Profit Aware Ad Ranking with Relevance Constraint

Manavender Malgireddy
mmalgireddy@wayfair.com
Wayfair LLC
Seattle, WA, USA

Dongyue Xie
dxie@wayfair.com
Wayfair LLC
Mountain View, CA, USA

Sergey Kolbin
skolbin@wayfair.com
Wayfair LLC
Seattle, WA, USA

Kurt Zimmer
kzimmer@wayfair.com
Wayfair LLC
Boston, MA, USA

Masoum Mosmer
mmosmer@wayfair.com
Wayfair LLC
Mountain View, CA, USA

Patrick Phelps
pphelps@wayfair.com
Wayfair LLC
Boston, MA, USA

Abstract

Wayfair Sponsored Products (WSP) is a cost-per-click (CPC) advertising program aimed at improving product discovery and driving sales. At the heart of WSP is a ranking function that determines the placement of ads on a page. In this paper, we introduce a new ranking function derived from an optimization framework that maximizes overall profitability by jointly considering advertising revenue from clicks and profit from resulting sales. To maintain a positive customer experience, the formulation includes a relevance constraint. We evaluate the approach through offline simulations using counterfactual estimates of clicks and orders, analyzing the impact of different parameter settings on key metrics such as ad revenue, total profit, and estimated conversion rate (CVR). Finally, we share results from a successful online test, that validated the effectiveness of the proposed method, leading to its full deployment in production, and highlight key learnings from the study.

CCS Concepts

• **Information systems → Sponsored search advertising**.

Keywords

Sponsored Products Advertising, Ranking Optimization, Profit Maximization, Revenue Maximization, Relevance Constraint, Profit-Relevance Tradeoff

ACM Reference Format:
Manavender Malgireddy, Dongyue Xie, Sergey Kolbin, Kurt Zimmer, Masoum Mosmer, and Patrick Phelps. 2025. Profit Aware Ad Ranking with Relevance Constraint. In *Proceedings of KDD '25*. ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 Introduction

In recent years, e-commerce platforms have increasingly leveraged sponsored product advertising to enhance product visibility, boost customer engagement, and increase overall profitability [1, 4]. WSP exemplifies this trend as a cost-per-click (CPC) advertising solution designed explicitly to facilitate product discovery and drive incremental sales on Wayfair's expansive marketplace. Central to the effectiveness of WSP is its underlying ranking function, responsible for determining the order of sponsored products shown to consumers.

This paper introduces a new ranking function developed as a solution to a carefully formulated optimization problem. Unlike, most traditional methods that focus primarily on maximizing immediate advertising revenue from clicks [11], our proposed approach adopts a holistic view of profitability [10]. Specifically, it simultaneously optimizes for both advertising revenue generated through clicks and incremental profit derived from actual product sales. To ensure sustainable engagement and maintain user satisfaction, the optimization framework incorporates a relevance constraint, thereby safeguarding a positive customer experience and preventing irrelevant or overly aggressive advertising placements.

To evaluate and validate our proposed ranking function, we employed offline simulation methodologies leveraging counterfactual estimates of user clicks and purchase behavior. Through these simulations, we systematically explored the influence of different parameter settings on critical performance indicators such as advertising revenue, total profit, and estimated conversion rate (CVR). Our analysis delves deeply into understanding how the sensitivity of ranking positions responds to variations in profit-weighting and the strictness of relevance constraints. Additionally, we address the inherent challenges in striking an appropriate balance between maximizing profitability and preserving customer-centric relevance.

We also present results from an online test of our ranking function on Wayfair, highlighting key benefits and trade-offs of the proposed optimization approach, and providing insights for future improvements.

Key Contributions

(1) **Profit-Relevance Ranking Framework:** A novel ranking function derived from a constrained optimization problem that jointly optimizes total profit and relevance, supported by an efficient parameter optimization algorithm.
(2) **Robust Constraint Selection Method:** The relevance constraint threshold is defined in terms of optimal expected conversion rate providing a robust and interpretable way to adapt the threshold to data/model drift

KDD '25, Toronto, Canada

ACM ISBN 978-x-xxxx-xxxx-x/YYYY/MM
https://doi.org/10.1145/nnnnnnn.nnnnnnn

(3) **Simulation-Based Analysis:** Analyze the trade-off between relevance threshold and key performance metrics through simulations, guiding threshold selection for deployment.
(4) **Empirical Validation via Online A/B Testing:** Extensive online experiments validate the effectiveness of the proposed approach, demonstrating gains in both total profit and relevance metrics, as well as highlighting the contribution of the VCD (sales profit) component. Furthermore, the results highlight the critical role of calibration with respect to VCD.

2 Related Work

Most existing research on enhancing ad ranking focuses on improving click-through rate (CTR) models [1, 2, 4]. Yang et al. [9] have conducted an extensive survey of CTR prediction methods, particularly emphasizing modeling frameworks. More recently, Yan et al. [8] presented an architecture designed to ensemble various feature interaction modules. This architecture incorporates a depth selection controller, which dynamically determines the optimal layer for early exits, thus enhancing prediction accuracy. Additionally, another recent work [6] introduces a multi-task framework tailored for early-stage learning, effectively combining several final-stage ranking factors, such as ad clicks and ad-quality events. This approach has demonstrated improvements in overall ad recall and ranking consistency.

Revenue maximization represents another critical research direction for improving ad ranking. Approaches in this area broadly fall into two categories: i) integrating revenue signals during the training of underlying models, either through weighting or joint learning, and ii) directly using revenue metrics within the ranking formula. In the first category, authors in [7] introduced a framework that jointly models clicks and purchases to optimize sales revenue. Another approach, proposed in [12], involves independently training ranking models for relevance and revenue maximization and introduces a trade-off strategy using constrained optimization.

Our work aligns closely with the second category, wherein we propose a ranking function designed to maximize total profit while adhering to a relevance constraint. The most closely related study to ours is by Ge et al. [10], who introduced a scoring function that incorporates a weight balancing ad revenue against organic revenue. Their approach optimizes this balance per impression through grid search, subject to a constraint based on KL-divergence. Another relevant study [3] introduces distinct weights for ad revenue and advertiser ROI, tuning these weights through split testing with equally-sized buckets. They employ an online learning approach following an epsilon-first strategy, beginning with pure exploration and subsequently transitioning to pure exploitation.

3 Proposed Approach

3.1 Optimization problem

Our objective is to maximize the expected total profit while enforcing a relevance constraint to ensure a high-quality customer experience. We formulate this objective as a constrained optimization problem, whose solution yields a scoring function used to rank ads. To proceed, we first define the following key terms that will support the formulation of this optimization problem.

Decision variables: Let $x_{rs} = 1$ indicate that the SKU (Product) s is selected to be shown as an impression for request r; otherwise $x_{rs} = 0$. The full set of decision variables is defined as:

$$x = \{x_{rs} \in \{0, 1\} : r \in R, s \in S_r\} \tag{1}$$

Expected Total Profit: Captures the expected profit from both ad clicks and product sales:

$$eTotalProfit(x) = \sum_{r \in R} \sum_{s \in S_r} \Big((pClick_{rs} \cdot bid_s) + W_1 (pPurchase_{rs} \cdot VCD_{rs}) \Big) x_{rs}, \tag{2}$$

We assume access to $pClick_{rs}$, $pPurchase_{rs}$, and VCD_{rs} as given inputs derived from upstream machine learning models. The details of these models are beyond the scope of this paper.

where:

- $pClick_{rs}$: Probability of user clicking on a SKU s for request r
- $pPurchase_{rs}$: Probability of user purchasing SKU s for request r
- VCD_{rs}: Predicted variable contribution dollar on SKU s for request r. Sales profit term used in our objective.
- bid_s: Advertiser bid for SKU s
- $W_1 \in [0, 1]$: Weight assigned to profit from sales

Expected Conversion Rate:

$$eCVR(x) = \frac{\sum_{r \in R} \sum_{s \in S_r} pPurchase_{rs} \cdot x_{rs}}{\sum_{r \in R} \sum_{s \in S_r} x_{rs}} \tag{3}$$

Optimization Problem The goal is to maximize expected total profit while maintaining a minimum threshold of relevance, enforced through a constraint on expected conversion rate:

$$\begin{aligned} &\max_{x} eTotalProfit(x) \\ &s.t. \;\; eCVR(x) \geq b_0, \end{aligned} \tag{4}$$

where b_0 is a lower bound on expected conversion rate, used as a relevance constraint to help preserve user experience and engagement.

3.2 Scoring function

We solve the optimization problem in Equation (4) by converting it into a dual problem using the Lagrangian method. The Lagrangian function after regrouping the terms involving x_{rs} is defined as:

$$L(x, \lambda) = \sum_{r \in R} \sum_{s \in S_r} \Big[(pClick_{rs} \cdot bid_s) + W_1 \cdot (pPurchase_{rs} \cdot VCD_{rs}) + \lambda_0 \cdot pPurchase_{rs} - \lambda_0 \cdot b_0 \Big] x_{rs} \tag{5}$$

Here, λ_0 is the Lagrange multiplier associated with the relevance constraint, which enforces a minimum expected conversion rate (eCVR(x) $\geq b_0$). This term reflects the marginal value of satisfying the constraint: a higher value of λ_0 increases the importance of purchase probability in the ranking function, encouraging the selection of more relevant SKUs.

The Lagrangian dual function is defined as:

$$F(\lambda) = \max_{x} L(x, \lambda) \tag{6}$$

and the corresponding dual optimization problem is:

$$\begin{aligned} &\min_{\lambda} F(\lambda) \\ &s.t.\ \ \lambda_0 \geq 0 \end{aligned} \tag{7}$$

Let us denote the expression inside the summation of equation (5) as the scoring function:

$$\begin{aligned} Score(r, s) =& (pClick_{rs} \cdot bid_s) + W_1(pPurchase_{rs} \cdot VCD_{rs}) \\ &+ \lambda_0 \cdot pPurchase_{rs} - \lambda_0 \cdot b_0 \end{aligned} \tag{8}$$

We observe that, for a given value of λ_0, the optimal solution to Equation (6) is obtained by setting:

$$x_{rs} = \begin{cases} 1, & \text{if } Score(r, s) > 0 \\ 0, & \text{otherwise} \end{cases} \tag{9}$$

If we add a constraint—such as selecting exactly 12 items per request—the solution is to sort the SKUs by their scoring values and choose the top 12. More generally, solving Equation (6) involves ranking SKUs in descending order of their scores and selecting the desired number to show as impressions.

Since the term $\lambda_0 \cdot b_0$ is constant across all SKUs for a given request, it doesn't affect their ranking. We can safely remove it from the scoring function to simplify computation. The resulting scoring function is:

$$\begin{aligned} Score(r, s) =& (pClick_{rs} \cdot bid_s) + W_1(pPurchase_{rs} \cdot VCD_{rs}) \\ &+ \lambda_0 \cdot pPurchase_{rs} \end{aligned} \tag{10}$$

This scoring function offers a practical and interpretable approach to ranking SKUs by jointly considering ad revenue, purchase-driven profit, and relevance. The term $\lambda_0 \cdot pPurchase_{rs}$ serves as a relevance incentive, promoting SKUs with higher purchase likelihood. Tuning λ_0 allows control over the trade-off between monetization (through ad revenue and sales profit) and relevance (via estimated conversion rate).

3.3 Policy for picking threshold b_0

The threshold b_0 in equation (4) represents the target average pPurchase value we aim to achieve from the displayed SKUs. To simplify the optimization process, we compute this average using only the top 12 SKUs per page request, as these typically account for the vast majority of ad impressions.

Determining an absolute value for b_0 based on historical data is challenging due to variability in user behavior, changes in the pPurchase model, and other dynamic factors. Instead, we define b_0 relative to the optimal expected conversion rate, ensuring consistency even as the model or data evolves:

$$b_0 = \alpha \cdot eCVR^* \tag{11}$$

Here, $\alpha \in [0, 1]$ is a tunable parameter representing the desired proportion of the optimal performance.
$eCVR^*$ denotes the maximum achievable expected conversion rate, obtained by ranking SKUs purely by their predicted pPurchase values.

For instance, setting $\alpha = 0.90$ implies a deliberate choice to target at least 90% of the optimal $eCVR$. This allows for a controlled trade-off, enabling other priorities such as business objectives, user experience, or content diversity.

This formulation using α enhances interpretability and maintains stability across shifts in data distribution or model updates, while providing flexibility to adapt b_0 over time.

3.4 Solving for λ_0

We utilize historical auction data that includes all candidate SKUs considered in each auction, along with their associated scores—such as *pClick*, *pPurchase*, and *VCD*. For a given value of α, we first compute the corresponding threshold b_0 as described in Section 3.3. We then apply the iterative search procedure outlined in Algorithm 1 to identify the optimal value of λ_0 for the specified W_1 and b_0. Figure 1 illustrates the convergence behavior of the dual and primal objectives, as well as the progression of the λ_0 values during optimization.

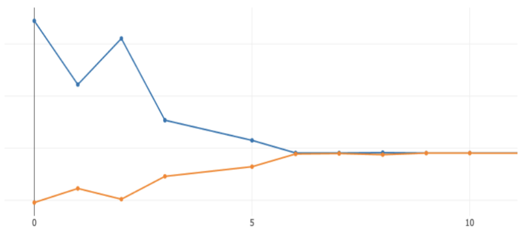
Algorithm 1: Iterative Dual Solver

Input: W_1, b_0, step size μ, iter threshold m, tol., λ
Output: λ, p_{best}, d_{best}
$\lambda \leftarrow 10^6$, $i, t \leftarrow 0$, $d_{\text{best}} \leftarrow \infty$, $p_{\text{best}} \leftarrow -\infty$;
repeat
  Solve Eq. 8 for solution x, dual d, primal p;
  if *constraint violated* **then**
    $\lambda \leftarrow \lambda + \lambda\mu$;
  else
    $\lambda \leftarrow \lambda - \lambda\mu$;
  end
  if $d < d_{best}$ **then**
    $d_{\text{best}} \leftarrow d$, $t \leftarrow 0$;
  else
    $t \leftarrow t + 1$;
  end
  $i \leftarrow i + 1$;
  if $t \geq m$ **then**
    $\mu \leftarrow \mu/2$, $t \leftarrow 0$;
  end
until $|p_{best} - d_{best}| \leq tol.$;

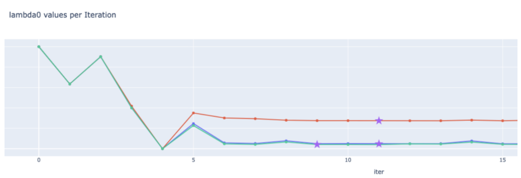
4 Offline Simulations

Before any online testing or deployment, we conduct offline simulations to select the optimal combination of parameters $\{W_1, \lambda_0\}$ in the scoring formula (10). These simulations serve as a critical step in understanding the impact of model parameters before deploying changes in a live environment.

The motivation for offline simulation is twofold: first, to simulate SKU rankings within each auction in order to analyze ranking shifts, SKU dynamics, as well as the ads competitiveness; and second, to validate that the proposed ranking method improves key business metrics, such as total profit and advertising revenue, without sacrificing user experience and business value. By leveraging counterfactual estimates of user clicks and purchases, we



(a) Convergence of Objectives: The Y-axis represents objective values, while the X-axis indicates iterations. The blue line (top) corresponds to the dual objective, and the orange line (bottom) represents the primal objective.



(b) Lambda Convergence: The Y-axis shows lambda values, and the X-axis represents iterations. Each line corresponds to a different platform (Web, MWeb, MApp), with a star indicating the optimal value.

Figure 1: Convergence plots for objectives and lambda

can approximate real-world outcomes across a range of parameter settings, enabling a deeper understanding of model sensitivity and trade-offs.

In this section, we describe the overall simulation methodology and highlight the metrics used to guide decision-making ahead of production rollout.

4.1 Approach

The offline simulation workflow consists of the following key steps:

(1) **Collect** logged auction data from production,
(2) **Re-rank** products using proposed parameter configurations and ranking method,
(3) **Compute** simulation metrics, and
(4) **Compare** results across different settings of the ranking function.

As discussed in the previous section, the value of the Lagrange multiplier λ_0 is determined by solving an optimization problem given a target relative conversion constraint α. For each parameter pair $\{W_1, \alpha\}$, we solve for λ_0, compute a ranking score for each product and sort SKUs in descending order of that score.

In practice, we choose W_1 values using a fixed grid ranging from 0 to 1, typically in steps of 0.01. The α values are chosen relative to the baseline "business-as-usual" (BAU) or in-production configuration. Usually the values are higher than the BAU level but upper bounded by 1. For all metrics requiring counterfactual estimates, we employ the methodology proposed by Nguyen et al. [5].

To comprehensively evaluate the impact of different parameter settings, we examine a range of metrics. Some of them are based on directly observable data, and others estimated through counterfactual modeling. Below, we group these into two these categories:

System-Driven Metrics:

- **Display Price**: The average listed price of products shown in ranked results.
- **Ratings and Review Count**: Aggregated indicators of product quality and popularity.
- **Bid Sensitivity**: Defined as the average change in a SKU's ranking position when its bid is increased by 50%. This metric captures how responsive the system is to bid adjustments. Higher sensitivity indicates a bid-driven system, while lower sensitivity suggests the ranking is more influenced by relevance or profitability factors.

Customer Interaction Metrics:

- **Ad Revenue**: Profit contribution from ad-driven clicks.
- **WSP Product VCD**: Profit from purchases attributed to ad impressions.
- **Total VCD**: Sum of Ad Revenue and WSP Product VCD.
- **ROAS**: Return on ad spend, defined as the ratio of ad revenue to ad cost.
- **CPC**: Cost per click.
- **CTR**: Click-through rate defined as ratio of clicks to impressions.
- **CVR**: Conversion rates defined as ratio of attributed orders to impressions

This structured evaluation allows for systematic tuning of ranking parameters and a better understanding of their impact on profitability, customer experience, and advertiser value.

4.2 Numerical examples

We show the simulation results on auctions from customer search pages. In Figure 2, we present the relative changes in key business metrics compared to the baseline (BAU) configuration. The results correspond to a fixed W_1 value, with the x-axis representing the α value used in the optimization, and the y-axis indicating metric lifts.

As α increases, the Lagrange multiplier λ_0 increases accordingly, reflecting stronger enforcement of the relevance constraint. Total VCD and ROAS both rise sharply beyond $\alpha \approx 0.7$, while ad profit from clicks starts to decline. This outcome suggests a trade-off between product-related profit and advertising revenue.

CTR and CVR also improve as α increases, especially when $\alpha > 0.7$, which indicates that user engagement benefits from more relevant product rankings. At the same time, CPC drops significantly with larger α values. This reduction is due to the decreasing influence of the bidding component in the ranking score. Bid sensitivity also declines, since the ranking becomes less responsive to changes in bids when relevance dominates the scoring logic.

Although CTR increases, it does not fully compensate for the decline in CPC. As a result, ad profit continues to fall as α grows. Display price also trends downward at higher α values. This is consistent with customer preference for more affordable products, which the relevance-focused ranking tends to prioritize.

Overall, these results underscore a fundamental trade-off between profitability and relevance. Tuning the α parameter provides

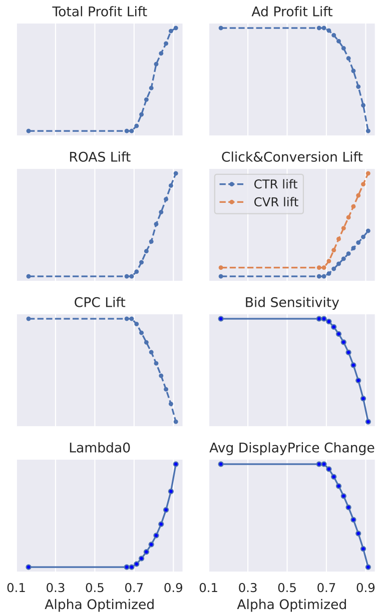


Figure 2: Simulation results for fixed W_1: X-axis shows optimization parameter α, Y-axis shows relative metric changes vs. BAU. Last three plots (Lambda0, bid sensitivity, price change) are *system-driven metrics*.

a practical way to manage this balance and align the ranking strategy with specific business objectives.

Based on simulation results, we selected two variants for online testing. One variant employed conservative parameter values for weight W_1 and α, while the other used more aggressive settings to effectively gauge performance differences. Table 1 summarizes the simulation outcomes for both variants. Subsequently, we'll compare these simulation results to the online test findings in Table 2, highlighting how effectively the simulations guided our choice of variants for testing.

Metric	Sim V1 (Conservative)	Sim V2 (Aggressive)
CTR	13.2	22.5
CVR	31.4	51.6
Total VCD	20.5	32.9
Ad Revenue	0.2	-7.7
ROAS	32.9	71.6
CPC	-11.4	-24.7
Display Price	4.8	36.7

Table 1: Simulation Results % lifts vs BAU.

5 Online Experiments

The baseline (BAU) model for comparison uses the ranking formula $pPurchase * maxBid$. In contrast, the test variants utilize the equation described in Equation (8), each with distinct settings for W_1 and λ_0. We conducted the test for four weeks; Table 2 summarizes the results, indicating that both variants demonstrated overall positive outcomes.

Metric	V1 (Conservative)	V2 (Aggressive)
CTR	13.4	24.1
CVR	22.8	37.7
Total VCD	14.3	24.8
Ad Revneue	1.3	-5.0
ROAS	23.4	42.8
CPC	-10.3	-23.3
Display Price	5.9	29.1

Table 2: Online A/B test results % lifts vs BAU. All numbers are statistically significant at 10%

Both test variants resulted in positive Total VCD, with the aggressive variant achieving a higher lift overall. However, despite its higher Total VCD, the aggressive variant showed negative Ad VCD due to a decline in CPC. This occurred because it assigns greater weight to relevance (*pPurchase*), promoting relevant products even if they have lower bids. Additionally, the aggressive variant caused a significant increase in display prices due to the higher emphasis on W1. Considering these factors, we opted to launch the conservative variant to maintain a balance between Ad VCD and Total VCD, and to minimize the customer impact associated with higher prices.

We now compare the results of the offline simulation with the outcomes of the online experiment. As shown in Tables 1 and 2, the simulation results directionally match the real A/B test outcomes across key metrics. Notably, trends in CTR, Ad VCD, CPC, and Display Price align closely, reinforcing the reliability of the simulation framework. Although other metrics exhibit minor deviations, they remain directionally consistent. These differences are expected because simulations exclusively focus on WSP products, whereas live pages incorporate organic listings, promotional banners, and other dynamic content. Moreover, metrics related to orders naturally exhibit higher variance, introducing additional noise into the simulation results. Overall, the observed consistency supports the use of offline simulations as a dependable proxy for assessing ranking impacts prior to online deployment.

5.1 VCD Sensitivity Test

In the previous test, each of the two variants differed in both the W_1 and λ_0 values, making it difficult to isolate the effect of the VCD term. To address this, we conducted a follow-up test where the control was the previously promoted variant (with both W_1 and λ_0 terms), and the test variant had W_1 set to 0. The results show a 2.9% drop in total VCD (Table 3), with all other ad metrics remaining neutral. This indicates that including the VCD term contributes to increased overall profit, aligning with the intended outcome of the proposed ranking.

5.2 Learnings from an Expansion Test

Following a successful launch in the US, we expanded the online test to the UK using the same methodology. However, we saw

Metrics	No-VCD
CTR	-0.88
CVR	0.90
Total VCD	**-2.94**
Ad Revenue	-0.31
ROAS	0.05
CPC	0.0

Table 3: VCD Sensitivty Test % Lifts vs BAU. Bold = significance at 10%

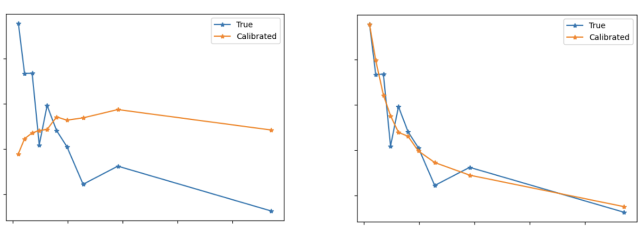
unexpected drops in key metrics like CTR and CVR. These declines are typically linked to either poor model calibration or a low α_0 value. We verified that overall calibration quality was good, and the treatment's α_0 was actually higher than the control, which should have improved CVR.

Further analysis revealed poor **CVR calibration with respect to VCD** (Figure 3a). Higher VCD values led to inflated *pPurchase* scores, causing high-priced SKUs to be ranked higher despite lower actual conversion rates. This misalignment explained the metric declines.

As a short-term fix, we applied a simple transformation to the *pPurchase* scores:

$$pPurchase' = c \cdot \frac{pPurchase}{(VCD + a)^b}$$

where a, b, c are tunable parameters. This adjustment led to an improved calibration (Figure 3b) and yielded better results in the simulations. We are currently in the process of re-running the online experiment in UK.



(a) Before Transformation **(b) After Transformation**

Figure 3: Calibration of pPurchase w.r.t VCD. X-axis is VCD bins and Y-axis is CVR.

6 Conclusion and Future Work

We introduced a new ad ranking function designed to maximize total profit while maintaining a balance with relevance. To support this, we proposed an elegant method for abstracting the relevance constraint using a more interpretable and robust parameter α_0. Through extensive simulations, we demonstrated how key performance metrics respond to variations in α_0 providing actionable guidance for selecting effective values in online experimentation. The approach delivered a 14% increase in total profit and a 22% improvement in CVR compared to the business-as-usual (BAU) baseline.

Moreover, insights from an expansion experiment highlighted the necessity for accurately calibrated pPurchase scores relative to VCD, underscoring calibration as a critical factor for the effectiveness of our proposed approach.

Future research directions include applying relevance constraints at finer granularities, such as individual queries or product classes, recognizing that different query intents and item categories may exhibit distinct user behaviors and varying trade-offs between relevance and profitability.

References

[1] Heng-Tze Cheng, Levent Koc, Jeremiah Harmsen, Tal Shaked, Tushar Chandra, Hrishi Aradhye, Glen Anderson, Greg Corrado, Wei Chai, Mustafa Ispir, Rohan Anil, Zakaria Haque, Lichan Hong, Vihan Jain, Xiaobing Liu, and Hemal Shah. 2016. Wide & Deep Learning for Recommender Systems. In *Proceedings of the 1st Workshop on Deep Learning for Recommender Systems* (Boston, MA, USA) *(DLRS 2016)*. Association for Computing Machinery, New York, NY, USA, 7–10. doi:10.1145/2988450.2988454

[2] Xinran He, Junfeng Pan, Ou Jin, Tianbing Xu, Bo Liu, Tao Xu, Yanxin Shi, Antoine Atallah, Ralf Herbrich, Stuart Bowers, and Joaquin Quiñonero Candela. 2014. Practical Lessons from Predicting Clicks on Ads at Facebook. In *Proceedings of the Eighth International Workshop on Data Mining for Online Advertising* (New York, NY, USA) *(ADKDD'14)*. Association for Computing Machinery, New York, NY, USA, 1–9. doi:10.1145/2648584.2648589

[3] Or Levi. 2017. Online Learning of a Ranking Formula for Revenue and Advertiser ROI Optimization. In *Proceedings of the 1st International Workshop on LEARning Next gEneration Rankers co-located with the 3rd ACM International Conference on the Theory of Information Retrieval (ICTIR 2017), Amsterdam, The Netherlands, October 1, 2017 (CEUR Workshop Proceedings, Vol. 2007)*, Nicola Ferro, Claudio Lucchese, Maria Maistro, and Raffaele Perego (Eds.). CEUR-WS.org. https://ceur-ws.org/Vol-2007/LEARNER2017_full_1.pdf

[4] H. Brendan McMahan, Gary Holt, D. Sculley, Michael Young, Dietmar Ebner, Julian Grady, Lan Nie, Todd Phillips, Eugene Davydov, Daniel Golovin, Sharat Chikkerur, Dan Liu, Martin Wattenberg, Arnar Mar Hrafnkelsson, Tom Boulos, and Jeremy Kubica. 2013. Ad click prediction: a view from the trenches. In *Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (Chicago, Illinois, USA) *(KDD '13)*. Association for Computing Machinery, New York, NY, USA, 1222–1230. doi:10.1145/2487575.2488200

[5] Phuong Ha Nguyen, Djordje Gligorijevic, Arnab Borah, Gajanan Adalinge, and Abraham Bagherjeiran. [n. d.]. Practical Budget Pacing Algorithms and Simulation Test Bed for eBay Marketplace Sponsored Search.

[6] Xuewei Wang, Qiang Jin, Shengyu Huang, Min Zhang, Xi Liu, Zhengli Zhao, Yukun Chen, Zhengyu Zhang, Jiyan Yang, Ellie Wen, Sagar Chordia, Wenlin Chen, and Qin Huang. 2023. Towards the Better Ranking Consistency: A Multi-task Learning Framework for Early Stage Ads Ranking. arXiv:2307.11096 [cs.IR] https://arxiv.org/abs/2307.11096

[7] Liang Wu, Diane Hu, Liangjie Hong, and Huan Liu. 2018. Turning Clicks into Purchases: Revenue Optimization for Product Search in E-Commerce. In *The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval* (Ann Arbor, MI, USA) *(SIGIR '18)*. Association for Computing Machinery, New York, NY, USA, 365–374. doi:10.1145/3209978.3209993

[8] YaChen Yan and Liubo Li. 2023. AdaEnsemble: Learning Adaptively Sparse Structured Ensemble Network for Click-Through Rate Prediction. arXiv:2301.08353 [cs.IR] https://arxiv.org/abs/2301.08353

[9] Yanwu Yang and Panyu Zhai. 2022. Click-through rate prediction in online advertising: A literature review. *Information Processing amp; Management* 59, 2 (March 2022), 102853. doi:10.1016/j.ipm.2021.102853

[10] Jesse Lute Zhen Ge, Wei Zhou and Adam Ilardi. 2021. Relevance Constrained Re-ranking in Sponsored Listing Recommendations. In *Proceedings of AdKDD (ADKDD'14)*. Association for Computing Machinery, New York, NY, USA.

[11] Han Zhu, Junqi Jin, Chang Tan, Fei Pan, Yifan Zeng, Han Li, and Kun Gai. 2017. Optimized Cost per Click in Taobao Display Advertising. *CoRR* abs/1703.02091 (2017). arXiv:1703.02091 http://arxiv.org/abs/1703.02091

[12] Yunzhang Zhu, Gang Wang, Junli Yang, Dakan Wang, Jun Yan, and Zheng Chen. 2009. Revenue optimization with relevance constraint in sponsored search *(ADKDD '09)*. Association for Computing Machinery, New York, NY, USA, 55–60. doi:10.1145/1592748.1592756