Fair Allocation Over Time, with Applications to Content Moderation

Amine Allouah
amine@allouah.com
Meta
Menlo Park, CA, USA

Christian Kroer
christian.kroer@columbia.edu
Columbia University
New York, NY, USA

Xuan Zhang
xz2569@columbia.edu
Meta
Menlo Park, CA, USA

Vashist Avadhanula
vavadhanula18@gsb.columbia.edu
Meta
Menlo Park, CA, USA

Nona Bohanon
cbohanon@meta.com
Meta
Menlo Park, CA, USA

Anil Dania
anildania@meta.com
Meta
Menlo Park, CA, USA

Caner Gocmen
caner@meta.com
Meta
Menlo Park, CA, USA

Sergey Pupyrev
spupyrev@meta.com
Meta
Menlo Park, CA, USA

Parikshit Shah
parikshit@meta.com
Meta
Menlo Park, CA, USA

Nicolas Stier-Moses
nstier@meta.com
Meta
Menlo Park, CA, USA

Ken Rodríguez Taarup
kentaarup@meta.com
Meta
Menlo Park, CA, USA

ABSTRACT

In today's digital world, interaction with online platforms is ubiquitous, and thus content moderation is important for protecting users from content that do not comply with pre-established community guidelines. Given the vast volume of content generated online daily, having an efficient content moderation system throughout every stage of planning is particularly important. We study the short-term planning problem of allocating human content reviewers to different harmful content categories. We use tools from fair division and study the application of competitive equilibrium and leximin allocation rules for addressing this problem. On top of the traditional Fisher market setup, we additionally incorporate novel aspects that are of practical importance. The first aspect is the forecasted workload of different content categories, which puts constraints on the allocation chosen by the planner. We show how a formulation that is inspired by the celebrated Eisenberg-Gale program allows us to find an allocation that not only satisfies the forecasted workload, but also fairly allocates the remaining working hours from the content reviewers among all content categories. A fair allocation of oversupply provides a guardrail in cases where the actual workload deviates from the predicted workload. The second practical consideration is time dependent allocation that is motivated by the fact that partners need scheduling guidance for the reviewers across days to achieve efficiency. To address the time component, we introduce new extensions of the various fair allocation approaches for the single-time period setting, and we show that many properties extend in essence, albeit with some modifications. Lastly, related to the time component, we additionally investigate how to satisfy markets' desire for smooth allocation (i.e, an allocation that does not vary much from time to time) so that the switch in staffing is minimized. We demonstrate the performance of our proposed approaches through real-world data obtained from Meta.

CCS CONCEPTS

• **Theory of computation → Market equilibria**.

KEYWORDS

Market Equilibrium, Fair Allocation, Content Moderation

ACM Reference Format:
Amine Allouah, Christian Kroer, Xuan Zhang, Vashist Avadhanula, Nona Bohanon, Anil Dania, Caner Gocmen, Sergey Pupyrev, Parikshit Shah, Nicolas Stier-Moses, and Ken Rodríguez Taarup. 2023. Fair Allocation Over Time, with Applications to Content Moderation. In *Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD '23), August 6–10, 2023, Long Beach, CA, USA.* ACM, New York, NY, USA, 12 pages. https://doi.org/10.1145/3580305.3599340

KDD '23, August 6–10, 2023, Long Beach, CA, USA

ACM ISBN 979-8-4007-0103-0/23/08…$15.00
https://doi.org/10.1145/3580305.3599340

1 INTRODUCTION

Content moderation is an important challenge faced by many online platforms, and it typically involves the analysis of content generated on the platform to verify whether it is compliant with the platform's content policies. Examples of problematic content categories include fake profiles, spam, hate speech, violent content, and harassment [8]. A key issue in this domain is the enormous amount of content that needs to be verified. For example, Meta

reported in Q3 of 2022: "the prevalence of bullying and harassment-related content [was] at 0.08% on Facebook" [25]; and Meta "review millions of pieces of content across the world every day" to reduce harmful content on its apps [19]. While artificial intelligence technology is necessarily used to handle this volume of content, it is not, at least currently, accurate enough to handle content moderation without human help. For this reason, Meta developed partnerships with highly reputable global partners who employs hundreds of thousands of reviewers to help analyze ambiguous content that requires human verification or investigation [22]. Across the internet industry, social media companies, such as Twitter, YouTube, and Facebook, spend billions of dollars every year on content moderation [26]. In this paper, we study the short-term planning problem of allocating reviewing capacity to different harmful content categories, while taking into account uncertainty about content volume as well as the skills (e.g., language requirement, training on community standards) of various partners supplying reviewers.

We introduce the *fair work allocation* (FWA) problem. We have n content categories, each with some predicted workload d_i, referred to as the *demand* of category i. We also have m partners for content reviewers, each partner with some capacity s_j for reviewing content; we use s_j because we see this capacity as analogous to supply in the context of markets. For every partner j, consider all and *only* the content categories for which the partner has the appropriate language, cultural context, and necessary training to work for. For each of these content category i, there is a value v_{ij} which we mostly deem as the rate of partner j to handle category i. The value can also model, for example, the accuracy of the work. The goal of the planner is to produce an allocation of workload capacity of partners to categories, that satisfies predicted demand and capacity constraint. In general, we are interested in the setting where work capacity exceeds demand, in which case we would like to distribute excess work capacity across the content categories "fairly", since each category has variable demand realizations, and have different stakeholders that care particularly about a given category.

To address the FWA problem, we propose the use of tools from fair division. In particular, we study the application of leximin, competitive equilibrium from equal incomes (CEEI) [27], and more generally the max Nash welfare (MNW) allocation rule [6], for approaching this problem. However, as we will discuss next, the FWA has many practical considerations that necessitate the development of extensions to these tools.

The first practical consideration of the FWA problem is the *demands* of the categories. In our model, these demands are hard constraints that must be satisfied, and the "utility" achieved by each category is the supply of excess reviewing capacity that it receives beyond the forecasted demand. We refer to this excess allocation as *over-allocation*, and the primary goal of the planner is to fairly over-allocate. This requires us to consider an allocation model that is similar to the Nash bargaining model [20], where each agent has a *disagreement* point, which is analogous to the demands of categories in our model. We introduce variations on market-equilibrium-based allocation (including CEEI) and leximin fairness for handling these demands. We then show that, analogous to results for the standard setting [3, 14], we have equivalence between leximin and CEEI when all categories have *binary valuations*, i.e. partners can either handle a category or not; a second useful market-equilibrium approach where the budgets are not uniform leads to a different type of allocation, however.

A second practical consideration is that of *time*. As discussed in prior work on other aspects of the content moderation problem, there is an important time component to this problem [18, 21]. The reason for this is that the time between the creation of a piece of content and its review can have a large impact on the platform quality. If harmful content is visible to users before it has been reviewed, then it negatively impacts users. Conversely, if a piece of policy-compliant content is left in the review queue for a long time, it negatively impacts both the content creator and users. To address this time component, we introduce new versions of the various fair allocation approaches that take into account this time component.

Finally, related to time is the desire for smooth allocation across time. In practice, partners prefer to have relatively smooth allocations from each category across time, as it would be easier for staffing. We introduce a new variation of the market equilibrium model that takes into account this smoothness component.

For each of the variations introduced above, we study the properties resulting from these new variations, and show that in most cases we can still give desirable guarantees, for example by showing that market-equilibrium interpretations of the MNW allocation are preserved for several of these extensions. We also perform extensive numerical tests on real-world content moderation data, and study the qualitative and quantitative results derived from applying each of the proposed models to the FWA problem.

While our paper is motivated by the FWA problem, the ideas apply more broadly to several related operational problems. Here we mention two applications that are quite similar to the content moderation problem. The first one is the staffing problem for customer support services. The buyers here are different types of issues that need to be resolved, and each one has its own (expected) workload. On the other hand, the sellers are customer service representatives, and each representative can only work on a subset of the issues due to training, and has an upper limit on how much they can work in a day or in a week. The second application is how to schedule workers to provide labels for content that will later be used for the training of machine learning algorithms. Again, the buyers here are different types of content, and sellers are the workers. It is not hard to see that for both applications, over-allocation and smooth allocation arise as issues that must be addressed.

Relationship to Market Equilibrium. So far we have discussed the FWA problem in its concrete context. However, for the remainder of the paper, we will discuss our results in a more general market equilibrium context, to be more consistent with existing literature. To that end, we now describe how the FWA problem maps to a market equilibrium problem. The set of content categories corresponds to the set of buyers in a market (who now possess demands). The set of partners corresponds to the set of items in the market, and capacities correspond to item supplies. The value v_{ij} that category i has for partner j maps directly to a valuation in the market.

1.1 Related Literature

Content moderation has gained its importance in the last decade as we spend more time browsing on the internet. Hence, many problems that arise from content moderation have been studied in

the literature. The first is on developing reliable machine learning (ML) algorithms to automate content moderation (see, e.g., [12, 16]). Given that ML alone cannot solve the content moderation problem [5], it is also important to efficiently use the human review capacity. Haimovich et al. [13] developed a framework to predict the popularity of social network content in real-time, which could be used to detect harmful viral content and ultimately enable timely content moderation. Nguyen et al. [21] developed a statistical framework that increases the accuracy for detecting harmful content, by combining the decisions from multiple reviewers as well as those from ML algorithms. Makhijani et al. [18] built simulation models for the large-scale review systems to understand and optimize the human review process and guide the platforms for operational decisions (e.g., hiring additional reviewers). Garcelon et al. [11] investigated a multi-arm bandit framework to calibrate the severity predictions of content based on various ML models.

To add to existing works on different aspects of content moderation, we study the problem of scheduling reviewers in advance given the anticipated amount of workload, which we model as an allocation problem. The problem of allocating *divisible* goods to buyers fairly was first solved by Varian et al. [27] via the so-called *competitive equilibrium from equal incomes* (CEEI) solution. The CEEI solution is fair in a sense that it is envy-free [9]. Moreover, the CEEI solutions coincide with allocations that maximize the *Nash social welfare* [1] and thus can be computed in polynomial time via the convex program given by Eisenberg and Gale [7].

In terms of utility optimization, besides maximizing Nash welfare (MNW), some work investigates an *egalitarian* notion of utilities, such as *max-min* or more generally *leximin-optimal*. In the case of binary additive valuation, the set of MNW allocations and leximin-optimal allocations are known to coincide [3, 14]. Moreover, leximin-optimal solutions have been shown to satisfy several fairness properties in the case of indivisible goods (see, e.g., [10, 17, 24]).

The over-allocation setting of our paper is closely related to the framework of Nash bargaining (see, e.g., [23]) with the hard demands being interpreted as the *disagreement point*. A key difference between our model and the Nash bargaining model is that in the latter, there is the additional constraint that every buyer is allocated with exactly one unit of items. Interestingly, such fractional matching type of constraints allow the equivalence relation between the MNW allocations and leximin-optimal allocations to be extended to a more general setting of *bi-valued* valuations [2]. However, we show in Section 3.3 that this result does not hold in our setting. The smoothness setting was also studied by [4] for the resource allocation problem. However, their objective is to optimize the max-min utilities where we study the MNW allocations.

2 FISHER MARKETS WITH OVER ALLOCATION AND DISCRETE TIME

We start by introducing the standard *Fisher Market* setting. There is a set of n *buyers* (indexed by i), and m *items* (indexed by j). Each buyer has some *budget* denoted by $B_i \in \mathbb{R}_{\geq 0}$, and each item has a certain *supply* denoted by $s_j \in \mathbb{R}_{\geq 0}$. Moreover, every buyer i has some utility function $u_i(x_i)$ denoting the utility that they derive from *bundle* $x_i = (x_{i1}, x_{i2}, \cdots, x_{im}) \in \mathbb{R}^m_{\geq 0}$. We mostly be interested in *linear utilities*, where $u_i(x_i) = \langle v_i, x_i \rangle := \sum_j v_{ij} x_{ij}$, and $v_i = (v_{i1}, v_{i2}, \cdots, v_{im}) \in \mathbb{R}^m_{\geq 0}$ is buyer i's *valuation*[1] of the items, although we also consider generalizations of linear utilities a few times in the paper. We use $x \in \mathbb{R}^{n \times m}_{\geq 0}$ to denote an *allocation* of items to buyers, with row x_i denoting the bundle allocated to buyer i, and x_{ij} denoting how much buyer i gets of item j. The *utility profile* of an allocation x is the *set* of utilities of all buyers. Note that for utility profile, the ordering of buyers does not matter.

Given *prices* $p \in \mathbb{R}^m_{\geq 0}$, the *demand set* of a buyer i is the set of optimal bundles within budget:

$$D_i(p) = \arg\max_{x_i \geq 0} \{u_i(x_i) | \langle p, x_i \rangle \leq B_i\}.$$

A *market equilibrium* is a price-allocation pair (p, x) such that:

- every buyer i receives a bundle from their demand set, i.e. $x_i \in D_i(p)$; and
- any item j with positive price $p_j > 0$ must be exactly allocated, i.e. $\sum_i x_{ij} = s_j$.

In the standard setting, a market equilibrium not only exists, but is also efficiently computable via convex programming. The celebrated *Eisenberg-Gale* convex program [7] achieves this:

$$\begin{array}{lll} \max\limits_{x \geq 0} & \sum\limits_i B_i \log u_i(x_i) & \text{dual var.} \\ \text{s.t.} & \sum\limits_i x_{ij} \leq 1, \quad \forall j = 1, \ldots, m & p_j \end{array} \tag{EG}$$

The market equilibrium arises by combining the optimal primal solution with its corresponding dual variables on the supply constraints. The former gives an allocation that maximizes the *budget-weighted geometric mean of buyers' utilities* (i.e., $[\prod_i u_i^{B_i}]^{\frac{1}{\sum_i B_i}}$), and the dual variables act as prices.

When every buyer has a budget of one (i.e., $B_i = 1$), the objective of (EG) is equivalent to maximizing the *Nash Social Welfare (NSW)*, which is defined as the geometric mean of buyers' utilities. Moreover, market equilibrium is used to define one of the premier methods for fairly allocating divisible goods: *competitive equilibrium from equal incomes* (CEEI). The reason is that allocation via CEEI leads to several attractive fairness properties:

- *Pareto optimality*: there is no other allocation where every buyer is weakly better off (in terms of utilities) and some buyer is strictly better off;
- *Envyfreeness*: every buyer i prefers their own bundle to that of any other buyer k. If budgets are not equal, then this holds after scaling by budgets: $u_i(x_i) \geq \frac{B_i}{B_k} u_i(x_k)$;
- *Proportionality*: every buyer i prefers their own bundle to the one where they receive $\frac{B_i}{\sum_k B_k}$ of every item.

In this paper, we are interested in two variations on the standard Fisher market setting.

Allocation with hard demands from buyers. Each buyer i has some *required* utility demand d_i. We are required to find an allocation where every buyer i gets utility at least d_i, and subject to feasibility, the goal is to *fairly* allocate the excess utility, where the utility by which we measure fairness is $u_i(x_i) = \langle v_i, x_i \rangle - d_i$. We discuss and compare different fairness solution concepts in Section 3.

[1] If buyer i is not eligible for item j (i.e., in the application of content moderation, if a partner does not have the appropriate language and cultural context or necessary training for a content category), valuation v_{ij} will be zero.

Allocation over multiple time periods. We are interested in settings where there is a temporal component: there are T time steps, and items have supply constraints both per time step, and across all time steps. In particular, let x_{ij}^t denote how much of item j buyer i receives at time t. Then, for every time step t, we must satisfy the *per-time-step supply constraint* $\sum_i x_{ij}^t \le s_j^t$, and additionally satisfy the *overall supply constraint* $\sum_{i,t} x_{ij}^t \le s_j$. Note that the temporal component only matters if $s_j < \sum_t s_j^t$, since otherwise we could split each item up into T separate items each with supply specified by its per-time-step supply. We discuss different ways to model buyers' overall utility in Section 4 and show how to satisfy markets' desire for smooth allocation overtime in Section 5.

3 FAIR ALLOCATION OF OVERSUPPLY

Here we assume that there's a *hard* demand d_i on utilities for each buyer i, and then they have additional *soft* demand for getting additional supply. The goal is to find an allocation that satisfies every hard demand, and furthermore fairly allocate the oversupply to buyers as soft demands. To model fair allocation of oversupply, we define the utility of buyer i given bundle x_i as

$$u_i(x_i) \coloneqq \sum_j v_{ij} x_{ij} - d_i.$$

In this section, we consider two solution concepts: market equilibrium and leximin rules.

3.1 Solution via Market Equilibrium

For market equilibrium, we can apply the same idea as in (EG) and give the formulation below, which maximizes the budget-weighted geometric mean of buyers' utilities. By scaling, we additionally assume without loss of generality that the supply $s_j = 1$ for all sellers.

$$\begin{array}{lll} \max\limits_{x \ge 0} & \sum\limits_i B_i \log\left(\sum\limits_j v_{ij} x_{ij} - d_i\right) & \text{dual var.} \\ s.t. & \sum\limits_i x_{ij} \le 1, \quad \forall j = 1, \dots, m & p_j \end{array} \tag{EG-demand}$$

The objective in the formulation resembles that for Nash bargaining (see, e.g., [23]) with the hard demands being interpreted as the *disagreement point*. However, our problem is fundamentally different from Nash bargaining, which enforces a fractional matching type of constraints and requires that every buyer is allocated to exactly one unit of goods. As a result, the solution spaces between our formulation and the Nash bargaining formulation are different, and results from the Nash bargaining setting do not immediately extend to our setting.

For the rest of the paper, we impose the following regularity assumption, which ensures that the maximum budget-weighted geometric mean of buyers' utilities is strictly positive, and the (EG-demand) problem does not take logarithm of zeros.

Assumption 1. *There exists an allocation x such that $u_i(x_i) > 0$ for all buyer i.*

To simplify the notation for the rest of the section, we let $u_i \coloneqq \sum_j v_{ij} x_{ij} - d_i$ be the utility of agent i. Then, by the stationarirty of the KKT conditions, we get that

$$B_i \frac{v_{ij}}{u_i} - p_j \ge 0 \quad \forall i = 1, \dots, n;\ j = 1, \dots, m, \tag{1}$$

where equality holds if $x_{ij} > 0$ due to complementary slackness.

One property of the traditional fisher market (i.e., demand $d_i = 0$ for all buyers) is that with price vector p, every buyer spends their entire budget. Here, we show a similar property, but with *demand-inflated* budgets.

Lemma 3.1. *The market prices (p_j) verify $B_i \left(1 + \frac{d_i}{u_i}\right) = \sum_j p_j x_{ij}$, for all buyer i.*

Proof. Multiplying both sides of Equation (1) by x_{ij} and summing over j, we get that $B_i \frac{\sum_j v_{ij} x_{ij}}{u_i} = \sum_j p_j x_{ij}$. Since $\sum_j v_{ij} x_{ij} = u_i + d_i$, we conclude the result. □

Note that for when buyers do not have hard demands (i.e., $d_i = 0$), the equality in Lemma 3.1 recovers the property that every buyer spends all their budget: $B_i = \sum_j p_j x_{ij}$.

We next show the primal and dual solution pair (x, p) to problem (EG-demand) forms a market equilibrium under a slightly different definition of demand sets as a consequence of Lemma 3.1. The proof can be found in the appendix.

Theorem 3.2. *Let x, p be a solution to* (EG-demand). *The pair (x, p) is a market equilibrium where the demand sets are defined with respect to the demand-inflated budgets – that is, for each buyer i,*

$$D_i(p) \coloneqq \arg\max_{x_i \ge 0} \left\{ u_i(x_i) \coloneqq \textstyle\sum_j v_{ij} x_{ij} - d_i \;\middle|\; \sum_j p_j x_{ij} \le B_i\left(1 + \frac{d_i}{u_i(x_i)}\right) \right\}.$$

Before we continue to the next solution concept, we would like to point out that there are two interesting choices for buyers' budgets and both result in solutions with interesting properties. The first one is when all buyers have budgets of one, and in this case, the objective of (EG-demand) is simply the Nash welfare. The second choice is when buyers' budgets are equal to their demands.

Proposition 3.3. *Assume $v_{ij} = 1$ for all buyers i and sellers j. Then,*

(i) when $B_i = 1$ for all buyers i, the difference $\sum_j x_{ij} - d_i$ is constant for all buyers i;

(ii) when $B_i = d_i$ for all buyers i, the ratio $\frac{\sum_j x_{ij}}{d_i}$ is constant for all buyers i.

Proof. First note that the fact every buyer is assigned with a bundle in their demand set, as we have shown in Theorem 3.2, implies that all items have the same price, which we denote by $\bar{p}$. Let $\bar{x}_i \coloneqq \sum_j x_{ij}$ denote the amount of items allocated to buyer i. Then the property in Lemma 3.1 becomes $B_i \frac{\bar{x}_i}{\bar{x}_i - d_i} = \bar{p}\bar{x}_i$ (‡). For (i), we can simplify (‡) as $\bar{x}_i - d_i = \frac{1}{\bar{p}}$ for all buyers i; and for (ii), we can simplify (‡) as $\frac{\bar{x}_i}{d_i} = 1 + \frac{1}{\bar{p}}$ for all buyers i. These two simplifications conclude the proof. □

For more general non-complete graphs, we anticipate similar results hold in a sense that the solution will try to be as fair as possible, additively when budgets are one and multiplicatively when budgets are equal to demands. Indeed, we observe this phenomenon in our experiments (see Section 3.4 for details).

3.2 Solution via Leximin Fairness Concept

In this section, we introduce another fairness solution concept, called the *leximin-optimal* allocation.

Consider two vectors of equal length $v = (v_1, v_2, \cdots, v_n)$ and $\tilde{v} = (\tilde{v}_1, \tilde{v}_2, \cdots, \tilde{v}_n)$. Let v^* denoted an ordered vector of v so that $v_1^* \leq v_2^* \leq \cdots v_n^*$ (ties breaking randomly), and similarly let $\tilde{v}^*$ be an ordered vector of $\tilde{v}$. We say v is *leximin-greater* than $\tilde{v}$, denoted by $v >_{\text{lex}} \tilde{v}$, if there exists $k \in [n]$ such that $v_i^* = \tilde{v}_i^*$ for all $i < k$, and $v_k^* > \tilde{v}_k^*$.

For any allocation x, let $r(x)$ be the vector of budget-weighted utilities, where $r_i(x) = u_i(x_i)^{B_i}$. An allocation x is said to be *leximin-optimal* if there is no other allocation $\tilde{x}$ such that $r(\tilde{x}) >_{\text{lex}} r(x)$. In other words, the leximin-optimal solution first maximizes the minimum component in the $r(x)$ vector, then among all allocations where the minimum component in the $r(x)$ vector is maximized, it maximizes the second minimum value, etc.

Leximin is a refinement of the *max-min fairness* and a leximin-optimal allocation can be found by iteratively solving a sequence of max-min convex problems [28]. See Algorithm 1 for details.

ALGORITHM 1: Leximin-optimal solution

1: $F \leftarrow [n]$ {set of free variables}
2: Initialize the optimization problem

$$\begin{array}{ll} \max\limits_{x \geq 0} & \min\limits_{i \in F} r_i(x) \\ s.t. & \sum_i x_{ij} \leq 1, \quad \forall j = 1, \ldots, m \end{array} \qquad \text{(leximin)}$$

3: **while** $F \neq \emptyset$ **do**
4: Solve (leximin) and let r^* be the optimal solution and let i^* be a minimizer.
5: Add to (leximin) constraints: $r_i(x) \geq r^*$ for all $i \in F$.
6: Remove i^* from F
7: **end while**
8: **return** the latest solution of (leximin)

When buyers have equal budget and *binary additive utilities*, the leximin-optimal solution is known to maximize NSW, both when the items are indivisible [3] and when the items are divisible [14]. Therefore, we are interested in the connection between the leximin-optimal solutions and the solutions to (EG-demand) (referred to as EG solutions thereafter) in our setting.

3.3 Comparison

We say two allocations are *equivalent* if they give the same utility profile. In this section, we investigate whether EG and leximin-optimal solution are equivalent under various scenarios.

In terms of valuation, we first consider the special case of *binary* valuations (i.e., $v_{ij} \in \{0, 1\}, \forall i, j$), which represents compatibility between buyers and sellers and is commonly used in practice. Before considering the case with general valuations, we consider a slightly more restricted generalization known as bi-valued valuations (i.e., $v_{ij} \in \{\alpha, \beta\}, \forall i, j$, for some $\alpha > \beta \geq 0$). The reason we consider the bi-valued setting is that in other domains such as Nash bargaining, many interesting results arise when valuations are bi-valued (see, e.g., [2]). However, as we show in the following, in our setting, the leximin-optimal solution and the EG solution do not coincide for bi-valued valuations. Hence, we no longer need to investigate the most general setting as these negative results directly apply.

	budget = 1	budget = demand
binary valuation	✓ (Theorem 3.6)	✗ (Example 3.5)
bi-valued valuation	✗ (Example 3.4)	✗

Table 1: This table summarizes if EG solutions are equivalent to leximin-optimal solutions under different cases.

Besides different cases of valuations, we also consider two cases for budgets. The first one is when all buyers have equal budgets, for which we assume wlog that every buyer has a budget of one; and the second is when buyers' budgets are equal to their demands. Note that for the first case, EG solutions coincide with the maximum Nash welfare (MNW) solutions, and the proof is similar to the one presented in [14] for binary valuations in the classical setting without hard demand constraints.

Example 3.4. This example shows that the EG solution is not the same as the leximin-optimal solution even for bi-valued valuations. Assume we have two items, and two buyers with zero demand and unit budgets. All pairs have valuation one except for one pair: $v_{1,1} = 2$ and $v_{1,2} = v_{2,1} = v_{2,2} = 1$. Both the EG solution and the leximin-optimal solution allocate the entire second item to the second buyer. However, for the first item, the EG solution allocates the whole item to the first buyer, whereas the leximin-optimal solution allocates only $\frac{2}{3}$ units of it to the first buyer.

Example 3.5. The following example shows that the EG solution is not the same as the leximin-optimal solution when budget equals to demand, even when the valuation is binary. Consider the instance with two buyers and two items. Each item has a supply of one. The buyers have demand 0.1 and 0.2 respectively. The first buyer only accepts the first item, while the second buyer accepts both items.

Under the EG solution, the amount of the first item allocated to the first buyer is the solution for $\max_{a \in [0,1]} (a - 0.1)(2 - a - 0.2)^2$. Therefore, the EG solution allocates $\frac{2}{3}$ units of the first item to the first buyer. However, under the leximin-optimal solution, the amount of the first item allocated to the first buyer is the solution for the equation $a - 0.1 = (2 - a - 0.2)^2$, with $a \in [0, 1]$ due to the AM-GM inequality. The solution here is clearly not $\frac{2}{3}$.

Theorem 3.6. *With binary valuations and unit budget, for problem* (EG-demand)*, the set of leximin-optimal solutions is the same as the set of MNW solutions under Assumption 1.*

We defer the proof to Section 4.1, where we study a more general setting where the same result holds. That is, Theorem 3.6 may be viewed as a corollary to the more general result in Theorem 4.3.

3.4 Application

The content quality and integrity of online content platforms play a central role for users experience and platform growth. For that, companies spend billions of dollars every year to contract with different partners that provide human content reviews [26]. In this and the following application sections, we demonstrate through real data acquired from Meta (formerly known as Facebook) how our proposed methods can be useful in allocating partners' capacity to different content categories.

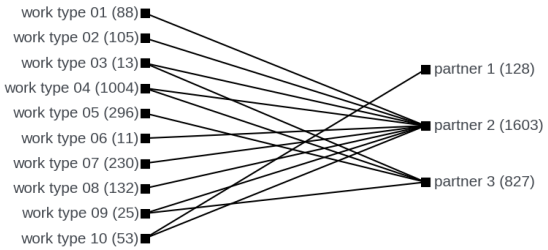
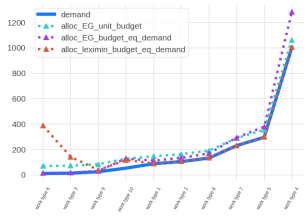


Figure 1: The graph displays buyers (reviewing work types) and sellers (reviewer partners) in a major market, where edges indicate whether a certain partner is able to work on a certain work type. The numbers in parentheses represent demands and supplies, and these numbers are generated through a certain transformation of the raw data.



(a) plot of demand and allocation

	buyer	demand	EG_B=1	EG_B=d	leximin_B=d
6	work type 6	10.83	69.45	13.83	387.28
3	work type 3	13.23	71.85	16.90	141.60
9	work type 9	24.63	83.25	31.46	38.20
1	work type 10	52.83	127.82	127.82	116.81
0	work type 1	87.64	146.26	111.94	89.72
2	work type 2	104.54	163.16	133.53	106.39
8	work type 8	132.12	190.74	168.76	133.75
7	work type 7	229.74	288.36	293.44	231.06
5	work type 5	296.32	354.94	378.49	297.56
4	work type 4	1003.56	1062.18	1281.83	1004.63

(b) table of demand and allocation

Figure 2: Allocation under different solution concepts and budget setups. In the table columns, "B=1" stands for the case where all buyers have budget of one; and "B=d" stands for the case where budgets are equal to demands. The buyers are sorted by their demands in an increasing order.

In this setting, sellers are *partners* that provide human review services and their supplies represent the amount of reviewer hours they can provide. The buyers are the different categories of content that need to be reviewed (hereafter referred to as *work types*), and their hard demands are the expected amount of human reviewing hours needed to review all the jobs (i.e., the *forecasted workloads*).

Since reviewers need to go through extensive training before being able to work on jobs in each work type, not all partners can work on all work types. As a result, we adopt the binary valuation structure to model compatibility between sellers and buyers. That is, if a partner can work on a certain work type, the pair has a valuation of one, and otherwise, it will have a value of zero.

For demonstration purposes, we focus on one particular medium-size market for in the main text of the paper. See Figure 1 for an example of buyers (and their demands), sellers (and their supplies), and the valuation structure for this market. We additionally include analyses for a large-size market in Appendix B. All takeaways from our analyses in the main text of the paper extend to the large-size market as well.

In the following, we delve into some specific reasons why a fair allocation is important for this business problem. First, since the demands are forecasted workload, the actually realized workload will unavoidably be different from the demand. Thus, being able to schedule the partners to handle more jobs than the forecasted workload provides a guardrail against natural fluctuation in the amount of workload as well as fluctuations due to headline events.

In addition, as the content review workforce is constantly evolving, occasionally, new work types are created and would have zero demands in the system. However, such zero demand is actually due to the fact that the machine learning models do not have adequate historical data to predict the workload yet. In this scenario, review hours should still be allocated to these new work types regardless of their predicted workload.

We investigate the allocations obtained from both the EG solutions and the leximin-optimal solutions. In addition, we consider two cases for buyers' budget: we first assume buyers have unit budgets, and then assume that buyers' budgets are equal to their demands. Note that in the former case, due to Theorem 3.6, the EG solutions and the leximin-optimal solutions coincide, and thus we omit one in the presentation in Figure 2.

We start by examining the EG solutions and we in particular observe that they satisfy properties similar to those stated in Proposition 3.3. When budgets are equal to ones, the oversupply from partners is equally divided among the work types. That is, the oversupply is distributed fairly in an *additive* sense. More specifically, except for "work type 10" (which is the only one that is compatible with the "partner 1"), all other work types are allocated an additional 60 hours on top of their demands.

However, when buyers' budgets are set to be the same as their demand, we notice that the oversupply is allocated fairly in a more *multiplicative fashion* under the EG solution. That is, work types with higher demand will receive higher amounts of allocation exceeding their demand. In our experiment, we observe that again except for the "'work type 10", all other work types receive an allocation that is about 128% of their demands.

Lastly, the leximin-optimal solution when budgets are equal to demands does not perform well as it allocates a significant amount of oversupply to the buyer with the smallest amount of demand (or equivalently budget). The reason for this is that under leximin optimality objective, in order to maximize the minimum budget adjusted utility, the buyer with the smallest budget (or demand) will end up with the highest amount of utility (i.e., allocation). However, this solution, although is fair from an egalitarian point of view, does not align well with the business purposes here.

4 MULTIPLE TIME PERIODS

In the application of short term planning for content reviewers, we would typically have constraints at several levels of the problem: a given review site may have an overall maximum capacity that they can allocate across the planning period, but then they also typically have tighter per-time-period staffing constraints. To accommodate such constraints, we extend the basic Fisher market model to allow each item to have time-based constraints. More specifically, we will consider a model with T discrete time steps, and each item j will have an *overall* supply s_j, as well as a *per-time period* supply s_j^t.

The allocation vector now has a time component and we denote by x_{ij}^t the amount of item j allocated to buyer i at time t. Similarly, the demand vector now also a time component, where d_i^t is the

demand of buyer i at time t, and we denote $d_i := \sum_t d_i^t$ as the total demand of buyer i.

Similar to the previous section, we may consider two solution concepts and investigate their equivalence under the special case of unit budget and binary valuation. In terms of buyers' overall utilities, we consider two possibility to combine their utilities at each time period, where we denote the utility of buyer i at time t as

$$u_i^t(x_i^t) = \sum_j v_{ij} x_{ij}^t - d_i^t.$$

In the first approach, the utility of each buyer is simply the sum of the utilities at different times, whereas in the second approach, the utility of each buyer is the geometric mean of the utilities at different times. In terms of fairness, the first approach is a direct extension of the previous section, aiming to be fair to all the buyers; and the second approach targets fairness at a more granular level, aiming to be fair to all the buyers across all the time periods.

4.1 Total utility as the sum across time periods

For one possibility, we assume the overall utility of buyer i is

$$u_i(x_i) = \sum_t u_i^t(x_i^t).$$

Hence, we have the following program, which maximizes the budget-weighted utilities over the set of feasible allocations under time-based constraints. Note that this formulation also applies to the special case where each buyer only has an overall demand throughout all time periods (i.e., no time-specific demand).

$$\begin{array}{lll|l} \max\limits_{x\geq 0} & \sum\limits_i B_i \log(\sum\limits_j v_{ij}(\sum\limits_t x_{ij}^t) - d_i) & & \text{dual var.} \\ s.t. & \sum\limits_i x_{ij}^t \leq s_j^t, & \forall j \in [m], t \in [T] & \lambda_j^t \\ & \sum\limits_{t,i} x_{ij}^t \leq s_j, & \forall j \in [m] & \lambda_j \\ & \sum\limits_j v_{ij} x_{ij}^t \geq d_i^t, & \forall i, t & \end{array} \qquad \text{(EG-Time-sum)}$$

It follows from stationarity of the KKT conditions that the following must hold:

$$B_i \frac{v_{ij}}{u_i} \leq \lambda_j^t + \lambda_j, \quad \forall i, j, t \tag{2}$$

Note that here, we omit the dual variables corresponding to the non-negative constraints and thus, the above must hold with equality if $x_{ij}^t > 0$ due to complementary slackness.

Lemma 4.1. *Let x, λ be a solution to* (EG-Time-sum)*. Consider the pair (x, p), where the price $p_j^t = \lambda_j^t + \lambda_j$ on each item j at a given time period t. Then, every buyer spends their entire demand-inflated budget. That is, $\sum_{t,j} x_{ij}^t p_j^t = B_i(1 + \frac{d_i}{u_i})$ for every buyer i.*

Proof. Multiplying (2) by x_{ij}^t and summing over t and j, we get that for all buyer i,

$$\sum_{t,j} x_{ij}^t p_j^t = \sum_{t,j} x_{ij}^t (\lambda_j^t + \lambda_j) = B_i \sum_{t,j} x_{ij}^t \frac{v_{ij}}{u_i} = B_i(\frac{u_i + d_i}{u_i}) = B_i(1 + \frac{d_i}{u_i}).$$

□

Thus, when buyers have no hard demand, the market clears if we set the price of item j at time t equal to $p_j^t = \lambda_j^t + \lambda_j$.

To show that the allocation-price pair (x, p) is a market equilibrium, we need to show the condition that an item has price p_j^t strictly greater than 0 only if $\sum_{i,t} x_{ij}^t = s_j$. This is not guaranteed in our setting. Instead, we have that $p_j^t > 0$ only if *either* $\sum_i x_{ij}^t = s_j$ *or* $\sum_{t,i} x_{ij}^t = s_j^t$. Thus, we get a market equilibrium if we generalize the price complementarity requirement to be that an item must have a nonzero price only if it is supply-constrained at some layer of its supply hierarchy. More formally, we have the following theorem. The proof can be found in the appendix.

Theorem 4.2. *Let x, λ be a solution to* (EG-Time-sum)*. Consider the pair (x, p), where the price $p_j^t = \lambda_j^t + \lambda_j$ on each item j at a given time period t. This pair is a market equilibrium under the condition that an item has $p_j^t > 0$ only if* either $\sum_i x_{ij}^t = s_j^t$ or $\sum_{t,i} x_{ij}^t = s_j$*, and under the following definition of demand sets:*

$$D_i(p) := \arg\max_{x_i \geq 0} \left\{ u_i(x_i) := \textstyle\sum_t u_i^t(x_i^t) \;\middle|\; \sum_{j,t} p_j^t x_{ij}^t \leq B_i(1 + \frac{d_i}{u_i(x_i)}) \right\}.$$

A very related theorem is shown by Jain and Vazirani [15]: the EG program associated to any market which can be formulated as a network flow problem with a single source, and a sink for each buyer, where the utility of a buyer is the sum of mass on paths to the corresponding sink, and with the price on a path being equal to the sum of prices on the edges in the path, can be solved by maximizing the sum of the logs of these buyer utilities, and this yields a market equilibrium. (EG-Time-sum) fits in this framework, except that we have heterogeneous edge weights that go into the utility function.

Algorithm 1 can be easily adjusted to obtain the leximin-optimal solution in this setting with time-based constraints. In the following, we investigate whether results in Section 3.3 extend to the multiple-time setting. Since the multiple-time setting is a generalization, the negative results in Section 3.3 immediately extend. Thus, we focus on the positive result and in particular we show that when all buyers have budgets of ones and when valuations are binary, the set of MNW solutions and the set of leximin-optimal solutions coincides.

Theorem 4.3. *With binary valuations and unit budget, for problem* (EG-Time-sum)*, the set of leximin-optimal solutions is the same as the set of MNW solutions under Assumption 1.*

Proof. Let x be an allocation that is an EG solution (or equivalently an MNW solution since all buyers have budgets of one), we want to show that it is also leximin-optimal. Assume by contradiction that it is not. Let $\tilde{x}$ be a leximin-optimal solution. Let $r := (u_i(x_i))_{i\in[n]}$ and $\tilde{r} := (u_i(\tilde{x}_i))_{i\in[n]}$ denote the vectors of budget-weighted utilities corresponding allocation x and $\tilde{x}$, receptively. First observe that by optimality we must have the following.

Claim 1. $\sum_i r_i(x_i) = \sum_i \tilde{r}_i(\tilde{x}_i) = \sum_j \min(s_j, \sum_t s_j^t) - \sum_i d_i$.

Then, by Assumption 1, we can apply the same proof idea as given in [14]. Let $o_1, o_2, \cdots, o_n$ be an enumeration of $[n]$ such that $r_{o_1} \leq r_{o_2} \leq \cdots \leq r_{o_n}$, and similarly let $\tilde{o}_1, \tilde{o}_2, \cdots, \tilde{o}_n$ be an enumeration such that $\tilde{r}_{\tilde{o}_1} \leq \tilde{r}_{\tilde{o}_2} \leq \cdots \leq \tilde{r}_{\tilde{o}_n}$. Since $\tilde{r} \succ_{\text{lex}} r$ by assumption, we can let $k \in [n]$ be the smallest index so that $\tilde{r}_{\tilde{o}_k} > r_{o_k}$. Then by Claim 1, we know there must be another index $i > k$ so that $\tilde{r}_{\tilde{o}_i} < r_{o_i}$ and we let k' be the smallest such index. Let $N := \{o_1, o_2, \cdots, o_{k'-1}\}$. Then, by definition of the enumeration o

and our choice of k', we have $\sum_{i\in N} r_i < \sum_{i\in N} \tilde{r}_i$. Thus, there exists buyers $i \in N$, $i' \in [n] \setminus N$, and item $j \in [m]$ such that $v_{ij} = v_{i'j} = 1$ and $x^t_{i'j} > 0$ for some $t \in [T]$. Also, note that we must have $r_i < r_{i'}$ by definition of N. Hence, by Assumption 1, one can re-allocate some item j from buyer i' to i at time period t without affecting feasibility. However, this re-allocation will increase the Nash Social Welfare, which contradicts the assumption that x is an MNW solution. □

4.2 Total utility as the geometric mean across time periods

For the second possibility, we assume the overall utility of buyer i is

$$u_i(x_i) = \big(\prod_t u_i^t(x_i^t)\big)^{\frac{1}{T}}.$$

To maximize the budget-weighted utilities over the set of feasible allocations under time-based constraints, we have the following program.

$$\begin{array}{lll|ll} \max\limits_{x\geq 0} & \sum\limits_i B_i \cdot \frac{1}{T}\sum\limits_t \log(\sum\limits_j v_{ij}x^t_{ij} - d^t_i) & & \text{dual var.} & \\ s.t. & \sum\limits_i x^t_{ij} \leq s^t_j, & \forall j \in [m], \forall t \in [T] & \lambda^t_j & \text{(EG-Time-geo-mean)} \\ & \sum\limits_{t,i} x^t_{ij} \leq s_j, & \forall j \in [m] & \lambda_j & \end{array}$$

In this setting, stationarity of the KKT conditions gives us the following relationship:

$$B_i \cdot \frac{1}{T} \cdot \frac{v_{ij}}{u^t_i} \leq \lambda^t_j + \lambda_j, \quad \forall i, j, t \tag{3}$$

where equality holds whenever $x^t_{ij} > 0$ due to complementary slackness.

Lemma 4.4. *Let x, λ be a solution to* (EG-Time-geo-mean). *Consider the pair (x, p), where the price $p^t_j = \lambda^t_j + \lambda_j$ on each item j at a given time period t. Then, every buyer spends their entire demand-inflated budget. That is,*

$$\sum_{t,j} x^t_{ij} p^t_j = B_i \left[\frac{1}{T}\sum_t (1 + \frac{d^t_i}{u^t_i})\right], \quad \forall i = 1, \ldots, n$$

Proof. Multiplying (3) by x^t_{ij} and summing over t and j, we get that for all buyer i,

$$\textstyle\sum_{t,j} x^t_{ij}p^t_j = B_i \cdot \frac{1}{T}\sum_{t,j} x^t_{ij}\frac{v_{ij}}{u^t_i} = B_i \cdot \frac{1}{T}\sum_t (\frac{u^t_i + d^t_i}{u^t_i}) = B_i\left[\frac{1}{T}\sum_t (1 + \frac{d^t_i}{u^t_i})\right].$$

□

Thus, when buyers have no hard demand at all time periods (i.e., $d^t_i = 0$ for all i, t), the market clears if we set the price of item j at time t equal to $p^t_j = \lambda^t_j + \lambda_j$.

To show that the allocation-price pair (x, p) is a market equilibrium, not only do we need to modify the price complementarity requirement as in the previous subsection, we also need to modify the definition of demand sets due to Lemma 4.4. The proof of the following theorem can be found in the appendix.

Theorem 4.5. *Let x, λ be a solution to* (EG-Time-geo-mean). *Consider the pair (x, p), where the price $p^t_j = \lambda^t_j + \lambda_j$ on each item j at a given time period t. This pair is a market equilibrium under the condition that an item has $p^t_j > 0$ only if* either $\sum_i x^t_{ij} = s^t_j$ or $\sum_{t,i} x^t_{ij} = s_j$, *and under the following definition of demand sets:*

$$D^t_i(p) = \arg\max_{x_i \geq 0} \left\{u_i(x_i) := \left(\prod_t (u^t_i(x^t_i))\right)^{\frac{1}{T}} \middle| \sum_{j,t} p^t_j x^t_{ij} \leq B_i\left[\tfrac{1}{T}\sum_t (1 + \tfrac{d^t_i}{u^t_i(x^t_i)})\right]\right\}.$$

In the following, we define what makes a solution leximin-optimal in this setting. Instead of having a vector of buyers' budget-weighted utilities, we now have a vector of buyers' budget-weighted utilities at all time periods. More specifically, for any allocation x, the vector $r(x)$ now has a time component: $r^t_i(x) = u^t_i(x^t_i)^{B_i}$; and an allocation x is *leximin-optimal* if there is no other allocation $\tilde{x}$ such that $r(\tilde{x}) >_{\text{lex}} r(x)$. With this definition, when all buyers have unit budgets, (EG-Time-geo-mean) is the same as (EG-demand) if we treat each buyer at each time period as an individual buyer. Thus, Theorem 3.6 extends naturally.

Corollary 4.6. *With binary valuations and unit budget, for* (EG-Time-geo-mean), *the set of leximin-optimal solution is the same as the set of MNW solutions under Assumption 1.*

4.3 Application

For application, we continue our investigation on the same market. Instead of a single time period, we now focus on four consecutive time periods. We observe from our data that the supply from our partners is constant across the four time periods. This is expected given that partners usually have a fixed amount of facilities to accommodate their employees. On the other hand, we observe that for some work types, their forecasted demands vary from one time period to the next.

Due to the fact that variations are on the demand side, we use the second utility definition (as in Section 4.2) for our experiments. Moreover, for the following, we only focus on the case where buyers have a budget of one. The reason for this simplification is that the qualitative results are similar to those under the case where budgets are equal to demands.

The allocation results are shown in Figure 3. We omit "partner 1" since it can only work on "work type 10", and thus intuitively all of its reviewing capacity should be allocated toward "work type 10". As one can see from Figure 3, for the most part, the allocation for each buyer-seller pair is rather smooth from time to time. However, for "work type 05", its allocation from "partner 3" partner goes down, but it naturally follows from the fact that its demand has a downward trend; and for "work type 04", although its demand is relatively constant, its allocation from both partners vary, with one having an upward trend and the other having a downward trend.

Because of the variability in allocation from one to the next, the solution here has the drawback that partners have to shift their staffing requirements from time to time, which could be costly from an operational point of view. As a result, it would be beneficial to the partners if we can obtain a solution that minimizes staffing change as much as possible. We discuss more in detail on markets' desire for smooth allocation in the next section.

5 SMOOTH ALLOCATION ACROSS MULTIPLE DISCRETE TIME PERIODS

In the previous section, we saw how to allow each item to have per-time-period and overall supply. However, the resulting allocations

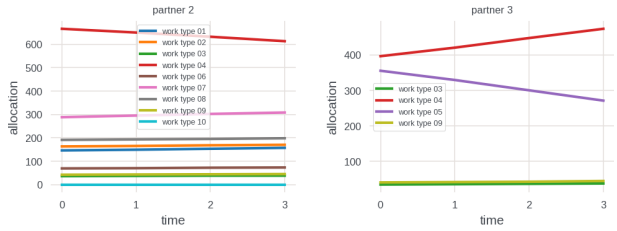


Figure 3: For each partner in the market, the plot shows the allocation to its compatible work types from all time periods. We omit "partner 1" since its allocation is immediate as it is only compatible to one work type.

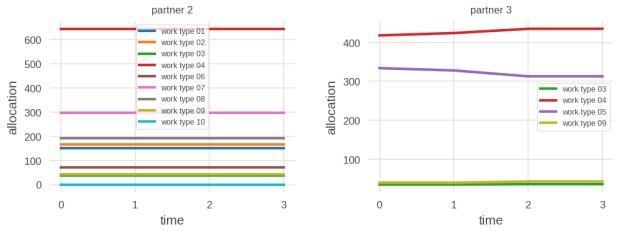


Figure 4: This is a similar plot to Figure 3, but this figure plots the allocation with absolute difference penalty function in the objective when the smoothness parameter $\gamma = 0.005$.

may still exhibit an effect which is often considered undesirable in practice: the allocation may have large swings across time, even when the utility from this is low. Consider the following content review example:

Example 5.1. There are two review categories (buyers) and two review sites (items), with three time steps. Each review site has per-period supply of one, and no cross-time supply constraint. All valuations are one. There are many optimal allocations, but we will consider three in particular:

- *high variation* one: $x_{11} = x_{22} = (1, 0, 1), x_{12} = x_{21} = (0, 1, 0)$;
- *low variation* one: $x'_{11} = x'_{22} = (1, 1, 1)$ and $x'_{12} = x'_{21} = (0, 0, 0)$;
- another low-variation one: $\hat{x}_{ij} = (0.5, 0.5, 0.5)$ for all i, j.

In the high-variation allocation, each partner gets a non-smooth allocation: in time step 2 and 3, they flip their entire review category from one to the other. In practice, this is inefficient due to staff context switching. This can lead to less efficient reviewing for a given content category. In contrast, the two low-variation allocations give each review site a nice constant amount of work from each review category. We do not necessarily prefer $\hat{x}$ or x' to each other; both are equally "smooth".

From the example above, we can clearly see the problem with the market equilibrium solutions, which is that they do not take into account market's preferences for low-variation solutions.

To model the preference for low variation, one approach is to add constraints forcing the variation to be within a certain range. For instance, if we want the allocation from one time period to the next to be within 5%, we can add constraints

$$x_{ij}^{t+1} \geq 0.95 \cdot x_{ij}^{t}, \quad \forall i \in [n], j \in [m], t \in [T-1]$$
$$x_{ij}^{t+1} \leq 1.05 \cdot x_{ij}^{t}, \quad \forall i \in [n], j \in [m], t \in [T-1].$$

However, it is not hard to see that one issue with the constraints approach is that when demands are high (i.e., small amount of oversupply to distribute) and when variation in demand from time to time is also high, the problem could become infeasible.

Alternatively, we can include (or subtract, to be more precise) a penalty term in the objective. We assume the penalty term takes the following form $\gamma \sum_{i,j,t} R(x_{ij}^{t+1}, x_{ij}^{t})$, where R_{jt} is any convex function that measures the discrepancy between two quantities and γ is a hyperparameter that captures the strength in terms of the market's desire for smooth allocation. Since R_{jt} is the penalty function for non-smooth allocation, we further require that $R_{jt}(x_{ij}^{t+1}, x_{ij}^{t}) = 0$ whenever $x_{ij}^{t+1} = x_{ij}^{t}$. Hence, we have the following program.

$$\begin{aligned} \max_{x \geq 0} \quad & \sum_i B_i \sum_t \log\Big(\sum_j x_{ij}^t v_{ij} - d_i^t\Big) - \gamma \sum_{i,j,t} R(x_{ij}^{t+1}, x_{ij}^t) \\ \text{s.t.} \quad & \sum_i x_{ij}^t \leq s_j^t, \quad \forall j = 1, \ldots, m, \forall t = 1, \ldots, T, \\ & \sum_{t,i} x_{ij}^t \leq s_j, \quad \forall j = 1, \ldots, m, \end{aligned} \tag{EG-Smooth}$$

Below, we list a couple of interesting choices for the penalty function $R_{jt}(x_{i,j}^{t+1}, x_{i,j}^{t})$

- *Absolute deviation*: $|x_{i,j}^{t+1} - x_{i,j}^{t}|$.
- *Kullback-Leibler*: $\max(x_{i,j}^{t+1} \log \frac{x_{i,j}^{t+1}}{x_{i,j}^{t}}, x_{i,j}^{t} \log \frac{x_{i,j}^{t}}{x_{i,j}^{t+1}})$.

5.1 Application

To continue our experiment in Section 4.3, we now include an additional penalty term in the objective. We show the results in Figure 4 using the absolute difference penalty function with appropriate smoothness parameters. We would like to point out the qualitative results using the KL divergence penalty function is similar and we omit them here.

Note that now for both "work type 04" and "work type 05", the allocations to "partner 3" are much smoother as compared those previously obtained as shown in Figure 3.

As the smoothness parameter increases, the dominating component in the objective of (EG-Smooth) shifts from the utility term to the smoothness penalty term. Hence, if we were to set $\gamma = 0.01$, the resulting allocation becomes constant for each (work type, partner) pair. However, such a smooth allocation is at the expense of wasting partners' reviewing capacity. Thus, the smoothness parameter should be carefully chosen to balance different business objectives.

6 CONCLUSION

We study the planning problem of allocating human content reviewers (supply) to different harmful content categories (demand) across multiple time periods with over-allocation and smoothness considerations. We model the problem as an resource allocation problem, provide an optimization formulation, and show its connection to market equilibria. Moreover, we demonstrate the performance of the proposed solution through data obtained from Meta.

REFERENCES

[1] Kenneth Arrow and Michael Intriligator. 2005. *Handbook of mathematical economics.* Technical Report. Elsevier.

[2] Haris Aziz and Ethan Brown. 2020. Random Assignment Under Bi-Valued Utilities: Analyzing Hylland-Zeckhauser, Nash-Bargaining, and other Rules. *arXiv preprint arXiv:2006.15747* (2020).

[3] Haris Aziz and Simon Rey. 2020. Almost group envy-free allocation of indivisible goods and chores. *Proceedings of the 29th International Joint Conference on Artificial Intelligence (IJCAI)* (2020).

[4] Evripidis Bampis, Bruno Escoffier, and Sasa Mladenovic. 2018. Fair resource allocation over time. In *AAMAS 2018-17th International Conference on Autonomous Agents and MultiAgent Systems*. International Foundation for Autonomous Agents and Multiagent Systems, 766–773.

[5] Caitlin Bassett. 2022. Will AI take over content moderation? *Mind Matters News* (2022). https://mindmatters.ai/2022/01/will-ai-take-over-content-moderation/

[6] Ioannis Caragiannis, David Kurokawa, Hervé Moulin, Ariel D Procaccia, Nisarg Shah, and Junxing Wang. 2019. The unreasonable fairness of maximum Nash welfare. *ACM Transactions on Economics and Computation (TEAC)* 7, 3 (2019), 1–32.

[7] Edmund Eisenberg and David Gale. 1959. Consensus of subjective probabilities: The pari-mutuel method. *The Annals of Mathematical Statistics* 30, 1 (1959), 165–168.

[8] Facebook. 2022. Facebook Community Standards. https://transparency.fb.com/policies/community-standards/ Online; accessed 7-February-2022.

[9] Duncan Karl Foley. 1966. *Resource allocation and the public sector.* Yale University.

[10] Rupert Freeman, Sujoy Sikdar, Rohit Vaish, and Lirong Xia. 2019. Equitable allocations of indivisible goods. *arXiv preprint arXiv:1905.10656* (2019).

[11] Evrard Garcelon, Vashist Avadhanula, Alessandro Lazaric, et al. 2021. Top K Ranking for Multi-Armed Bandit with Noisy Evaluations. *arXiv preprint arXiv:2112.06517* (2021).

[12] Tarleton Gillespie. 2020. Content moderation, AI, and the question of scale. *Big Data & Society* 7, 2 (2020), 2053951720943234.

[13] Daniel Haimovich, Dima Karamshuk, Thomas J Leeper, Evgeniy Riabenko, and Milan Vojnovic. 2020. Scalable Prediction of Information Cascades over Arbitrary Time Horizons. *arXiv preprint arXiv:2009.02092* (2020).

[14] Daniel Halpern, Ariel D Procaccia, Alexandros Psomas, and Nisarg Shah. 2020. Fair division with binary valuations: One rule to rule them all. In *International Conference on Web and Internet Economics*. Springer, 370–383.

[15] Kamal Jain and Vijay V Vazirani. 2010. Eisenberg–Gale markets: algorithms and game-theoretic properties. *Games and Economic Behavior* 70, 1 (2010), 84–106.

[16] Shagun Jhaver, Iris Birman, Eric Gilbert, and Amy Bruckman. 2019. Human-machine collaboration for content regulation: The case of Reddit Automoderator. *ACM Transactions on Computer-Human Interaction (TOCHI)* 26, 5 (2019), 1–35.

[17] David Kurokawa, Ariel D Procaccia, and Nisarg Shah. 2018. Leximin allocations in the real world. *ACM Transactions on Economics and Computation (TEAC)* 6, 3-4 (2018), 1–24.

[18] Rahul Makhijani, Parikshit Shah, Vashist Avadhanula, Caner Gocmen, Nicolás E Stier-Moses, and Julián Mestre. 2021. QUEST: Queue Simulation for Content Moderation at Scale. *arXiv preprint arXiv:2103.16816* (2021).

[19] Meta. 2022. Community Standards Enforcement Report. https://transparency.fb.com/data/community-standards-enforcement/ Online; accessed 31-January-2023.

[20] John Nash. 1953. Two-Person Cooperative Games. *Econometrica* 21, 1 (1953), 128–140.

[21] Viet-An Nguyen, Peibei Shi, Jagdish Ramakrishnan, Udi Weinsberg, Henry C Lin, Steve Metz, Neil Chandra, Jane Jing, and Dimitris Kalimeris. 2020. CLARA: confidence of labels and raters. In *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. 2542–2552.

[22] Justin Osofsky. 2019. Our Commitment to Our Content Reviewers. (2019). https://about.fb.com/news/2019/02/commitment-to-content-reviewers/

[23] Ioannis Panageas, Thorben Tröbst, and Vijay V Vazirani. 2021. Combinatorial Algorithms for Matching Markets via Nash Bargaining: One-Sided, Two-Sided and Non-Bipartite. *arXiv preprint arXiv:2106.02024* (2021).

[24] Benjamin Plaut and Tim Roughgarden. 2020. Almost envy-freeness with general valuations. *SIAM Journal on Discrete Mathematics* 34, 2 (2020), 1039–1068.

[25] Guy Rosen. 2022. Community Standards Enforcement Report, Third Quarter 2022. (2022). https://about.fb.com/news/2022/11/integrity-and-transparency-reports-q3-2022/

[26] Katie Schoolov. 2021. Why content moderation costs billions and is so tricky for Facebook, Twitter, YouTube and others. *CNBC* (2021). https://www.cnbc.com/2021/02/27/content-moderation-on-social-media.html

[27] Hal R Varian et al. 1974. Equity, envy, and efficiency. *Journal of Economic Theory* 9, 1 (1974), 63–91.

[28] Stephen J Wilson. 1998. Fair division using linear programming. *preprint, Departement of Mathematics, Iowa State University* (1998).

A MISSING PROOF

Proof of Theorem 3.2. The second condition in the definition of market equilibrium is satisfied due to complementary slackness. To show the first condition, which is every buyer i receives a bundle from their demand set, we show that every item in the bundle x_i maximizes the utility per dollar for buyer i. By dual feasibility of problem EG-demand, we have $\lambda_{ij} \geq 0$. Hence, rearranging the terms of (1), we have

$$\frac{v_{ij}}{p_j} \leq \frac{u_i}{B_i} \quad \forall i = 1, \ldots, n;\ j = 1, \ldots, m,$$

where equality holds when $x_{ij} > 0$ due to the complementary slackness condition $x_{ij}\lambda_{ij} = 0$. This implies that the utility per dollar is maximized by the buyer's bundle, concluding the proof. □

Proof of Theorem 4.2. Rearranging the terms of (2), we have $\frac{v_{ij}}{p_j^t} \leq \frac{u_i}{B_i}$ for all i, j, t, where equality holds when $x_{ij}^t > 0$ due to the complementary slackness condition on x_{ij}^t. Thus, we get that buyers only buy items that give maximal bang-per-buck. It follows that each buyer gets a bundle in their demand set. In addition, by complementary slackness, we have that $\lambda_j^t = 0$ unless $\sum_i x_{ij}^t = s_j^t$, and similarly $\lambda_j = 0$ unless $\sum_{t,i} x_{ij}^t = s_j$. It follows that $p_j^t = \lambda_j + \lambda_j^t > 0$ only if *either* $\sum_i x_{ij}^t = s_j^t$ *or* $\sum_{t,i} x_{ij}^t = s_j$. □

Proof of Theorem 4.5. Note that the utility function is no longer linear but is concave since the geometric mean function is concave. To show that every buyer is allocated a bundle that is in their demand set, we instead want to show that $x_{ij}^t > 0$ only if it has the maximal *marginal* bang-per-buck, where the marginal utility of item j for buyer i at time t is

$$\frac{\partial u_i(x_i)}{\partial x_{ij}^t} = \left(\prod_{t':t' \neq t} u_i^{t'}(x_i^{t'}) \right) v_{ij}^t = \frac{u_i(x_i)}{u_i^t(x_i^t)} v_{ij}^t.$$

Rearranging the terms of (3), we have

$$\frac{v_{ij}}{p_j^t} \leq \frac{u_i^t}{B_i} \Rightarrow \frac{u_i}{u_i^t} v_{ij}/p_j^t \leq \frac{u_i}{B_i}, \quad \forall i, j, t,$$

where equality holds when $x_{ij}^t > 0$ due to the complementary slackness condition on x_{ij}^t. Thus, we get that buyers only buy items that give max marginal bang-per-buck, as desired. The proof for the second condition is the same as that in the proof of Theorem 4.2. □

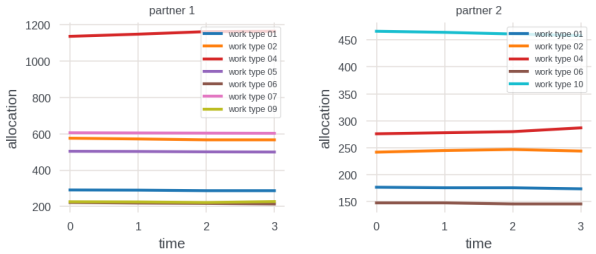


Figure 7: For "partner 1" and "partner 2" in the market, the plot shows the allocation to its compatible work types from all time periods.

B ADDITIONAL EXPERIMENTS

In this section, we include our experimental results on a large-size market, with more partners than the market shown in Figure 1. The structure of work types and partners is shown in Figure 5.

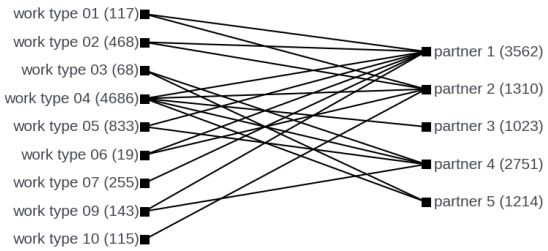
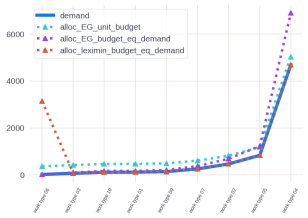


Figure 5: The graph displays buyers (reviewing work types) and sellers (reviewer partners) in a major market, where edges indicate whether a certain partner is able to work on a certain work type. The numbers in parentheses represent demands and supplies, and these numbers are generated through a certain transformation of the raw data.

Similar to Figure 2, we show, in Figure 6, the allocations obtained from both the EG solutions and the leximin-optimal solutions. In addition, we consider two cases for buyers' budget, with buyers either have unit budgets or have budgets that are equal to their demands.



(a) plot of demand and allocation

	buyer	demand	EG_B=1	EG_B=d	leximin_B=d
5	work type 06	19.22	369.89	28.27	3149.50
2	work type 03	67.81	418.48	99.73	77.60
8	work type 10	115.28	465.95	169.54	119.11
0	work type 01	116.96	467.63	172.02	120.71
7	work type 09	142.96	493.64	210.25	145.91
6	work type 07	255.15	605.81	375.25	256.99
1	work type 02	468.01	818.68	688.31	469.41
4	work type 05	833.43	1184.09	1225.74	834.63
3	work type 04	4685.97	5036.68	6891.74	4687.01

(b) table of demand and allocation

Figure 6: Allocation under different solution concepts and budget setups. In the table columns, "B=1" stands for the case where all buyers have budget of one; and "B=d" stands for the case where budgets are equal to demands. The buyers are sorted by their demands in an increasing order.

We observe similar patterns in the allocation under the EG solutions as for the market investigated in the main text. In particular, when buyers' budgets are set to be the same as their demand, the oversupply is allocated fairly *multiplicatively* (about 47% more); and when budgets are equal to ones, the oversupply is allocated fairly *additively* (about 350 units more). Further, when budgets equal to demands, the leximin solution does not perform well.

Lastly, for allocations involving multiple time periods, we similarly observe that without the smoothness penalty term, the allocation for each buyer-seller pair could vary from time to time (see Figure 7). Although minimal, this variation means that the partners have to shift their staffing schedules from time to time.

Again, with penalty terms, we can smooth out the allocation to minimize staffing switches for the partners. The allocation obtained with an absolute deviation penalty term is shown in Figure 8.

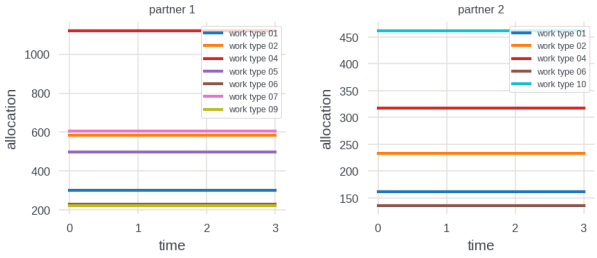


Figure 8: This figure plots the allocation with absolute difference penalty function in the objective when the smoothness parameter γ is set to be 0.005.