UNIVERSITY OF AMSTERDAM

UvA-DARE (Digital Academic Repository)

Understanding Accuracy-Fairness Trade-offs in Re-ranking through Elasticity in Economics

Xu, Chen; Zhao, Jujia; Wang, Wenjie; Pang, Liang; Xu, Jun; Chua, Tat-Seng; de Rijke, Maarten

DOI
10.1145/3726302.3730106

Publication date
2025

Document Version
Final published version

Published in
SIGIR '25

License
CC BY

Link to publication

Citation for published version (APA):
Xu, C., Zhao, J., Wang, W., Pang, L., Xu, J., Chua, T.-S., & de Rijke, M. (2025). Understanding Accuracy-Fairness Trade-offs in Re-ranking through Elasticity in Economics. In *SIGIR '25: Proceedings of the 48th International ACM SIGIR Conference on Research and Development in Information Retrieval : July 13-18, 2025, Padua, Italy* (pp. 539-548). Association for Computing Machinery. https://doi.org/10.1145/3726302.3730106

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, P.O. Box 19185, 1000 GD Amsterdam, The Netherlands. You will be contacted as soon as possible.
UvA-DARE is a service provided by the library of the University of Amsterdam (https://dare.uva.nl)

Download date:13 Oct 2026

Understanding Accuracy-Fairness Trade-offs in Re-ranking through Elasticity in Economics

Chen Xu
Gaoling School of Artificial Intelligence
Renmin University of China
Beijing, China
xc_chen@ruc.edu.cn

Jujia Zhao
Leiden University
Leiden, The Netherlands
j.zhao@liacs.leidenuniv.nl

Wenjie Wang
University of Science and Technology of China
Hefei, China
wenjiewang96@gmail.com

Liang Pang
Institute of Computing Technology
Chinese Academy of Sciences
Beijing, China
pangliang@ict.ac.cn

Jun Xu*
Gaoling School of Artificial Intelligence
Renmin University of China
Beijing, China
junxu@ruc.edu.cn

Tat-Seng Chua
National University of Singapore
Singapore
chuats@comp.nus.edu.sg

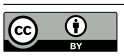
Maarten de Rijke
University of Amsterdam
Amsterdam, The Netherlands
m.derijke@uva.nl

Abstract

Fairness is an increasingly important factor in re-ranking tasks. Prior work has identified a trade-off between ranking accuracy and item fairness. However, the underlying mechanisms are still not fully understood. An analogy can be drawn between re-ranking and the dynamics of economic transactions. The accuracy-fairness trade-off parallels the coupling of the commodity tax transfer process. Fairness considerations in re-ranking, similar to a commodity tax on suppliers, ultimately translate into a cost passed on to consumers. Analogously, item-side fairness constraints result in a decline in user-side accuracy. In economics, the extent to which commodity tax on the supplier (item fairness) transfers to commodity tax on users (accuracy loss) is formalized using the notion of elasticity. The re-ranking fairness-accuracy trade-off is similarly governed by the elasticity of utility between item groups. This insight underscores the limitations of current fair re-ranking evaluations, which often rely solely on a single fairness metric, hindering comprehensive assessment of fair re-ranking algorithms.

Centered around the concept of elasticity, this work presents two significant contributions. We introduce the Elastic Fairness Curve (EF-Curve) as an evaluation framework. This framework enables a comparative analysis of algorithm performance across different elasticity levels, facilitating the selection of the most suitable approach. Furthermore, we propose ElasticRank, a fair re-ranking algorithm that employs elasticity calculations to adjust inter-item distances within a curved space. Experiments on three widely used ranking datasets demonstrate its effectiveness and efficiency.

*Jun Xu is the corresponding author.



This work is licensed under a Creative Commons Attribution 4.0 International License.

SIGIR '25, Padua, Italy

© 2025 Copyright held by the owner/author(s).

ACM ISBN 979-8-4007-1592-1/2025/07
https://doi.org/10.1145/3726302.3730106

CCS Concepts

• **Information systems** → **Information retrieval**.

Keywords

Re-ranking, Fairness, Elasticity in Economics

ACM Reference Format:
Chen Xu, Jujia Zhao, Wenjie Wang, Liang Pang, Jun Xu, Tat-Seng Chua, and Maarten de Rijke. 2025. Understanding Accuracy-Fairness Trade-offs in Re-ranking through Elasticity in Economics. In *Proceedings of the 48th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '25), July 13–18, 2025, Padua, Italy*. ACM, New York, NY, USA, 10 pages. https://doi.org/10.1145/3726302.3730106

1 Introduction

Over the past decade, fairness has become an increasingly important and urgent topic on the information retrieval (IR) agenda [9, 20]. Previous work proposes diverse fairness objectives to ensure a healthy ecosystem from economic or social perspectives [28, 38, 41]. However, fair re-ranking often entails a trade-off between ranking accuracy and item fairness, where improving one typically comes at the expense of the other [9, 20, 41]. Various methods [26, 28, 38] have been proposed to mitigate this trade-off, but the underlying mechanisms of this phenomenon remain insufficiently understood.

An economic perspective on accuracy-fairness trade-off. The accuracy-fairness trade-off mirrors the coupling relationship of the commodity tax transfer process [29]. In economics, promoting fairness is achieved through taxation [27]. When a commodity tax is imposed on a supplier, the tax rate is not entirely absorbed by the supplier but is partially transferred to consumers. When we relate re-ranking to economic transactions, we can view users as suppliers, item groups as consumers, and the ranking scores as the price; a more detailed correspondence can be seen in Table 1. In this way, fairness in re-ranking functions like a commodity tax on the item side [41], where the ranking scores adjusted by the fairness

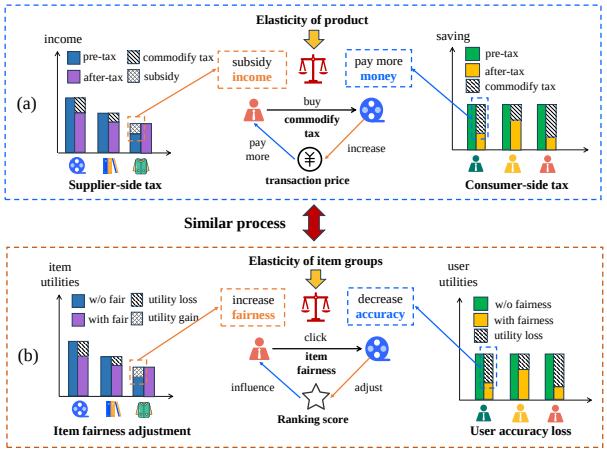


Figure 1: Parallels between (a) the commodity tax transfer process and (b) the accuracy-fairness trade-off in re-ranking.

function (commodity tax for the supplier) transform into the cost of accuracy loss (commodity tax transferred to users). We provide a more detailed theoretical analysis of this analogy in Section 5.

We use the example in Figure 1 to visualize this insight. In the economic transactions shown in Figure 1 (a), the imposition of a commodity tax on the supplier triggers a partial transfer of this tax burden to consumers in the form of a higher commodity price (a detailed example is in Section 3.2). Similarly, in Figure 1 (b), users will be exposed to certain items with high-ranking scores. When promoting fairness among items, poor item groups are assigned with higher ranking scores [41], which alters the utility for some users. Consequently, part of the fairness cost is shifted to the users as an accuracy loss. Given the parallels between the two processes, we hypothesize that tools designed for analyzing taxation can be applied to understand accuracy-fairness trade-offs in re-ranking.

Understanding re-ranking through elasticity. In economics, elasticity theory [31] provides a framework for analyzing the transfer of commodity tax burdens to consumers, with the degree of transfer dependent on the elasticity of goods. Drawing an analogy, we demonstrate in Section 4 that the fairness-accuracy trade-off in re-ranking is similarly governed by the elasticity of utility between item groups. This insight reveals a key limitation of current fair re-ranking evaluations: their reliance on single fairness metrics, which effectively assess tax transfers under only specific elasticity assumptions, thereby hindering a comprehensive assessment of fair re-ranking algorithms.

Grounded in elasticity theory, we introduce the *Elastic Fairness Curve* (EF-Curve), a framework for evaluating fair re-ranking algorithms. The EF-Curve visualizes algorithm performance across a spectrum of elasticities, with each point representing performance under a specific fairness metric. Intuitively, the EF-Curve illustrates that different fairness metrics measure the degree of support for item groups with varying levels of elasticity. This framework facilitates a comparative analysis of algorithms, enabling the selection of the most suitable approach for diverse deployment scenarios. Moreover, the area enclosed by the EF-Curve and the axes, termed EF, provides a quantitative measure of overall algorithm performance.

Furthermore, we introduce *ElasticRank*, a fair re-ranking algorithm designed to optimize the EF metric. Grounded in elasticity theory, ElasticRank models the re-ranking space as curved, where inter-group distances are dynamically adjusted based on their respective elasticities. This approach intuitively prioritizes fairness for item groups with higher elasticity, minimizing the associated accuracy loss. Importantly, ElasticRank achieves this with comparable complexity as standard ranking algorithms.

Main contributions. We summarize the major contributions:

(1) This research frames the re-ranking accuracy-fairness trade-off as a commodity taxation transfer problem. By employing elasticity theory from economics, we elucidate the intricate relationship between these competing objectives.
(2) Inspired by elasticity theory, we propose a novel evaluation framework for fair re-ranking algorithms, the EF-Curve. This framework facilitates comprehensive comparisons between different algorithms by visualizing their performance across a spectrum of fairness constraints.
(3) To optimize EF-Curve, we introduce ElasticRank, a novel fair re-ranking algorithm designed to optimize the EF-Curve. Rigorous empirical evaluation on three publicly available ranking datasets demonstrates that ElasticRank consistently surpasses state-of-the-art baselines.

2 Related Work

Fair re-ranking. Over the past decade, work on fair ranking tasks has rapidly grown in volume, driven by the need for a responsible and trustworthy ecosystem [9, 21, 23, 39]. Previous research often categorizes fair-aware methods into three categories based on ranking phases: pre-processing [7, 37], in-processing [34], and post-processing (*i.e.,* re-ranking tasks) [28, 38]. The re-ranking phase is regarded as the most easily adaptable and practical stage in optimizing ranking systems [28]. During the re-ranking phase, the concept of fairness in re-ranking depends on the stakeholders involved [1, 2]. Prior work has examined user-oriented fairness [3, 19] and item-oriented fairness [15, 28, 33, 38, 41, 43]. In this paper, we focus on item group fairness in re-ranking tasks.

Metrics and algorithms in fair re-ranking. Fairness metrics vary widely across works, with different studies optimizing distinct metrics. For instance, some work [28, 43] employs proportional fairness, Do and Usunier [11] focuses on the Gini Index, other work [38] prioritizes MMF, and TaxRank [41] optimizes α-fairness. However, these approaches rely on single fairness metrics, which limits their ability to provide a comprehensive evaluation.

Previous work on re-ranking methods to improve item fairness can be divided into (i) regularized methods, which use a multi–task optimization approach with a linear combination of accuracy and fairness loss functions, incorporating a trade-off coefficient λ [11, 26, 38], and (ii) constraint-based methods, which formulate the task as a constrained optimization problem to ensure that fairness metrics do not exceed a specified threshold [5, 28, 36, 46]. Despite achieving notable performance improvements, existing fairness intervention methods are often designed to optimize specific fairness metrics and typically involve high computational costs, making them challenging to adapt to real-world industrial systems.

Table 1: Correspondence between taxation elements in economics and fair re-ranking.

Economics	Fair re-ranking
Consumer (buy product)	Users $\mathcal{U}$ (click items)
Supplier (sell product)	Item groups $\mathcal{G}$ (provide items)
Commodity tax	Fairness constraint
Tax subsidies for the poor	Increase ranking score for the poor
Selling price (tax objective)	Ranking scores (fairness objective)
Elasticity on price E_e	Elasticity on utilities of item group $E_{r,p}$

An economic perspective on fair re-ranking. In economics, resource allocation typically occurs through processes of distribution and re-distribution [17]. Previous work [32] regards fair ranking as a resource allocation problem and formulated the problem related to Nash Social Welfare in economics, see also Biswas et al. [5], Patro et al. [28]. TaxRank [41] regards fair re-ranking as a taxation process, which often serves as a key mechanism in the re-distribution process, enabling wealth reallocation and addressing income inequality [13, 27]. However, they merely use economic objectives to define different fairness metrics, without understanding how fairness-accuracy trade-offs occur under different metrics.

3 Problem Formulation

We begin by defining the fair re-ranking task, followed by introducing the concept of elasticity in economics.

3.1 Fair re-ranking

In re-ranking tasks, let $\mathcal{U}$ denote the set of users, $\mathcal{I}$ the set of items, and each item $i \in \mathcal{I}$ belongs to a unique group $g \in \mathcal{G}$. The set of items within a specific group g is represented as $\mathcal{I}_g$. When a user $u \in \mathcal{U}$ accesses the re-ranking system, the system will re-rank items more fairly according to a given candidate ranked list of size K, denoted as $L_K(u) \in \mathcal{I}^K$. In each ranked list $L_K(u)$, we will get the ranking scores $s_{u,i}, i \in L_K(u)$ generated by ranking models. The ranking score can usually be regarded as the probability of a user clicking on an item (*i.e.,* click-through-rate (CTR) value [24, 44]).

Then, we will define the user and item utilities for a certain group in re-ranking tasks. The item group utility v_g and user utility w_u in re-ranking tasks is typically defined as the accumulated utilities of item group g across all ranked lists:

$$v_g = \sum_{u \in \mathcal{U}} \sum_{i \in L_K(u)} s_{u,i} I(i \in \mathcal{I}_g), \quad w_u = \sum_{i \in L_K(u)} s_{u,i}, \tag{1}$$

where $I(\cdot)$ is the indicator function.

The goal of fair re-ranking f is to maximize the overall user utilities ($\sum_{u \in \mathcal{U}} w_u$), while simultaneously striving to equalize utilities across item groups ($v_r \approx v_p, \forall i \in \mathcal{I}_r, \forall j \in \mathcal{I}_p$)

3.2 Elasticity theory

In this section, we will first introduce elasticity theory from economics. Then we relate elasticity theory to re-ranking tasks.

3.2.1 Elasticity in economics. Elasticity is a measure of the responsiveness of one variable to changes in another variable [31]. In economics, the price elasticity of demand is defined as:

$$E_e = \frac{\partial Q/Q}{\partial P/P},$$

where Q is the user demand quantity of an item and P is the item price. Elastic E_e reflects how sensitive consumers are to changes in price or other variables, like the price of related goods. Specifically, Elastic E_e measures the percentage decrease in the quantity Q of items users are willing to purchase when the current price P increases by 1%.

For example, bread has low elasticity since bread is a necessity for most people. If the price of bread increases slightly, consumers are unlikely to stop buying it. This small change in demand despite a large price change illustrates low elasticity. On the other hand, diamonds are not a necessity and there are many alternatives. If the price of a pair increases slightly, consumers might decide not to buy it and instead look for a cheaper option. This large change in demand with a small price change demonstrates high elasticity.

This transfer rate of commodity tax burden depends on the price elasticity of demand E_d in taxation theory [13]. If consumers are less sensitive to price changes, a larger share of the commodity tax burden will fall on consumers because they will continue buying the product even if the price increases due to the tax. On the other hand, if consumers are highly sensitive to price changes, a larger share of the tax burden falls on suppliers because consumers will significantly reduce purchases if prices rise.

Therefore, when imposing commodity tax, it is generally more effective to tax products with high elasticity, as this minimizes the extent to which the tax burden is transferred to users.

3.2.2 Elasticity in re-ranking. When we relate re-ranking to the economic transaction process, we can relate the re-ranking elements to taxation elements in Table 1. Then the utility elasticity of different two groups (group r and group p) of ranking is defined as:

$$E_{r,p} = \frac{\partial v_r}{\partial v_p}. \tag{2}$$

The elasticity term $E_{r,p}$ quantifies the sensitivity of item group r's utility to changes in item group p's utility, capturing the interdependence between the two items in the ranking system. [1]

Intuitively, applying the economics analogy of bread and diamonds to re-ranking, we find a striking similarity: for a relatively rich item group r_1 and a less "rich" item group r_2, they resemble diamonds and bread (with the former having higher elasticity and the latter lower elasticity compared to same group p). Suppose we slightly reduce the exposure of r_1 to subsidize the poor item group p, then p's utility would increase significantly with minimal accuracy loss. In contrast, reducing some exposure from r_2 to subsidize p would lead to a smaller increase in p's utility but would transfer most of the "tax" to users, resulting in more accuracy loss.

Therefore, in fair re-ranking, we should reduce more exposure from item groups with higher elasticity to subsidize the poorer groups. Detailed theoretical analysis can be seen in the next section.

4 Fairness Evaluation in Re-ranking

We first analyze the fairness objectives of fair re-ranking using elasticity theory. Then, we propose a new fairness evaluation metric EF by introducing the EF-Curve.

[1] Note that in ranking systems, we focus more on the absolute value rather than the percentage value; thus, the elastic derivation differs slightly from economic elasticity.

4.1 Analysis of fairness metrics

We aim to demonstrate that different fairness objectives fundamentally alter the elasticity of different item groups, reflecting their varying adaptability to fairness constraints.

Firstly, we define the fairness objective. Let $\boldsymbol{v} = [v_1, v_2, \ldots, v_{|\mathcal{G}|}]$ be the utility vector of different item groups. The fairness objective involves defining a function $f(\boldsymbol{v})$, where the output represents a fairness metric that quantifies the inequity among item group utilities. The function $f(\boldsymbol{v})$ should increase when different utilities become more equal. $f(\boldsymbol{v})$ usually has many different forms in previous work, such as max-min fairness [38], entropy fairness [8], α-fairness [41], proportional fairness [35], p-norm [4] and Renyi Entropy [30]. Previous studies often adopted different fairness optimization objectives and evaluation metrics without fully understanding the distinctions between these metrics.

4.1.1 General form of fairness metric. From the perspective of taxation [41], fairness in re-ranking objectives imposes taxes on richer item groups and redistributes these as subsidies to poorer item groups. Given the requirements of the taxation [27], the fairness function $f(\boldsymbol{v})$ should be: (i) continuous on $\mathbb{R}_+$; (ii) scale-invariant: $f(\boldsymbol{v}) = f(c\boldsymbol{v}), \forall c > 0$; and (iii) independent of number of item groups. Also, given the requirements of ranking systems, the fairness optimization objective should be distributed [40], allowing the data to be partitioned across different servers and aggregated results to ensure scalability and efficiency.

According to these requirements, let $\bar{v}_g$ be the normalized utility: $\bar{v}_g = v_g / \sum_{g=1}^{|\mathcal{G}|} v_g$, then fairness metrics have a general form:

$$f(\boldsymbol{v}; t) = \text{sign}(1-t) \left(\sum_{g=1}^{|\mathcal{G}|} \bar{v}_g^{1-t} \right)^{\frac{1}{t}}, \tag{3}$$

where $\text{sign}(\cdot)$ is the symbolic function and t is the tax base, illustrated in the following theorem:

Theorem 1. *The $f(\boldsymbol{v}; t)$ is the unique form of $f(\boldsymbol{v})$. When t takes on different values, as shown in Figure 2(b), $f(\boldsymbol{v}; t)$ will be generalized to different fairness metrics, especially, $\lim_{t \to 0} f(\boldsymbol{v}; t) = e^{H(\bar{\boldsymbol{v}})}$, where $H(\bar{\boldsymbol{v}})$ is the entropy fairness: $H(\bar{\boldsymbol{v}}) = -\sum_{g=1}^{|\mathcal{G}|} \bar{v}_g \log \bar{v}_g$.*

A detailed proof of Theorem 1 is provided in Appendix A. Eq. (3) presents a general form of fairness metric, where the parameter t can be adapted to represent various fairness metrics, as illustrated in Figure 2(b). The parameter t can be understood as the tax base, which will be further interpreted through the elastic theory described in the next section. Intuitively, as the absolute value $|t|$ increases, the fairness metric places greater emphasis on item groups with lower utility values.

4.1.2 Elasticity theory for analyzing the general form. In this section, we apply the elasticity theory introduced in Section 3.2 to analyze the general form of the fairness metric described in Eq. (3) through the following theorem:

Theorem 2. *The parameter t in Eq. (3) represents the tax base, where, upon adding the next item to the ranked list, the rich group r compared to the poor group p will be subjected to a commodity tax with a rate of:*

$$E_{r,p} = \frac{\partial v_p}{\partial v_r} = \left(\frac{\bar{v}_r}{\bar{v}_p} \right)^{-|t|}, \tag{4}$$

which means the elastic value as defined in Eq. (2). Meanwhile, the rich and poor item group threshold is: $\theta = \left(\frac{\sum_{g=1}^{|\mathcal{G}|} v_g}{\sum_{g=1}^{|\mathcal{G}|} v_g^{1-t}} \right)^{\frac{1}{t}}$, where $v_g > \theta$, the group g is a rich group, otherwise, it will be a poor group.

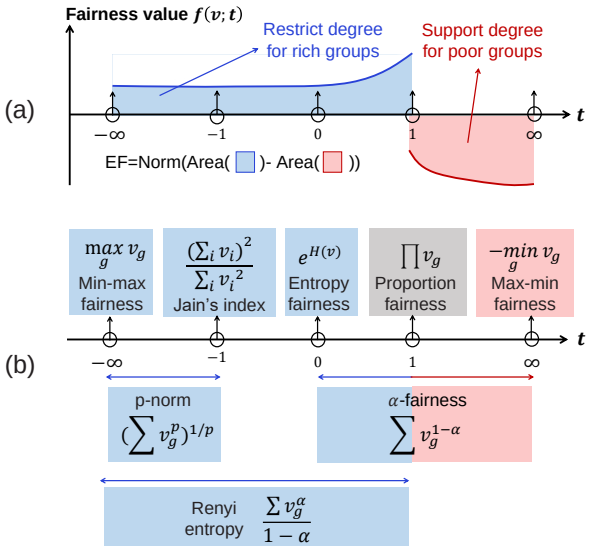


Figure 2: (a) The EF-Curve, where the x-axis is tax base t and the y-axis is the fairness metric $f(v;t)$. EF-Curve describes the restrict/support degree for different item groups. (b) The specific fairness metrics corresponding to particular t values.

A detailed proof can be found in Appendix B. Theorem 2 means that different fairness metrics change the elasticity between the poor item group and the rich item group.

Intuitively, different fairness metrics in re-ranking tasks are determined by adjusting the elasticity between groups. From a taxation perspective, these metrics impose $E_{r,p}$ times more commodity tax to the rich groups r and subsidy to the poor group p.

4.2 EF-Curve and EF metric

After analyzing fairness metrics from an economic perspective, we observe that previous fair re-ranking evaluations rely on single fairness metrics (a single t value), which assesses tax transfers only under specific elasticity conditions.

To provide a comprehensive fair re-ranking evaluation, we use the general form $f(\boldsymbol{v}; t)$ in Eq. (3) to design the EF-curve (shown in Figure 2 (a)). On the EF-curve, each point on the horizontal axis corresponds to a different tax base t, and each point on the vertical axis reflects the fairness metric $f(\boldsymbol{v}; t)$ under the respective metric.

Comparing the EF-Curve of different algorithms reveals their performance across varying elasticities, helping to identify the most suitable algorithm for online deployment based on specific application requirements. Meanwhile, we propose to utilize the area enclosed by the EF-curve (shown in Figure 2 (a)) and the axes

(called EF):

$$\text{EF} = \int_{1-M}^{1+M} \frac{f(\boldsymbol{v}; t)}{Z} dt, \tag{5}$$

where $Z = 2M|\mathcal{G}|$ is the normalized factor for the area computation (*i.e.,* the Norm operation in Figure 2 (a)), $M \geq 0$ is used to approximate infinity integral values. Intuitively, Eq. (5) measures the averaged fairness performances across different fairness metrics.

5 Accuracy-Fairness Optimization

We first analyze the accuracy-fairness ranking trade-off objective from an economic perspective. Then we propose a new fair re-ranking algorithm named ElasticRank.

5.1 Analysis of the accuracy-fairness trade-off

Given the general form of the fairness metric in Theorem 1, this section aims to analyze the objective of maximizing ranking accuracy while balancing the trade-off with the fairness function. We aim to show that the trade-off is just like the commodity tax transfer process in economics, using elasticity theory.

In ranking tasks [26, 38], previous work often adapts a linear trade-off between the fairness and accuracy function:

$$v^* = \arg\max_{\boldsymbol{v}} \sum_{u \in \mathcal{U}} \boldsymbol{w}_u + \lambda f(\boldsymbol{v}), \tag{6}$$

where the $\boldsymbol{w}_u$ is defined as Eq. (1) and $\lambda \in [0, \infty)$ is the trade-off coefficient and the first part is the accuracy part while the second part is the fairness objective. Next, we will use the following theorem to rewrite and analyze the trade-off function.

Theorem 3. *Eq. (6) can equivalently be optimized as:*

$$v^* = \arg\max_{\boldsymbol{v}} L = \arg\max_{\boldsymbol{v}} f(\boldsymbol{v}; |t|)^{|t|} \cdot a(\boldsymbol{w})^{1-|t|}, \tag{7}$$

where $a(\boldsymbol{w}) = \sum_{u \in \mathcal{U}} \boldsymbol{w}_u$ is the accuracy function.

Let $\eta = \mathbf{1} - \bar{\boldsymbol{v}}$ be the gradient direction of nearest fairness (moving to the averaged utility) and $\alpha = \nabla_{\boldsymbol{v}} \boldsymbol{w}$ be the gradient of accuracy function. The transfer ratio between fairness (commodity tax on groups) and accuracy (commodity tax transferred to users) is:

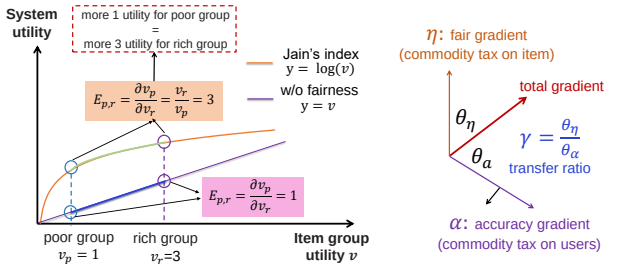
$$\gamma = \frac{\langle \nabla_{\boldsymbol{v}} L, \eta \rangle}{\langle \nabla_{\boldsymbol{v}} L, \alpha \rangle} = 1 - \frac{1}{1 + k(E_{r,p})}, \; k(E_{r,p}) = \frac{\sum_{p \in \mathcal{G}} \sum_{r \neq p} \boldsymbol{v}_p^{1-|t|} E_{r,p}}{\sum_{p \in \mathcal{G}} \boldsymbol{v}_p^{1-|t|}}. \tag{8}$$

The transfer ratio γ is determined by the elasticity between any two item groups, where the ratio can also be interpreted as the extent to which the commodity tax is transferred to the users.

A detailed proof is provided in Appendix C.

In Figure 3 (a), we give a more intuitive example to understand how elasticity works for the commodity tax transfer. For example, Jain's index can naturally be regarded as changing the elasticity $E_{r,p} = 1$ (w/o fairness) of two different item groups as $\left(\frac{\bar{v}_r}{\bar{v}_p}\right)^{-1}$. Suppose there are two item groups, p with a utility of 1 and r with a utility of 3. In that case, Jain's index indicates that adding the next item from group p to the ranked list, compared to group r, will have its utility weighted three times more to support the poorer group (changing elasticity $E_{r,p} = 3$).

From a taxation perspective, as shown in Figure 3 (b), Jain's index will give three times more commodity tax to the rich groups than



(a) Elasticity curve (b) Commodity tax transfer

Figure 3: (a) The Elasticity curve when optimizing Jain's index and objective without fairness constraint. (b) Illustration of how the commodity tax (item fairness) is being transferred to users (accuracy loss).

to the poor group. However, the commodity tax (item fairness) will inevitably be transferred to accuracy loss. Intuitively, when $E_{r,p}$ is small, poor items act like necessities, adding exposure yields limited systemic utility and maximizes accuracy loss (lower commodity tax transfer rate). When $E_{r,p}$ is large, poor items resemble luxuries, adding exposure greatly boosts systemic utility but increases accuracy loss (higher commodity tax transfer rate).

5.2 ElasticRank

Inspired by the aforementioned analysis, we propose an efficient and effective re-ranking algorithm, ElasticRank.

From the proof of Theorem 3, we observe that the fairness function influences the optimal gradient direction, steering it to balance the trade-off between fairness and accuracy. Moreover, we can observe that as the elasticity value increases, the ratio of commodity tax transfer also rises, reflecting a stronger responsiveness of the tax structure to changes in elasticity. From an optimization perspective, the elasticity in economics essentially curves the optimization space, increasing the separation between two item groups, as represented by their geodesic distance [6]. Therefore, we propose ElasticRank, which measures item distances in the elasticity-curved space.

The overall workflow can be seen in Algorithm 1. Formally, the ElasticRank re-ranked list for each user u can be defined as

$$L_K(u) = \arg\max_{i \in I^k} \sum_{i \in I} \left[s_{u,i} + d\left(g(i), a\right) \right], \tag{9}$$

where $g(i)$ is the group of item i and $d(g(i), a)$ represents the additional distance in the curved space between the group $g(i)$ and an anchor group a. The anchor group a can be any group in order to reduce the computational complexity. In our algorithm, a is chosen as the group that has the last $\eta\%$ utility when the user u arrives in the systems: $a = \text{argsort}(\boldsymbol{v})[b]$, $b = \eta\%|\mathcal{G}|$.

The additional distance $d\left(g(i), a\right)$ is computed through the curve distance:

$$d\left(g(i), a\right) = \int_{v_{g(i)}}^{v_a} \sqrt{1 + \left(\frac{\partial \boldsymbol{v}_a}{\partial x}\right)^2} dx \approx (1-t)\boldsymbol{v}_{g(i)}^{-t}(\boldsymbol{v}_a^{1-t} - \boldsymbol{v}_{g(i)}^{1-t}), \tag{10}$$

where the term $\frac{\partial \boldsymbol{v}_a}{\partial x}$ measures the variation in the slope of the curve (i.e., the elasticity $E_{a,g(i)}$ of group a with respect to group $g(i)$ moves along the curve). Intuitively, if the utility of group $g(i)$

Algorithm 1: Learning algorithm of ElasticRank

Input: User set $\mathcal{U}$, item set $\mathcal{I}$, group set $\mathcal{G}$, ranking size K, tax rate t, anchor group index $b = \eta\%|\mathcal{G}|$, user-item ranking score $s_{u,i}, \forall u \in \mathcal{U}, \forall i \in \mathcal{I}$

Set $v_g = 1, \forall g \in \mathcal{G}$
for $u \in \mathcal{U}$ **do**
  Choose anchor group $a = \text{argsort}(v)[b]$
  $d\left(g(i), a\right) = (1-t)v_{g(i)}^{-t}\left(v_a^{1-t} - v_{g(i)}^{1-t}\right)$
  $L_K(u) = \arg\max_{i \in \mathcal{I}^k} \sum_{i \in \mathcal{I}} \left[s_{u,i} + d\left(g(i), a\right)\right]$
  $v_g = v_g + \sum_{i \in \mathcal{I}_g} s_{u,i}, \forall g \in \mathcal{G}$
end for

has a huge gap with the anchor group a, then the distance $d\left(g(i), a\right)$ will be larger to close the utility of such two groups.

Intuitively, taxing higher elasticity item groups minimizes accuracy loss. ElasticRank leverages elasticity calculations to adjust item distances in a curved space, boosting fairness scores for high-elasticity groups and reducing accuracy loss. Meanwhile, as shown in Algorithm 1, ElasticRank does not introduce additional computational overhead, ensuring that its complexity aligns with that of standard ranking algorithms.

6 Experiments

We evaluate ElasticRank using three publicly available ranking datasets, and the source code is shared at GitHub https://github.com/XuChen0427/ElasticRank.

6.1 Experimental settings

Dataset. Our experiments are based on three large-scale, publicly available ranking applications, including:

• **Steam** [16]: a ranking dataset for games on the Steam platform. We use the data for games played for more than 10 hours in our experiments. The publishers of games are considered item groups. It has 169,030 samples, which contains 4,446 users, 1,238 items, and 43 item groups.[2]

• **Amazon-Digital-Music** [14]: a subset (digital music domains) of Amazon Product dataset. After the pre-processing steps, it has 11,320 samples, with 3,175 users, 3,766 items, and 26 item groups.[3]

• **Yelp**: a large-scale businesses recommendation dataset. The categories of items are considered as item groups. After the pre-processing steps, it has 702,457 samples, which contains 8,198 users, 6,429 items, and 64 item groups.[4]

During the pre-processing step, users and items that have interactions with fewer than L items or users are excluded from the entire dataset to mitigate the issue of extreme sparsity. For Yelp, we set $L = 10$ and for the other two datasets, we set $L = 5$. Following Xu et al. [40], we consider groups with fewer than 10 items as a single group, which we name the "infrequent group".

Following [26, 38], we sort all interactions by time and use the first 80% of the interactions as data to train the base ranking model (i.e., MF ranking model [42]). The remaining 20% of interactions are used as the test data for re-ranking tasks.

Evaluation. The performance of the models is evaluated from two aspects: re-ranking accuracy and fairness degree. For the accuracy, following [36, 38], we use NDCG@K and Loss@K:

$$\text{NDCG@K} = \frac{1}{|\mathcal{U}|}\sum_{u \in \mathcal{U}} \frac{\sum_{i \in \mathbf{L}_K^F(u)} s_{u,i}/\log(\text{rank}_i^F + 1)}{\sum_{i \in \mathbf{L}_K(u)} s_{u,i}/\log(\text{rank}_i + 1)}, \tag{11}$$

$$\text{Loss@K} = \frac{\sum_{u \in \mathcal{U}}\sum_{i \in \mathbf{L}_K(u)} s_{u,i} - \sum_{u \in \mathcal{U}}\sum_{i \in \mathbf{L}_K^F(u)} s_{u,i}}{|\mathcal{U}|K}, \tag{12}$$

where $\mathbf{L}_K(u_t)$ is the original ranked list and $\mathbf{L}_K^F(u_t)$ is the fair-aware re-ranked list, and rank_i and rank_i^F are the ranking positions of the item i in $\mathbf{L}_K(u_t)$ and $\mathbf{L}_K^F(u_t)$, respectively. The improved re-ranking accuracy results in a higher NDCG@K value and a lower Loss@K value, indicating better re-ranking quality.

For the fairness degree, we use the EF metric defined in Eq. (5). The improved fairness results in a higher EF@K value.

Baselines. The following representative item fairness re-ranking models were chosen as baselines: **FairRec** [28] and **FairRec+** [5] propose to ensure Max-Min Share of exposure for different items. We also compare **Welf** [10], which uses the Frank-Wolfe algorithm to optimize two-sided fairness. **CPFair** [26] formulates the re-ranking problem as a knapsack problem and solves it greedily. **min-regularizer** [38]: adds a regularizer that penalizes the exposure gaps between the target provider and the groups that have worst-off utilities. **P-MMF** [38] uses the mirror gradient descent method to improve the worst-off item group's utility. **Tax-Rank** [41][5] solves the re-ranking utilizing the optimal transportation (OT) techniques.

Implementation details. Our experiments are implemented using Python 3.9. All experiments are conducted on a server with Ubuntu 18.04. As for the hyper-parameters in all models, the anchor group radio $\eta\%$ is tuned among $[50\%, 95\%]$. The tax base t is tuned among $[1, 2]$. The range of the integral bounds, denoted as M, is set to 50. The ranking base model is the most commonly used MF [42].

6.2 Experimental results

We report on the performance of ElasticRank and other baselines.

6.2.1 Fairness performance comparison. To enable fair comparisons, we conduct experiments to show the performance of ElasticRank and other baselines by tuning the accuracy (NDCG@K) close to 99% to test the accuracy-fairness trade-off performances in re-ranking tasks. Since an accuracy loss within 1% can be considered acceptable for ranking models, as the platform's profit fluctuation remains within a normal range [45]. Testing fairness improvements within this range is more meaningful and instructive for real-world applications. Table 2 presents the experimental outcomes for our ElasticRank model and the baseline methods across all datasets and all ranking sizes $K = 5, 10, 20$.

From Table 2, we first observe that **FairRec** [28], **FairRec+** [5], and **TaxRank** [41], as item-level fairness methods, struggle to effectively balance accuracy and group-level fairness, often resulting in

[2] http://cseweb.ucsd.edu/~wckang/Steam_games.json.gz.
[3] http://jmcauley.ucsd.edu/data/amazon/.
[4] https://www.yelp.com/dataset.

[5] Note that the evaluation of **Tax-Rank** employs a probability sampling method for re-ranking, whereas we only adopt one sample operation for fairly comparison.

Table 2: Performance comparison between ElasticRank and the baselines on Steam, Amazon, and Yelp, where we tuned various re-ranking models to achieve accuracy performances (NDCG) close to 99%, and evaluated fairness across three cut-offs K to assess effectiveness. Bold numbers mean the best fairness performance for the models controlling NDCG close to 99% (except for FairRec, FairRec+, and TaxRank); ↑ and ↓ indicate whether a higher or lower value of the metric is better, respectively. $*$ indicates that the improvements over the baselines are statistically significant (t-tests and p-value < 0.05).

	Top-K	K=5			K=10			K=20		
	Models	Loss@5↓	NDCG@5↑	EF@5↑	Loss@10↓	NDCG@10↑	EF@10↑	Loss@20↓	NDCG@20↑	EF@20↑
Steam	FairRec	0.1724	0.9395	-2.0529	0.0833	0.9670	-1.9659	0.0740	0.9653	-2.2483
	FairRec+	0.1500	0.9470	-2.1579	0.0836	0.9668	-1.9799	0.0785	0.9632	-2.2416
	TaxRank	0.1105	0.9468	-1.2284	0.1183	0.9374	-1.0655	0.1236	0.9249	-0.7544
	Welf	0.0491	0.9760	-1.1725	0.0409	0.9776	-1.0188	0.0283	0.9812	-0.8818
	CPFair	0.0251	0.9891	-1.3717	0.0174	0.9911	-1.5670	0.0109	0.9938	-1.3234
	P-MMF	0.0286	0.9875	-1.0435	0.0222	0.9890	-0.9228	0.0221	0.9879	-0.7744
	min-regularizer	0.0219	0.9903	-1.4547	0.0368	0.9797	-1.3490	0.0374	0.9747	-0.9249
	ElasticRank (Ours)	0.0174	0.9924	**-0.9147***	0.0112	0.9948	**-0.8283***	0.0136	0.9931	**-0.7310***
Amazon	FairRec	0.0000	1.0000	-26.5456	0.0000	1.0000	-12.9110	0.0012	0.9990	-7.7857
	FairRec+	0.0000	1.0000	-26.5456	0.0000	1.0000	-12.9110	0.0013	0.9990	-8.1340
	TaxRank	0.2827	0.7098	-0.8586	0.1526	0.7992	-0.9156	0.0508	0.9150	-1.1968
	Welf	0.0315	0.9734	-3.6578	0.0326	0.9679	-0.9395	0.0096	0.9902	-0.8497
	CPFair	0.0112	0.9910	-2.7632	0.0064	0.9942	-1.1135	0.0068	0.9929	-0.8907
	P-MMF	0.0107	0.9916	-2.4519	0.0092	0.9917	-1.0574	0.0091	0.9907	-0.8741
	min-regularizer	0.0108	0.9906	-2.9959	0.0380	0.9543	-1.0571	0.0147	0.9813	-0.8615
	ElasticRank (Ours)	0.0102	0.9905	**-2.4391***	0.0061	0.9940	**-1.0370***	0.0095	0.9898	**-0.8334***
Yelp	FairRec	0.0413	0.9715	-7.5726	0.0176	0.9883	-7.6010	0.0145	0.9905	-7.5270
	FairRec+	0.0064	0.9962	-14.2024	0.0048	0.9972	-11.4978	0.0047	0.9971	-10.0469
	TaxRank	0.0665	0.9331	-0.5671	0.0363	0.9588	-0.6460	0.0319	0.9657	-0.7348
	Welf	0.0147	0.9870	-0.5316	0.0086	0.9932	-0.4623	0.0123	0.9906	-0.4011
	CPFair	0.0173	0.9885	-0.5281	0.0091	0.9927	-0.4584	0.0128	0.9901	-0.3951
	P-MMF	0.0085	0.9933	-0.5302	0.0098	0.9923	-0.4379	0.0159	0.9876	-0.3350
	min-regularizer	0.0122	0.9888	-0.5544	0.0110	0.9900	-0.4438	0.0158	0.9876	-0.3357
	ElasticRank (Ours)	0.0116	0.9908	**-0.5257***	0.0108	0.9918	**-0.4211***	0.0154	0.9901	**-0.3046***

either insufficient accuracy or inadequate fairness. For the remaining baselines, the experimental results clearly demonstrate that ElasticRank achieves superior fairness performance at the same accuracy level, highlighting the effectiveness of our model.

Next, we test the performance of ElasticRank and the baselines for different accuracy-fairness trade-off degrees and various fairness metrics by conducting experiments on the Steam dataset. Similar trends can be observed in the other two datasets.

6.2.2 Performance on different tax base t. Figure 4 presents the Pareto frontiers [41] for the accuracy metric (NDCG@K) and fairness metric (EF@K) at $K = 10$ and $K = 20$. These Pareto frontiers [25] are derived by systematically adjusting different model parameters and selecting the points that optimize both NDCG@K and EF@K, thus achieving an ideal trade-off between item fairness and total utility. Note that we only compare the best trade-off baselines, excluding FairRec, FairRec+, and TaxRank.

Firstly, it is evident that a trade-off exists between re-ranking accuracy metrics (NDCG@K) and item fairness metrics (EF@K) with respect to the tax base t. When the tax base t is small, FairTax prioritizes ranking accuracy (where item distances are computed in the flat space). Still, when the tax base t increases, ElasticRank emphasizes item fairness by boosting fairness scores for high-elasticity groups while minimizing accuracy loss.

Moreover, compared to the baseline methods, it is evident that the proposed ElasticRank method consistently outperforms them, as indicated by the ElasticRank curves occupying the upper right corner of the Pareto front. This Pareto dominance demonstrates that,

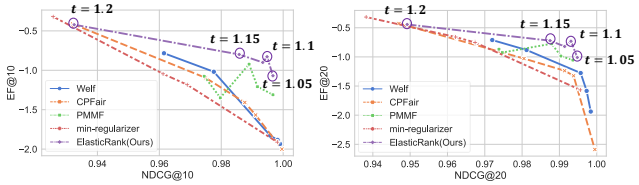


Figure 4: Pareto frontier with different size K under Steam.

for a given NDCG@K level, ElasticRank achieves superior EF@K values, and for a given EF@K level, it delivers better NDCG@K performance. These results highlight the significant advantage of ElasticRank over the baseline methods.

6.2.3 Performances on EF-Curve. Figure 5 presents the EF-Curve described in Section 4 on Steam under $K = 10$. The experiments were also conducted for controlling NDCG close to 99% (see Table 2). Intuitively, each point on the EF-Curve represents a fairness metric that evaluates the level of fairness, with these metrics reflecting the varying degrees of support provided to different item groups, each characterized by different elasticity. When t approaches $-\infty$, the function $f(v;t)$ measures the utility of the richest groups while when t approaches $+\infty$, the function $f(v;t)$ measures the utility of the poorest groups.

Firstly, we observe that different fair re-ranking algorithms enhance fairness from distinct perspectives. For instance, P-MMF often aims to support the poor groups (as indicated by the highest EF-curve of baselines when $t < 0$), while struggling to restrict the utility of the rich groups (as shown by the lowest EF-curve when

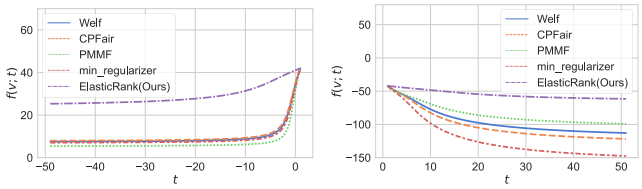


Figure 5: EF-Curve for different models with cut-off size $K = 10$ under Steam.

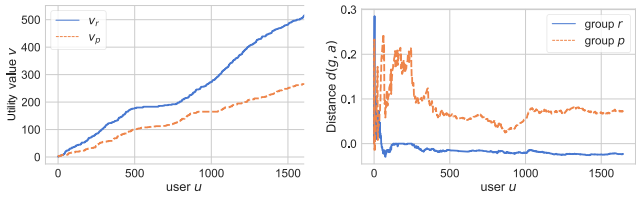
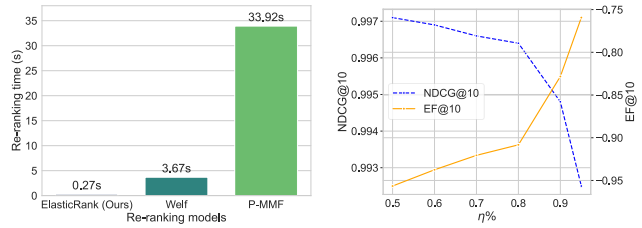


Figure 6: Group utility v_g and distance $d(g, a)$ *w.r.t.* the user u arriving numbers, where g is one rich group r and poor group p; $\eta\%$ is set to 90% and t to 1.05.



(a) Inference speed comparison (b) Ablation study on η

Figure 7: (a) Inference speed comparison between ElasticRank and two best-performing baselines. (b) NDCG@K and EF@K performance *w.r.t.* the anchor group radio $\eta\%$.

$t < 0$). Similarly, we can observe that CPFair is good at restricting the utility of the rich groups but fails to support the poor groups. Through the EF-Curve, we can observe how each algorithm impacts fairness, allowing us to select the most suitable algorithm based on the specific requirements of different applications.

Moreover, compared to the baseline methods, it is clear that ElasticRank consistently outperforms them at every point on the EF-curve. These results demonstrate that ElasticRank not only effectively restricts the utility of the rich groups (as shown by the highest curve when $t > 0$), but also better supports the poor groups (indicated by the highest EF-curve when $t > 0$). These findings highlight the significant advantage of ElasticRank over the baseline methods across all fairness metrics.

6.3 Experimental analysis

We also conduct experiments to analyze ElasticRank on the Steam dataset under $K = 10$. Similar phenomena can be observed in the other two datasets and other K values.

6.3.1 Visualization of distances between different groups. In this section, we investigate how ElasticRank manages the accuracy-fairness trade-off by visualizing the curved distances between different groups. Figure 6 presents the group utility v_g and distance $d(g, a)$ *w.r.t.* the user u arriving numbers, where g is one rich group r and poor group p. The experiments also are conducted on the Steam dataset with $K = 10$. Intuitively, the distance $d(g, a)$ can be interpreted as a fairness score: smaller distances indicate that the system is less likely to provide exposure to the group, whereas larger distances suggest a higher likelihood of exposure for that group. The anchor radio $\eta\%$ is 90%.

From Figure 6, we can first observe that the rich group r exhibits a higher utility than the poor group p, which helps ensure the accuracy performance of the system. However, to preserve fairness, the group r is assigned smaller distances (fairness scores) relative to the anchor group a, indicating that ElasticRank is less likely to expose r during the re-ranking process. Conversely, the group p is assigned larger distances (fairness scores) relative to the anchor group a, highlighting ElasticRank's support for such a group.

6.3.2 Inference speed. In Figure 7 (a), we compare the inference time of the two best-performing baseline models, Welf and P-MMF (see performances in Figure 4), with our model, ElasticRank, by testing the inference time across all users.

From Figure 7 (a), we can observe that ElasticRank is an order of magnitude faster than the Welf model in re-ranking and two orders of magnitude faster than P-MMF. This is because our model uses elasticity to dynamically compute fairness scores, rather than using optimization algorithms to compute gradients. This approach not only yields better results but also provides a significant improvement in efficiency, enabling rapid deployment in any industrial-level IR system. The computational complexity is the same as the standard sorting algorithms.

6.3.3 Ablation study on choosing anchor group. In this section, we conduct an ablation study for the anchor group setting radio $\eta\%$ (Eq. (10) and Figure 7 (b)) illustrates how the ranking accuracy (NDCG) and fairness metric (EF) *w.r.t.* $\eta\%$ from 50%-95%.

From Figure 7 (b), we found that η can trade off accuracy and fairness, as increasing η often improves fairness performance while reducing accuracy. Upon further investigation, we discovered that a larger anchor group ratio increases the gap between poorer groups and the anchor (richer) group, which results in more support for the poorer group to improve fairness. However, this approach tends to harm accuracy. Therefore, in practical applications, it is important to select different anchor group ratios based on specific needs to balance accuracy and fairness.

7 Conclusion and Discussion

Conclusion. In this paper, we understand accuracy-fairness trade-offs in re-ranking by framing them as a commodity taxation transfer problem. By leveraging elasticity theory from economics, we reveal that these trade-offs are determined by the elasticity between inter-groups. Inspired by elasticity theory, we introduce the EF-Curve, a evaluation framework for fair re-ranking, alongside ElasticRank, an algorithm that consistently outperforms state-of-the-art baselines in both efficiency and effectiveness.

Discussion. Although we leverage the elasticity theory to analyze fairness in re-ranking, there are significant differences between them. In economics, the commodity tax transfer problem is inherently more complex because the tax revenue is often allocated to public goods, benefiting all users collectively. Additionally, savings

mechanisms can restrict the impact of commodity taxes on users. In contrast, re-ranking corresponds to a simpler, static setup, making it less intricate than its economic counterpart. In the future, the concept of commodity transfer could be extended to analyze and design more complex fair-aware IR applications.

Acknowledgments

This work was funded by the National Key R&D Program of China (2023YFA1008704), the National Natural Science Foundation of China (No. 62472426, 62276248), the Youth Innovation Promotion Association CAS (No. 2023111). This research was (partially) supported by the Dutch Research Council (NWO), under project numbers 024.004.022, NWA.1389.20.183, and KICH3.LTP.20.006, and the European Union's Horizon Europe program under grant agreement No 101070212. All content represents the opinion of the authors, which is not necessarily shared or endorsed by their respective employers and/or sponsors.

Appendix

A Proof of Theorem 1

Proof. First, we can use the normalized utility $\bar{v}_g = v_g / \sum_{g=1}^{|\mathcal{G}|} v_g$ to make the function $f(\cdot)$ meet the second requirements scale-invariants: $\bar{v}_g = cv_g / \sum_{g=1}^{|\mathcal{G}|} cv_g$.

Then, since $f(\cdot)$ should meet the third requirement, we can write: $f(\mathbf{1}_n/n) = f(\mathbf{1}_m/m), \forall n, m > 0$ where the $\mathbf{1}_n$ means the n-th dimensional vector with all 1 elements. Therefore, we can write

$$\lim_{n\to\infty} \frac{f(\mathbf{1}_{n+1}/(n+1))}{f(\mathbf{1}_n/n)} = 1. \tag{13}$$

Then, to meet the requirements for fairness and optimization objectives that should be distributed, we can make any data $\boldsymbol{x} \in \mathbb{R}^m$ partition into $[x_1, x_2, \cdots, x_n]$, where n is the partition number. Then $f(x)$ can be aggregated as: $f(x) = h(f(x_1), f(x_2), \cdots, f(x_n))$, where h is a Kolmogorov-Nagumo mean function [12] and combining the requirements and the Equation (13) into the Theorem 2 of [18], we can write h as $h = k^{-1}(\sum_{g=1}^{|\mathcal{G}|} k(\bar{v}_g))$, where the function $k(\cdot)$ should only be the power family logarithm or power generator function: $k(x) = \log x$ or $k(x) = x^t$.

Therefore, the function is generated by the power function:

$$f(\boldsymbol{v}; t) = \text{sign}(1-t) \left[\sum_{g=1}^{|\mathcal{G}|} \bar{v}_g^{1-t} \right]^{\frac{1}{t}}.$$

When $t \to 0$, the function f becomes logarithm:

$$\lim_{t\to 0} \log f(\boldsymbol{v}; t) = \lim_{t\to 0} \log \left[\sum_{g=1}^{|\mathcal{G}|} \bar{v}_g^{1-t} \right]^{\frac{1}{t}} = \lim_{t\to 0} \frac{\left[\sum_{g=1}^{|\mathcal{G}|} \bar{v}_g^{1-t} \right]}{t}$$
$$= \lim_{t\to 0} -\sum_{g=1}^{|\mathcal{G}|} (\log \bar{v}_g) \bar{v}_g^{1-t} = -\sum_{g=1}^{|\mathcal{G}|} \bar{v}_g \log \bar{v}_g = H(\bar{v}),$$

which will be reduced to the entropy fairness $H(\boldsymbol{v})$. Therefore, we have $\lim_{t\to 0} f(\boldsymbol{v}; t) = e^{H(\bar{v})}$. When $t \to \infty$, the fairness formulation simplifies to the infinity norm, effectively reducing to max-min fairness, while other types of fairness can be easily mapped to their corresponding relationships. □

B Proof of Theorem 2

Proof. According to elasticity definition in Equation (2):

$$\frac{\partial f(\boldsymbol{v}; t)}{\partial \bar{v}_g} = \frac{|1-t|}{t} \bar{v}_g^{-t} \left[\sum_{g=1}^{|\mathcal{G}|} \bar{v}_g^{1-t} \right]^{\frac{1}{t}-1}. \tag{14}$$

Then, the elasticity can be measured through:

$$E_{r,p} = \frac{\partial v_p}{\partial v_r} = \frac{\frac{\partial f(\boldsymbol{v};t)}{\partial v_r}}{\frac{\partial f(\boldsymbol{v};t)}{\partial v_p}} = \left(\frac{\bar{v}_r}{\bar{v}_p} \right)^{-|t|},$$

where the absolute value of $1-t$ means the fairness metric typically taxes the rich group and subsidizes the poor group to ensure proper redistribution.

$$\frac{\partial f(\boldsymbol{v}; t)}{\partial t} \text{sign}(1-t) \geq 0. \tag{15}$$

Since let $t_1 > t_2 > 1$, $\phi(y) = y^{\frac{t_2}{t_1}}$ is concave, we have:

$$f(\boldsymbol{v}; t_2) = \geq -\left[\sum_{g=1}^{|\mathcal{G}|} \phi(\bar{v}_g^{-t_1}) \right]^{\frac{1}{t_2}} \geq -\left[\phi\left(\sum_{g=1}^{|\mathcal{G}|} \bar{v}_g^{-t_1} \right) \right]^{\frac{1}{t_2}} \geq f(\boldsymbol{v}; t_1),$$

where the third step follows from Jensen's inequality.

Next, we will prove how to distinguish the poor and rich item groups. we check the $\frac{\partial f(\boldsymbol{v};t)}{\boldsymbol{v}_g} = 0$ has a single root of $\boldsymbol{v}_g^* = \left(\frac{\sum_{g=1}^{|\mathcal{G}|} v_g}{\sum_{g=1}^{|\mathcal{G}|} v_g^{1-t}} \right)^{\frac{1}{t}}$, where for $t \neq 1$, $\frac{\partial f(\boldsymbol{v};t)}{\partial \boldsymbol{v}_g} > 0$, if $\boldsymbol{v}_g > \theta$, otherwise, $\frac{\partial f(\boldsymbol{v};t)}{\partial \boldsymbol{v}_g} < 0$. □

C Proof of Theorem 3

Proof. First, we re-write the Equation (6) as $L_1 = \sum_{g\in\mathcal{G}} \boldsymbol{v}_g + \lambda f(\boldsymbol{v})$. Then we can write:

$$L = \left[f(\boldsymbol{v}; |t|)^{|t|} \cdot a(\boldsymbol{w})^{1-|t|} \right] \propto \sum_{g\in\mathcal{G}} \boldsymbol{v}_g^{1-|t|}.$$

Let $l(\boldsymbol{v}; r) = \sum_{g\in\mathcal{G}} \boldsymbol{v}_g^{1-r}$, $r \geq 0$, then $L_1 = l(\boldsymbol{v}; 0) + l(\boldsymbol{v}; |t|)$. Since $l(\boldsymbol{v}; r)$ is continuous *w.r.t.* r and the feasible region of $\boldsymbol{v}$ is convex and continuous (because $\boldsymbol{v}$ is the linear transformation over a simplex space [22]), there exists a constant number $\lambda \geq 0$ s.t. $\sum_{g\in\mathcal{G}} \boldsymbol{v}_g + \lambda f(\boldsymbol{v}; |t|) = L_1$. Therefore, the $\arg\max_{\boldsymbol{v}} L_1 = \arg\max_{\boldsymbol{v}} L$.

Then, we define the accuracy gradient $\alpha = \mathbf{1}$ and fairness gradient $\eta = \mathbf{1}/|\mathcal{G}| - \bar{v}$. Then we can write:

$$\langle \nabla_{\boldsymbol{v}} L, \alpha \rangle = \sum_{g\in\mathcal{G}} \boldsymbol{v}_g^{-|t|}, \langle \nabla_{\boldsymbol{v}} L, \eta \rangle = \sum_{g\in\mathcal{G}} \left(\boldsymbol{v}_g^{-|t|} (\mathbf{1} - \frac{\boldsymbol{v}_g}{\sum_g \boldsymbol{v}_g}) \right).$$

Therefore,

$$\gamma = \frac{\langle \nabla_{\boldsymbol{v}} L, \eta \rangle}{\langle \nabla_{\boldsymbol{v}} L, \alpha \rangle} = 1 - \frac{\sum_g \bar{v} \boldsymbol{v}_g^{-|t|}}{\sum_g \boldsymbol{v}_g^{-|t|}}$$
$$= 1 - \frac{\sum_g \boldsymbol{v}_g^{1-|t|}}{\sum_g \boldsymbol{v}_g \sum_g \boldsymbol{v}_g^{|t|}} = 1 - \frac{\sum_g \boldsymbol{v}_g^{1-|t|}}{\sum_g \boldsymbol{v}_g^{1-|t|} + \sum_g \sum_{r\neq g} \boldsymbol{v}_g^{-|t|} \boldsymbol{v}_r}$$
$$= 1 - \frac{1}{1 + \frac{\sum_{p\in\mathcal{G}} \sum_{r\neq p} \boldsymbol{v}_p^{1-|t|} E_{r,p}}{\sum_{p\in\mathcal{G}} \boldsymbol{v}_p^{1-|t|}}} = 1 - \frac{1}{1 + k(E_{r,p})}.$$ □

References

[1] Himan Abdollahpouri, Gediminas Adomavicius, Robin Burke, Ido Guy, Dietmar Jannach, Toshihiro Kamishima, Jan Krasnodebski, and Luiz Pizzato. 2020. Multi-stakeholder Recommendation: Survey and Research Directions. *User Modeling and User-Adapted Interaction* 30, 1 (2020), 127–158.

[2] Himan Abdollahpouri and Robin Burke. 2019. Multi-stakeholder Recommendation and its Connection to Multi-sided Fairness. *arXiv preprint arXiv:1907.13158* (2019).

[3] Himan Abdollahpouri, Masoud Mansoury, Robin Burke, and Bamshad Mobasher. 2019. The Unfairness of Popularity Bias in Recommendation. *arXiv preprint arXiv:1907.13286* (2019).

[4] Tolga Bektaş and Adam N Letchford. 2020. Using ℓp-Norms for Fairness in Combinatorial Optimisation. *Computers & Operations Research* 120 (2020), 104975.

[5] Arpita Biswas, Gourab K. Patro, Niloy Ganguly, Krishna P Gummadi, and Abhijnan Chakraborty. 2021. Toward Fair Recommendation in Two-sided Platforms. *ACM Transactions on the Web (TWEB)* 16, 2 (2021), 1–34.

[6] Jérémie Bouttier, Philippe Di Francesco, and Emmanuel Guitter. 2003. Geodesic Distance in Planar Graphs. *Nuclear physics B* 663, 3 (2003), 535–567.

[7] Flavio P. Calmon, Dennis Wei, Bhanukiran Vinzamuri, Karthikeyan Natesan Ramamurthy, and Kush R. Varshney. 2017. Optimized Pre-processing for Discrimination Prevention. In *Proceedings of the 31st International Conference on Neural Information Processing Systems* (Long Beach, California, USA) *(NIPS'17)*. Curran Associates Inc., Red Hook, NY, USA, 3995–4004.

[8] Yashar Deldjoo, Vito Walter Anelli, Hamed Zamani, Alejandro Bellogín, and Tommaso Di Noia. 2019. Recommender Systems Fairness Evaluation via Generalized Cross Entropy. *arXiv preprint arXiv:1908.06708* (2019).

[9] Yashar Deldjoo, Dietmar Jannach, Alejandro Bellogin, Alessandro Difonzo, and Dario Zanzonelli. 2022. A Survey of Research on Fair Recommender Systems. *arXiv preprint arXiv:2205.11127* (2022).

[10] Virginie Do, Sam Corbett-Davies, Jamal Atif, and Nicolas Usunier. 2021. Two-sided Fairness in Rankings via Lorenz Dominance. *Advances in Neural Information Processing Systems* 34 (2021), 8596–8608.

[11] Virginie Do and Nicolas Usunier. 2022. Optimizing Generalized Gini Indices for Fairness in Rankings. In *Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval*. 737–747.

[12] Ambedkar Dukkipati. 2010. On Kolmogorov-Nagumo Averages and Nonextensive Entropy. In *2010 International Symposium On Information Theory & Its Applications*. IEEE, 446–451.

[13] Michelle Hanlon and Shane Heitzman. 2010. A Review of Tax Research. *Journal of accounting and Economics* 50, 2-3 (2010), 127–178.

[14] Ruining He and Julian McAuley. 2016. Ups and downs: Modeling the visual evolution of fashion trends with one-class collaborative filtering. In *proceedings of the 25th international conference on world wide web*. 507–517.

[15] Thomas Jaenich, Graham McDonald, and Iadh Ounis. 2024. Fairness-Aware Exposure Allocation via Adaptive Reranking. In *Proceedings of the 47th International ACM SIGIR Conference on Research and Development in Information Retrieval*. 1504–1513.

[16] Wang-Cheng Kang and Julian McAuley. 2018. Self-attentive Sequential Recommendation. In *2018 IEEE international conference on data mining (ICDM)*. IEEE, 197–206.

[17] Peter J. Lambert. 1992. *The Distribution and Redistribution of Income*. Springer.

[18] Tian Lan, David Kao, Mung Chiang, and Ashutosh Sabharwal. 2010. An Axiomatic Theory of Fairness in Network Resource Allocation. In *Proceedings IEEE INFOCOM*. 1–9.

[19] Yunqi Li, Hanxiong Chen, Zuohui Fu, Yingqiang Ge, and Yongfeng Zhang. 2021. User-oriented Fairness in Recommendation. In *Proceedings of the Web Conference*. 624–632.

[20] Yunqi Li, Hanxiong Chen, Shuyuan Xu, Yingqiang Ge, Juntao Tan, Shuchang Liu, and Yongfeng Zhang. 2022. Fairness in Recommendation: A Survey. *arXiv preprint arXiv:2205.13619* (2022).

[21] Yunqi Li, Hanxiong Chen, Shuyuan Xu, Yingqiang Ge, Juntao Tan, Shuchang Liu, and Yongfeng Zhang. 2023. Fairness in Recommendation: Foundations, Methods and Applications. *ACM Transactions on Intelligent Systems and Technology* 14, 5 (2023), Article No.: 95.

[22] Joram Lindenstrauss, Gunnar Olsen, and Yaki Sternfeld. 1978. The Poulsen Simplex. In *Annales de l'Institut Fourier*, Vol. 28. 91–114.

[23] Aldo Lipani. 2016. Fairness in Information Retrieval. In *Proceedings of the 39th International ACM SIGIR conference on Research and Development in Information Retrieval*. 1171–1171.

[24] Xiangyu Liu, Chuan Yu, Zhilin Zhang, Zhenzhe Zheng, Yu Rong, Hongtao Lv, Da Huo, Yiqing Wang, Dagui Chen, Jian Xu, et al. 2021. Neural Auction: End-to-end Learning of Auction Mechanisms for E-commerce Advertising. In *Proceedings of the 27th ACM Conference on Knowledge Discovery & Data Mining*. 3354–3364.

[25] Alexander V Lotov and Kaisa Miettinen. 2008. Visualizing the Pareto frontier. In *Multiobjective optimization*. Springer, 213–243.

[26] Mohammadmehdi Naghiaei, Hossein A Rahmani, and Yashar Deldjoo. 2022. CP-Fair: Personalized Consumer and Producer Fairness Re-ranking for Recommender Systems. *arXiv preprint arXiv:2204.08085* (2022).

[27] Birger Nerré. 2001. The Concept of Tax Culture. In *Proceedings Annual Conference on Taxation and Minutes of the Annual Meeting of the National Tax Association*, Vol. 94. JSTOR, 288–295.

[28] Gourab K. Patro, Arpita Biswas, Niloy Ganguly, Krishna P. Gummadi, and Abhijnan Chakraborty. 2020. FairRec: Two-sided Fairness for Personalized Recommendations in Two-sided Platforms. In *Proceedings of the Web Conference*. 1194–1204.

[29] Frank P. Ramsey. 1927. A Contribution to the Theory of Taxation. *The Economic Journal* 37, 145 (1927), 47–61.

[30] Renato Renner and Stefan Wolf. 2004. Smooth Rényi Entropy and Applications. In *Proceedings International Symposium on Information Theory. ISIT 2004*. IEEE, 233.

[31] Lionel Robbins. 1997. On the Elasticity of Demand for Income in Terms of Effort. In *Economic Science and Political Economy: Selected Articles*. Springer, 79–84.

[32] Yuta Saito and Thorsten Joachims. 2022. Fair Ranking as Fair Division: Impact-Based Individual Fairness in Ranking. In *Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*. 1514–1524.

[33] Ashudeep Singh and Thorsten Joachims. 2019. Policy Learning for Fairness in Ranking. *Advances in Neural Information Processing Systems* 32 (2019).

[34] Jiakai Tang, Shiqi Shen, Zhipeng Wang, Zhi Gong, Jingsen Zhang, and Xu Chen. 2023. When Fairness meets Bias: a Debiased Framework for Fairness aware Top-N Recommendation. In *Proceedings of the 17th ACM Conference on Recommender Systems* (Singapore, Singapore) *(RecSys '23)*. Association for Computing Machinery, New York, NY, USA, 200–210.

[35] Dong Wei, Md Mouinul Islam, Baruch Schieber, and Senjuti Basu Roy. 2022. Rank Aggregation with Proportionate Fairness. In *Proceedings of the 2022 international conference on management of data*. 262–275.

[36] Yao Wu, Jian Cao, Guandong Xu, and Yudong Tan. 2021. TFROM: A Two-sided Fairness-aware Recommendation Model for both Customers and Providers. In *Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval*. 1013–1022.

[37] Zikai Xiong, Niccolò Dalmasso, Alan Mishler, Vamsi K Potluru, Tucker Balch, and Manuela Veloso. 2024. Fairwasp: Fast and optimal fair wasserstein preprocessing. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 38. 16120–16128.

[38] Chen Xu, Sirui Chen, Jun Xu, Weiran Shen, Xiao Zhang, Gang Wang, and Zhenhua Dong. 2023. P-MMF: Provider Max-min Fairness Re-ranking in Recommender System. In *Proceedings of the ACM Web Conference 2023*. 3701–3711.

[39] Chen Xu, Zhirui Deng, Clara Rus, Xiaopeng Ye, Yuanna Liu, Jun Xu, Zhicheng Dou, Ji-Rong Wen, and Maarten de Rijke. 2025. FairDiverse: A Comprehensive Toolkit for Fair and Diverse Information Retrieval Algorithms. arXiv:2502.11883 [cs.IR]

[40] Chen Xu, Jun Xu, Yiming Ding, Xiao Zhang, and Qi Qi. 2024. FairSync: Ensuring Amortized Group Exposure in Distributed Recommendation Retrieval. In *Proceedings of the ACM Web Conference 2024* (Singapore, Singapore) *(WWW '24)*. Association for Computing Machinery, New York, NY, USA, 1092–1102.

[41] Chen Xu, Xiaopeng Ye, Wenjie Wang, Liang Pang, Jun Xu, and Tat-Seng Chua. 2024. A Taxation Perspective for Fair Re-ranking. In *Proceedings of the 47th International ACM SIGIR Conference on Research and Development in Information Retrieval* (Washington DC, USA) *(SIGIR '24)*. Association for Computing Machinery, New York, NY, USA, 1494–1503.

[42] Hong-Jian Xue, Xinyu Dai, Jianbing Zhang, Shujian Huang, and Jiajun Chen. 2017. Deep Matrix Factorization Models for Recommender Systems.. In *IJCAI*, Vol. 17. Melbourne, Australia, 3203–3209.

[43] Tao Yang, Zhichao Xu, and Qingyao Ai. 2023. Vertical Allocation-based Fair Exposure Amortizing in Ranking. In *Proceedings of the Annual International ACM SIGIR Conference on Research and Development in Information Retrieval in the Asia Pacific Region* (Beijing, China) *(SIGIR-AP '23)*. Association for Computing Machinery, New York, NY, USA, 234–244.

[44] Xun Yang, Yasong Li, Hao Wang, Di Wu, Qing Tan, Jian Xu, and Kun Gai. 2019. Bid Optimization by Multivariable Control in Display Advertising. In *Proceedings of the 25th ACM international conference on knowledge discovery & data mining*. 1966–1974.

[45] Xiaopeng Ye, Chen Xu, Jun Xu, Xuyang Xie, Gang Wang, and Zhenhua Dong. 2024. Guaranteeing Accuracy and Fairness under Fluctuating User Traffic: A Bankruptcy-Inspired Re-ranking Approach *(CIKM '24)*. Association for Computing Machinery, New York, NY, USA, 2991–3001.

[46] Muhammad Bilal Zafar, Isabel Valera, Manuel Gomez-Rodriguez, and Krishna P Gummadi. 2019. Fairness Constraints: A Flexible Approach for Fair Classification. *The Journal of Machine Learning Research* 20, 1 (2019), 2737–2778.