_u u$, and the equilibrium markup is the markup that would occur if the noise were only $\sigma_u u$.

10 Appendix C. Variety of Bayesian Inferences

We consider the case where consumer a has a true utility for good i, $V_{ia} = v_i - p_i + \sigma_T u_{ia}$, but receives a noisy signal,

$$U_{ia} = v_i - p_i + \sigma_T u_{ia} + \sigma_N \varepsilon_{ia}.$$

Here $\sigma_T u_{ia}$ is true variation in tastes and $\sigma_N \varepsilon_{ia}$ is just evaluation noise. In the body of the paper, we assume that consumer choose the good with the largest U_{ia}, because we view this an a useful benchmark. This section discusses the choices of a Bayesian consumer[31]. We first explore cases in which all the v_i and σ_N are equal.

10.1 Case Where Just One Signal is Observed

Suppose the consumer observes just U_{ia} but not p_i (for instance, the price itself is unclear, as in a checking account or a mutual fund, or the quality to price ratio itself is unclear). The consumer should rationally choose the good i that maximizes $E[V_{ia} \mid U_{ia}]$. Under weak conditions[32], this is equivalent to maximizing U_{ia}.

10.2 Case Where Just a Quality Signal and a Price Signal are Observed

We take the case of Gaussian noise. $E[V_{ia} \mid U_{ia}, p_i] = v_i - p_i + \lambda(\sigma_T u_{ia} + \sigma_N \varepsilon_{ia})$, where λ is the traditional signal-extraction factor: $\lambda = 1/\left(1 + \sigma_{e,N}^2/\sigma_{e,T}^2\right)$ with $\sigma_{e,N}^2/\sigma_{e,T}^2$ the ratio of noise to true differentiation, as perceived by the customer. The decision utilities are therefore:

$$E[V_{ia} \mid U_{ia}, p_i] = v_i - p_i + \nu w_{ia} \tag{57}$$

$$\nu = \lambda\sqrt{\sigma_T^2 + \sigma_N^2} \tag{58}$$

and w_{ia} is a standard normal. The treatment of the paper follows, with the effective noise ν. For instance, one has, in the Bertrand case: $p - c = b_n \nu$ where b_n is the value calculated for the Gaussian noise.

10.2.1 Case of Market-Unspecific Expectations

If consumers do not have a theory of the market specific noise, they will use a signal-extraction λ that is independent of σ_N^2/σ_T^2. Then, we have $p - c = b_n \lambda\sqrt{\sigma_T^2 + \sigma_N^2}$. So that again, the markup

[31] Anantham and Ben-Shoham (2004) examine the strategy of a monopolist with Bayesian consumers.

[32] We thank Avinash Dixit for providing us with a simple sufficient condition. It is that the distribution of V_{ia} should undergo a first-order stochastic dominant shift to the right as U_{ia} increases. So if $F(V_{ia} \mid U_{ia})$ is the cumulative distribution function of V_{ia} given U_{ia}, the condition is $\partial F(V_{ia} \mid U_{ia})/\partial U_{ia} \leq 0$. Here all firms have the same amount of noise σ.

will increase as a function of the noise σ_N, and the conclusions of our analysis go through.

10.2.2 Case of Market-Specific Expectations

An example where confusion decreases profits under rational expectations In this case $\sigma_{e,N}^2/\sigma_{e,T}^2 = \sigma_N/\sigma_T^2$. Then, (57) gives

$$\nu = \sigma_T^2/\sqrt{\sigma_T^2 + \sigma_N^2}. \tag{59}$$

Thus, when the noise σ_N increases, the markup decreases. To understand this effect and assess its robustness, consider the case where $\sigma_T = 0$, $\sigma_N > 0$. There is no true heterogeneity in quality, so consumers should just use the rule "pick the lowest price, irrespective of quality," and the markup is 0, as in (59). This is not, however, a stable equilibrium once we allow for entry with arbitrary qualities. Firms would take advantage of this rule to enter with very low quality products, and so the rule "pick the lowest price" would not be valid any more. Consumers would have also to pay attention to the perceived quality of the product. Here the consumer noise σ_N would determine the dispersion in the quality of the product. We will not present a general analysis of the case with heterogenous qualities. Instead, we present an example where confusion increases profits.

An example where confusion increases profits under rational expectations There are three firms indexed by $i = 1, 2, 3$. The qualities are $q_1 = q_2 = 6$, $q_3 = 0$, and the unit costs are $c_1 = c_2 = 2$, $c_3 = 1$. The consumer will buy at most one good. He has utility $q^e - p$ if he buys one good with expected value q^e and price p, and 0 otherwise. Firm i can post a price $p_i \geq 0$, or exit the market.

If there is no confusion, then in equilibrium $p_1 = p_2 = 2$ and firm 3 can post any positive price but does not sell anything.[33] We have a Bertrand equilibrium where the aggregate profit is

$$\pi^{\text{No Confusion}} = \sum_{i=1}^{3} \pi_i = 0. \tag{60}$$

For simplicity, to study the effects of confusion we now consider the case of infinite confusion.

[33] We could postulate that firm 3 has to pay a small positive price to stay in the market. Then, there is unique equilibrium where firm 3 exits the market, firms 1 and 2 stay in the market, and $p_1 = p_2 = 2$.

We assume that consumers cannot evaluate qualities at all, though they can see prices. We assume rational expectations. The equilibrium decision rule picks a product randomly from the set of goods with the lowest price. Call the equilibrium price p. We shall see that $p = 2$. Indeed, if $p > 2$, there is an incentive for firm 1 or 2 to charge a slightly lower price $p' = p - \delta$, $\delta > 0$ small; it would get the whole market and thus see a discrete increase its profit from $(p - 2)/3$ to $(p - 2 - \delta)$. So, we cannot have $p > 2$.

If $p < 2$ firms 1 and 2 make negative profits $(p - 2)/3$ if they are still in the market. So the market is just composed of firm 3, and the rational expectation on the quality is $q^e = 0$. To get non-negative surplus $q^e - p$, we need $p = 0$, but then firm 3 makes losses as $c_3 > 0$, and thus should exit the market. So $p < 2$ is not possible either. We conclude that the equilibrium is characterized by $p = 2$. The rational expectation of the quality is $q^e = \left(\sum_{i=1}^{3} v_i\right)/3 = 4$, which yields a positive surplus for the consumer. It is easy to see that no firm has an incentive to deviate. Each firm has 1/3 of the market, and the aggregate profits are:

$$\pi^{\text{Full Confusion}} = \sum_{i=1}^{3} \pi_i = \sum_{i=1}^{3} \frac{1}{3} (p - c_i) = \frac{1}{3}. \tag{61}$$

The aggregate profit under full confusion, (61), is higher than the profit under no confusion (60). The economic origin of this is that the low quality firm takes advantage of the confusion to get some market share and obtain positive profits. The two high quality firms still are at a Bertrand, zero-profit state ($p = c_1 = c_2 = 2$), as in the situation with no confusion.

We conclude that the phenomenon of confusion-induced increases in profits is consistent with rational expectations in at least some cases. In our second example confusion decreases profits. It would be useful to have a complete characterization of rational expectation equilibria with confusion. The general analysis would be rather involved, including analysis of non-symmetric Bayesian equilibria and some conceptually problematic equilibrium selection issues. Such a characterization is a mathematically challenging topic for further research.

11 References

Alexander, Gordon, Jonathan D. Jones, and Peter J. Nigro (1998). "Mutual Fund Shareholders: Characteristics, Investor Knowledge, and Sources of Information." *Financial Services Review*, 7, 301-16.

Anantham, Siva and Assaf Ben-Shoham (2004) "Quality Uncertainty and Monopolistic Pricing," Harvard Mimeo.

Anderson, Simon, André de Palma and Jacques-Francois Thisse, *Discrete Choice Theory of Product Differentiation*, MIT Press, 1992.

Anderson, Simon, Jacob Goeree and Charles Holt, "Rent Seeking with Bounded Rationality: An Analysis of the All-Pay Auction," *Journal of Political Economy* (1998), 828-853.

Barber, Brad M., Terrance Odean, and Lu Zheng (2002). "Out of sight, out of mind: The effects of expenses on mutual fund flows." Mimeo, U.C. Davis.

Bénabou, Roland, and Robert Gertner "Search with Learning from Prices: Does Increased Inflationary Uncertainty Lead to Higher Markups?," *Review of Economic Studies*, 60 (1993) 69-95.

Benartzi, Shlomo, and Richard Thaler, "Naive Diversification Strategies in Retirement Saving Plans," *American Economic Review* 91 (2001) 79-98.

Benartzi, Shlomo, and Richard Thaler, "How Much Is Investor Autonomy Worth?" *Journal of Finance* 57 (2002), pp. 1593-1616.

Bertrand, Marianne, Dean Karlan, Sendhil Mullainathan, Eldar Shafir and Jon Zinman (2004) "Pricing Psychology: A Field Experiment". Mimeo.

Caplin, A., and B. Nalebuff, "Aggregation and Imperfect Competition," *Econometrica*, 1991, 25-60.

Choi, James, Xavier Gabaix, David Laibson, and Brigitte Madrian (2004). "Measuring Noise." Harvard mimeo.

Cronqvist, Henrik and Richard H. Thaler (2004) "Design Choices in Privatized Social Security Systems: Learning from the Swedish Experience," *American Economic Review (Papers and Proceedings).*

DellaVigna, Stefano, and Ulrike Malmendier (2002) "Overestimating Self-Control: Evidence from the Health Club Industry," mimeo.

De Palma, André, Gordon M. Myers, and Yorgos Y. Papageorgiou, "Rational Choice Under an Imperfect Ability To Choose," *The American Economic Review*, Vol. 84, No. 3. (Jun., 1994), pp. 419-440.

Diamond, Peter, "A Model of Price Adjustment," *Journal of Economic Theory*, 1971, 156-68.

Duflo, Esther and Emmanuel Saez, "The Role of Information and Social Interactions in Retirement Plan Decisions: Evidence from a Randomized Experiment," *Quarterly Journal of Economics*, Vol. 118 (3), 2003.

Ellison, Glenn and Drew Fudenberg, "Rules of Thumb for Social Learning," *Journal of Political Economy* 101 (1993), 612-643.

Embrechts, P., Kluppelberg, C., and Mikosch, T. *Modelling Extremal Events* (Springer Verlag, New York, 1997).

Gabaix, Xavier, Parameswaran Gopikrishnan, Vasiliki Plerou, H. Eugene Stanley "A Theory of Power Law Distributions in Financial Market Fluctuations," *Nature*, 423 (2003), 267–30.

Gabaix, Xavier, David Laibson (2004) "Shrouded Attributes and Information Suppression in Competitive Markets," Harvard and MIT Manuscript.

Glaeser, Edward (2003) "The Political Economy of Hatred," Harvard mimeo.

Goldberg, Pinelopi. "Product Differentiation and Oligopoly in International Markets: The Case of the U.S. Automobile Industry," *Econometrica*, July 1995, pp. 891-951.

Hall, Robert (1997) "The Inkjet Aftermarket: An Economic Analysis," Stanford mimeo.

Heidhues, Paul and Botond Koszegi. "Loss Aversion, Price Stability, and Sales". Berkeley manuscript (2004).

Hortacsu, Ali and Chad Syverson (2004) "Product Differentiation, Search Costs and Competition in the Mutual Fund Industry: A Case Study of S&P 500 Index Funds," *Quarterly Journal of Economics.*

Investment Company Institute, *Mutual fund fact book*, 43rd edition (2003).

Judd, Kenneth and Mike Riordan (1994), "Price and Quality in a New Product Monopoly," *Review of Economic Studies*, 61, 773-789.

Kahneman, Daniel, and Amos Tversky, "Prospect Theory: An Analysis of Decision under Risk", *Econometrica* 47 (1979), 263-291.

Koszegi, Botond and Rabin, Matthew (2003) "Modeling Reference-Dependent Preferences," Berkeley mimeo.

Luce, R. D., *Individual Choice Behavior*, New York: John Wiley, 1959.

Madrian, Brigitte C. and Dennis F. Shea (2001), "The Power of Suggestion: Inertia in 401(k) Participation and Savings Behavior," *Quarterly Journal of Economics*, 116(4), p.1149-1187.

McFadden, D. "Econometric models of probabilistic choice." in C. Manski and D. McFadden, eds., *Structural analysis of discrete data*, pp. 198-272, Cambridge, MIT Press, 1981.

Mullainathan, Sendhil and Andrei Shleifer (2003) "The Market for News". NBER WP.

Nevo, Aviv. "Measuring Market Power in the Ready-to-Eat Cereal Industry," *Econometrica*, 69(2), 307-342, 2001.

Oster, Sharon M. and Fiona M. Scott Morton (2004) "Behavioral Decision-Making: An Application to the Setting of Magazine Subscription Prices". Yale mimeo.

Perloff, Jeffrey M. and Steven C. Salop. "Equilibrium with Product Differentiation," *The Review of Economic Studies*, Vol. 52, No. 1. (Jan., 1985), pp. 107-120.

Reiss, R., *Approximate Distributions of Order Statistics*, (Berlin, Germany: Springer Verlag, 1989).

Resnick, Sidney. *Extreme Values, Regular Variation, and Point Processes*. Springer Verlag, 1987.

Rosenthal, Robert "A Model in which an Increase in the Number of Sellers Leads to a Higher Price," *Econometrica*, 48(6), p.1575-9, 1980.

Salop, Steven and Stliglitz, Joseph "Bargains and Ripoffs: A Model of Monopolistically Competitive Price," *Review of Economic Studies*, 44(3), 1977, pp.493–510.

Sheshinsksi, Eytan (2003) "Optimal Policy to Influence Individual Choice Probabilities," Manuscript.

Spector, David, "The Noisy Duopolist," *Contributions to Theoretical Economics*, 2(1), 2002, Article 4.

Spiegler, Ran (2003) "The Market for Quacks," mimeo, Tel Aviv University.

Stein, Jeremy (2003) "Why Are Most Funds Open-End? Competition and the Limits of Arbitrage". Harvard mimeo.

Woodward, Susan (2003) "Consumer Confusion In the Morgage Market," Sand Hill Econometrics resesarch paper.