ECONSTOR
Make Your Publications Visible.

A Service of
ZBW
Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

Koenker, Roger

Working Paper

The ignorant monopolist redux

cemmap working paper, No. CWP57/19

Provided in Cooperation with:
Institute for Fiscal Studies (IFS), London

Suggested Citation: Koenker, Roger (2019) : The ignorant monopolist redux, cemmap working paper, No. CWP57/19, Centre for Microdata Methods and Practice (cemmap), London, https://doi.org/10.1920/wp.cem.2019.5719

This Version is available at:
https://hdl.handle.net/10419/211150

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

WWW.ECONSTOR.EU

Mitglied der Leibniz-Gemeinschaft

cemmap

centre for microdata methods and practice

The ignorant monopolist redux

Roger Koenker

The Institute for Fiscal Studies
Department of Economics,
UCL

cemmap working paper CWP57/19

E·S·R·C ECONOMIC & SOCIAL RESEARCH COUNCIL

An ESRC Research Centre

Manuscript submitted to *The Econometrics Journal*, pp. 1–7.

The Ignorant Monopolist Redux

Roger Koenker†

† *Department of Economics, University College London, 30 Gordon Street, London, WC1H 0AX, UK*
E-mail: r.koenker@ucl.ac.uk

Summary The classical problem of the monopolist faced with an unknown demand curve is considered in a simple stochastic setting. Sequential pricing strategies designed to maximize discounted profits are shown to converge sufficiently rapidly that they leave the monopolist ignorant about all but the most local features of demand. The failure of the monopolist to "learn" his demand curve would seem to call into question some standard assumptions about agents' grasp of their economic environment.

Keywords: *Stochastic approximation, Statistical learning, Demand analysis*

1. INTRODUCTION

Clower (1959) describes the dilemma of the monopolist facing an unknown, but non-stochastic, demand function. By a sequence of Newton-like steps he gradually finds his way to a profit maximizing output/pricing policy, and yet is left ignorant of the global features of his economic environment, learning only about its local appearance near the profit maximizing price. Since Clower there have been many investigations that introduced stochastic features into this monopoly predicament. The monopolist is always torn between the desire to maximize his (presumably discounted) stream of profits and his wish to learn more about his economic environment. Learning is presumably helpful in the relentless quest for better profits, but too much experimentation leads to sacrifices in valuable early profits.

Following Grossman et al. (1977) most of the literature has posed such problems in the context of a dynamic programming formulation with Bayesian updating of the monopolist's beliefs about demand and cost. See, e.g. Easley and Kiefer (1988) and Nyarko (1991). Even in very simple environments with known parametric forms for demand and cost, this approach yields computationally challenging problems. In this brief note I would like to consider a much simpler class of sequential design strategies employing the stochastic approximation methods introduced by Kiefer and Wolfowitz (1952). I will show that although these methods can achieve a certain optimality from a profitability perspective, they fail miserably in the secondary objective of quenching the monopolist's curiosity about the precise form of his demand curve.

2. THE MODEL

To make life relatively easy for our monopolist we will suppose that he knows the parametric form of his log-quadratic demand function, but not its parameters,

$$D(p, u) = \exp\{\alpha_0 + \alpha_1 \log(p) + \alpha_2 (\log(p))^2 + u\}.$$

For added simplicity we will take the random demand shocks, u, to be iid Gaussian with mean zero, and constant variance, σ^2. Cost will be assumed to be non-stochastic and

again log-quadratic,

$$C(x) = \exp\{\beta_0 + \beta_1 \log(x) + \beta_2 (\log(x))^2\},$$

The monopolist's objective is to choose a sequence of prices, $p = \{p_1, p_2, \cdots\}$, to maximize expected discounted profit,

$$\mathbb{E}R(p) = \mathbb{E}\sum_{t=1}^{\infty} \delta^t \pi(p_t) \equiv \mathbb{E}\sum_{t=1}^{\infty} \delta^t (p_t X_t - C(X_t)),$$

where $X_t = D(p_t, u_t)$. Since cost is nonstochastic it is easy to learn its parameters, so we will treat them as known from the outset and focus on the unknown demand parameters and the choice of a profit maximizing pricing sequence.

3. STOCHASTIC APPROXIMATION

Robbins and Monro (1951) introduced a recursive strategy for finding a zero crossing of the expected value of a random function; Kiefer and Wolfowitz (1952) noted that similar methods could be employed to find the argmax of an unknown regression function. The recursion may be viewed as an adaptive variant of the classical Newton method. Given an initial guess p_0 of $p^* = \text{argmax}\ \mathbb{E}\pi(p)$ we allow our monopolist "two bites of the apple" in each period, that is he is allowed to see independent realizations of profits at both $p_t + c_t$ and $p_t - c_t$ and then update according to the strategy,

$$p_{t+1} = p_t + a_t(\pi(p_t + c_t) - \pi(p_t - c_t))/(2c_t).$$

Allowing the monopolist to do central differencing introduces an additional element of symmetry that simplifies the analysis of the pricing strategies somewhat. Note that the form of the chosen pricing strategy does not exploit the particular parametric form of the demand function, but relies only on smoothness of the profit function. In this respect it differs from more formal Bayesian updating schemes that would exploit this parametric structure and will be will be briefly touched upon later.

The sequences $\{a_t\}$ and $\{c_t\}$, as shown by Fabian (1971), are optimally taken to be $a_t = a_0 t^{-1}$ and $c_t = c_0 t^{-1/6}$ respectively. Assuming that expected profit,

$$\pi_0(p) = \mathbb{E}\pi(p),$$

has a continuous third derivative in a neighborhood of the profit maximizing price, p^*, Fabian notes that with these choices,

$$n^{1/3}(p_n - p^*) \rightsquigarrow \mathcal{N}(\mu(p^*), \sigma^2(p^*)),$$

where

$$\mu(p) = \frac{-a_0 c_0^2 \pi_0^{'''}(p)}{3(a_0 \pi_0^{''}(p) - 1/3)}$$

and

$$\sigma^2(p) = \frac{a_0^2 \sigma^2}{2c_0^2 (a_0 \pi_0^{''}(p) - 1/3)}.$$

Optimal choices of the constants, a_0 and c_0 were shown by Abdelhamid (1973) to be, $a_0 = (\pi_0^{''}(p^*))^{-1}$ and $c_0 = ((3\sigma^2)/(2(\pi_0^{'''}(p^*)^2)))^{1/6}$. Improved rates can be achieved, as shown by Fabian, by exploiting stronger smoothness assumptions and requiring more observations per update, but this would only accentuate the dilemma that we would like

to highlight here so we will restrict our attention to the classical form of the Kiefer-Wolfowitz proposal.

The cube root convergence of the pricing sequence is obviously unfortunate from a profit maximizing perspective; the stereotypic greedy monopolist may feel entitled to the parametric rate $O(1/\sqrt{n})$. Fabian's results reveal that even after adopting the optimal choices of the $\{a_t\}$ and $\{c_t\}$ sequences, we obtain a biased limiting distribution, after normalization, for the $\{p_t\}$ sequence. This is an inevitable consequence of the non-parametric nature of problem of finding a root of $\pi_0'(p) = 0$; in the original Robbins-Monro setting one has $\sqrt{n}$ convergence to a normal limit without the possible asymptotic bias effect induced when $\pi_0'''(p^*) \neq 0$. For our monopolist both bias and variance of the pricing sequence decline like $n^{-1/3}$ as p^* is approached. Stochastic approximation must, however, be gauged a huge success for profit maximization compared to its ability to reveal the precise form of the demand curve.

4. LEARNING ABOUT DEMAND

Even the limited success we can achieve in finding p^* poses some serious problems for the curious monopolist who would like to learn more about his demand function. After n steps of the stochastic approximation iteration, our monopolist has $3n$ observations on the demand curve, but the experimental design of these observations makes them less than ideal for learning about demand.

Before considering the quadratic specification, suppose that the demand were known to be linear in price. In this case the precision of the slope parameter in a demand regression is proportional to the sum of squares of the centered design observations. Let $z_t = \log(p_t/p^*)$, since z_t is converging to zero at the rate, $n^{-1/3}$, we have

$$\sum_{t=1}^{n} z_t^2 \approx K \sum_{t=1}^{n} t^{-2/3} \approx K n^{1/3},$$

and consequently our estimated slope parameter of the demand function converges at the phlegmatic rate of $O_p(n^{-1/6})$.

The quadratic case is even more troubling since the quadratic coefficient of the demand curve is estimated with precision proportional to the fourth moment of $\log p$, and thus,

$$\sum_{t=1}^{n} z_t^4 \approx K \sum_{t=1}^{n} t^{-4/3} \approx 3.60K < \infty,$$

and consequently the quadratic coefficient *cannot* be consistently estimated, as explained in further detail in the Appendix. Despite the conscientious efforts of the monopolist to profit maximize, or indeed because of these efforts, he is doomed to ignorance about his own demand environment. For those imagining a more ambitious nonparametric demand environment the situation is even more bleak.

4.1. A Numerical Example

To illustrate the foregoing situation we consider the following explicit version of the model:

$$D(p, u) = \exp\{5 - 0.5\log(p) - 0.35(\log(p))^2 + u\},$$

where u is iid Gaussian with mean zero, and constant variance, $\sigma^2 = 0.0025$. Cost is

$$C(x) = \exp\{2 + 0.3\log(x) + 0.1(\log(x))^2\},$$

In Figure 1 we see a scatter plot of 1000 realizations of profit evaluated at uniformly distributed p_t over the interval $[3, 60]$. Superimposed is an estimate of the expected profit function, $\mu_0(p)$. Of course, any self-respecting monopolist would be crazy to spread his pricing strategy so thinly over such a wide range of prices far from the optimum, but we not are so constrained.

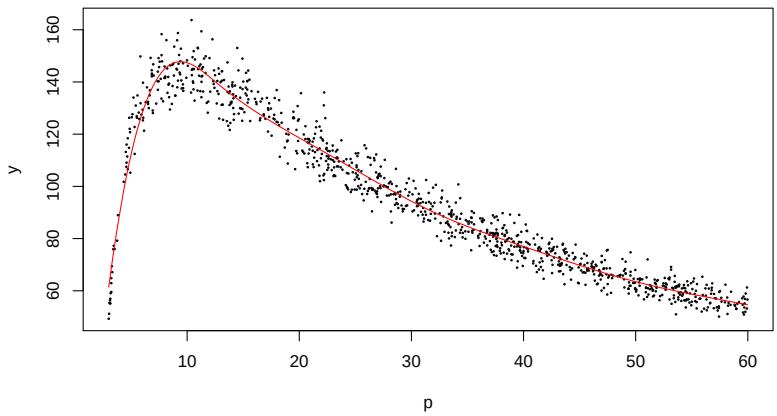

Figure 1. Realized profit for 1000 uniformly distributed p_t on the interval $[3, 60]$. An estimate of the expected profit function is superimposed on the scatterplot.

In Figure 2 we illustrate a more realistic sequence of prices and their corresponding demand realizations; these prices arise from 100 iterations of the basic Kiefer-Wolfowitz procedure. As described above, every third p_t is a "real" iterate and the intermediate evaluations of demand are used to compute the estimated derivative. In the figure we have connected the former points to illustrate the Kiefer-Wolfowitz trajectory. Clearly there is a rapid ascent to the profit maximizing value, and the isolated points on either side of this value illustrate the continued effort to refine this estimate with very gradually declining bandwidths for the derivative estimates. The starting value for this example was $p_0 = 15$, so we are able to see reasonably well what demand looks like for $p \in [8, 15]$, but this is not sufficient to allow us to estimate well the demand curve. Least squares on the full sample of 300 observations yields:

$$\log x_t = \underset{(0.48)}{4.58} - \underset{(0.41)}{0.16} \log p_t - \underset{(0.09)}{0.25} (\log p_t)^2.$$

Inference based on these estimates is fine, the true parameters are covered by their respective confidence intervals, but even after extensive continued iteration, because of

the concentration of the design points around p^*, we fail to get convergent estimates of the demand parameters.

Results in this section can be reproduced by downloading the R package `monopoly` from `http://www.econ.uiuc.edu/~roger/research/monopoly/mono.html`. In particular, Figures 1 and 2 can be reproduced by running `demo(fig1)` and `demo(fig2)`.

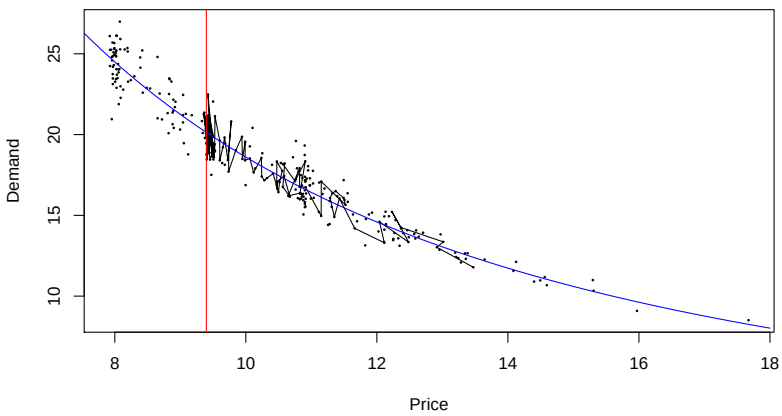

Figure 2. Realized price and demand for 100 iterations of the Kiefer-Wolfowitz process. The trajectory of the central iterates is depicted in black and the isolated points are those required to compute the difference quotients. A least squares estimate of the demand curve is over plotted in blue, and the final estimate of the profit maximizing price is depicted by the vertical red line.

5. RAMIFICATIONS

Fabian (1967) has proposed a more elaborate variant of the Kiefer-Wolfowitz procedure whose convergence rate approaches the parametric, $O_p(1/\sqrt{n})$ ideal provided that the underlying $\mu_0(p)$ function is sufficiently smooth near p^*. This can be regarded as a stochastic approximation variant of the more familiar use of higher order kernels to do bias removal for non-parametric density estimation. Ruppert (1991) notes that there is little practical (simulation) experience regarding this procedure. It may be noted that any "improved" performance of the sequence $\{p_t\}$ from the profitability perspective, accentuates the difficulties from a demand analysis perspective. In this case we have,

$$\sum_{t=1}^{n} z_t^2 = K \sum_{t=1}^{n} t^{-1} \approx K \log n,$$

and even the slope coefficient for demand is barely estimable, converging at the rate $O_p(1/\sqrt{\log n})$.

Of course knowledge of the parametric form of cost and demand, iid error for the

random component of demand, stationarity of the demand process and a variety of other simplifying assumptions can be easily called into question. But relaxing such conditions only leaves the ignorant monopolist deeper in the dark.

Since stochastic approximation purports to achieve optimal rates of convergence for the profit maximizing sequence of prices in problems of the foregoing type, it would be of interest to see whether Bayesian updating could improve upon the simple methods described above, and if so by how much. Bayesian updating could presumably exploit knowledge of the parametric form of the demand function, however to the extent that more profitable changes in the monopolist's pricing strategy were made, they may only further deepen his ignorance about the global nature of demand.

ACKNOWLEDGEMENTS

This note began life as an attempt to introduce some experimental design ideas into my graduate econometric theory course. It was partially supported by NSF grant SES-08-50060. Thanks to the referee and editor whose comments helped to clarify the exposition.

REFERENCES

Abdelhamid, S. (1973). Transformation of observations in stochastic approximation. *Annals of Statistics 1*, 1158–1174.

Clower, R. (1959). Some theory of an ignorant monopolist. *Economic Journal 69*, 705–716.

Easley, D. and N. Kiefer (1988). Controlling a stochastic process with unknown parameters. *Econometrica 56*, 1045–1064.

Fabian, V. (1967). Stochastic approximation of minima with improved asymptotic speed. *Annals of Mathematical Statistics 38*, 191–200.

Fabian, V. (1971). Stochastic approximation. In J. Rustagi (Ed.), *Optimizing Methods in Statistics*, pp. 439–470. Academic Press: New York.

Grossman, S., R. Kihlstrom, and L. Mirman (1977). A Bayesian approach to the production of information and learning by doing. *The Review of Economic Studies 44*, 533–547.

Kiefer, J. and J. Wolfowitz (1952). Stochastic estimation of the maximum of a regression function. *Annals of Mathematical Statistics 23*, 462–466.

Lai, T. and H. Robbins (1977). Strong consistency of least-squares estimates in regression models. *PNAS 74*, 2667–2669.

Nyarko, Y. (1991). On the convergence of Bayesian posterior processes in linear economic models. *J. of Economic Dynamics and Control 15*, 687–713.

Robbins, H. and S. Monro (1951). A stochastic approximation method. *Annals of Mathematical Statistics 22*, 400–407.

Ruppert, D. (1991). Stochastic approximation. In B. Ghosh and P. Sen (Eds.), *Handbook in Sequential Analysis*, pp. 503–529. Dekker: New York.

A. DESIGN CONDITIONS FOR CONVERGENCE OF THE LEAST SQUARES REGRESSION ESTIMATOR

Consider the bivariate linear regression model,

$$y_i = \alpha + \beta x_i + u_i \quad i = 1, \cdots, n,$$

with iid $\{u_i\}$ and $\mathbb{E}u_1 = 0$ and $\mathbb{E}u_1^2 \log(1+|u_i|^r) < \infty$ for some $r > 1$. A necessary as well as sufficient condition for the least squares estimator $\hat{\beta}_n$ to be consistent for β is that,

$$\lim_{n\to\infty} A_n \equiv \lim_{n\to\infty} \sum (x_i - \bar{x})^2 = \infty.$$

When this condition holds, Lai and Robbins (1977) prove that for every $\delta > 0$,

$$\lim_{n\to\infty} \frac{\sqrt{A_n}(\hat{\beta} - \beta)}{(\log A_n)^{(1+\delta)/2}} = 0,$$

with probability one, so $\hat{\beta}$ converges to β essentially at rate $O_p(1/\sqrt{A_n})$. This can be strengthened to limiting normality for $\sqrt{A_n}(\hat{\beta} - \beta)$ provided that $\max |x_i|/\sqrt{A_n} \to 0$, which assures the Lindeberg condition.

For the quadratic model,

$$y_i = \alpha + \beta x_i + \gamma x_i^2 + u_i \quad i = 1, \cdots, n,$$

we can project $z = (x_i^2)$ onto the orthogonal complement of $X = (1, x_i)$, say $\tilde{z} = M_X z$, thus reducing the question of consistency of the least squares $\hat{\gamma}$ to the requirement that $\|\tilde{z}\|^2 \to \infty$. But

$$\|\tilde{z}\|^2 \leq \|z\|^2 \equiv \sum_{i=1}^{n} x_i^4,$$

so, for example, when $x_i = x_0 + \xi_i$ with $\xi_i = O_p(i^{-1/3})$, as in the monopoly problem described above, $\|\tilde{z}\|^2$ fails to diverge, and consequently $\hat{\gamma}$ is inconsistent.