Technical Appendix – Price Optimisation Newsletter

Variance rstimation for the two estimators, with illustrative magnitudes

All figures below are illustrative orders of magnitude for the medium-sized PCW broker of the essay: roughly $N = 1.8 \times 10^8$ quote-board appearances per year, of which 5×10^5 land in top position, yielding 5×10^4 click-throughs and 10^4sales. Mean premium is taken as £600, and where segmentation is needed we assume $S = 50$ pricing segments and weekly repricing (52 windows per year). Outcomes vary by firm; the orders, not the exact digits, carry the argument.

A.1 Notation and the two estimands

Let $x$denote risk features, $p$the broker's retail price ($p = c + m$, net rate plus margin), $Y \in \{0,1\}$ the sale, and $M = \min_{j} P_j(x)$ the best competing price on the board, an order statistic of the competitor quote vector. Define the win indicator $W = \mathbf{1}\{p < M\}$ and the survival function of the market minimum

$$F_{\text{win}}(p \mid x) = \Pr(M > p \mid x), f_{\text{win}}(p \mid x) = -\partial F_{\text{win}} / \partial p,$$

so $F_{\text{win}}$is read directly as the probability of winning the board at price p, and f_{win} as the rate at which that probability erodes per pound of price.

Traditional estimand. A logit demand model $q(p,x) = \Pr(Y = 1 \mid p,x) = \Lambda(\alpha_x + \beta p)$, with elasticity $\varepsilon = \beta\, p\, (1 - q)$. The optimiser consumes β(equivalently ε).

Proposed estimand. The factorisation

$$q(p,x) = F_{\text{win}}(p \mid x) \cdot g_{\text{click}}(x,p), g_{\text{click}} = \Pr(Y = 1 \mid \text{top}, x, p),$$

taking, for clarity, the top-of-board event as the surrogate; the multi-rank generalisation adds terms but changes nothing structural. Stage one is estimated from the 1.8×10^8 boards; stage two from the 5×10^5 wins.

A.2 Fisher information of the conversion estimator

For a logit model the Fisher information for β is $I(\beta) = \sum_i q_i\,(1 - q_i)\,\tilde{p}_i^{\,2}$, where $\tilde{p}_i$ is the price *residual* after the risk features have absorbed their share. Hence

$$\text{Var}(\hat{\beta}) \approx \frac{1}{n\, q_{\text{avg}}(1 - q_{\text{avg}})\, \sigma_p^2}.$$

Two facts about this expression do all the damage:

(i) The rare-event limit. At board level, $q_{\text{avg}} = 10^4/1.8 \times 10^8 \approx 5.6 \times 10^{-5}$, so $n\, q_{\text{avg}}(1 - q_{\text{avg}}) \approx n\, q_{\text{avg}} = 10^4$: for rare outcomes the information is governed by the *event count*, not the sample size. The hundred and eighty million rows collapse to ten thousand effective observations. Modelling at click level instead ($n = 5 \times 10^4$, $q_{\text{avg}} = 0.2$) gives $n\, q_{\text{avg}}(1 - q_{\text{avg}}) = 8{,}000$, essentially the same number by a different route. There is no clever choice of denominator that recovers the lost information; the sales are the sales.

(ii) The identification term. σ_p^2 is the variance of *exogenous* own-price variation conditional on x. A deterministic rating engine sets this to zero exactly; what rescues it in practice is drift in your rates relative to the market across time. Grant, generously, a residual standard deviation of 3% of premium, $\sigma_p = £18$. Then

$$\mathrm{SE}(\hat{\beta}) = (10^4 \times 18^2)^{-1/2} \approx 5.6 \times 10^{-4} \text{ per £}.$$

Converting to elasticity via $\varepsilon = \beta p(1 - q)$ with $p = £600$: $\mathrm{SE}(\hat{\varepsilon}) \approx 0.33$. Against a true PCW-channel elasticity of order -4, the *portfolio-level, full-year* coefficient of variation is a respectable 8%. But nobody prices a portfolio with one number. Divide the information across $S = 50$ segments: the SE inflates by $\sqrt{50}$ to ≈ 2.4, a segment-level cv of roughly **60%**. Ask additionally for currency within a quarter (the market reprices weekly, so a year-old elasticity is a fossil) and the cv passes 100%. Per segment per *week*, the dataset contains $10^4/(50 \times 52) \approx 3.8$ sales. One does not estimate a derivative from four coin flips; one consults one's priors and calls it calibration.

A.3 Information content of the board

A board observation does not merely report the Bernoulli outcome W at your realised price; it reveals the realised value of M itself, i.e. one draw from the entire object being estimated. The empirical win curve on a segment-week cell with $n_{\text{cell}} = 1.8 \times 10^8/(50 \times 52) \approx 69{,}000$ boards satisfies

$$\mathrm{Var}\ (\hat{F}_{\text{win}}(p)) = \frac{F_{\text{win}}(1 - F_{\text{win}})}{n_{\text{cell}}} \ \leq\ \frac{1}{4\, n_{\text{cell}}}, \mathrm{SE} \leq 0.0019.$$

So the win probability at any candidate price is pinned to within ±0.4 percentage points (95%) *per segment, per week*, nonparametrically, before any smoothing borrows strength across cells.

The optimiser also needs the derivative $f_{\text{win}}(p)$. A kernel estimate with bandwidth $h = £10$ at the operating point uses roughly $2h\, f_{\text{win}}\, n_{\text{cell}}$ local observations. With the market minimum dispersed with a standard deviation of order £60 (10% of premium, an illustrative figure), $f_{\text{win}} \approx 0.007$ per £ near the centre of the action, giving $\approx 9{,}000$ effective local observations per cell and

$$\frac{\mathrm{SE}(\hat{f}_{\text{win}})}{f_{\text{win}}} \;\approx\; \frac{1}{\sqrt{9{,}000}} \;\approx\; 1\%.$$

Set the two gradients side by side. The quantity the optimiser actually differentiates is estimated with relative error of order **1% per segment-week** under the market model, versus roughly **60% per segment-year** under the conversion model. As a variance ratio at matched granularity this is a factor in the thousands; at matched *timeliness* (the market model refreshes weekly into a weekly-moving market) the comparison stops being quantitative and becomes anatomical.

A.4 The composite, and a cancellation

The two-stage predictor is $\hat{q} = \hat{F}_{\text{win}} \cdot \hat{g}_{\text{click}}$, with independent samples, so by the delta method

$$\frac{\mathrm{Var}(\hat{q})}{q^2} \;\approx\; \frac{\mathrm{Var}(\hat{F}_{\text{win}})}{F_{\text{win}}^2} + \frac{\mathrm{Var}(\hat{g}_{\text{click}})}{g_{\text{click}}^2}.$$

Stage two is fed by the funnel's middle: per segment-year, 10^4 top positions and 200 sales, so $\hat{g}_{\text{click}} = 0.02$ carries $\mathrm{SE} = \sqrt{0.02 \times 0.98/10^4} \approx 0.0014$, a relative error of about 7% (shrinkable well below that by partial pooling, since $g_{\text{click}}$is slowly varying by construction). The composite volume prediction is therefore accurate to a few per cent, dominated by the *level* term g_{click}.

But now observe what the margin decision requires. Expected profit per board is

$$\Pi(m) \;=\; F_{\text{win}}(c+m) \;\cdot\; g_{\text{click}} \;\cdot\; V(m),$$

and the first-order condition $\Pi'(m_{\text{opt}}) = 0$ reads $F'_{\text{win}}\, V + F_{\text{win}}\, V' = 0$: **the factor** g_{click} **cancels**. To first order, the optimal margin is a functional of the market model alone, i.e. of the well-estimated factor; the scarce-data factor sets the scale of profit but not its argmax. It re-enters only through any residual price-dependence $\frac{\partial g_{\text{click}}}{\partial p}$ (the post-click, on-screen price effect), a second-order correction which the 5×10^4clicks are perfectly adequate to bound. The architecture thus routes the abundant data to the derivative and the scarce data to a level, which is precisely the opposite of the traditional routing, where four coin flips a week were asked to supply a derivative.

With $V(m) = m$, the FOC gives the inverse-hazard rule

$$m_{\text{opt}} \;=\; \frac{F_{\text{win}}(p_{\text{opt}})}{f_{\text{win}}(p_{\text{opt}})}, p_{\text{opt}} = c + m_{\text{opt}},$$

so the relative error of $\widehat{m}_{\text{opt}}$ is, to first order, $\left[\text{relvar}(\hat{F}_{\text{win}}) + \text{relvar}(\hat{f}_{\text{win}})\right]^{1/2} \approx$ **1–2%**. Under the traditional method the analogous rule is $m_{\text{opt}} \approx -p/\varepsilon \cdot (1 + o(1))$, so the relative error of the margin inherits the relative error of the elasticity: roughly **60%** at segment level, before staleness.

A.5 What the variance costs in money

Near the optimum the profit function is locally quadratic, so an unbiased margin error $\delta = \Delta m/m_{\text{opt}}$ forfeits expected profit $\approx \frac{\kappa}{2}\,\mathbb{E}[\delta^2] = \frac{\kappa}{2}\,\text{cv}^2$, with curvature κ of order unity for win curves with exponential-type tails (for the exact exponential, $\Pi \propto m\, e^{-m/m_{\text{opt}}}$, the shortfall is $1 - (1+\delta)e^{-\delta} \approx \delta^2/2$). Plugging in:

	cv of $\widehat{m}_{\text{opt}}$	**expected profit forgone**
Conversion model, segment level	~0.6	~15–18% of attainable margin
Market model, segment-week	~0.015	~0.01%

This is the essay's claim in one line: the traditional estimator's *sampling variance alone*, before any bias, taxes the broker on the order of a sixth of his optimal margin, and the tax is invisible because sampling error never appears as a line item.

A.6 Honest residuals

Three terms survive on the other:

Surrogacy bias: $\hat{q}$ inherits a bias b from imperfect exclusion (price visible beside rank, brand effects, funnel leakage); the MSE comparison is $b^2 +$ var, and since stage two *estimates* the link rather than assuming it, b is the residual after calibration, plausibly a few per cent of q, against a variance term two orders larger under the alternative.

Nonstationarity: every quantity above carries an implicit shelf life of days; the large N is what makes the weekly refresh statistically affordable, but the estimand is a snapshot of an equilibrium, not a constant of nature.

Selection on winning: the 5×10^5 wins are not a random draw of boards; $\mathbb{E}[\text{value} \mid W = 1] \neq \mathbb{E}[\text{value}]$, so $V(m)$ must be estimated conditional on the win event, from exactly the sample the funnel provides for it. None of these three erodes the central ratio, which is not a modelling refinement but a change of measurement: one method observes the target variable four times a week, and the other observes it sixty-nine thousand times.