Supplement to 'Semi-Nonparametric Estimation of Consumer Search Costs' *

José Luis Moraga-González[a,b], Zsolt Sándor †[c], and Matthijs R. Wildenbeest[d]

[a]*Faculty of Economics and Business Administration, VU University, Amsterdam, The Netherlands*
[b]*Faculty of Economics and Business, University of Groningen, Groningen, The Netherlands*
[c]*Faculty of Economic and Human Sciences, Sapientia University, Miercurea Ciuc, Romania*
[d]*Kelley School of Business, Indiana University, Bloomington, IN, USA*

Abstract

This supplement consists of two sections. The first section presents the results of several Monte Carlo exercises that study the behavior of our SNP estimator in small samples. The second section provides condition on the unknown density under which the estimator is consistent using i.i.d. observations.

*Moraga-González is also affiliated with the Tinbergen Institute, the CEPR, and the Public-Private Sector Research Center (IESE, Barcelona).

†Correspondence to: Zsolt Sándor, Faculty of Economic and Human Sciences, Sapientia University, Piaţa Libertăţii 1, 530104 Miercurea Ciuc, Romania. E-mail: zsosan@gmail.com

1 Monte Carlo study

In addition to studying the small sample properties of the estimator, the Monte Carlo study serves three other goals: (i) we study how using cross-validation to pick the number of polynomial terms in the SNP density function performs in our setting; (ii) we compare the performance of our estimator to an estimator that does not directly link different markets but instead estimates search costs market-by-market (based on Moraga-González and Wildenbeest, 2008); and (iii) we estimate the coverage probability of our method of obtaining the bootstrap confidence interval of the estimated search cost density. We focus on the estimation of the following search cost density:

$$g_0(c) = 0.5 \cdot \text{lognormal}(c, 2, 10) + 0.5 \cdot \text{lognormal}(c, 3, 0.5), \tag{S1}$$

where $\text{lognormal}(c, a, b)$ refers to the densities of the lognormal distribution with parameters a and b, respectively. To make sure we are working in an environment that is not very different from the one used in our application in Section 4 of the main text we take $M = 10$ markets. Each market has the same search cost distribution but a different valuation net of marginal cost, $v^m - r^m$. The 10 values we use for $v^m - r^m$ are $\{40, 80, \ldots, 400\}$. For each market m, we set K^m, the maximum number of prices a consumer can observe, equal to 35.[1] With the parameters of a market m at hand, we compute the market equilibrium by numerically solving the system of equations (8) in the main text. Given the cutoff values for a market m, we construct the equilibrium price distribution in that market m using equation (9) in the main text. Next, we randomly draw 35 prices from each equilibrium price distribution F^m and use all 350 prices as an input for the SNP estimation procedure. The estimation is replicated 100 times.[2]

[1] Typically, the number of firms operating in a market will vary from market to market. Though this constitutes an additional source of variation, we do not need to use it here since we are assuming that the valuation net of marginal cost is different across markets.

[2] To gain computing time we use the empirical distribution of prices in each market to estimate the c_k's. Although consistency of the estimator is preserved, this is likely to lead to less precise estimates, so our results should be seen as a lower bound on the performance of the estimator when using equation (8) in the main text instead.

1.1 Cross-validation

In practice the number of polynomial terms N has to be chosen in an optimal way. For this, we can build on the cross-validation method of Coppejans and Gallant (2002). The essence of their cross-validation is to determine N for the data at hand by minimizing some loss function. Let f denote the true price density function and $\widehat{f}_N$ the price density function estimate computed as

$$\widehat{f}_N(p) = f(p|\widehat{g}_N),$$

where $\widehat{g}_N$ is the estimated search cost density with N polynomial terms. A standard way of choosing N is by minimizing the integrated squared error (ISE), which in our case is

$$\int_{\underline{p}}^{\overline{p}} \left(\widehat{f}_N(p) - f(p)\right)^2 dp.$$

Since the true distribution $f(p)$ is not known, the ISE needs to be approximated.

There are various problem-specific ways to approximate the ISE; we proceed as follows. First write

$$\int_{\underline{p}}^{\overline{p}} \left(\widehat{f}_N(p) - f(p)\right)^2 dp = \int_{\underline{p}}^{\overline{p}} \widehat{f}_N^2(p)\,dp - 2\int_{\underline{p}}^{\overline{p}} \widehat{f}_N(p) f(p)\,dp + \int_{\underline{p}}^{\overline{p}} f^2(p)\,dp$$

and note that on the RHS only the first two terms depend on N. The first term $\int_{\underline{p}}^{\overline{p}} \widehat{f}_N^2(p)\,dp$ can be estimated (for example) by Monte Carlo simulations by drawing a sample from the uniform distribution on $\left[\underline{p}, \overline{p}\right]$. The integral from the second term can be written as

$$\int_{\underline{p}}^{\overline{p}} \widehat{f}_N(p) f(p)\,dp = E_P\left[\widehat{f}_N(p)\right],$$

which can be estimated by using the price observations in one market, i.e.,

$$\int_{\underline{p}}^{\overline{p}} \widehat{f}_N(p) f(p)\,dp \approx \frac{1}{K}\sum_{k=1}^{K} \widehat{f}_N(p_k).$$

In the empirical example the prices from different markets have different distributions, so we take the approximation of the average ISE across markets

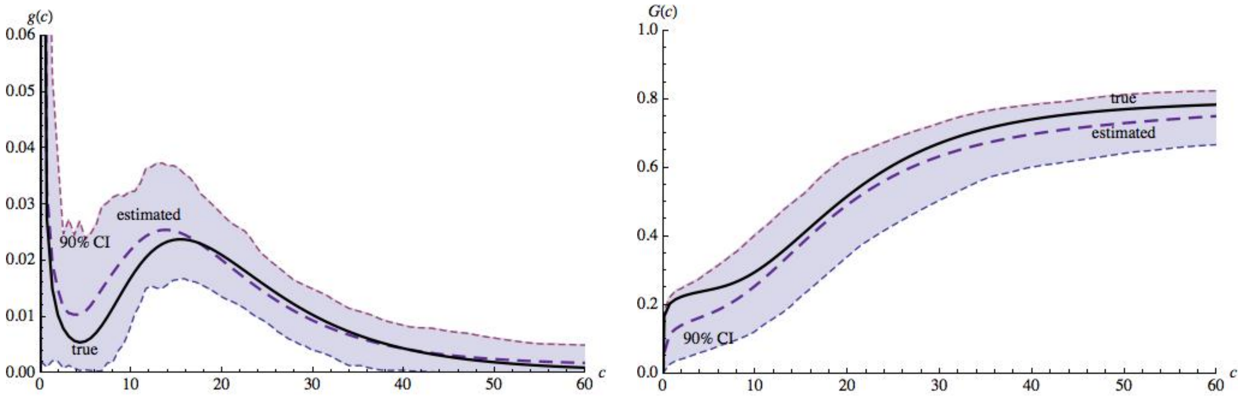
$$\frac{1}{M}\sum_{m=1}^{M}\int_{\underline{p}}^{\overline{p}}\left(\widehat{f}_N^m(p)-f^m(p)\right)^2 dp.$$

Table 1 shows the outcome of the Monte Carlo simulations for various values of N. The approximated (feasible) estimate of ISE selects $N = 8$, while the true (unfeasible) criterion selects $N = 6$ as the optimal number of polynomial terms. Search costs, however, are closest to the true search cost distribution when using $N = 8$, as shown in the last column of Table 1. This suggests our method works reasonably well.

Table 1: Monte Carlo results

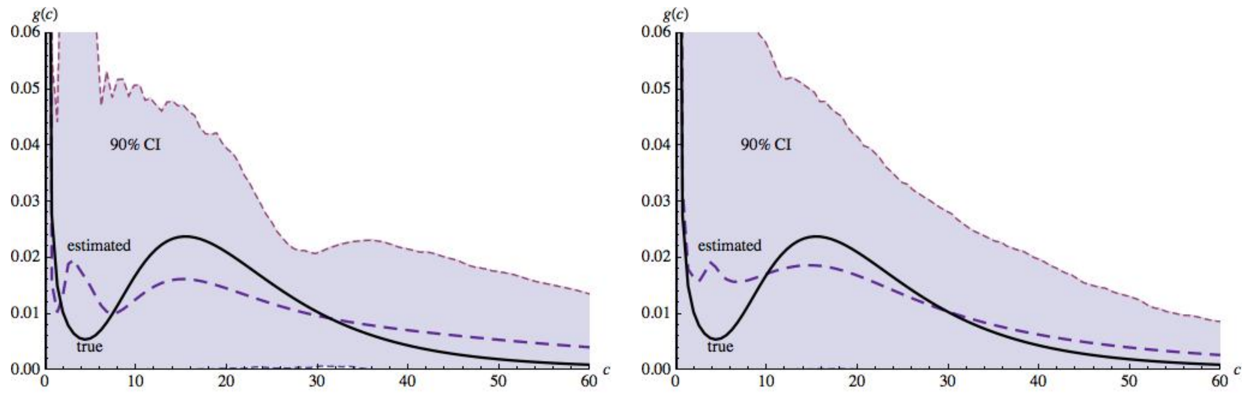
	Prices		Search costs
	ISE (approx.)	ISE (true) $\times 10^{-2}$	ISE (true) $\times 10^{-4}$
$N=1$	-0.097	1.031	1.625
$N=2$	-0.235	0.838	0.181
$N=3$	-0.236	0.841	0.147
$N=4$	-0.244	0.844	0.143
$N=5$	-0.241	0.829	0.121
$N=6$	-0.242	0.822	0.097
$N=7$	-0.243	0.824	0.092
$N=8$	-0.245	0.823	0.077
$N=9$	-0.244	0.839	0.078

Notes: ISE values are calculated for the mean price and search cost densities of the 100 replications.



(a) Search cost PDF (b) Search cost CDF

Figure 1: Monte Carlo results: estimated search costs for $N = 8$



(a) One market ($v = 400$)

(b) All markets ($N = 8$)

Figure 2: Monte Carlo results: estimated search costs (market-by-market)

Figures 1(a) and 1(b) show the estimated search cost distribution for $N = 8$. We report the mean and the 90 percent confidence interval of the 100 replications. Figure 1(a) corresponds to the search cost PDF, while Figure 1(b) corresponds to the search cost CDF. In both graphs, the solid curve represents the true search cost distribution, while the thick dashed curve shows the mean of the 100 estimations. The 90 percent confidence interval is given by the shaded area between the thin dashed curves. In spite of the relatively small number of markets and observations per market, the figures illustrate that our estimation procedure performs fairly well. The estimates mimic the true shape of the search cost PDF as well as CDF relatively well at most quantiles. Note that if we were to add more markets with relatively high valuation to our data set the number of search cost cutoffs would increase, which would improve the outcome of the estimation.

1.2 Comparison to market-by-market estimators

Existing approaches to estimate search costs (e.g., Hong and Shum, 2006; Moraga-González and Wildenbeest, 2008) are designed to estimate search costs market-by-market, while our SNP estimation procedure is specifically set up to maximize the joint likelihood from all markets. Figure 2(a) shows the estimated search cost PDF when we take the existing approach and use data for only one market.[3] Not only are the differences between the true search cost distribution

[3]We use prices for the market with $v = 400$ to make sure the maximum identifiable search cost value is the same as in our main specification. To obtain a parametric estimate of the search cost density we fit a SNP density function with $N = 8$ polynomial terms through the identified points on the search cost distribution,

(solid curves) and the mean of the 100 fitted distributions (thick dashed curves) larger than when using our multi-market SNP estimation procedure, also the 90 percent confidence interval (shaded area) is much wider. The search costs ISE confirms our visual findings: when taking data from just one market, the ISE takes on value 0.294×10^{-4}, which is almost four times as large as the corresponding ISE value for our SNP estimation procedure. If, alternatively, we use the data from all the markets and after estimating search costs market-by-market we take the average search cost density as an estimate of the overall search cost distribution, our SNP estimation procedure still outperforms the market-by-market approach, as illustrated in Figure 2(b). Although the search cost ISE in this case is slightly lowered to 0.229×10^{-4}, the 90 percent confidence interval widens. In sum, Figures 2(a) and 2(b) provide evidence that the market-by-market approach underperforms vis-à-vis our multi-market SNP approach. It is less efficient because search costs are only constrained to be similar across markets after search costs have already been estimated for each market separately. However, we note that since the market-by-market approach is designed to maximize the likelihood function in each market separately it does an equally good job in terms of fitting the model to observed prices.

1.3 Bootstrap confidence interval

For the SNP estimator g_N of the search cost density g, for a given search cost value c, asymptotically

$$\frac{g_N(c) - E[g_N(c)]}{\sqrt{\mathrm{var}(g_N(c))}} \sim N(0,1).$$

This should be true at least when we take the number of polynomial terms in g_N fixed. If in addition $E[g_N(c)] \underset{N\to\infty}{\longrightarrow} g(c)$, then asymptotically

$$\frac{g_N(c) - g(c)}{\sqrt{\mathrm{var}(g_N(c))}} \sim N(0,1).$$

This yields the 95%-confidence interval

$$\left(g_N(c) - 1.96\sqrt{\mathrm{var}(g_N(c))}, g_N(c) + 1.96\sqrt{\mathrm{var}(g_N(c))}\right).$$

which are obtained using the approach in Moraga-González and Wildenbeest (2008).

We use the following steps to estimate var($g_N(c)$) by bootstrap resampling.

1. Draw M markets randomly with replacement.
2. Within each drawn market m draw K^m prices from the estimated price distribution.
3. Keeping the estimated values for v and r fixed, use the drawn prices to estimate g_N.
4. Repeat steps 1-3 a large number of times to obtain a sample of g_N estimates.
5. For each search cost value c from the grid estimate var($g_N(c)$); this yields a confidence band.

To study the coverage probability of the above method to estimate var($g_N(c)$) by bootstrap resampling we construct the 95%-confidence interval for each of the 100 replications, using 100 repetitions. Next, for each replication we check if the true value of the search cost density is within the confidence interval. Table 2 gives the percentage of replications for which this is the case for selected search cost values. Ideally, this happens in 95% of the replications; Table 2 shows that for most search cost values the percentages for the search cost PDF are not too far from 95%. Note that even though the bootstrap confidence interval is derived for $g_N(c)$, the last two columns of Table 2 show that it does reasonably well for the search cost CDF as well, at least for search costs that are not too small.

Table 2: Results bootstrap confidence interval

	PDF		CDF	
search cost	true	within bounds	true	within bounds
1	0.0196	98%	0.2104	57%
5	0.0057	92%	0.2436	70%
10	0.0171	97%	0.2968	87%
15	0.0238	99%	0.4039	92%
25	0.0153	97%	0.6089	94%
35	0.0067	95%	0.7142	94%
50	0.0019	99%	0.7709	95%

Notes: Results are based on 100 replications.

2 Consistency search cost density estimator

In this section of the Supplement we adapt the general conditions in Gallant and Nychka (1987, henceforth GN) for the consistency of the search cost density estimator and discuss some primitive conditions specific to our model. Since the price observations in our model come from multiple markets that may be heterogeneous in valuations, firms' costs and number of firms, the price observations may not be i.i.d. In order to be able to treat the prices as i.i.d., we will regard these conditioning variables as random. This is not restrictive since it is just a matter of interpretation; in fact it is analogous to treating the covariates in a regression as random, in order to have i.i.d. dependent variables.

For this purpose, let us first modify the notation of the price density to $f\left(p|g;v^m,r^m,K^m\right)$ in order to make explicit the dependence on valuations, firms' costs and number of firms. Then

$$L_M\left(g\right)=\frac{1}{M}\sum_{m=1}^{M}\left(\frac{1}{K^m}\sum_{i=1}^{K^m}\log f\left(p_i^m|g;v^m,r^m,K^m\right)\right),$$

is the log-likelihood presented above, where for simpler notation we ignore the price vectors on the LHS. We regard the triplets $\left(v^m,r^m,K^m\right)_{m=1}^M$ as an i.i.d. sample of random variables. Then by Kolmogorov's strong law of large numbers, $L_M\left(g\right)\underset{M\to\infty}{\overset{a.s.}{\longrightarrow}}L\left(g\right)\equiv E\left[\log f\left(p|g;v,r,K\right)\right]$, provided that $E\left[\log f\left(p|g;v,r,K\right)\right]<\infty$ (this condition will follow from Lemma S.1 below).[4]

In order to state sufficient conditions for the consistency of our search cost density estimator, we introduce some further notation. Recall that the search costs c we consider are exponential transformations of the random variables x from GN, that is, $c=\exp\left(\gamma+\sigma x\right)$. The density of

[4]Note that

$$L\left(g\right)=E\left[\frac{1}{M}\sum_{m=1}^{M}\left(\frac{1}{K^m}\sum_{i=1}^{K^m}\log f\left(p_i|g;v^m,r^m,K^m\right)\right)\right].$$

Indeed, since $\left(\frac{1}{K^m}\sum_{i=1}^{K^m}\log f\left(p_i|g;v^m,r^m,K^m\right)\right)_{m=1}^M$ is an i.i.d. sample, by the law of iterative expectation,

$$\begin{aligned}E\left[\frac{1}{M}\sum_{m=1}^{M}\left(\frac{1}{K^m}\sum_{i=1}^{K^m}\log f\left(p_i|g;v^m,r^m,K^m\right)\right)\right] &= E\left[\frac{1}{K^m}\sum_{i=1}^{K^m}E\left[\log f\left(p_i|g;v^m,r^m,K^m\right)|\,K^m\right]\right]\\ &= E\left[E\left[\log f\left(p_i|g;v^m,r^m,K^m\right)|\,K^m\right]\right],\end{aligned}$$

where the last equality holds because in each market m the prices $\left(p_i\right)_{i=1}^{K^m}$ are i.i.d.. Then by the law of iterative expectation the statement follows.

c is $g(c) = \frac{1}{\sigma c} h\left(\frac{\log c - \gamma}{\sigma}\right)$, where h denotes the density of x. Let

$$\mathcal{G} = \left\{ g : g(c) = \frac{1}{\sigma c} h\left(\frac{\log c - \gamma}{\sigma}\right), \gamma \in \mathbb{R}, \sigma > 0, h \in \mathcal{H} \right\}$$

denote the set of admissible search cost densities, where $\mathcal{H}$ is the set of admissible densities defined by GN (p.369). For each $\gamma \in \mathbb{R}$, $\sigma > 0$ define the operator $\|\cdot\| : \mathcal{G} \to \mathbb{R}$ such that $\|g\| = \|h\|_{GN}$, where $\|\cdot\|_{GN}$ is the consistency norm from GN (p.371), and define the operator $T : \mathcal{H} \to \mathcal{G}$ with $T(h)(c) = \frac{1}{\sigma c} h\left(\frac{\log c - \gamma}{\sigma}\right)$. Then $\|\cdot\|$ is a norm on $\mathcal{G}$ and T is a homeomorphism between the normed spaces $(\mathcal{H}, \|\cdot\|_{GN})$ and $(\mathcal{G}, \|\cdot\|)$.

Let $g_0 \in \mathcal{G}$ be the true search cost density and $\mathcal{G}_N = \{g_N(\cdot; \gamma, \sigma, \theta) : \gamma \in \mathbb{R}, \sigma > 0, \theta \in \Theta_N\}$ the space of SNP estimators, where $g_N(\cdot; \gamma, \sigma, \theta)$ is defined in equation (11) in the main text. Denote the SNP estimator of g_0 by $\widehat{g}$, let the number of observations be n.

Proposition S.1 *Under the following conditions:*

(a) Compactness: The closure of $\mathcal{G}$ is compact,

(b) Denseness: $\cup_{N \geq 1} \mathcal{G}_N$ is dense in $\mathcal{G}$ and $\mathcal{G}_N \subset \mathcal{G}_{N+1}$,

(c) Continuity: $E[\log f(p|g; v, r, K)]$ is continuous in g,

(d) Dominance: There is a function $B(p; v, r, K) > 0$ with $E[B(p; v, r, K)] < \infty$ such that $\log f(p|g; v, r, K) \leq B(p; v, r, K)$ for any g and any $(p; v, r, K)$,

(e) Identification: For any density g with support $(0, \infty)$ such that

$$E[\log f(p|g; v, r, K)] \geq E[\log f(p|g_0; v, r, K)]$$

$g = g_0$ must hold, $\lim_{n\to\infty} \|\widehat{g} - g_0\| = 0$ almost surely, provided that $N \equiv N_n \to \infty$.

This result is a modified version of Theorem 0 in GN. The modification consists of replacing uniform convergence of the objective function by a one-sided uniform convergence implied by Condition (d) and partially by Condition (c), which is possible for maximum likelihood estimators, as shown by Wald (1949).

In the sequel we discuss briefly how Conditions (a)-(e) can be verified for our problem. Condition (a) follows from Theorem 1 in GN that states that the closure of $\mathcal{H}$ is compact, which is homeomorphic to $\mathcal{G}$ for given γ, σ and by assuming that the location and scale parameters γ, σ are in a compact subset of $\mathbb{R} \times (0, \infty)$. Condition (b) follows from Theorem 2 in GN by using the homeomorphism between $\mathcal{H}$ and $\mathcal{G}$. Whether Condition (c) is satisfied or not depends on whether the price density $f(p|g; v, r, K)$ is continuous in g. This mild condition appears to be very difficult to verify due to the implicit nature of the price distribution and the nonlinearity of the system of equations that determines the price distribution. Condition (d) is a one-sided dominance condition for which we provide primitive conditions in Lemma S.1 below. These primitive conditions are sufficient for the case when firms' marginal cost r is estimated from an additional data source, so we can regard the valuations and marginal costs in every market as known by the econometrician.[5] Condition (e) can be verified under the conditions of our identification result in Proposition 2 in the main text by using the (Shannon-Kolmogorov) Information Inequality.

Lemma S.1 *(1) For any density g with support $(0, \infty)$ and any $(p; v, r, K)$*

$$\log f(p|g; v, r, K) \leq |\log(v - r)| + |\log(p - r)| + |\log(v - p)| \equiv B(p; v, r, K).$$

(2) Assume that g_0 and the joint distribution of (v, r, K) satisfy the following conditions: (i) $f(v, r, K)$ is bounded; (ii) either (A) g_0 has at least polynomial upper tail, i.e., there is $\alpha > 0$, $L > 0$, $\bar{c} > 1/2$ such that $g_0(c) \geq Lc^{-1-\alpha}$ for any $c > \bar{c}$ and $\int v^{\alpha+1} |\log v| f(v) dv < \infty$, or (B) g_0 has at least exponential upper tail, i.e., there is $\alpha > 0$, $L > 0$, $\bar{c} > 1/2$ such that $g_0(c) \geq Le^{-\alpha c}$ for any $c > \bar{c}$ and the distribution of valuations has at most exponential upper tail, i.e., there is $\alpha' > 0$, $L' > 0$, $\bar{c}' > 1/2$ such that $g_0(c) < L'e^{-\alpha' c}$ for any $c > \bar{c}'$ with $\alpha' > \alpha$.

We note that conditions (i) and (ii) are somewhat restrictive, but they still allow the one-sided dominance condition to hold for a large class of search cost and valuation distributions. Condition (ii) suggests that there is a trade-off between the restrictions on the tails of the search cost and valuation distributions.

[5] In this respect our results are incomplete, but we believe they are still interesting because they serve as an illustration of how one can verify the dominance condition (d) in a structural model so highly nonlinear as ours.

Proof of Lemma S.1. For notational simplicity let us drop the conditioning variables v, r, K from $f(p|g; v, r, K)$. From equations (10) and (4) in the main text

$$f(p|g) = \frac{\mu_1 (v-r)}{(p-r)^2 \sum_{k=2}^{K} k(k-1)\mu_k (1-F(p|g))^{k-2}}, \tag{S2}$$

and since

$$\sum_{k=2}^{K} k(k-1)\mu_k (1-F(p|g))^{k-2} \geq \sum_{k=2}^{K} k\mu_k (1-F(p|g))^{k-1},$$

we obtain

$$\begin{aligned}
f(p|g) &\leq \frac{\mu_1 (v-r)}{(p-r)^2 \sum_{k=2}^{K} k\mu_k (1-F(p|g))^{k-1}} \\
&= \frac{\mu_1 (v-r)}{(p-r)^2 \sum_{k=1}^{K} k\mu_k (1-F(p|g))^{k-1} - \mu_1 (p-r)^2} \\
&= \frac{\mu_1 (v-r)}{\mu_1 (p-r)(v-r) - \mu_1 (p-r)^2} = \frac{(v-r)}{(p-r)(v-p)},
\end{aligned}$$

where the last-but-one equality follows from equation (4) in the main text. That is,

$$\log f(p|g) \leq \log\left[\frac{(v-r)}{(p-r)(v-p)}\right] \leq |\log(v-r)| + |\log(p-r)| + |\log(v-p)| = B(p; v, r, K).$$

This establishes (1). In what follows we prove (2). We have

$$\begin{aligned}
E[B(p; v, r, K)] &= \int (|\log(v-r)| + |\log(p-r)| + |\log(v-p)|) f(p|g_0) f(v, r, K)\, dp d(v, r, K) \\
&= \int \left[\int_{\underline{p}_0}^{v} (|\log(v-r)| + |\log(p-r)| + |\log(v-p)|) f(p|g_0)\, dp\right] f(v, r, K)\, d(v, r, K) \\
&\equiv I_1 + I_2 + I_3,
\end{aligned} \tag{S3}$$

where $f(v, r, K)$ is the joint density of (v, r, K). Below we prove that the three integrals $I_1,\ I_2,\ I_3$ are finite.

Bounding I_1. We have

$$I_1 = \int |\log(v-r)| \left[\int_{\underline{p}_0}^{v} f(p|g_0)\, dp\right] f(v, r, K)\, d(v, r, K) = \int |\log(v-r)| f(v, r, K)\, d(v, r, K).$$

This can be split such that

$$\int |\log(v-r)| f(v,r,K)\, d(v,r,K) = \int_{v-r\leq 1} |\log(v-r)| f(v,r,K)\, d(v,r,K) + \int_{v-r>1} \log(v-r) f(v,r,K)\, d(v,r,K).$$

The first term is finite by Condition (i) and the fact that $\int_0^1 |\log x|\, dx = 1$. The second term is also finite because

$$\begin{aligned}\int_{v-r>1} \log(v-r) f(v,r,K)\, d(v,r,K) &< \int_{v>1} \log(v) f(v,r,K)\, d(v,r,K) = \int_{v>1} \log(v) f(v)\, dv \\ &< \int v f(v)\, dv,\end{aligned}$$

which is finite by Condition $(ii$,A). Here and throughout this proof $f(v)$ denotes the marginal density of v.

Bounding I_2. We have

$$I_2 = \int \left[\int_{\underline{p}_0}^{v} |\log(p-r)| f(p|g_0)\, dp\right] f(v,r,K)\, d(v,r,K);$$

First focus on the integral in the brackets. Since

$$\sum_{k=2}^{K} k(k-1)\mu_k (1-F(p|g))^{k-2} \geq 2\mu_2$$

from equation (S2) we obtain,

$$f(p|g) \leq \frac{\mu_1 (v-r)}{2(p-r)^2 \mu_2} = \left(\frac{v-r}{p-r}\right)^2 f(p|g)|_{p=v}. \tag{S4}$$

Then

$$\begin{aligned}\int_{\underline{p}_0}^{v} |\log(p-r)| f(p|g_0)\, dp &\leq \int_{\underline{p}_0}^{v} |\log(p-r)| \left(\frac{v-r}{p-r}\right)^2 f(v|g_0)\, dp \\ &= (v-r) f(v|g_0) \int_{\underline{p}_0}^{v} |\log(p-r)| \frac{v-r}{(p-r)^2} dp,\end{aligned}$$

where

$$\begin{aligned}\int_{\underline{p}_0}^{v} |\log (p-r)| \frac{v-r}{(p-r)^2} dp &= \int_1^{\frac{v-r}{\underline{p}_0 - r}} \left|\log\left(\frac{v-r}{x}\right)\right| dx \\ &\leq \int_1^{\frac{v-r}{\underline{p}_0 - r}} |\log (v-r)| \, dx + \int_1^{\frac{v-r}{\underline{p}_0 - r}} \log x dx \\ &= |\log (v-r)| \left(\frac{v-r}{\underline{p}_0 - r} - 1\right) + \frac{v-r}{\underline{p}_0 - r}\left(\log \frac{v-r}{\underline{p}_0 - r} - 1\right) + 1 \\ &\leq |\log (v-r)| \frac{v-r}{\underline{p}_0 - r} + \frac{v-r}{\underline{p}_0 - r} \log \frac{v-r}{\underline{p}_0 - r} + 1. \qquad \text{(S5)}\end{aligned}$$

So

$$\int_{\underline{p}_0}^{v} |\log (p-r)| f(p|g_0) \, dp \leq (v-r) f(p|g_0)|_{p=v} \left[|\log (v-r)| \frac{v-r}{\underline{p}_0 - r} + \frac{v-r}{\underline{p}_0 - r} \log \frac{v-r}{\underline{p}_0 - r} + 1\right].$$

Based on this, we need to show that

$$J_1 = \int (v-r) f(v|g_0) |\log (v-r)| \frac{v-r}{\underline{p}_0 - r} f(v,r,K) \, d(v,r,K) < \infty, \qquad \text{(S6)}$$

$$J_2 = \int (v-r) f(v|g_0) \frac{v-r}{\underline{p}_0 - r} \log \frac{v-r}{\underline{p}_0 - r} f(v,r,K) \, d(v,r,K) < \infty, \qquad \text{(S7)}$$

$$J_3 = \int (v-r) f(v|g_0) f(v,r,K) \, d(v,r,K) < \infty. \qquad \text{(S8)}$$

We expect that $f(v|g_0) < M$ for some appropriate M for any (v,r,K) because $f(v|g_0) \to 0$ when $v \to \infty$, since $f(v|g_0)$ is the density at the upper bound of its support, although we find it difficult to prove this formally. Further, by equation (5) in the main text

$$\frac{v-r}{\underline{p}_0 - r} = \frac{\sum_{k=1}^{K} k\mu_{k0}}{\mu_{10}} > 1,$$

where $(\mu_{k0})_{k=1}^{K}$ correspond to the true g_0. The numerator is bounded, in fact $\sum_{k=1}^{K} k\mu_{k0} \in [1, K]$ for any g_0. By equation $(3a)$ from the main text,

$$\frac{1}{\mu_{10}} = \frac{1}{1 - G_0(c_{10})} \leq \frac{1}{1 - G_0\left(\frac{v-r}{2}\right)}$$

because G_0 is increasing and from equation (8) from the main text

$$c_{10} = \int_0^1 \left(\frac{\mu_1(v-r)}{\sum_{k=1}^K k\mu_k(1-z)^{k-1}} + r \right) (2z-1)\, dz \leq \mu_1(v-r) \int_0^1 \frac{|2z-1|}{\sum_{k=1}^K k\mu_k(1-z)^{k-1}} dz$$
$$\leq (v-r) \int_0^1 |2z-1|\, dz = \frac{v-r}{2},$$

where the latter inequality follows by taking $z = 1$ in the denominator. Therefore,

$$\frac{v-r}{\underline{p}_0 - r} \leq \frac{K}{1 - G_0\left(\frac{v-r}{2}\right)}. \tag{S9}$$

Now we proceed by proving equations (S6)-(S8). Applying equation (S9), we have

$$\begin{aligned} J_1 &\leq MK \int \frac{(v-r)\left|\log(v-r)\right|}{1-G_0\left(\frac{v-r}{2}\right)} f(v,r,K)\, d(v,r,K) \\ &= MK \int_{v-r \leq 2\overline{c}} \frac{(v-r)\left|\log(v-r)\right|}{1-G_0\left(\frac{v-r}{2}\right)} f(v,r,K)\, d(v,r,K) \\ &+ MK \int_{v-r > 2\overline{c}} \frac{(v-r)\log(v-r)}{1-G_0\left(\frac{v-r}{2}\right)} f(v,r,K)\, d(v,r,K). \end{aligned}$$

The first term is finite because the function $x \log x$ is bounded on any bounded interval and $1/\left[1 - G_0\left(\frac{v-r}{2}\right)\right] \leq 1/\left[1 - G_0(\overline{c})\right]$. For the second term we note that $x \log x / \left[1 - G_0\left(\frac{x}{2}\right)\right]$ is an increasing function in x, so

$$\begin{aligned} \int_{v-r>2\overline{c}} \frac{(v-r)\log(v-r)}{1-G_0\left(\frac{v-r}{2}\right)} f(v,r,K)\, d(v,r,K) &< \int_{v>2\overline{c}} \frac{v\log v}{1-G_0\left(\frac{v}{2}\right)} f(v,r,K)\, d(v,r,K) \\ &= \int_{v>2\overline{c}} \frac{v\log v}{1-G_0\left(\frac{v}{2}\right)} f(v)\, dv. \end{aligned}$$

Under Condition (ii,A), for $c > \overline{c}$

$$1 - G_0(c) = \int_c^\infty g_0(x)\, dx \geq \int_c^\infty L x^{-1-\alpha} dx = L\frac{c^{-\alpha}}{\alpha},$$

so for $v > 2\overline{c}$

$$1 - G_0\left(\frac{v}{2}\right) \geq 2^\alpha L \frac{v^{-\alpha}}{\alpha}. \tag{S10}$$

Therefore the second integral term bounding J_1 is less than

$$\alpha 2^{-\alpha} L^{-1} MK \int_{v>2\overline{c}} v^{\alpha+1} \log(v) f(v) \, dv < \alpha 2^{-\alpha} L^{-1} MK \int v^{\alpha+1} |\log v| f(v) \, dv < \infty,$$

the latter inequality by the second part of Condition (ii,A). This proves $J_1 < \infty$.

Under Condition (ii,B), for $c > \overline{c}$

$$1 - G_0(c) = \int_c^\infty g_0(x) \, dx \geq \int_c^\infty L e^{-\alpha x} dx = L \frac{e^{-\alpha c}}{\alpha},$$

so for $v > 2\overline{c}$

$$1 - G_0\left(\frac{v}{2}\right) \geq \frac{L}{\alpha} e^{-\frac{\alpha v}{2}}. \tag{S11}$$

Therefore the second integral term bounding J_1 is less than

$$\alpha L^{-1} MK \int_{v>2\overline{c}} v \log(v) e^{\frac{\alpha v}{2}} f(v) \, dv < \alpha L^{-1} L' MK \int_{v>2\overline{c}} v \log(v) e^{-\left(\alpha' - \frac{\alpha}{2}\right) v} dv < \infty,$$

where the former inequality follows from the second part of Condition (ii,B). This proves $J_1 < \infty$.

Now, applying again equation (S9), we have

$$J_2 \leq MK \int \frac{(v-r)}{1 - G_0\left(\frac{v-r}{2}\right)} \log\left(\frac{K}{1 - G_0\left(\frac{v-r}{2}\right)}\right) f(v, r, K) \, d(v, r, K). \tag{S12}$$

This can be split into the sum of two integrals:

$$\begin{aligned} &MK \int_{v-r\leq 2\overline{c}} \frac{(v-r)}{1 - G_0\left(\frac{v-r}{2}\right)} \log\left(\frac{K}{1 - G_0\left(\frac{v-r}{2}\right)}\right) f(v, r, K) \, d(v, r, K) \\ &+ MK \int_{v-r> 2\overline{c}} \frac{(v-r)}{1 - G_0\left(\frac{v-r}{2}\right)} \log\left(\frac{K}{1 - G_0\left(\frac{v-r}{2}\right)}\right) f(v, r, K) \, d(v, r, K). \end{aligned}$$

The first integral is less than

$$MK \frac{2\overline{c}}{1 - G_0(\overline{c})} \log\left(\frac{K}{1 - G_0(\overline{c})}\right) < \infty.$$

The second integral can be bounded in a way similar to the second integral term of J_1. We obtain

$$\begin{aligned}
&\int_{v-r>2\bar{c}} \frac{(v-r)}{1-G_0\left(\frac{v-r}{2}\right)} \log\left(\frac{K}{1-G_0\left(\frac{v-r}{2}\right)}\right) f(v,r,K)\, d(v,r,K) \\
&< \int_{v>2\bar{c}} \frac{v}{1-G_0\left(\frac{v}{2}\right)} \log\left(\frac{K}{1-G_0\left(\frac{v}{2}\right)}\right) f(v,r,K)\, d(v,r,K) \\
&= \int_{v>2\bar{c}} \frac{v}{1-G_0\left(\frac{v}{2}\right)} \log\left(\frac{K}{1-G_0\left(\frac{v}{2}\right)}\right) f(v)\, dv. \qquad \text{(S13)}
\end{aligned}$$

Under Condition (ii,A), from equation (S10) this is less than

$$\begin{aligned}
&\alpha 2^{-\alpha} L^{-1} \int_{v>2\bar{c}} v^{\alpha+1} \log\left(\frac{\alpha K}{2^{\alpha} L} v^{\alpha}\right) f(v)\, dv \\
&= \alpha 2^{-\alpha} L^{-1} \int_{v>2\bar{c}} v^{\alpha+1} (\log a + \alpha \log v) f(v)\, dv \\
&= \alpha 2^{-\alpha} L^{-1} \log a \int_{v>2\bar{c}} v^{\alpha+1} f(v)\, dv + \alpha^2 2^{-\alpha} L^{-1} \int_{v>2\bar{c}} v^{\alpha+1} |\log v|\, f(v)\, dv \\
&< \left(\alpha 2^{-\alpha} L^{-1} \log a + \alpha^2 2^{-\alpha} L^{-1}\right) \int_{v>2\bar{c}} v^{\alpha+1} |\log v|\, f(v)\, dv,
\end{aligned}$$

where $a = \alpha 2^{-\alpha} K L^{-1}$. Consequently, Condition (ii,A) implies that this is finite, and therefore $J_2 < \infty$.

Under Condition (ii,B), from equation (S11) the expression in equation (S13) is less than

$$\begin{aligned}
&\alpha L^{-1} \int_{v>2\bar{c}} v e^{\frac{\alpha v}{2}} \log\left(\frac{\alpha K}{L} e^{\frac{\alpha v}{2}}\right) f(v)\, dv \\
&= \alpha L^{-1} \int_{v>2\bar{c}} v e^{\frac{\alpha v}{2}} \left(\log a + \frac{\alpha v}{2}\right) f(v)\, dv \\
&= \alpha L^{-1} \log a \int_{v>2\bar{c}} v e^{\frac{\alpha v}{2}} f(v)\, dv + \frac{\alpha^2 L^{-1}}{2} \int_{v>2\bar{c}} v^2 e^{\frac{\alpha v}{2}} f(v)\, dv \\
&< \alpha L^{-1} L' \log a \int_{v>2\bar{c}} v e^{-\left(\alpha' - \frac{\alpha}{2}\right) v} dv + \frac{\alpha^2 L^{-1} L'}{2} \int_{v>2\bar{c}} v^2 e^{-\left(\alpha' - \frac{\alpha}{2}\right) v} dv < \infty,
\end{aligned}$$

where $a = \alpha K L^{-1}$. Consequently, $J_2 < \infty$.

The statement in equation (S8) follows easily from the second part of Condition (ii,A). This completes the proof of $I_2 < \infty$.

Bounding I_3. We have

$$I_3 = \int \left[\int_{\underline{p}_0}^{v} |\log (v-p)| \, f(p|g_0) \, dp \right] f(v,r,K) \, d(v,r,K) \, .$$

The integral in the brackets is

$$\begin{aligned}
\int_{\underline{p}_0}^{v} |\log (v-p)| \, f(p|g_0) \, dp &\leq \int_{\underline{p}_0}^{v} |\log (v-p)| \left(\frac{v-r}{p-r} \right)^2 f(p|g_0) \, |_{p=v} dp \\
&= (v-r) \, f(v|g_0) \int_{\underline{p}_0}^{v} |\log (v-p)| \, \frac{v-r}{(p-r)^2} dp,
\end{aligned}$$

where

$$\begin{aligned}
\int_{\underline{p}_0}^{v} |\log (v-p)| \, \frac{v-r}{(p-r)^2} dp &= \int_1^{\frac{v-r}{\underline{p}_0-r}} \left| \log \left[(v-r) \, \frac{x-1}{x} \right] \right| dx \\
&\leq \int_1^{\frac{v-r}{\underline{p}_0-r}} |\log (v-r)| \, dx + \int_1^{\frac{v-r}{\underline{p}_0-r}} \log x dx - \int_1^{\frac{v-r}{\underline{p}_0-r}} \log (x-1) \, dx \\
&= |\log (v-r)| \left(\frac{v-r}{\underline{p}_0-r} - 1 \right) + \frac{v-r}{\underline{p}_0-r} \log \frac{v-r}{\underline{p}_0-r} \\
&- \left(\frac{v-r}{\underline{p}_0-r} - 1 \right) \log \left(\frac{v-r}{\underline{p}_0-r} - 1 \right) \\
&\leq |\log (v-r)| \, \frac{v-r}{\underline{p}_0-r} + \frac{v-r}{\underline{p}_0-r} \log \frac{v-r}{\underline{p}_0-r} \qquad \text{(S14)} \\
&- \left(\frac{v-r}{\underline{p}_0-r} - 1 \right) \log \left(\frac{v-r}{\underline{p}_0-r} - 1 \right).
\end{aligned}$$

So we need to show that

$$\begin{aligned}
H_1 &= \int (v-r) \, f(v|g_0) \, |\log (v-r)| \, \frac{v-r}{\underline{p}_0-r} f(v,r,K) \, d(v,r,K) < \infty, \qquad \text{(S15)} \\
H_2 &= \int (v-r) \, f(v|g_0) \left[\frac{v-r}{\underline{p}_0-r} \log \frac{v-r}{\underline{p}_0-r} \right. \\
&\left. - \left(\frac{v-r}{\underline{p}_0-r} - 1 \right) \log \left(\frac{v-r}{\underline{p}_0-r} - 1 \right) \right] f(v,r,K) \, d(v,r,K) < \infty. \qquad \text{(S16)}
\end{aligned}$$

The first statement is proved in equation (S6). For the second statement we note that the function $x \log x - (x-1)\log(x-1)$ is increasing in x. Therefore, by equation (S9)

$$\begin{aligned} H_2 < M \int (v-r) \Bigg[& \frac{K}{1-G_0\left(\frac{v-r}{2}\right)} \log \frac{K}{1-G_0\left(\frac{v-r}{2}\right)} \\ & - \left(\frac{K}{1-G_0\left(\frac{v-r}{2}\right)} - 1 \right) \log \left(\frac{K}{1-G_0\left(\frac{v-r}{2}\right)} - 1 \right) \Bigg] f(v,r,K)\, d(v,r,K) \\ < MK \int & \frac{(v-r)}{1-G_0\left(\frac{v-r}{2}\right)} \log \frac{K}{1-G_0\left(\frac{v-r}{2}\right)} f(v,r,K)\, d(v,r,K) . \end{aligned}$$

The latter expression is the same as the RHS expression in equation (S12), which we have already proved to be finite. Consequently, $I_3 < \infty$. This completes the proof that $E[B(p;v,r,K)] < \infty$. ■

References

Coppejans M, Gallant AR. 2002. Cross-validated SNP density estimates. *Journal of Econometrics* **110**: 27–65.

Gallant AR, Nychka DW. 1987. Semi-nonparametric maximum likelihood estimation. *Econometrica* **55**: 363–390.

Hong H, Shum M. 2006. Using price distributions to estimate search costs. *RAND Journal of Economics* **37**: 257–275.

Moraga-González JL, Wildenbeest MR. 2008. Maximum likelihood estimation of search costs. *European Economic Review* **52**: 820–48.

Wald A. 1949. Note on the consistency of the maximum likelihood estimate. *Annals of Mathematical Statistics* **20**: 595–601.