Extreme Value Theory and the Effects of Competition on Profits

Xavier Gabaix
MIT and NBER

David Laibson
Harvard University and NBER

Hongyi Li
NBER

Current Draft: March 7, 2005*

Abstract

It is likely that consumers make some errors and have at least some noise in their calculations. Using the framework of Perloff and Salop (1985), we ask how such noise effects markups. Because of problems with analytical tractability, past research has only analyzed special cases. To study the general case, we use results from extreme value theory. We show that markups are asymptotically proportional to $\left(nF'\left[F^{-1}\left(1-n^{-1}\right)\right]\right)^{-1}$, where n is the number of competing firms, and F is the distribution function for noise. We show that the asymptotic markup is proportional to the expected gap between the highest draw and second highest draw in a sample of n draws. This formula implies that for realistic distributions of noise, the markup is insensitive to the number of firms. For example, for the Gaussian case asymptotic markups are proportional to $1/\sqrt{\ln n}$, implying a zero asymptotic elasticity of the markup with respect to the number of firms. We show that these results generalize to cases with bounded support. For realistic noise distributions, competition only produces weak pressure on prices.

JEL classification: D00, D80.

Keywords: extreme value theory, discrete choice theory.

---

*For useful suggestions we thank Simon Anderson, Abhijit Banerjee, Roland Bénabou, Douglas Bernheim, Andrew Caplin, Victor Chernozhukov, Casper de Vries, Avinash Dixit, Glenn Ellison, Paul Embrechts, Edward Glaeser, Robert Hall, Sergei Izmalkov, Filip Lindskog, Sidney Resnick, Nancy Rose, José Scheinkman, Andrei Shleifer, Ximing Wu, 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, the 2004 SEDS 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. Hongyi Li: NBER, 1050 Massachusetts Ave., Cambridge, MA, 02138, hongyi@nber.org.

1 Introduction

Consumers may not know the exact value of the products they buy (Luce 1959, McFadden 1981). Even the most sophisticated consumers make small errors and hence have at least a little noise in their calculations. Motivated by such consumer errors or by heterogeneity in true preferences, economists have included noise in models of consumer choice. Perloff and Salop (1985)[1] derive a closed-form expression for equilibrium markups in a random utility framework. Their expression includes integrals that are generally analytically intractable. Drawing from extreme value theory, the current paper develops tools that solve this tractability problem.

Analysis of the Perloff-Salop model has focused on a small number of tractable cases in which markups turn out to be either unresponsive to competition or highly responsive to competition. When noise has an exponential density or a logit (i.e. Gumbel) density, markups converge to a strictly positive value. Hence, asymptotic markups have a zero elasticity with respect to n, the number of competing firms in an industry (Perloff and Salop 1985, Anderson et al. 1992). However, when noise is uniformly distributed, markups are proportional to $1/n$, so markups have a unit elasticity and hence a strong negative relationship with n (Perloff and Salop 1985).

All three of these illustrative distributions — exponential, logit, and uniform — are appealing for their analytic tractability rather than their realism. The exponential and logit cases have relatively fat tails while the uniform case has no tails. *We would like to know how markups respond to competition when the noise follows more general distributions, including the Gaussian.*

In this paper we apply extreme value theory to develop an asymptotic result that can be used to analytically characterize the equilibrium consequences of general noise distributions (Proposition 4 and 2). As in Perloff and Salop (1985), our analysis is applicable whether the noise reflects consumer errors or heterogeneity in true preferences. We show that markups are asymptotically proportional to $1/\left(nF'\left[F^{-1}\left(1-n^{-1}\right)\right]\right)$, where F is the distribution function for noise. Moreover, we show that this turns out to be a limit pricing result. For the most important class of distributions, the Perloff-Salop markup is asymptotically equal to the expected gap between the highest draw and second highest draw in a sample of n draws.

---

[1] In this model n identical firms pick prices. Then consumers with i.i.d. taste shocks buy from the firm that offers the highest perceived net surplus. Perloff and Salop analyze the symmetric equilibrium and express the equilibrium markup as a function of the number of competing firms, n, and the density function of consumer noise.

We pay particular attention to the Gaussian case because it is a good approximation of natural phenomena. For the Gaussian case we show that asymptotic markups are proportional to $1/\sqrt{\ln n}$, where n is the number of competing firms. This formula implies that mark-ups fall extremely slowly as n rises. The elasticity of the markup with respect to n converges to 0. Hence, the Gaussian case turns out to behave much more like the exponential and logit cases than like the uniform case. Our analysis implies that rising competition in an environment with a Gaussian noise distribution will only produce weak downward pressure on prices.

The tools that we develop also enable us to characterize markups for several other distributions including two fat-tailed cases — log-normal and power-law — in which mark-ups *increase* as the number of competing firms increase.

Finally, we show that our results do not depend on the unbounded tails of the distributions that we study. Our results are preserved when we truncate these distributions, as long as the truncation point is large.

We conclude that markups associated with noise are remarkably robust: they do not decline rapidly as competition increases. The rest of this paper formalizes these claims. Section 2 presents the Perloff-Salop model and our extreme value results. Section 3 applies these results to derive markups for nine noise distributions. Section 4 discusses extensions including truncation and a formal statement of our limit pricing result. Section 5 concludes.

2 The main result

In the Perloff-Salop (1985) model identical firms pick prices and consumers with i.i.d. shocks — "noise" — choose among the firms. This noise could either represent true taste shocks or evaluation errors.

To establish notation, assume that firm i picks price p_i. Assume that a particular consumer receives net utility $\sigma\varepsilon_i - p_i$ by purchasing the good of firm i, where ε_i is i.i.d. across firms and consumers. Without loss of generality, ε_i has zero mean and unit standard deviation.

In a symmetric-price equilibrium,[2] the demand function of firm i is the probability that the

---

[2] If the logarithm of the density of ε is concave, the existence of the equilibrium is ensured by Caplin and Nalebuff (1991), Theorem 2 and Proposition 7. The question of the equilibrium when the density is not concave is an open one. The Technical Appendix to this paper, available on the authors' web page, discusses the existence of the equilibrium

consumer's surplus at firm i, $\sigma\varepsilon_i - p_i$, exceeds the consumer's surplus at all other firms, which charge $p_j = p$,

$$\mathcal{D}(p_i, p) = P\left(\sigma\varepsilon_i - p_i > \max_{j\neq i} \sigma\varepsilon_j - p\right) = P\left(p - p_i > \max_{j\neq i} \sigma\varepsilon_j - \sigma\varepsilon_i\right) \equiv D(p - p_i).$$

The simplified demand function, D, takes as its argument the average surplus x of firm i relative to its competitors. Firms maximize profit, π_i, by setting their price equal to $\arg\max_{p_i} \pi_i \equiv (p_i - c)\,\mathcal{D}(p_i, p)$, where c is the marginal cost of production. Perloff and Salop (1985) show that this (normed) equilibrium markup is

$$\begin{aligned} p - c &= \mu_n \sigma \qquad (1) \\ \mu_n &= \frac{1}{n(n-1)\int F(x)^{n-2} f(x)^2\, dx}, \qquad (2) \end{aligned}$$

where n is the number of firms.

To interpret the Perloff-Salop markup equation, call M_{n-1} the largest of $n-1$ noise realizations: $M_{n-1} \equiv \max_{j\in\{1,\ldots,n\}, j\neq i} \varepsilon_i$. Then, $D(x) = P\left(\varepsilon_i > \frac{-x+\sigma M_{n-1}}{\sigma}\right)$, so

$$D(x) = E\left[\overline{F}\left(\frac{-x}{\sigma} + M_{n-1}\right)\right], \qquad (3)$$

where $\overline{F}(x) = \int_x^\infty f(y)\, dy$ is the countercumulative distribution function. This formulation emphasizes that the demand for good i is driven by the properties of the right-hand tail of the countercumulative distribution function, $\overline{F}$. We can also confirm that the Perloff-Salop markup is $p - c = D(0)/D'(0)$.

The properties of the symmetric equilibrium can be derived from the behavior of $D(x)$ at $x = 0$. Specifically, (3) gives:

Lemma 1 *In a symmetric Bertrand equilibrium,*

$$p - c = \frac{\sigma}{nE[f(M_{n-1})]} \qquad (4)$$

---

for the distributions used in this paper that are not log-concave: the lognormal distribution and the unbounded power law distributions.

where M_{n-1} *is a random variable with cumulative density function* $P(M_{n-1} \le x) = F(x)^{n-1}$. *This is a rewriting of Perloff-Salop (1985)'s formula (1).*

Now we simply need an asymptotic expression for $E[f(M_{n-1})]$. It turns out that we can prove a more general result. We will refer to the class S of well-behaved functions, which is characterized in Definition 14. $j(.) \in S$ means that $\forall u > 0, \lim_{t \to 0} j(ut)/j(t) = u^{\rho}$ (written as $j(t) \in RV_{\rho}^{0}$) for some $\rho > -1$, and that $j(.)$ is bounded on $(\varepsilon, 1)$ for all $\varepsilon \in (0,1)$.

Proposition 2 *Suppose that* $j(t) = J\left(\overline{F}^{-1}(t)\right) \in S \cap RV_{\rho}^{0}$. *Let* $W_{n-r:n-1}$ *be the r-th largest realization of* $n-1$ *i.i.d. random variables with CDF* F. *Set* $A_n = \overline{F}^{-1}(1/n)$, *where* $\overline{F}(x) \equiv 1 - F(x)$. *Then:*

$$E[J(W_{n-r:n-1})] \sim \frac{J(A_n)\Gamma(\rho + r)}{(r-1)!} \tag{5}$$

and

$$\frac{dE[J(W_{n-r:n-1})]}{dn} \sim \frac{J'(A_n)\Gamma(\rho + r)}{f(A_n)n^2(r-1)!}. \tag{6}$$

Proof. See Appendix B. ■

The asymptotic Perloff-Salop markup follows as a simple corollary of Proposition 2.

Let us first define the class of regular distributions. Essentially all continuous distributions used in economics belong to this class, including the exponential, Fréchet, Gaussian, Gumbell, loggamma, lognormal, pareto, uniform, and Weibull (Embrechts et al. 1997, p.153-7).

Definition 3 *The distribution function* F *is in the class of regular distributions (written* $F \in R$*) iff the following conditions are satisfied:*

1. f *is differentiable in a neighborhood of* $F^{-1}(1)$, *and the following characteristic index of distribution* F *exists and is finite:*

$$\xi = \lim_{x \to F^{-1}(1)} \left(\overline{F}/f\right)'(x). \tag{7}$$

2. f *is bounded.*

Under such a restriction, the asymptotic price markup can be characterized:

Proposition 4 *Suppose that F is a regular distribution with characteristic index ξ. In a symmetric Bertrand equilibrium:*

$$\begin{aligned} p-c &= \mu_n \sigma \text{ with} \\ \mu_n &\sim \frac{1}{nf\left(A_n\right)\Gamma\left(2+\xi\right)} \end{aligned} \tag{8}$$

where $A_n = F^{-1}\left(1-1/n\right)$ and Γ is the Gamma function.

Proof. See Appendix B. ■

Proposition 4 yields useful formulae, since the key mathematical objects, A_n, $f\left(A_n\right)$, and ξ are easy to calculate for most distributions of interest.

There is an intuitive interpretation for Proposition 4. First, recall that M_{n-1} is the maximum value of $n-1$ draws. 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 suggests that if we define[3]

$$A_n \equiv \overline{F}^{-1}\left(1/n\right), \tag{9}$$

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

Call S_{n-1} the second-highest draw. $E\left[\overline{F}\left(S_{n-1}\right)\right] = 2/n$, so it is likely that $S_{n-1} \simeq \overline{F}^{-1}\left(2/n\right)$. Intuitively, to set its optimum price, a firm conditions on its getting the largest draw, then evaluates the likely draw of the second highest firm, and engages in limit pricing, where it charges a markup equal to the difference between its draw and the next highest draw. This heuristic reasoning suggests:

$$\begin{aligned} p-c &\simeq M_{n-1} - S_{n-1} \simeq \overline{F}^{-1}\left(1/n\right) - \overline{F}^{-1}\left(2/n\right) = \overline{F}^{-1}\left(1/n\right) - \overline{F}^{-1}\left(1/n + 1/n\right) \\ &\simeq -\frac{d\overline{F}^{-1}\left(x\right)}{dx}_{|x=1/n} \cdot \frac{1}{n} \text{ by Taylor expansion} \\ &= \frac{1}{nf\left(A_n\right)} \end{aligned}$$

Proposition 4 shows that the heuristic argument generates the right approximation for the

---

[3] We use the usual convention (see Resnick 1987) that $F^{-1}\left(t\right) = \inf\left\{x : F\left(x\right) \geq t\right\}$, and $\overline{F}^{-1}\left(t\right) = F^{-1}\left(1-t\right)$.

Gaussian, logit (Gumbel), exponential, and lognormal distributions, and that the approximation remains accurate up to a corrective constant $\Gamma(2+\xi)$ in other cases. A rigorous counterpart of this heuristic reasoning can be found in Proposition 8.

In addition, Proposition 4 shows that ξ has a concrete economic implication: it is the asymptotic elasticity of the markup with respect to the number of firms. In other terms, the markup behaves as $\mu_n \simeq kn^{\xi}$. We interpret n as a continuous variable in the expression of the markup, Eq. 2.

Proposition 5 *If F is a regular distribution with characteristic index $\xi = \lim_{x\to F^{-1}(1)} \left(\overline{F}/f\right)'(x)$, the asymptotic elasticity of the markup with respect to the number of firms is:*

$$\lim_{n\to\infty} \frac{n}{\mu_n}\frac{d\mu_n}{dn} = \xi. \tag{10}$$

The proof is in Appendix B.

There is an interesting consequence. We can call "regular distributions with a declining right tail" regular distributions such that $f'(x) \leq 0$ for x large enough, i.e. such that the density of the right tail is weakly declining. As $\left(\overline{F}/f\right)'(x) = -1 - \frac{\overline{F}}{f^2}f'(x) \geq -1$, this implies that $\xi = \lim_{x\to F^{-1}(1)} \left(\overline{F}/f\right)'(x) \geq -1$. Hence Proposition 5 implies that for distributions with a declining right tail, the mark-up μ_n falls more slowly than $1/n$. This is a sense in which the uniform density case is an extreme case: for distributions with a declining right tail, the markup declines (weakly) more slowly than for the uniform density. Actually, Lemma 13 shows that all regular distributions have $\xi \geq -1$.

3 Noise distributions and Markups

To analyze the impact of competition on markups, we examine the equilibrium markup for various noise distributions. We consider nine well-studied distributions. For our application, some of the distributions need to be shifted to produce a zero expected value. To simplify notation, we omit the shift term in the following equations.

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

$$f_{\text{Uniform}}(\varepsilon) = \frac{1}{2}1_{|\varepsilon|<1}. \tag{11}$$

This generalizes to a density in $[0,-1]$ that has a power law distribution near 0^-

$$f_{\text{Bounded power law}}(\varepsilon)=\alpha(-\varepsilon)^{\alpha-1}, \tag{12}$$

with $\alpha\geq 1$. Another paradigmatic example is the Weibull distribution, defined in $(-\infty,0)$ with $\alpha\geq 1$

$$f_{\text{Weibull}}(\varepsilon)=\alpha(-\varepsilon)^{\alpha-1}e^{(-\varepsilon)^{\alpha}}. \tag{13}$$

We also consider the Gaussian density,

$$f_{\text{Gaussian}}(\varepsilon)=\frac{1}{\sqrt{2\pi}}e^{-\varepsilon^2/2}, \tag{14}$$

the Gumbel density, which is also known as the logit density,

$$f_{\text{Gumbel}}(\varepsilon)=\exp\left(-e^{-\varepsilon}-\varepsilon\right), \tag{15}$$

the exponential density,

$$f_{\text{Exponential}}(\varepsilon)=e^{-\varepsilon}1_{\varepsilon>0}, \tag{16}$$

the log-normal density,

$$f_{\text{Lognormal}}(\varepsilon)=\frac{1}{\varepsilon\sqrt{2\pi}}e^{-\ln^2(\varepsilon)/2}1_{\varepsilon>0}, \tag{17}$$

and the power law[4] density on $[1,\infty)$

$$f_{\text{Power law }\zeta}(\varepsilon)\sim\zeta\varepsilon^{-\zeta-1}. \tag{18}$$

Another type of power law distribution is the Fréchet distribution, defined on $[0,\infty)$

$$f_{\text{Fréchet}}(\varepsilon)=\zeta\varepsilon^{-\zeta-1}e^{-\varepsilon^{-\zeta}}. \tag{19}$$

---

[4] From an empirical perspective, we do not know whether the fat-tailed case is relevant. We speculate that it might apply in markets with fat tailed distribution of sales – for instance, the book market. See Chevalier and Goolsbee (2004) and Sornette et al. (2003). Movies (De Vany 2004) also have power law distributions. Power laws generally arise in markets where word of mouth creates snowballing effects (Simon 1955, Gabaix 1999, and the survey in Gabaix and Ioannides 2003).

The densities are ranked from thinnest to fattest tails.[5]

We calculate the Bertrand outcome for the nine distributions discussed above. Table 1 reports values for the key ingredients in our calculations.[6] 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). Some of our calculations are asymptotic expansions, which hold for large n and small positive t. For application of Proposition 4, 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 and Weibull	$-n^{-1/\alpha} + o\left(n^{-1/\alpha}\right)$	$\sim \alpha t^{1-1/\alpha}$	$-1/\alpha$
Gaussian	$\sim \sqrt{2\ln n}$	$\sim t\sqrt{2\ln\frac{1}{t}}$	0
Logit (Gumbel)	$\sim \ln n$	$\sim t$	0
Exponential	$\ln n$	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 and Fréchet	$\sim n^{1/\zeta}$	$\sim \zeta t^{1+1/\zeta}$	$1/\zeta$

We now show how markups change as competition intensifies. Proposition 6 provides closed form expressions for the markups in different distributional cases for fixed σ and a fixed number of competitors, n.

Proposition 6 *The Bertrand equilibrium generates the following markups. For uniform noise (11),*

$$p - c = \frac{2}{n}\sigma. \tag{20}$$

---

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

[6] The proof is a consequence of e.g. Embrechts et al. (1997, p.155-7) and simple calculations.

For bounded power law noise (12) with $\alpha \geq 1$,

$$p - c = \frac{\Gamma\left(1 - 1/\alpha + n\right)}{\alpha\Gamma\left(2 - 1/\alpha\right)\Gamma\left(1 + n\right)}\sigma \sim \frac{1}{\alpha\Gamma\left(2 - 1/\alpha\right)} n^{-1/\alpha}\sigma. \tag{21}$$

For Weibull noise (13) with $\alpha \geq 1$,

$$p - c = \frac{1}{\alpha\Gamma\left(2 - 1/\alpha\right)} \frac{n^{1-1/\alpha}}{n - 1}\sigma \sim \frac{1}{\alpha\Gamma\left(2 - 1/\alpha\right)} n^{-1/\alpha}\sigma \tag{22}$$

For Gaussian noise (14),

$$p - c \sim \frac{1}{\sqrt{2 \ln n}}\sigma. \tag{23}$$

For Gumbel noise (15),

$$p - c = \frac{n}{n - 1}\sigma. \tag{24}$$

For exponential noise (16),

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

For log-normal noise (17),

$$p - c \sim e^{\sqrt{2 \ln n} - \frac{1}{2}\ln(2 \ln n)}\sigma. \tag{26}$$

For power-law noise (18) with exponent $\zeta > 1$,

$$p - c = \frac{\Gamma\left(1 + 1/\zeta + n\right)}{\zeta\Gamma\left(2 + 1/\zeta\right)\Gamma\left(1 + n\right)}\sigma \sim \frac{1}{\zeta\Gamma\left(2 + 1/\zeta\right)} n^{1/\zeta}\sigma. \tag{27}$$

For Fréchet noise (19) with exponent $\zeta > 1$,

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

The distributions in Proposition 6 are presented in increasing order of fatness of the tails. For the uniform distribution, which has the thinnest tails, the markup is proportional to $1/n$. This is the same equilibrium markup generated by the Cournot model. However the uniform/Cournot case is unrepresentative of the general picture. Proposition 6 implies that markups scale with $n^{-1/\alpha}$. For the distributions reported in Proposition 6 ξ is bounded above by one, so the uniform case is an extreme case.

For the distributions with the fattest tails, the markups paradoxically[7] *rise* as the number of competitors *increases*. 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 scale with $n^{1/\zeta}$ but sales per firm are proportional to $1/n$.

Thin-tailed distributions (e.g., uniform) and fat-tailed distributions (e.g., power-laws) are the extreme cases in Proposition 6. 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 $1/\sqrt{\ln n}$, and the elasticity of the markup with respect to n is $-1/\ln n$. So $p-c$ converges to 0, but this convergence proceeds at a glacial pace. Indeed, the elasticity of the markup with respect to n converges to zero.

To illustrate the slow convergence, we normalize the markup at $n=10$ to be 1 and calculate the markup as the number of competitors expands by factors of 10. Table 2 shows that a highly competitive industry with $n=1,000,000$ firms will retain a third of the markup of a highly concentrated industry with only $n=10$ competitors. We also compare markups in the Perloff-Salop model to those in the Cournot model, which features markups proportional to $1/n$ and a markup elasticity w.r.t. n of -1.

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

---

[7] See Bénabou and Gertner (1993), Rosenthal (1980), Spector (2002) for perverse competitive effects generated by different microfoundations.

noise and uniform noise (Cournot competition).

n	Markup with Gaussian noise	Markup with Uniform noise
10	1.00	1.00
100	0.61	0.1
$1,000$	0.47	0.01
$10,000$	0.40	0.001
$100,000$	0.35	0.0001
$1,000,000$	0.32	0.00001

We normalize the markup for $n = 10$. We integrate numerically Eq. (4). The asymptotic result (23) provides a good approximation for these exact results.

In cases with moderate fatness, such as the Gumbel (i.e., logit), exponential, and log-normal densities, the markup again shows little (or no) response to changes in n. Finally, the case of *bounded* power law noise (21) shows that an infinite support is not necessary for our results. In this case the markup is proportional to $1/n^{1/\alpha}$ and markup decay is slow for large α. In section 4 we show that *all* of our results can be reformulated for truncated distributions.

In practical terms, these results imply that in markets with noise we should not necessarily expect increased competition to dramatically reduce markups. The mutual fund industry may exemplify such stickiness. Currently 10,000 mutual funds are available in the U.S. and many of these funds offer similar portfolios. Even in a narrow class of homogenous products, such as medium capitalization value stocks or S&P 500 index funds it is normal to find 100 or more competing funds (Hortacsu and Syverson 2004). 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 homogeneous competing funds has increased by a factor of 10 over the past several decades.

4 Discussion and Extensions

4.1 Truncated distributions

This section will show that the assumption of unbounded support is *not* necessary for the property that the elasticity of the markup with respect to n may be small. We have already analyzed one case — bounded power law noise (21) — that illustrates this point. We extend the analysis to truncated distributions.

Intuitively, truncation need not matter since our markup calculations in Proposition 6 depend only on the properties of the density in a neighborhood of $A_n \equiv \overline{F}^{-1}(1/n)$. We would expect that the same equilibrium markups will apply to truncated noise distributions as long as n is large *and* the truncation point X is chosen such that $\overline{F}(X) \ll 1/n$. This section shows that this intuition is correct.

To formalize this, we define $F^{\wedge X}(x)$ to be the distribution F truncated on the right at point X, i.e. $F^{\wedge X}(x) = \min\left(F(x)/F(X), 1\right)$. We set an increasing series of lower bounds X_n which satisfy $n\bar{F}(X_n) \to 0$. Condition $n\bar{F}(X_n) \to 0$ is, in a sense, the broadest lower bound that allows convergence. If $n\bar{F}(X) \to \delta > 0$, convergence does not hold in general.[8]

Proposition 7 *Let F be a regular distribution with characteristic index ξ. We define $F^{\wedge X}$ to be the distribution F truncated on the right at point X, i.e. $F^{\wedge X}(x) = \min\left(F(x)/F(X), 1\right)$. Call $\mu_n(X)$ the Perloff-Salop markup (Eq. 2) with noise distribution $F^{\wedge X}$. Consider X_n s.t. $\overline{F}(X_n)\,n \to 0$. Then, $\mu_n(X_n)$ is asymptotically equal to the untruncated markup of Proposition 4:*

$$\mu_n(X_n) \sim \frac{1}{nf(A_n)\Gamma(2+\xi)} \tag{29}$$

The convergence is uniform for $X \geq X_n$:

$$\forall \eta > 0, \exists n_0 : \forall n > n_0, \forall X \geq X_n, \left|\mu_n(X)\, nf(A_n)\Gamma(2+\xi) - 1\right| \leq \eta. \tag{30}$$

Concretely, Proposition 7 means that Proposition 4 holds even for distributions truncated to the

---

[8] Consider $F(x) = 1 - e^{-x}$, so that $F^{\wedge X}(x) = (1 - e^{-x})/\left(1 - e^{-X}\right)$ for $x \in [0, X]$. By direct calculation $\mu_n(X_n) = \left[1 - e^{-X_n}\right]/\left[1 + (n-1)e^{-X_n}\right]$. So if $n\overline{F}(X_n) \to \delta$, then $\lim \mu_n = 1$ iff $\delta = 0$.

right by some X, as long as $n\overline{F}(X)$ is small. This type of "intermediate asymptotics" where a result holds in the domain "$1 \ll n \ll 1/\overline{F}(X)$" is used very often in mathematical physics (Barenblatt 1996), but less in economics.

For instance, suppose we take Gaussian noise truncated at $X = 6$, i.e. truncated six standard deviations into the tails. Proposition 7 suggests the markup will be essentially the same for the bounded and the unbounded Gaussian as long as $1 \ll n \ll 1/\overline{F}(X) \simeq 10^9$. Indeed, we verified numerically that the markup was the same, to a difference of less than 0.1%, for $n = 10, 10^2, ..., 10^6$.

4.2 The limit pricing interpretation of our asymptotic approximation

Let M_n and S_n represent the highest and second highest draws of n i.i.d. signals. We represent the expected difference between the two as $\delta_n \equiv E[M_n - S_n]$. The following Proposition characterizes the expected value of this gap and shows that it is equal to the normed markup in the Perloff-Salop model for distributions with $\xi = 0$ (i.e., the Gaussian, Gumbel, exponential, and log-normal distributions).

Proposition 8 *Call $\delta_n \equiv E[M_n - S_n]$ the expected value of the difference between the largest value and the second value of n i.i.d. draws. Assume that F is regular. Then, for large n,*

$$\delta_n \sim \Gamma(1-\xi)\Gamma(2+\xi)\mu_n \sim \frac{\Gamma(1-\xi)}{nf(A_n)}. \tag{31}$$

In particular, for the Gaussian, Gumbel (i.e., logit), exponential, and log-normal distributions $\delta_n \sim \mu_n$.

This result has an interpretation in auction theory. Consider a second-price auction with n buyers who have independent valuations ε. The winner of the auction has valuation M_n and pays the second price S_n. So his profit is $M_n - S_n$. Hence δ_n is the expected profit of the winner in a second price auction[9].

So the economics of the Perloff-Salop model with Gaussian, Gumbel, exponential, or log-normal distributions asymptotically matches the economics of the second-price auction model. Asymptot-

---

[9] Additionally, the proof of Proposition 8 allows us easily to calculate the expected revenue of the seller, $s_n = E[S_n]$, in an auction with n buyers with independent valuation. One finds $s_n \sim \Gamma(2-\xi) A_n$ for $\xi \geq 0$, and $s_n \to F^{-1}(0)$ for $\xi < 0$. See also Kremer and Skrzypacz (2003) for a recent application of order statistics to auction theory.

ically, the two models – Perloff Salop and limit pricing/second price auction — yield isomorphic results.

4.3 Implications for consumer surplus

Sometimes the random utility framework is criticized as generating a too high consumer surplus. Indeed, if the distribution is unbounded, the total surplus goes to ∞ as the number of firms increases. Our analytical results allow us to examine this criticism. We consider distributions in the domain of attraction of the Gumbel and the Fréchet, i.e. with $\xi \geq 0$.

Expected surplus is $\sigma E\left[M_n\right] = \sigma m_n$, where M_n is the highest of n draws. Proposition 8 shows that $m_n = E\left[M_n\right] \sim \Gamma\left(1-\xi\right) A_n$ for $\xi \geq 0$. The value of A_n for a truncated distribution is bounded above by the value of A_n for the analogous non-truncated distribution. So we study the latter case for simplicity.

For all the distributions that we study except the unbounded power law case, A_n rises only slowly with n. Hence, even for unbounded distributions, and large numbers of producers, consumer surplus is quite small.

For example, for the case of Gaussian noise when consumer preferences have a standard deviation of \$1, $A_n \sim \sqrt{2\ln n}$. So with a million toothpaste producers consumer surplus averages only \$5.25 per tube. Hence, our framework — even with unbounded distributions — does not generate counterfactual predictions about consumer surplus or counterfactual predictions about the prices that cartels would set.

5 Conclusion

Even the most sophisticated consumers make small errors and hence have some noise in their calculations. We have shown that this noise generates equilibrium markups that are not likely to fall quickly as competition increases.

Using extreme value theory, we characterize markups in the Perloff-Salop (1985) model. We derive an asymptotic approximation for the Perloff-Salop markup and show that for realistic distributions the markup has a natural limit pricing property; the asymptotic Perloff-Salop markup is equal to the expected gap between the highest draw and second highest draw in a sample of n draws, where n is the number of competing firms.

Previous authors have characterized extreme — hence, analytically tractable — noise distributions. These extreme distributions yielded contradictory implications. For the exponential and logit densities, markups do not decay with competition (Perloff and Salop 1985, Anderson et al. 1992). For the uniform density, markups are proportional to $1/n$, and hence have a unit elasticity (Perloff and Salop, 1985).

Using our asymptotic results, we characterize *general* noise distributions. We find that most distributions yield markup elasticities that are close to the exponential and logit cases. For example, for the Gaussian case asymptotic markups are proportional to $1/\sqrt{\ln n}$, implying a zero asymptotic elasticity. Increasing competition in an environment with Gaussian noise (even truncated Gaussian noise) will only produce weak pressure on prices. We conclude that markups due to noise are highly robust and are unlikely to decay with competition. The tools we describe in this study (Proposition 8 and 2) may be useful for other applications where it is the extreme right tail of distributions that matters, for instance in auction theory.

6 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 9 *A function g defined in a right neighborhood of 0 has regular variation at* 0 *(written $g \in RV_\rho^0$) with exponent ρ if*

$$\forall u > 0, \lim_{t \to 0^+} g(ut)/g(t) = u^\rho. \tag{32}$$

A function g defined in a neighborhood of ∞ has regular variation at ∞ (written $g \in RV_\rho^\infty$) with exponent ρ if

$$\forall u > 0, \lim_{t \to \infty} g(ut)/g(t) = u^\rho. \tag{33}$$

Clearly, (32) and (33) are related: $g(t) \in RV_\rho^0 \iff g(t^{-1}) \in RV_{-\rho}^\infty$. If $g \in RV_\rho^0$, then $g(t)$ behaves like t^ρ around 0, perhaps up to a constant or slowly varying function. For instance, (32) holds if $g(t) = t^\rho$ and $g(t) = t^\rho [\ln(1/t)]^\alpha$ for some α.

Three Types of Distributions

In extreme value theory, there are three types of distributions. They are classified by the degree of fatness of their right tail. A useful indicator of their fatness is the characteristic index ξ defined in Eq. 7. As Table 1 indicates, distributions with fatter tails have a weakly larger ξ.

The formal classification is as follows.

Fact 10 *(Resnick 1987, Prop. 1.11, Prop. 1.13, Prop. 1.4; p. 43-59.)*
$F \in D(\Phi)$ *iff* $F^{-1}(1) = \infty$ *and* $\overline{F}(t) \in RV_{-\zeta}^\infty, -\zeta < 0$.
$F \in D(\Psi)$ *iff* $F^{-1}(1) < \infty$ *and* $\overline{F}(F^{-1}(1) - t) \in RV_\alpha^0, \alpha > 0$.
$F \in D(\Lambda)$ *iff* $\overline{F}(t) = b(x) \exp\left\{ -\int_{t_0}^t \frac{1}{a(x)} dx \right\}, b(x) \to b > 0, a'(x) \to 0$ *with* $a(.)$ *absolutely continuous.*

In other terms, $D(\Phi)$, the Domain of Attraction of the Fréchet, comprises power law tail distributions of the type (18). Their support is unbounded on the right. $D(\Psi)$, the Domain of Attraction of the Weibull, comprises very “thin tailed” distributions of the type (12). Their support has an upper bound. $D(\Lambda)$, the Domain of Attraction of the Gumbel, comprises distributions of medium thinness, such as the Gaussian, Gumbel, Exponential, and Gamma distributions. Their support may or may not be bounded on the right.

Definition 3 details two useful and mild restrictions for our "regular" definitions. It is easy to verify that all the distributions in Table 1 are regular. Also, Reiss (1989, p. 159) ensures that a regular distribution lies in one of the 3 domains of attraction. In fact, the sign of the characteristic index determines the domain of attraction of a distribution:

Fact 11 *(Reiss, 1989, p. 159-160) Let* F *be a regular distribution. Then* $F \in D(\Phi) \cup D(\Psi) \cup D(\Lambda)$*. In addition,*

$$\begin{aligned} \xi = 1/\zeta > 0 \Leftrightarrow &\quad F \in D(\Phi), \overline{F}(t) \in RV_{-\zeta}^{\infty}; \\ \xi = -1/\alpha < 0 \Leftrightarrow &\quad F \in D(\Psi), \overline{F}(F^{-1}(1) - t) \in RV_{\alpha}^{0}; \\ \xi = 0 \Leftrightarrow &\quad F \in D(\Lambda). \end{aligned}$$

7 Appendix B: Proofs

7.1 Some useful lemmas

Lemma 12 *For a regular distribution* F *with characteristic index* ξ*,*

$$f(\overline{F}^{-1}(t)) \in RV_{\xi+1}^{0}.$$

Proof. First note that by substituting $t = 1/x, j(t) = U(1/x)$ into Resnick (1987, Prop. 0.7.a, p. 21), $\lim_{t\to 0} tj'(t)/j(t) = \rho$ implies $j \in RV_{\rho}^{0}$. Resnick (1987, Prop. 1.18, p. 66) shows, with $x = \overline{F}^{-1}(t), j(t) = f(\overline{F}^{-1}(t))$: $tj'(t)/j(t) = -tf'(\overline{F}^{-1}(t))/f(\overline{F}^{-1}(t))^2 = -\overline{F}(x)f'(x)/f(x)^2 = \left(\overline{F}/f\right)'(x) + 1$, so $\lim_{t\to 0} tj'(t)/j(t) = \lim_{x\to F^{-1}(1)} \left(\overline{F}/f\right)'(x) + 1 = \xi + 1$. ■

Lemma 13 *For a regular distribution,* $\xi \geq -1$*.*

Proof. In the Fréchet and Gumbel domains, $\xi \geq 0$. In the Weibull domain, recall that regularity requires $\lim_{t\to 0} f(\overline{F}^{-1}(t)) = \lim_{x\to F^{-1}(1)} f(x) < \infty$. By Resnick (1987, Prop. 0.8.ii, p. 22), this implies that $f(\overline{F}^{-1}(t)) \in RV_{\rho}^{0}, \rho \geq 0$. Lemma 12 allows us to conclude $\xi + 1 \geq 0$. ■

7.2 The main approximation result

We will prove Proposition 2 and apply it to the proof of Proposition 4. Essentially, Proposition 2 allows us to replace the integral $E\left[J\left(W_{n-r:n-1}\right)\right]$ by the deterministic expression $J\left(A_n\right)\Gamma\left(\rho+r\right)/(r-1)!$.

First, some preliminaries:

Definition 14 *A function j defined on $(0,1)$ is in the class S of well-behaved functions iff it satisfies the following hypotheses:*

1. $j \in RV_\rho^0, \rho > -1$.
2. *For all $\varepsilon \in (0,1)$, $j(t)$ is bounded on $(\varepsilon, 1)$.*

Lemma 15 *If $F \in R$, then $j(t) = f\left(\overline{F}^{-1}(t)\right) \in S$.*

Proof. Condition 1 is immediate from Lemmas 12 and 13. Condition 2 holds because by Definition 3, $f(.)$ is bounded on $\left(F^{-1}(0), F^{-1}(1)\right)$. ■

7.3 Proof of Proposition 2

Let us first prove that

$$E\left[J\left(W_{n-r:n-1}\right)\right] \sim \frac{J\left(A_n\right)\Gamma\left(\rho+r\right)}{(r-1)!}. \tag{34}$$

For $1 \leq r \leq n-1$, denote by $w_{r:n-1}$ the r-th order statistic of $n-1$ i.i.d. uniform $[0,1]$ variables y_i. By Reiss (1989, Thm. 1.3.2, p. 21), its probability distribution function is:

$$f_{r:n-1}^{(uniform)}(t) = (n-1)!\frac{t^{r-1}(1-t)^{n-r-1}}{(r-1)!(n-r-1)!} \tag{35}$$

Now, $W_{n-r:n-1}$ has the same distribution as $\overline{F}^{-1}\left(w_{r:n-1}\right)$. Indeed:

$$\begin{aligned} P\left(\overline{F}^{-1}\left(w_{r:n-1}\right) < X\right) &= P\left(w_{r:n-1} > \overline{F}\left(X\right)\right) = P\left(w_{n-r:n-1} < F\left(X\right)\right) \text{ by symmetry} \\ &= P\left(W_{n-r:n-1} < X\right). \end{aligned}$$

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

$$\begin{aligned}
I_{n:r} &\equiv E\left[J\left(W_{n-r:n-1}\right)\right]=E\left[J\left(\overline{F}^{-1}\left(w_{r:n-1}\right)\right)\right]=E\left[j\left(w_{r:n-1}\right)\right] &(36)\\
&= \int_0^1 j\left(x\right)\frac{(n-1)!x^{r-1}(1-x)^{n-r-1}}{(r-1)!(n-r-1)!}dx &(37)\\
&= \int_0^u j\left(\frac{u}{n}\right)\frac{(n-1)!}{n^r(r-1)!(n-r-1)!}u^{r-1}(1-\frac{u}{n})^{n-r-1}du. &(38)
\end{aligned}$$

We can expect to hold the following series of "heuristic asymptotic equalities", signalled by $\simeq$

$$\begin{aligned}
I_{n:r} &= \frac{(n-1)!}{n^r(r-1)!(n-r-1)!}\int_0^n j\left(\frac{u}{n}\right)u^{r-1}(1-\frac{u}{n})^{n-r-1}du\\
&\simeq \frac{1}{(r-1)!}\int_0^n j\left(\frac{u}{n}\right)u^{r-1}e^{-u}du=\frac{j\left(\frac{1}{n}\right)}{(r-1)!}\int_0^n\frac{j\left(u/n\right)}{j\left(1/n\right)}u^{r-1}e^{-u}du\\
&\simeq \frac{j\left(\frac{1}{n}\right)}{(r-1)!}\int_0^n u^{\rho+r-1}e^{-u}du\simeq\frac{j\left(\frac{1}{n}\right)}{(r-1)!}\int_0^\infty u^{\rho+r-1}e^{-u}du=\frac{j\left(\frac{1}{n}\right)\Gamma\left(\rho+r\right)}{(r-1)!}=\frac{J\left(A_n\right)\Gamma\left(\rho+r\right)}{(r-1)!}.
\end{aligned}$$

We now proceed to a rigorous proof. Let $\eta>0$, define $K_{n:r}=\frac{n^r(r-1)!(n-r-1)!}{(n-1)!}\frac{I_{n:r}}{j(1/n)}$. We will show that for n large enough, $\left|K_{n:r}-\Gamma\left(\rho+r\right)\right|<4\eta$. Since $\frac{n^r(r-1)!(n-r-1)!}{(n-1)!}\sim(r-1)!$, this will prove the Proposition. To do this, we specify $0<\varepsilon<L$, and starting from (36), we use the following decomposition:

$$\begin{aligned}
K_{n:r} &= \frac{n^r(r-1)!(n-r-1)!}{(n-1)!}\frac{I_{n:r}}{j\left(1/n\right)}=\gamma_{n:r}+\kappa_{n:r}+\nu_{n:r}+\phi_{n:r} &(39)\\
\gamma_{n:r} &= \int_0^\varepsilon\chi_{n:r}\left(u\right)du,\ \kappa_{n:r}=\int_\varepsilon^L\chi_{n:r}\left(u\right)du,\ \nu_{n:r}=\int_L^{\varepsilon n}\chi_{n:r}\left(u\right)du,\ \phi_{n:r}=\int_{\varepsilon n}^n\chi_{n:r}\left(u\right)du\\
\chi_{n:r}\left(u\right) &= \frac{j\left(u/n\right)}{j\left(1/n\right)}u^{r-1}\left(1-\frac{u}{n}\right)^{n-r-1}.
\end{aligned}$$

We study each term in turn. Let us start by analyzing $\gamma_{n:r}$. Pick some $\delta\in(0,1+\rho)$. Then by Resnick (1987, Prop. 0.8.ii, p. 22), with $U(x)=j(1/x)\in RV_{-\rho}^{\infty}$, we can pick n_1 large enough so that for $t>n_1,x\geq1:(1-\delta)x^{-\rho-\delta}<\frac{U(tx)}{U(t)}<(1+\delta)x^{-\rho+\delta}$. Substituting $n=t,u=1/x$, we get, for

$u \leq 1, n > n_1 : (1-\delta)u^{\rho+\delta} < \frac{j(u/n)}{j(1/n)} < (1+\delta)u^{\rho-\delta}$. So for $n > n_1$:

$$\begin{aligned} |\gamma_{n:r}| &= \left| \int_0^\varepsilon \frac{j\left(u/n\right)}{j\left(1/n\right)} u^{r-1} \left(1 - \frac{u}{n}\right)^{n-r-1} du \right| \leq \int_0^\varepsilon (1+\delta) u^{\rho+r-\delta-1} \left(1 - \frac{u}{n}\right)^{n-r-1} du \\ &\leq \int_0^\varepsilon (1+\delta) u^{\rho+r-\delta-1} du = \frac{1+\delta}{\rho + r - \delta} \varepsilon^{\rho+r-\delta} < \eta. \end{aligned}$$

The last term vanishes as $\varepsilon \to 0$. It follows that for small enough ε and $n > n_1$,

$$|\gamma_{n:r}| < \eta. \tag{40}$$

We now turn to $\kappa_{n:r}$.

$$\kappa_{n:r} = \int_\varepsilon^L \frac{j\left(u/n\right) u^{-\rho}}{j\left(1/n\right)} u^{\rho+r-1} \left(1 - \frac{u}{n}\right)^{n-r-1} du.$$

By Resnick (1987, Prop. 0.5, p.17), Hyp. 1 implies that $\frac{j(u\lambda)}{j(\lambda)} u^{-\rho} \to 1$ locally uniformly for $u \in (0, \infty)$. This means that, for a given η', there is a $\lambda_1 > 0$ such that $\forall \lambda \in (0, \lambda_1), \forall u \in [\varepsilon, L], \left| \frac{j(u\lambda)}{j(\lambda)} u^{-\rho} - 1 \right| \leq \eta'$. This implies, for $n > n_2 = 1/\lambda_1$,

$$1 - \eta' < \frac{\kappa_{n:r}}{\int_\varepsilon^L u^{\rho+r-1} \left(1 - \frac{u}{n}\right)^{n-r-1} du} < 1 + \eta'.$$

Because $\left(1 - \frac{u}{n}\right)^{n-r-1} \to e^{-u}$ uniformly in $[\varepsilon, L]$ (Dieudonné, p.111), $\int_\varepsilon^L u^{\rho+r-1} \left(1 - \frac{u}{n}\right)^{n-r-1} du \to \int_\varepsilon^L u^{\rho+r-1} e^{-u} du$, which implies that for n greater than some n_3,

$$1 - \eta' < \frac{\int_\varepsilon^L u^{\rho+r-1} \left(1 - \frac{u}{n}\right)^{n-r-1} du}{\int_\varepsilon^L u^{\rho+r-1} e^{-u} du} < 1 + \eta'.$$

Also, as $\Gamma\left(1+\rho\right) = \int_0^\infty u^\rho e^{-u} du$, if we choose ε small enough and L large enough,

$$1 - \eta' < \frac{\int_\varepsilon^L u^{\rho+r-1} e^{-u} du}{\Gamma\left(\rho + r\right)} < 1 + \eta'.$$

The last 3 displayed Eqs. imply

$$\left(1-\eta'\right)^3 < \frac{\kappa_{n:r}}{\Gamma\left(\rho+r\right)} < \left(1+\eta'\right)^3.$$

If we choose η' small enough, we ensure

$$\left|\kappa_n - \Gamma\left(\rho+r\right)\right| < \eta. \tag{41}$$

We now study $\nu_{n:r}$. Choose $\rho' > \rho$, so $j(x)x^{-\rho'} \in RV^0_{\rho-\rho'}$. By Gulek (1987, Prop. 1.7.3, p. 9), $j(x)x^{-\rho'}$ is asymptotic to some decreasing function $\tilde{j}(x)$. Choose $0 < \sigma < 1$ and ε so that for $x < \varepsilon, 1-\sigma < \frac{j(x)x^{-\rho'}}{\tilde{j}(x)} < 1+\sigma$. Then for $1/n < u/n < \varepsilon$, $\frac{j(u/n)(u/n)^{-\rho'}}{j(1/n)(1/n)^{-\rho'}} < \frac{(1+\sigma)\tilde{j}(u/n)}{(1-\sigma)\tilde{j}(1/n)} < \frac{(1+\sigma)}{(1-\sigma)}$, and for $L > 1$,

$$\left|\nu_{n:r}\right| = \left|\int_L^{\varepsilon n} \frac{j\left(u/n\right)}{j\left(1/n\right)} u^{r-1}\left(1-\frac{u}{n}\right)^{n-r-1} du\right| \le \int_L^{\varepsilon n} \frac{\left(1+\sigma\right)}{\left(1-\sigma\right)} u^{r-\rho'-1}\left(1-\frac{u}{n}\right)^{n-r-1} du.$$

Also, inequality $1-x \le e^{-x}$ implies $\left(1-\frac{u}{n}\right)^{n-r-1} \le e^{-\frac{n-r-1}{n}u} \le e^{-u/(r+2)}$ for $n \ge n_4 = r+2$, which implies

$$\left|\nu_{n:r}\right| \le \int_L^{\varepsilon n} \frac{\left(1+\sigma\right)}{\left(1-\sigma\right)} u^{r-\rho'-1} e^{-u/(r+2)} du \le \int_L^{\infty} \frac{\left(1+\sigma\right)}{\left(1-\sigma\right)} u^{r-\rho'-1} e^{-u/(r+2)} du.$$

The right hand side of the last equation goes to 0 as $L \to \infty$. So if L is large enough, it is less than η. We conclude that for $n \ge n_4$,

$$\left|\nu_{n:r}\right| < \eta \tag{42}$$

We finally study $\phi_{n:r}$. By Resnick (1987, Proposition 0.8.ii, p.23), Hyp. 1 implies that $\left|j\left(1/n\right)\right| > \left(1/n\right)^{\rho'}$ for some $\rho' > \rho$ and n large enough. Also, by Hyp. 2 we have $\left|j(u)\right| < j_0$ for some j_0 and $u \in (0,1)$, so

$$\begin{aligned} \left|\phi_{n:r}\right| &= \left|\int_{\varepsilon n}^{n} \frac{j\left(u/n\right)}{j\left(1/n\right)} u^{r-1}\left(1-\frac{u}{n}\right)^{n-r-1} du\right| = \frac{n^r}{\left|j(1/n)\right|}\left|\int_{\varepsilon}^{1} j(u)\left(1-u\right)^{n-r-1} n du\right| \\ &< \frac{n^r}{\left|j(1/n)\right|}\left|j_0 \int_{\varepsilon}^{1}\left(1-u\right)^{n-r-1} n du\right| \le j_0 n^{r+\rho'} \frac{n(1-\varepsilon)^{n-r}}{n-r}. \end{aligned}$$

The last expression goes to 0 when $n \to \infty$. So for all n above a certain n_5,

$$|\phi_{n:r}| < \eta. \tag{43}$$

Combining (40)-(43), we conclude that, for $n > \max(n_0, ..., n_5)$, $|K_{n:r} - \Gamma(1+\rho)| < 4\eta$, which proves[10] (34).

Now we proceed to prove that

$$\frac{dE\left[J\left(W_{n-r:n-1}\right)\right]}{dn} \sim \frac{J'(A_n)\Gamma(\rho+r)}{f(A_n)n^2(r-1)!}. \tag{44}$$

This is an intuitive result; since $E\left[J\left(W_{n-r:n-1}\right)\right] \sim J\left(\overline{F}^{-1}(\frac{1}{n})\right)\Gamma(\rho+r)/(r-1)!$, we expect that

$$\frac{dE\left[J\left(W_{n-r:n-1}\right)\right]}{dn} \sim \frac{\Gamma(\rho+r)}{(r-1)!}\frac{d}{dn}J\left(\overline{F}^{-1}(\frac{1}{n})\right) = \frac{J'\left(\overline{F}^{-1}(\frac{1}{n})\right)\Gamma(\rho+r)}{f(\overline{F}^{-1}(\frac{1}{n}))n^2(r-1)!}$$

as in (44). With $t = \overline{F}^{-1}(x)$, we have

$$\begin{aligned} E\left[J\left(W_{n-r:n-1}\right)\right] &= \int_0^1 j(t)\frac{(n-1)!t^{r-1}(1-t)^{n-r-1}}{(r-1)!(n-r-1)!}dt \\ &= \frac{\Gamma(n)}{n^r\Gamma(r)\Gamma(n-r)}\int_0^n j(\frac{u}{n})u^{r-1}(1-\frac{u}{n})^{n-r-1}du. \end{aligned}$$

Let us define

$$\nu_r(n) = \frac{\Gamma(n)}{n^r\Gamma(r)\Gamma(n-r)}, \pi_{r,u}(n) = j(\frac{u}{n}), \iota_{r,u}(n) = u^{r-1}(1-\frac{u}{n})^{n-r-1}$$

so that $E\left[J\left(W_{n-r:n-1}\right)\right] = \nu_r(n)\int_0^n \pi_{r,u}(n)\iota_{r,u}(n)du$. Our choice of subscripts is meant to emphasize that n is the variable being differentiated. Treating n as a continuous variable,

$$\frac{dE\left[J\left(W_{n-r:n-1}\right)\right]}{dn} = \nu_r'(n)\int_0^n \pi_{r,u}(n)\iota_{r,u}(n)du + \nu_r(n)\int_0^n \pi_{r,u}'(n)\iota_{r,u}(n)du + \nu_r(n)\int_0^n \pi_{r,u}(n)\iota_{r,u}'(n)du.$$

---

[10] Sidney Resnick has suggested a somewhat simpler method of proof that uses Karamata's Tauberian Theorem for regularly varying functions. This method allows us to prove (5), but does not easily accomodate the cases (6) and (29).

Now,

$$
\begin{aligned}
\nu'_r(n) &= \frac{1}{\Gamma(r)}\frac{d}{dn}\prod_{k=1}^{r}(\frac{n-k}{n}) = \frac{1}{\Gamma(r)}\sum_{j=1}^{r}\left(-\frac{\prod_{k=1}^{r}(1-\frac{k}{n})}{1-\frac{j}{n}}\frac{j}{n^2}\right) \sim \frac{r^2}{2n^2\Gamma(r)} \\
\pi'_{r,u}(n) &= -j'(\frac{u}{n})\frac{u}{n^2} \\
\iota'_{r,u}(n) &= u^{r-1}(1-\frac{u}{n})^{n-r-1}\left(\frac{u(n-r-1)}{n(n-u)} + \ln\left(1-\frac{u}{n}\right)\right)
\end{aligned}
$$

so

$$
\begin{aligned}
\frac{dE\left[J\left(W_{n-r:n-1}\right)\right]}{dn} &= \frac{1}{\Gamma(r)}\sum_{j=1}^{r}\left(-\frac{\prod_{k=1}^{r}(1-\frac{k}{n})}{1-\frac{j}{n}}\frac{j}{n^2}\right)\int_0^n j(\frac{u}{n})u^{r-1}(1-\frac{u}{n})^{n-r-1}du \\
&- \frac{\Gamma(n)}{n^r\Gamma(r)\Gamma(n-r)}\int_0^n j'(\frac{u}{n})\frac{u^r}{n^2}(1-\frac{u}{n})^{n-r-1}du \\
&+ \frac{\Gamma(n)}{n^r\Gamma(r)\Gamma(n-r)}\int_0^n j(\frac{u}{n})u^{r-1}\left(\frac{u(n-r-1)}{n(n-u)} + \ln\left(1-\frac{u}{n}\right)\right)(1-\frac{u}{n})^{n-r-1}du
\end{aligned}
$$

To evaluate each of these integrals, we will appeal to (34). Now, by Resnick (1987, Thm. 0.6.a, p.17), $j'(t) \in RV^0_{\rho-1}$, which gives us

$$
\begin{aligned}
& -\frac{\Gamma(n)}{n^r\Gamma(r)\Gamma(n-r)}\int_0^n j'(\frac{u}{n})\frac{u^r}{n^2}(1-\frac{u}{n})^{n-r-1}du \\
\sim\ & -\frac{j'(\frac{1}{n})\Gamma(n)\Gamma(\rho+r)}{n^{r+2}\Gamma(r)\Gamma(n-r)} \sim -\frac{j'(\frac{1}{n})\Gamma(\rho+r)}{n^2\Gamma(r)}.
\end{aligned}
$$

Also,

$$
\begin{aligned}
& \frac{1}{\Gamma(r)}\sum_{j=1}^{r}\left(-\frac{\prod_{k=1}^{r}(1-\frac{k}{n})}{1-\frac{j}{n}}\frac{j}{n^2}\right)\int_0^n j(\frac{u}{n})u^{r-1}(1-\frac{u}{n})^{n-r-1}du \\
\sim\ & \frac{r^2\Gamma\left(\rho+r\right)j(\frac{1}{n})}{2n^2\Gamma(r)} = o\left(-\frac{j'(\frac{1}{n})\Gamma(\rho+r)}{n^2\Gamma(r)}\right).
\end{aligned}
$$

since $j(1/n) = o(j'(1/n))$. Finally, since

$$
\frac{u(n-r-1)}{n(n-u)} + \ln\left(1-\frac{u}{n}\right) = -\frac{1+r}{n^2}u + \frac{1}{2n^2}u^2 + o(u^2n^{-2}),
$$

we have

$$\begin{aligned}
&\frac{\Gamma(n)}{n^r\Gamma(r)\Gamma(n-r)}\int_0^n j(\frac{u}{n})u^{r-1}\left(\frac{u(n-r-1)}{n(n-u)}+\ln\left(1-\frac{u}{n}\right)\right)(1-\frac{u}{n})^{n-r-1}du\\
=\;&\frac{\Gamma(n)}{n^r\Gamma(r)\Gamma(n-r)}\int_0^n j(\frac{u}{n})u^{r-1}\left(-\frac{1+r}{n^2}u+\frac{1}{2n^2}u^2+o(n^{-2})\right)(1-\frac{u}{n})^{n-r-1}du\\
\sim\;&\frac{\Gamma(n)j(\frac{1}{n})}{n^r\Gamma(r)\Gamma(n-r)}\left(-\frac{\Gamma(\rho+r+1)(1+r)}{n^2}+\frac{\Gamma(\rho+r+2)}{2n^2}+o(n^{-2})\right)\\
=\;&o\left(\frac{\Gamma\left(\rho+r\right)j(\frac{1}{n})}{n\Gamma(r)}\right).
\end{aligned}$$

Thus

$$\frac{dE\left[J\left(W_{n-r:n-1}\right)\right]}{dn}\sim-\frac{j'(\frac{1}{n})\Gamma(\rho+r)}{n^2\Gamma(r)}=\frac{J'(A_n)\Gamma(\rho+r)}{f(A_n)n^2\Gamma(r)}.$$

7.4 Proof of Proposition 4

First, from Proposition 12, the coefficient of regular variation at zero of $f\left(\overline{F}^{-1}(t)\right)$ is $1+\xi$. Applying Proposition 2 to $J(x)=f(x)$ gives: $E\left[f\left(M_{n-1}\right)\right]\sim f\left(A_n\right)\Gamma\left(2+\xi\right)$. Substituting this into (4) we get (8).

7.5 Proof of Proposition 6

For approximate results we use Eq. (8) in Proposition 4. Note that we need to know the asymptotic behaviour of $f(A_n)$, which we have included in Table 1. Although the calculation is not difficult for most well-known distributions, there is no general algorithm. As an illustration, we will calculate the Gaussian markup explicitly.

For Gaussian noise, $f(\varepsilon)=\frac{1}{\sqrt{2\pi}}e^{-\varepsilon^2/2}$. Let $\tilde{F}(\varepsilon)=f(\varepsilon)/\varepsilon$. L'Hopital's rule reveals that $\overline{F}(\varepsilon)/\tilde{F}(\varepsilon)\sim 1$. Thus $\overline{F}(\varepsilon)=f(\varepsilon)t(\varepsilon)/\varepsilon$ for some $t(\varepsilon)\sim 1$. Let $y=\overline{F}^{-1}(x)$. Then we have $f(y)t(y)/y=x$; taking logarithms on both sides, $-y^2/2-\log(y)-\log(t(y))=\log(x)$, or $-y^2/2+o(-y^2/2)=\log(x)$. It follows that $y^2/(-2\log(x))\sim 1$, and thus $\overline{F}^{-1}(x)\sim\sqrt{2\log(1/x)}$. Returning to $\overline{F}(x)/\tilde{F}(x)\sim 1$, we find that $\overline{F}(y)/\tilde{F}(y)=x/(f(\overline{F}^{-1}(x))/\overline{F}^{-1}(x))\sim 1$, or $f(\overline{F}^{-1}(x))\sim x\overline{F}^{-1}(x)\sim x\sqrt{2\log(1/x)}$. We infer from this that $\xi=0, p-c\sim\frac{1}{\sqrt{2\ln n}}\sigma$.

To obtain exact results we use Eq. (2) in Lemma 1, and identity $\int_0^1 t^{a-1}\left(1-t\right)^{b-1}dt=\Gamma\left(a\right)\Gamma\left(b\right)/\Gamma\left(a+b\right)$. For the Weibull and Fréchet distributions, Proposition 1 and the fact that $M_n=^d n^{\xi}\varepsilon$ (Embrechts et al. 1997, p.124) offer a nice way to simplify the calculations.

7.6 Proof of Proposition 5

Using Eq. 2, and treating n as a continuous variable,

$$-\frac{d\ln\mu_n}{d\ln n}=1+\frac{n}{n-1}+\frac{(n-1)\,n\int F(x)^{n-2}f(x)^2\ln F(x)\,dx}{(n-1)\int F(x)^{n-2}f(x)^2\,dx}=\frac{2n-1}{n-1}+\frac{nE\left[f(M_{n-1})\ln F(M_{n-1})\right]}{E\left[f(M_{n-1})\right]}.$$

From Lemma 12, $f\left(\overline{F}^{-1}(t)\right)\in RV^0_{1+\xi}$. Since $\ln F\left(\overline{F}^{-1}(t)\right)=\ln(1-t)\sim -t$, we have $-\ln F\left(\overline{F}^{-1}(t)\right)\in RV^0_1$ by (32). It follows that $-f\left(\overline{F}^{-1}(t)\right)\ln F\left(\overline{F}^{-1}(t)\right)\in RV^0_{2+\xi}$. Before proceeding, we verify that $-f\left(\overline{F}^{-1}(t)\right)\ln F\left(\overline{F}^{-1}(t)\right)\in S$. Since $-f\left(\overline{F}^{-1}(t)\right)\ln F\left(\overline{F}^{-1}(t)\right)\in RV^0_{2+\xi}$ and $2+\xi>0$, Condition 1 in Definition 14 holds. Now, since F is regular, f is bounded. Also, with $\varepsilon\in(0,1)$, $\ln(1-\varepsilon)<\ln(1-t)<0$ for $t\in(0,\varepsilon)$; so $f\left(\overline{F}^{-1}(t)\right)\ln F\left(\overline{F}^{-1}(t)\right)=f\left(\overline{F}^{-1}(t)\right)\ln(1-t)$ is bounded on $(0,\varepsilon)$, and Condition 2 holds. We can thus apply Proposition 2 to get:

$$\frac{nE\left[f(M_{n-1})\ln F(M_{n-1})\right]}{E\left[f(M_{n-1})\right]}\sim\frac{n\Gamma(3+\xi)f(A_n)\ln F(A_n)}{\Gamma(2+\xi)f(A_n)}=\frac{n\Gamma(2+\xi)(2+\xi)\ln(1-1/n)}{\Gamma(2+\xi)}\sim-(2+\xi),$$

so we conclude that $\lim_{n\to\infty}-\frac{d\ln\mu_n}{d\ln n}=2-(2+\xi)=-\xi$.

7.7 Proof of Proposition 8

For regular distributions, the largest and second largest order statistics, appropriately normalized, converge to universal distributions $G_{i,\xi^{-1},1}$ and $G_{i,\xi^{-1},2}$ which differ for each domain. We will not prove the following proposition, which is found in Reiss (1989, p.160-161).

Proposition 16 *Let F be a regular distribution with characteristic index ξ. Let M_n and S_n be the first and second largest order statistics respectively. Then*

$$\begin{aligned}P\left(b_n^{-1}(M_n-a_n)\le x\right)&\to G_{i,\xi^{-1},1}(x)\\P\left(b_n^{-1}(S_n-a_n)\le x\right)&\to G_{i,\xi^{-1},2}(x)\end{aligned}$$

Type $i=1,2,3$ correspond respectively to the domain of attraction of the Fréchet, Weibull and

Gumbel, i.e. $\xi > 0$, $\xi < 0$ *and* $\xi = 0$. *The cumulative distribution functions* $G_{i,\xi^{-1},k}$ *are given by:*

$$\begin{aligned}
\textit{(Fréchet) } G_{1,\xi^{-1},k}(x) &= \exp\left(-x^{-1/\xi}\right)\sum_{j=0}^{k-1}\frac{x^{-j/\xi}}{j!}, x > 0;\\
\textit{(Weibull) } G_{2,\xi^{-1},k}(x) &= \exp\left(-(-x)^{-1/\xi}\right)\sum_{j=0}^{k-1}\frac{(-x)^{-j/\xi}}{j!}, x < 0;\\
\textit{(Gumbel) } G_{3,\infty,k}(x) &= \exp\left(-e^{-x}\right)\sum_{j=0}^{k-1}\frac{e^{-jx}}{j!}, -\infty < x < \infty.
\end{aligned}$$

The centering constants are: For the Fréchet, $a_n = 0$, $b_n = A_n \sim kn^{\xi}$, *for the Weibull:* $a_n = F^{-1}(1), b_n = a_n - A_n \sim kn^{\xi}$, *and for the Gumbel:* $a_n = A_n, b_n = 1/(nf(A_n))$.

Let $\overline{G}_{i,\xi,k,n}(x)$ be the exact df for large order statistics; that is,

$$\begin{aligned}
P\left(b_n^{-1}(M_n - a_n) \le x\right) &= \overline{G}_{i,\xi^{-1},1,n}(x)\\
P\left(b_n^{-1}(S_n - a_n) \le x\right) &= \overline{G}_{i,\xi^{-1},2,n}(x)
\end{aligned}$$

From Proposition 16, $\overline{G}_{i,\xi^{-1},k,n}(x) \to G_{i,\xi^{-1},k}(x)$. Using the natural limits for the 3 types, $(U_i, V_i) = (0,\infty), (-\infty,0), (-\infty,\infty)$ respectively, this implies:

$$\begin{aligned}
b_n^{-1}\delta_n &= E\left[\left(b_n^{-1}(M_n - a_n)\right) - b_n^{-1}(S_n - a_n)\right]\\
&= \int_{U_i}^{V_i} x\left(\overline{G}'_{i,\xi^{-1},1,n} - \overline{G}'_{i,\xi^{-1},2,n}\right)dx = \left[x\left(\overline{G}_{i,\xi^{-1},1,n} - \overline{G}_{i,\xi^{-1},2,n}\right)\right]_{U_i}^{V_i} - \int_{U_i}^{V_i}\left(\overline{G}_{i,\xi^{-1},1,n} - \overline{G}_{i,\xi^{-1},2,n}\right)dx\\
&\to -\int_{U_i}^{V_i}\left(G_{i,\xi^{-1},1} - G_{i,\xi^{-1},2}\right)dx = I_i,
\end{aligned}$$

where the asymptotic relation holds by dominated convergence, and

$$\begin{aligned}
I_1 &= \int_0^{\infty}\exp\left(-x^{-1/\xi}\right)x^{-1/\xi}dx = \xi\Gamma(1-\xi) \text{ by change of variable } y = x^{-1/\xi}\\
I_2 &= -\xi\Gamma(1-\xi) \text{ likewise}\\
I_3 &= \int_{-\infty}^{\infty}\exp\left(-e^{-x}\right)e^{-x}dx = \left[\exp\left(-e^{-x}\right)\right]_{-\infty}^{\infty} = 1.
\end{aligned}$$

That is, $\delta_n = b_n I_i$ for each of the domains. Now note that in the Fréchet domain, using Resnick (1987, Prop. 0.7.b, p. 21) and setting $x = A_n : b_n n f(A_n) = A_n n f(A_n) = x f(x)/\overline{F}(x) \to -\zeta = 1/\xi$; so $b_n \sim 1/\left(\xi n f(A_n)\right)$. A similar calculation reveals, for the Weibull domain, $b_n \sim -1/\left(\xi n f(A_n)\right)$. Since $b_n = 1/\left(n f(A_n)\right)$ for the Gumbel, we have, in all 3 domains, $\delta_n \sim \frac{\Gamma(1-\xi)}{n f(A_n)}$.□

7.8 Proof of Proposition 7

We study

$$Q_n(X) = \frac{1}{\mu_n\left(X\right) n f(A_n) \Gamma\left(2+\xi\right)} = \frac{n(n-1) \int_{F^{-1}(0)}^{X} \left(\frac{F(x)}{F(X)}\right)^{n-2} \left(\frac{f(x)}{F(X)}\right)^2 dx}{n f(A_n) \Gamma\left(2+\xi\right)}.$$

Given $\eta > 0$, we will show that for n large enough and $X > X_n$, $|Q_n(X) - 1| < 3\eta$. This is equivalent to our proposition. We first decompose $Q_n\left(X\right)$ as follows:

$$\begin{aligned} Q_n\left(X\right) &= \varphi_n + \psi_n(X) - \omega_n(X), \\ \varphi_n &= \frac{n(n-1) \int_{F^{-1}(0)}^{F^{-1}(1)} F\left(x\right)^{n-2} f\left(x\right)^2 dx}{n f(A_n) \Gamma\left(2+\xi\right)}, \\ \psi_n(X) &= \left(F(X)^{-n} - 1\right) \varphi_n, \\ \omega_n(X) &= F(X)^{-n} \frac{n(n-1) \int_{X}^{F^{-1}(1)} F\left(x\right)^{n-2} f\left(x\right)^2 dx}{n f(A_n) \Gamma\left(2+\xi\right)}. \end{aligned}$$

Let us study each term in turn. We will show that ψ_n and ω_n vanish, while $\varphi_n \to 1$. Throughout the proof, we shall use the fact that $F(X_n)^{-n} \to 1$, which is clear as $F(X_n)^{-n} = \left(1 - \overline{F}(X_n)\right)^{-n} = e^{-n\left[\overline{F}(X_n) + o\left(\overline{F}(X_n)\right)\right]} = e^{-n\overline{F}(X_n) + o\left(n\overline{F}(X_n)\right)} \to 1$.

To study φ_n, we first note that $\mu_n(F^{-1}(1))$ is the untruncated markup. By Proposition 4, it converges to $\left(n f(A_n) \Gamma\left(2+\xi\right)\right)^{-1}$. So there exists n_1 such that for $n > n_1$, we have

$$|\varphi_n - 1| = \left| \frac{1}{\mu_n(F^{-1}(1)) n f(A_n) \Gamma\left(2+\xi\right)} - 1 \right| < \eta. \tag{45}$$

We now analyze $\psi_n(X)$. For $X \geq X_n$,

$$0 \leq \psi_n\left(X\right) = \left(F(X)^{-n} - 1\right) \varphi_n \leq \left(F(X_n)^{-n} - 1\right) \varphi_n.$$

The last expression converges to 0, as $F(X_n)^{-n} \to 1$ and $\varphi_n \to 1$. So there is a $n_2 > n_1$ s.t. for $n > n_2$, $(F(X_n)^{-n} - 1)\,\varphi_n < \eta$. This implies that for $n > n_2$ and $X \geq X_n$,

$$0 \leq \psi_n\left(X\right) < \eta. \tag{46}$$

We now proceed to $\omega_n(X)$. Because F is non-decreasing, and $X > X_n$, $\omega_n\left(X\right) \leq \omega_n\left(X_n\right)$. Set $j(t) = f(F^{-1}(t))$ and $\varepsilon_n = n\overline{F}(X_n)$. Using the notation of the proof of Proposition 2,

$$\omega_n(X_n) = \Gamma\left(2+\xi\right)^{-1} F(X_n)^{-n} \frac{(n-1)}{n} \int_0^{\varepsilon_n} \frac{j(u/n)}{j(1/n)} (1-\frac{u}{n})^{n-2} du.$$

Pick some $0 < \delta < 1$. Then by Resnick (1987, Prop. 0.8.ii, p. 22), with $U(x) = j(1/x) \in RV_{-\rho}^{\infty}$, we can pick n_3 so that for $t > n_3, x \geq 1 : (1-\delta)x^{-\rho-\delta} < \frac{U(tx)}{U(t)} < (1+\delta)x^{-\rho+\delta}$. Substituting $n = t, u = 1/x$, we get, for $u \leq 1, n > n_3 : (1-\delta)u^{\rho+\delta} < \frac{j(u/n)}{j(1/n)} < (1+\delta)u^{\rho-\delta}$. With $n > n_3 :$

$$\begin{aligned} \left|\int_0^{\varepsilon_n} \frac{j\left(u/n\right)}{j\left(1/n\right)} \left(1-\frac{u}{n}\right)^{n-2} du\right| &\leq \int_0^{\varepsilon_n} (1+\delta)u^{\rho-\delta} \left(1-\frac{u}{n}\right)^{n-2} du \\ &\leq \int_0^{\varepsilon_n} (1+\delta)u^{\rho-\delta} du = \frac{1+\delta}{1+\rho-\delta}\varepsilon_n^{1+\rho-\delta}. \end{aligned}$$

So, for $n > n_3$,

$$\omega_n(X_n) \leq \Gamma\left(2+\xi\right)^{-1} F(X_n)^{-n} \frac{(n-1)}{n} \frac{1+\delta}{1+\rho-\delta} \varepsilon_n^{1+\rho-\delta}.$$

This expression converges to 0 when $n \to \infty$, since $\varepsilon_n \to 0$ and $F(X_n)^{-n} \to 1$. So there is a $n_4 > n_3$ such that, for $n > n_4$, $\omega_n(X_n) < \eta$. This implies for $n > n_4$ and $X > X_n$,

$$|\omega_n(X)| < \eta. \tag{47}$$

By (45)-(47), for $n > n_0 = \max\left(n_1, n_2, n_3, n_4\right)$, and $X > X_n$, $|Q_n(X) - 1| < 3\eta$, which is what we set out to prove.

8 References

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

Anderson, Simon P, André de Palma, Yurii Nesterov, 1995. "Oligopolistic Competition and the Optimal Provision of Products," *Econometrica* 63(6), p.1281-1301.

Barenblatt, G. (1996) *Scaling, Self-similarity, and Intermediate Asymptotics*, Cambridge University Press.

Becker, Gary (1957). *The Economics of Discrimination*, University of Chicago Press.

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.

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

Dieudonné, Jean Alexandre, *Infinitesimal Calculus* (Boston, Houghton Mifflin, 1971).

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

Gulek, J.L. and L. de Haan, Regular Variation, Extensions and Tauberian Theorems (CWI Tract, Amsterdam, 1987).

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.*

Kremer, Ilan and Andrzej Skrzypacz "Information Aggregation and the Information Content of Order Statistics", Stanford GSB mimeo.

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

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.

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.

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