Fair School Allocation

Citation

Permanent link
https://nrs.harvard.edu/URN-3:HUL.INSTREPOS:37376429

Terms of Use
This article was downloaded from Harvard University’s DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA

Share Your Story
The Harvard community has made this article openly available. Please share how this access benefits you. Submit a story.

Accessibility
Fair School Allocation

A dissertation presented by Isaac A. Robinson to The Department of Computer Science in partial fulfillment of the requirements for the degree of Bachelors of Arts with Honors in the subject of Computer Science and Mathematics

Harvard University Cambridge, Massachusetts May 2023
Fair School Allocation

Abstract

A well-functioning public education system is essential to a well-functioning society. However, deciding which child gets enrolled in which school is a difficult task. Specifically, with complex geographies and unequal access to high-quality schools as well as housing equity issues, it is not always clear how best to assign students to schools such that different demographic groups are given roughly equal opportunity. In this work we explore the feasibility of assigning students to schools in order to satisfy existing fairness constraints from the fair division literature. We provide impossibility results for the most popular fairness guarantees in existing fair division literature for this setting, namely proportionality-based systems and envy-based systems. Finally, we present a new fairness guarantee, and prove that an satisfying allocation always exists and can be found in polynomial time.
# Contents

0 Introduction  
1 Background  
2 Settings and Notation  
3 Fairness in the Indivisible Case  
4 Fairness in the Divisible Case  
5 A New Definition for Constrained Fairness  
6 Discussion  
Appendix A Some extra stuff  
References
To public schools
Acknowledgments

This work was a joint effort between Professor Ariel Procaccia, third-year computer science PhD student Jamie Tucker-Foltz, and myself.

To my thesis advisor Ariel, thank you for responding a cold email from a random sophomore who was incredibly excited not only to take your class but by the idea that there was a space in the world where people are doing the type of work he found invigorating. I am thankful for your patience, support, and presence in my life and the life of all the students you inspire. I certainly would not be on the path I am without your guidance and I hope to go forth into the world and expand the fields that we spend hours talking about.

To Jamie, thank you for your time which I know is valuable and your mentorship through this thesis, my courses, and my graduate school application season. You breathed life into this project at every turn and I look forward to the day someday soon where you are running your own program and inspiring the next generation of researchers.

To Professor Yiling Chen, thank you for helping me to realize research is a work in progress by definition. I have long been paralyzed by progress that is not perfect, but your guidance and approach to research in your courses has helped me do the most important thing - start, somewhere, anywhere. I look forward to hearing your feedback, and continuing this project as a work in progress.

Finally, to Professor David C. Parkes, who introduced me to all of these incredible people. Even when not here physically, your presence is felt throughout campus, in the halls of the SEC, on the chalkboards in sever, and certainly on the overleaf servers of the students you inspired. Words cannot describe my immense respect and appreciation I have for the Professor who opened up my world view, and somehow, somehow, gave me just barely enough confidence to think that I could be one of those smart people who “do research”. Clearly I have a long way to go to recuperate on the investment you and all of your colleagues have put into me, but I hope that I am at least on my way.

Thank you to my friends, family, and the members of Harvard Rowing for putting up with even less sleep than usual. To my thesis buddies Silvia, Kat, and Rachel for the long nights in the library. It was well worth the sacrifice.
Public schools are foundational to a functioning democracy, providing quality education for all inhabitants regardless of ability to pay or socio-economic background. Higher quality education, specifically at the K-12 level, has repeatedly been linked to improved earning potential throughout a person’s livelihood [Heckman et al., 2016] [Card & Krueger, 1992]. Moreover studies have shown that even controlling for higher earning, quality education has been shown to improve outcomes across categories including health, happiness, future earnings, or social capital [Cutler & Lleras-Muney, 2006].
However, educational quality is not constant even within states and cities. In states like Illinois, a Nevada, New Hampshire and North Carolina, reports show that school districts with a poverty rate of 30 percent receive at least 20 percent less money per pupil than districts with a 10 percent poverty rate [Porter, 2013]. Such inequalities can occur despite relatively small geographic differences. Take states like Texas where students are assigned spots in high schools based on their residential district. According to US News, Highland Park High School is ranked number 49 in Texas and 281 out of approximately 17,000 public high schools in the States, putting it in the 98th percentile. Yet, approximately 3 miles away lies Thomas Jefferson High School, ranked 11,689 in National Rankings putting it only in the 33rd percentile. While the criteria under which these rankings are established can be questioned, what is clear is that school quality can vary wildly even within a town or city. Given the immense impact that early education has on future outcomes, and the lasting impact of historical redlining that has led to minority populations being concentrated in certain residential neighborhoods, school assignment based on location of domicile is a recipe for disaster from a fairness and equality perspective.

As such, in this paper we use traditional fair division concepts and apply them to the context of assigning students to spots in public schools. In Chapter 2 we begin by introducing our model, in which we place blocks of students and schools with open spots on a graph. We connect students and schools in our graph if and only if those students can be assigned to those schools. The decision of which schools are feasible for which students can be based on real-life distance or travel time. For instance, feasible schools for a block of students would include all schools within 5 miles or a 15-minute average daily commute.

Using this setting, we then in Chapters 3 and 4 examine existing algorithms and notions of fairness from fair division including optimization algorithms and the round-robin protocol. We show why these solutions are untenable given the constraints of our problem - specifically, that we want fairness across groups of agents who have different feasible allocations, and we want to make sure
essentially all students are allocated a school.

We then expand this analysis and show in general that envy-freeness and proportional allocations need not exist, and that determining their existence in this context is a computationally difficult problem.

Finally, in Chapter 5 we provide a new definition of fairness constrained proportionality. We demonstrate that for fairness across 2 groups, a constrained proportional allocation can always be found in polynomial time.

Throughout, we address both the indivisible case, where schools have integer capacity and the divisible case, where schools can be arbitrarily divided. While the former most closely matches the real-world scenario which we model, the latter approximates contexts where school capacities and block sizes are sufficiently large as to appear divisible. As such, impossibility results for the latter provide intuition as to why even approximations of proportionality and envy-freeness are unlikely to always be attainable.
"The whole people must take upon themselves the education of the whole people, and must be willing to bear the expenses of it. There should not be a district of one mile square, without a school in it, not founded by a charitable individual, but maintained at the public expense of the people."

John Adams

While the literature on school matching is vast, the literature on school assignment remains relatively scarce. Here, we examine several auxiliary fields that are related to the work we are doing.

**School Zoning - Theoretical and Empirical**

The literature on redistricting and fair partitioning is vast, though when it comes to schooling the focus is largely on crafting fairer school zones rather than changing the school assignment system.
[?] define a Fair Partitioning algorithm which aims to minimize spatial inequality in district-level funding. Similar work by [Mawene & Bal, 2020] examined the empirical impact of the redrawing of elementary school attendance zones on the concentration of racially and economically minoritized students in a specific elementary public school. Others still have studied the empirical impact of alternative systems like vouchers and charter schools and their impact on educational disparity across protected groups [Ash, 2013]. Other theoretical work on coordinated assignment and minimizing justified envy are related to our work in spirit, but differ in that they are assuming the existence of individual preference rankings [Pathak et al., 2017, Abdulkadiroglu et al., 2017]. As such these fall into the realm of school choice, a vast literature that is distinct from the category of problem we detail here. Namely, we are focused on group outcomes rather than individual outcomes, and do not assume the existence of different preference rankings across students and schools.

Constrained Fair Division

Perhaps more relevant to our problem is theoretical work done in the field of constrained fair division. This sub-field of fair division deals with various additional caveats to the traditional problem of dividing some good amongst dome set of agents in the fairest way possible. The sub-field has been around for decades, with early constraints on the connectivity of partitions of divisible goods being an area of particular interest. Walter Stromquist of the Stromquist moving-knives procedure famously evangelized the necessity of fair division with constraints being, saying in a 1980 piece for the American Mathematical Monthly that with no constraints, "a player who hopes only for a modest interval of cake may be presented instead with a countable union of crumbs [Stromquist, 1980]."

As such, many different types of constraints have been explored throughout the decades including capacity constraints, cardinality constraints, and connectivity constraints [Shoshan et al., 2023, Suksompong, 2021]. Perhaps most relevant to our work is fair division on graphs, such as the the
work of [Bouveret et al., 2017] on fair allocation of indivisible items on graphs, and of [Horev & Halevi, 2022] on envy-free bipartite matching. However, these works remain relatively distinct from ours, as our definitions of feasibility and envy-freeness differ drastically. [Bouveret et al., 2017]’s work differs from ours in that instead of using the graph to demonstrate which items individually are feasible for each agent, they group items together by placing them on an undirected graph and adding the constraint that each agent’s share must form a connected subgraph of this graph. [Horev & Halevi, 2022] on the other hand uses a purely match-based and individual definition of envy - namely, that a vertex is envious if it is not matched but is adjacent to a vertex that was matched. This type of envy-freeness is more closely related to school choice and stable matching than it is to our setting, though we take inspiration by thinking of our allocations as matchings on a bipartite graph.
Settings and Notation

Fair division problems are traditionally modeled using a set of agents $N = \{n_1, n_2, \ldots\}$, a set of goods $G$ and a set of utility functions $U_N : G \rightarrow R$ where $U_n(g)$ for some $g \in G$ is the utility that agent $n$ gets from being allocated the bundle of goods $g$. In this case, we have a set of blocks of students $b$, and our agents consists of groups of these agents which we call demographic groups. Each block contains one or more students and can be thought of as a neighborhood or census blocks.
any small groupings of students such that for every student in block $b$ has the same feasible allocation of schools. We define a function $C_B : b \mapsto \mathbb{Z}$ that maps each block $b \in b \in b$ to some integer representing the number of students who live in that block and who need to be allocated to a school. For simplicity, in the divisible case we let every block node have unit capacity, and use multiple nodes with the same set of neighbors to represent blocks of larger sizes.

Our schools correspond to the set of goods $G$ that we are allocating to each of our blocks. We let $s$ be a set of schools and define a function $C : s \mapsto \mathbb{Z}$ that for a given school $s \in s$, returns the number of spots available in that school.

In order to represent which schools are available to which students we adopt a graphical structure. Namely, for any set of schools $s$ and set of blocks $b$, let each $s \in s$ and $b \in b$ be a vertex on a graph $G = (V, E)$. If $s$ is in the feasible allocation of $b$, meaning students in $b$ can be allocated to $s$, then we let the edge $(s, b) \in G$. We call a specific assignment of blocks to schools an allocation.

**Definition 1.** Let $b = \{b_1, b_2, ..., b_{k_1}\}$ be a set of blocks and $s = \{s_1, s_2, ..., s_{k_2}\}$ a set of schools.

We define an allocation $x(b, s)$ as an $k \times s$-dimensional array of real numbers. Specifically, we have $x(b, s)[i] \in \{1, 2, ..., k_1\} \times \{1, 2, ..., k_2\} \times \mathbb{R}$ such that $x(b, s)[i][j] = t$ indicates that block $b_i$ gets $t$ spots in school $s_j$. We can alternatively index this as $x(b, s)[b_i][s_j] = t$.

<table>
<thead>
<tr>
<th>$b_i \backslash s_j$</th>
<th>$s_1$</th>
<th>$s_2$</th>
<th>...</th>
<th>$s_{k_2}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$b_1$</td>
<td>$r_{11}$</td>
<td>$r_{12}$</td>
<td>...</td>
<td>$r_{1k_2}$</td>
</tr>
<tr>
<td>$b_2$</td>
<td>$r_{21}$</td>
<td>$r_{22}$</td>
<td>...</td>
<td>$r_{2k_2}$</td>
</tr>
<tr>
<td>$\vdots$</td>
<td>$\vdots$</td>
<td>$\vdots$</td>
<td>...</td>
<td>$\vdots$</td>
</tr>
<tr>
<td>$b_{k_1}$</td>
<td>$r_{k_11}$</td>
<td>$r_{k_12}$</td>
<td>...</td>
<td>$r_{k_1k_2}$</td>
</tr>
</tbody>
</table>

**Table 2.1:** An example of an allocation $x(b)$, meaning a table with rows labeled $b_1, b_2, ..., b_{k_1}$ and columns labeled $s_1, s_2, ..., s_{k_2}$, filled with entries $r_{i,j}$ which represent the number of students from $b_i$ allocated to $s_j$.

Here we consider the restriction in our setting to *additive agent utilities* and *homogenous preferences*. 
Figure 2.1: This is an example of a set of blocks and their feasible schools. In this case, students in block 1 can be assigned to schools 1 or 3, students in block 2 can be assigned to schools 1 or 3 and students in block 3 can only be assigned to school 1.

**Definition 2.** Let $S$ be some finite set of items. We call a utility function $U : 2^S \mapsto \mathbb{R}$ additive if for any $s_1, s_2 \subseteq S$ we have

$$U(s_1) + U(s_2) = U(s_1 \cup s_2) - U(s_1 \cap s_2).$$

Intuitively, additive utility functions means items are independent. For example, winning an auction for a new car should not make you more or less happy if you have previously won an auction for a drum set. By contrast, winning an auction for a drum set after having won an auction for drum lessons may make you happier than if you had won the drum set without winning the lessons. This is an example of super-additive valuation.

Homogeneous preferences, on the other hand, mean that for any two agents or blocks $b_i, b_j$, for any school $s$ in the feasible allocation of $b_i$ and $b_j$ we have $U_{b_i}(s) = U_{b_j}(s)$. In other words, we state that there is some agreed-upon metric by which a spot in a given school is valued. In practice, these metrics could be approximations for utility such as a student to faculty ratio, the number of advanced placement classes a school offers, or the rating of the school by a third-party company like
the US News and World Report or Niche.

As such we simply let $U$ be our universal utility function across all agents. For convenience we overload $U$ and define it over the set of schools $s$ and the set of partial allocations $a_k$ in addition to over the set of allocations.

1. $U((x(b, s))[b_i])$ for some block $b_i \in b$ represent the total utility of block $b_i$ under $x(b, s)$,

   $$\sum_{s_j \in s} x(b, s)[b_i][s_j]$$

2. $U(s)$ for some school $s \in s$ represents the utility value of a school $s$

Finally, we define formally the groups with respect to which we want to guarantee fairness. We let $D$ be our set of demographic groups and define the injective function $G_D : V \mapsto D$ as the map which takes each vertex to their assigned demographic according to $G_D$. While we will discuss traditional notions of fairness in chapter 3, traditionally each such notion is defined either with respect to individual agents or with respect to groups of agents, with the latter tending to be a stronger notion of fairness than the former. In our case, we are considering the allocation of students from different neighborhoods to schools, and we want to ensure fairness across demographic groups that may span multiple neighborhoods. As such, given two demographic groups $N$ and $M$, we define our blocks such that each block contains only students of one demographic type. If we have a set of students who all have the same feasible allocation but who are mixed between the two demographics, we separate them into individual blocks of homogeneous type who each have the same neighboring nodes.

1. For ease of notation, for some set of demographic grouping functions $D = \{d_1, d_2, \cdots d_k\}$ we can think of $d_i = \{b, b', b'', \cdots \}$ as a set of blocks which correspond to the demographic group $d_i$.

2. $U_{d_k}(x(b, s))$ for some allocation $x(b, s)$ represents the total utility of group $d_k$ under $x(b, s)$,
Figure 2.2: This is an example of a set of blocks and their feasible schools. In this case, students in block 1 and block 2 both have the same set of feasible schools they can be allocated to. However, since one set of students is part of the red demographic group and the other is part of the blue demographic group, we separate them into different blocks $b_1$ and $b_2$.

$$\sum_{b_i \in d_k, s_j \in s} \pi(b, s)[b_i][s_j]$$

3. $U_{d_k}(s) = U(s)$ for some school $s \in s$ represents the utility value of a school $s$

We examine fairness criteria across these demographic groups under the assumption that given a set of spots in schools, each demographic group assigns their own students to those schools in a utility maximizing way. One area of possible further research relaxes this assumption, and looks for allocations of school spots to groups such that we have a fairness guarantee across the groups under the assumption that individuals in the group compete greedily for the spots given to that demographic group, or where local envy within demographic groups is also considered. In our scenario we are designing allocations that would in practice be made by a central planner, which in the case of early-education in the United States is prevalent.

One objection to the model as proposed may be the inclusion of hard constraints in the first place. Namely, several major U.S. cities including New Orleans, Washington DC, Denver, Indianapolis, and San Francisco all have “city-wide” districts where students are allowed to enter lotteries for any school regardless of geographic location. The benefits of such a system are clear - namely, that given a set of demographic groups we can use existing results from fair division to guarantee all...
kinds of equality and fairness across those groups in this setting.

The issue with this approach is twofold. One, the assumption that some objective measure of school quality is sufficient as a proxy for a student’s or demographic group’s utility becomes less justifiable. Having to attend a cross-town school as part of a city-wide zoning plan, and as a result having a long commute to and from school, can negatively impact a student’s performance. A study on 120,000 New York City busing students showed not only that minority students are disproportionately more likely to experience long or very long bus rides than students at the same school but also that students with long bus rides have attendance rates that are half a percentage points lower and are two percentage points more likely to be chronically absent than bus riders on the same route with short bus rides [Cordes et al., 2022]. As such, an allocation where two demographic group have students that on average attend the same quality of school but where one group has to be bussed across the city and the other gets to walk to school remains sub-optimal from a fairness perspective.

One alternative solution to the above issue is to construct utility functions for the schools where, for a given block of students, we weigh the quality of the school by how difficult it is for a student from said block to access said school. Such a solution would allow for exact notions of fairness such as demographic parity.

One issue with this setting is that it would require a more complex process for accurately defining utility functions beyond simply assessing the quality of a school. In this setting we now have to define a utility function over all blocks instead of over all schools. Moreover, crafting accurate utility functions with appropriate decay rates as commutes grow longer and adjusted for specific students may prove difficult, as seemingly identical students may be more negatively or less impacted by long commute times - depending on whether or not their parent owns a car, for example.

*Note that demographic parity has it’s pitfalls. Namely, it is difficult to implement when agents are members of multiple different protected groups of different categories. However, for this introductory work, we don’t consider intersectional groups.
Furthermore, according to an interview with Lori Greenwood, Director of Assignment and Selective Admissions, transportation costs public school systems already put “astronomical” pressure on school budgets. Our setting places a limit on the possible complexity of any such busing system, and allows a central planner to address any number of the above issues by utilizing the idea of hard constraints. For all of these reasons, we address this novel setting with hard constraints on feasible school allocations, trading lower complexity of determining accurate models of real-world utility for weaker notions of guaranteed fairness.
In order to address the problem of fairly assigning students to schools we first apply traditional fairness solutions from the fair division literature. We first address the *indivisible* case, where schools are modeled as having integer capacity. A good is *indivisible* if it cannot be separated or split amongst agents. For example, the utility from a concert ticket cannot be split amongst multiple agents - an agent either reaps the whole benefit of the experience, or gets nothing. This is in contrast
to divisible goods which can be divided with the utility functions being defined on any arbitrary subsection of the aforementioned good. The quintessential example of a divisible good is a cake, where each agent has a utility function that is defined continuously over the entirety of the cake which can be split into arbitrarily many pieces.

We begin by discussing the indivisible case, where we model schools as having an integer number of spots available. We consider two types of algorithms for fair allocation of indivisible goods - round based methodologies, namely the round-robin protocol and envy-graph procedure, and optimization-based algorithms, namely envy minimization subject to our given constraints.

Specifically, we first show that the round-based methodologies which are use in the general fair division literature fail in our case. We then show that the alternative, an optimization-based algorithm, is computationally difficult and that in general, Prop, Prop-1, EF, and EF1 allocations need not always exist in our setting.

3.0.1 Background

Part of the motivation for developing relaxations of envy-freeness like EF1 comes from the fact that in general, it is difficult to guarantee that an envy-free allocation exists in the case of indivisible goods. In this section, we show that there exists graphs $G$ and demographic assignments $D$ such that in the indivisible case there does not exist an envy-free allocation - something common in the existing fair-division - and that checking if such an allocation exists is in general a computationally difficult problem. We then show in contrast to traditional settings that in school assignment, an envy-free up to one good allocation need not exist. We provide similar results for proportionality and proportionality up to one good, a related notion of fairness.

We begin by formally defining envy-freeness and proportionality across groups.

**Definition 3.** Let $D = \{d_1, d_2, \ldots d_n\}$ be a set of demographic groups, $b$ be a set of blocks and $s$
be a set of schools. Given an allocation $x(s,b)$ which we shorten to $x$, let $x[d_i]$ and $x[d_j]$ be the spots in schools cumulatively available to the blocks that compose $d_i$ and $d_j$ respectively. We say $d_i = \{b_{i1}, b_{i2}, \ldots\}$ envies $d_j = \{b_{j1}, b_{j2}, \ldots\}$ if under optimal assignment of students to available spots, $U_{d_i}(x[d_j]) > U_{d_i}(x[d_i])$.

**Definition 4.** Given a set of demographic groups $D = \{d_1, d_2, \ldots, d_n\}$ an allocation $x$ is called **envy-free** if there does not exist a pair of groups $d_i, d_j \in D$ such that $d_i$ envies $d_j$.

**Definition 5.** Let $D = \{d_1, d_2, \ldots, d_n\}$ be a set of demographic groups, $b$ be a set of blocks and $s$ be a set of schools. Given an allocation $x$ let $x[d_i]$ be the spots in schools $s$ available to $|d_i| = k$ and let $x_{opt} = \arg\max_x(U_{d_i}(x))$ be an allocation where $d_i$ is assigned the $k$ spots that maximize its cumulative utility. We say that an allocation $x$ satisfies proportionality if for all $d_i \in D$, we have $U_{d_i}(x) \geq \frac{1}{n}U_{d_i}(x_{opt})$.

### 3.0.2 Round-Robin and Envy-Graph

In traditional fair division settings, the Round-Robin protocol is a simple and effective algorithm that guarantees an allocation that is envy free up to one good - a relaxation of the general notion of envy freeness - when all agents have positive valuations for all goods. Consider a scenario with $n$ agents with additive valuations and $m$ goods. The Round-Robin protocol functions as follows:

1. Step 1: Randomly choose a permutation of the numbers $1 \ldots n$ and use that as an ordering for agents

2. Step i: Have agents choose items in accordance with this ordering until no items are left.

Intuitively, in the classical scenario where all items are available to all agents, an agent $Alice$ an only ever envy some other agent $Bob$ for the first item $Bob$ chooses. As soon as everyone chooses an item, then the second item that $Bob$ chooses must be worth less to $Alice$ than the first item that $Alice$
chose, since otherwise Alice would have chosen this different item. The same holds for the third item Bob chooses as compared to Alice’s second item and so on. As such, we end up with an allocation that up to the presence of a single good in an agent’s bundle, is envy free. For a full proof see A.

In our scenario, envy freeness up to one item translates to envy freeness up to one spot in a school.

**Definition 6.** We say an allocation \( x \) is envy free up to one good if for all pairs of groups \( d_i, d_j \) under optimal assignment of students to available spots, \( \min_{s_k \in x[d_j]} (U_d_i(x[d_j] \setminus s_k)) \leq U_d_i(x[d_i]) \) where the left hand side represents the utility of the allocation of \( d_j \) minus the value of a single spot in the highest utility school.

Round-robin takes a greedy approach to fairness and is widely used on operations scheduling and computing networking. This begs the question of if we can apply Round-Robin directly to our setting. In general, greedy approaches to problems like round-robin work well when problems have little structure - in our setting, when most goods are available to most people. In these cases, while a greedy approach may not find an allocation that gives every student a seat, by virtue of each group choosing what is best for them and every group wanting their students to receive a set, most students at least end up with a space even if it's not a desirable one.

We now show that Round-Robin fails to return a reasonable allocation for the school assignment problem when the graph is highly structured leading to a limited number of allocations that are feasible in that they assign most or all students to schools. In these cases, there is no guarantee that round robin will settle on one such reasonable allocation.
Figure 3.1: In this example, the red blocks only have access to good schools, whereas the blue blocks have access to both a good school and a bad school. We call each of the four sections of this graph a snowman. See appendix for explanation.

Feasibility of Round Robin

Suppose we have eight schools with unit capacity, and eight blocks with one student each split into two groups, the red group and the blue group. Each block’s feasible allocations is depicted in figure 4.3.

Running the Round-Robin protocol on this graph means alternating between the red group and the blue group, with each group choosing one spot for one of its blocks in each round. We assume each group is individually rational and aware of the graphical structure, meaning they may not always choose the school with the highest utility if it is uncontested, and may instead choose to secure a spot in a more competitive school.

In this case, the schools are of binary type - either a good school with utility 10 or a bad school with utility 0. Both groups are going to want to grab as many spots as possible in the good schools -
the red blocks because they are the only schools available, and the blue blocks because they are better than going to one of the bad schools.

As a result, the blue group will always choose one of the utility 10 schools if they are available. Even if the protocol randomly chooses red to go first, then, the blue group will end up taking two out of the four good schools, leaving two red students without a spot in any school. Assuming red goes first in the round robin, 3.2 shows one possible outcome of the Round-Robin protocol. The bidirectional arrows represent a group choosing to accept a spot for that block in that school.

Observe that given 4.3, there does exist an allocation of students to schools such that all students are allocated - namely, all the red students are assigned value 10 schools and all blue students are assigned value 0 schools. However, Round-Robin returns a result that leaves \( \frac{2}{5} = \frac{1}{4} \) of students unallocated. While the one feasible allocation results in a large gap in utilities, such an allocation is certainly superior to an allocation where a large portion of students are left with no school at all. We use this same argument to show that in the worst case, the number of students that round-robin leaves unallocated is linear in the number of overall students.

**Theorem 3.0.1.** There exists graphs \( G \), demographic sets \( D' \) and mappings \( G' \) such that the Round Robin procedure leaves \( \Omega(n) \) students unallocated.
In this example, the red blocks only have access to good schools, whereas the blue blocks have access to both a good school and a bad school.

Proof. Suppose we have $2n$ blocks each with one student, and $D = \{d_1, \ldots, d_{n+1}\}$ demographic groups. We have $n$ blocks of group type $d_1$, represented in figure 3.3 as the red grup. The remaining $n$ blocks each have their own type $d_2 \cdots d_{n+1}$. Finally, we have $|s| = 2n$ schools with unit capacity such that $u(s_1) = u(s_3) = u(s_5) = \cdots = u(s_{2n-1}) = a$ and $u(s_2) = u(s_4) = \cdots = u(s_{2n}) = b$ for $a > b$, $a, b \in \mathbb{R}$. Note that the $n$ good schools have odd index and the $n$ bad schools have even index.

Each block $b_i \in d_1$ has a feasible allocation of $s_i$, while for each $d_j, j \neq i$ the singular block belonging to $d_j$ has feasible allocation $s_j$ and $s_{j-1}$. For a depiction, see 3.3.

We first show that Round-Robin terminates in exactly two rounds. Suppose $d_1$ is assigned the $w$th pick in the Round-Robin cycle. Since $w \leq n$, we know there must exist at least one odd-indexed school with open capacity for $d_1$ to chose, which it must do so as those are the only schools available to $d_1$. Call the school it chooses $s_{k_{d_1}}$. Now, observe that each $d_j \neq 1$ prefers the odd-indexed schools to the even-indexed schools, meaning that if given the chance they will chose a spot in the odd-indexed schools. Since $d_1$ only gets one pick in the first round, after the first round,
we have that each of the $n$ demographic groups $d_i$ will have allocated exactly one student to an odd-indexed school leaving all odd-indexed schools full. In the second round, the only block with a feasible allocation and without a spot in a school will be the block $b_{k_{d_1}+1}$. It will only have one school available, $s_{k_{d_1}+1}$ and will thus chose it and the protocol will terminate. Now, observe that the above leaves $n - 1$ students from $d_1$ unallocated or in other words $\frac{n-1}{2n}$ fraction of all students. We have $\lim_{n \to \infty} \frac{n-1}{2n} = \frac{1}{2}$ meaning for large $n$, up to half of students may remain without schools.

Since we are treating groups of blocks as agents, it may be more reasonable to weigh the number of picks each group gets in each round by the number of students said group has. Even still, in the worst case $d_1$ gets the $n$ last picks in the first round, resulting in $\frac{n-1}{2n}$ being left unallocated. The fundamental issue is that the structure of the graph limits the number of reasonable allocations that exist, and Round-Robin cannot optimize over this subset of allocations. The envy-graph procedure suffers from this same curse of locality and the proof follows the same reasoning as above.

**Feasibility of Envy-Graph**

**Definition 7.** An Envy-Graph $G = (V, E)$ is a graph constructed as follows. For every agent $n$, let $v_n \in V$ be a vertex in a graph $G$. If given the current allocation $x$, agent $i$ envies agent $j$ then we include the edge $(v_i, v_j)$ in our envy-graph.

Using this definition we have the following procedure adopted to our setting from [Lipton et al., 2004].

**Definition 8.** The Envy-Graph Procedure is defined as follows. First decide on some ordering of the items. Then repeat the following set of steps:

1. **Step 1:** Find and eliminate all the directed cycles from the envy-graph.

2. **Step 2:** Give the next school to a group that no one envies. If that group has already assigned it’s students then it can choose to swap out the new good or keep it’s existing allocation.
The idea here is that if a cycle exists in the envy graph one can eliminate the cycle by trading goods backwards along said cycle, providing a pareto improvement. This then guarantees that in each iteration there is an agent that no one envies. As a result, since an item is always given to an unenvied agent, the envy of all other agents after that item is given to that unenvied agent is at most the value of that singular item, similarly to Round-Robin. (isaac) Lipton et. al use this algorithm to prove that in general there always exists an EF1 allocation, or more specifically in their words an allocation where the maximum difference in utilities between any two players is at most the maximal marginal utility of an additional good across all goods and all players.

However, when restricted to allocations that leave only some constant number of students unallocated, this argument no longer holds. Running his procedure on the above example, we see that once the majority group (the red group) gets allocated a single good school it becomes an envied group since there is some $d_j \neq 1$ that no longer has access to a good school and must thus send their one student to a bad school. As a result, the majority group will not be assigned another school until all other groups are assigned a school since groups without any school assignments will be unenvied. Thus we end up with the same bound linear lower bound on the number of students the algorithm can leave unallocated.

**Corollary 3.0.1.1.** There exists graphs $G$, a demographic set $D'$ and a mapping $G'_D$ such that as the number of students $n \to \infty$ the Envy-Graph procedure leaves $\frac{n}{2}$ students unallocated.

As such, it is clear that these existing algorithms do not always perform well in our setting. As a result, we now turn to the optimization problem where we can enforce a solution space in which every student is allocated.
3.0.3 Optimization Based Algorithm

In our setting, the cost of an imperfect matching where some blocks are left unallocated is extremely high. As such, we now explore the idea of envy minimization. [Lipton et al., 2004] provide two such objectives. The first is aiming to minimize envy directly, that is to say for two demographic groups $d_i$ and $d_j$, minimize the quantity $\max(0, U_{d_i}(x[d_i]) - U_{d_i}(x[d_j]))$. The second, which they find to be a more suitable objective function is the envy-ratio.

**Definition 9.** Given an allocation $x$ and two demographic groups $d_i$ and $d_j$, we define the envy-ratio to be

$$\text{EnvyRatio}(x, d_i, d_j) := \max(1, \frac{U_{d_i}(x[d_j])}{U_{d_i}(x[d_i])})$$

meaning the ratio increases as $i$ envies $j$ more and is 1 if $i$ does not envy $j$.

For our optimization, we also chose to use the envy-ratio as a multiplicative factor is easier to compare across situations and geographies *. As such we formally define the envy-minimization problem as follows.

**Definition 10.** Given two demographic groups $d_i$ and $d_j$, the envy-minimization problem seeks the feasible allocation such that

$$\max(\text{EnvyRatio}(x, d_i, d_j), \text{EnvyRatio}(x, d_j, d_i))$$

is minimized.

We now show that the problem of envy-minimization in this scenario is NP-Hard. We begin by showing that envy-free allocations necessarily have minimal envy.

**Lemma 3.0.1.1.** The envy-minimization problem is lower bounded by 1.

**Proof.** We have for demographic groups $d_i$ and $d_j$ that the envy-minimization problem returns an $x$

*We believe the same hardness result with a virtually identical similar argument holds for the difference minimization case
such that $x$ minimizes

$$\max(\text{EnvyRatio}(x, d_i, d_j), \text{EnvyRatio}(x, d_j, d_i))$$

$$= \max\left(\max\left(1, \frac{U_{d_i}(x[d_j])}{U_{d_i}(x[d_i])}\right), \max\left(1, \frac{U_{d_j}(x[d_i])}{U_{d_j}(x[d_j])}\right)\right)$$

$$\geq \max(1, 1)$$

$$\geq 1$$

\[ \square \]

**Lemma 3.0.1.2.** Given two demographic groups $d_i$ and $d_j$, an allocation $x$ is envy-free if and only if

$$\max(\text{EnvyRatio}(x, d_i, d_j), \text{EnvyRatio}(x, d_j, d_i)) = 1$$

**Proof.** First, suppose $x$ is an envy-free allocation. By definition we say $d_i$ envies $d_j$ if $U_{d_i}(x[d_j]) > U_{d_i}(x[d_i])$. Thus, if $d_i$ does not envy $d_j$ and $d_j$ does not envy $d_i$ we have $U_{d_i}(x[d_j]) \leq U_{d_i}(x[d_i])$ and $U_{d_j}(x[d_i]) \leq U_{d_j}(x[d_j])$. This means that

$$\frac{U_{d_i}(x[d_j])}{U_{d_i}(x[d_i])} \leq 1$$

$$\frac{U_{d_j}(x[d_i])}{U_{d_j}(x[d_j])} \leq 1$$

This gives us that

$$\max(\text{EnvyRatio}(x, d_i, d_j), \text{EnvyRatio}(x, d_j, d_i))$$

$$= \max\left(\max\left(1, \frac{U_{d_i}(x[d_j])}{U_{d_i}(x[d_i])}\right), \max\left(1, \frac{U_{d_j}(x[d_i])}{U_{d_j}(x[d_j])}\right)\right)$$

$$= \max(1, 1)$$
Second, suppose the envy-minization problem returns an allocation $x$ such that

$$\max(\text{EnvyRatio}(x, d_i, d_j), \text{EnvyRatio}(x, d_j, d_i)) = 1$$

This gives us that

$$\text{EnvyRatio}(x, d_i, d_j) \leq 1$$

$$\text{EnvyRatio}(x, d_j, d_i) \leq 1$$

$$\Rightarrow \max(1, \frac{U_{d_i}(x[d_j])}{U_{d_i}(x[d_i])}, \max(1, \frac{U_{d_j}(x[d_i])}{U_{d_j}(x[d_j]))} \leq 1$$

$$\Rightarrow \frac{U_{d_i}(x[d_j])}{U_{d_i}(x[d_i])}, \frac{U_{d_j}(x[d_i])}{U_{d_j}(x[d_j]))} \leq 1$$

$$\Rightarrow U_{d_i}(x[d_j]) \leq U_{d_i}(x[d_i]), U_{d_j}(x[d_i]) \leq U_{d_j}(x[d_j])$$

Which means by definition $d_i$ does not envy $d_j$ and vice versa.

We now show that the associated decision problem is NP-Hard, providing both a Turing and a Karp reduction and showing that in the case of $> 2$ demographic groups the problem is strongly NP-Hard.

**Theorem 3.0.2.** Checking if an envy-free allocation exists is NP-Hard.

**Proof.** Proof 1: In the case of 2 groups, we provide a turing reduction from the known NP-Hard problem of integer partition. Suppose we had an algorithm that given an instance of the assignment problem with 2 groups could return whether or not there exists an envy-free allocation.
Now, suppose we have a multiset of \( n \) integers \( S = \{n_1, n_2, \cdots, n_n\} \) and we want to determine whether there exists a partition of these numbers \( S_1 \) and \( S_2 \) such that \( \sum_1 S_1 = \sum_1 S_2 \). We construct an instance of the assignment problem as follows:

Let there be \( n \) schools, each one with utility corresponding to one of our integers \( s_i \in S \), 2 groups, and \( n \) blocks with one student each. All schools are feasible for all students and have capacity 1. We first assign 1 block to group \( A \) and all remaining \( n - 1 \) blocks to group \( B \) and run our algorithm \text{ExistsEnvyFree}. If our algorithm returns true then we return true, otherwise we assign 2 blocks to group \( A \) and all remaining \( n - 2 \) blocks to group \( B \) and run our algorithm again. We continue this iteration until we find an EF allocation or until we have \( |A| > n/2 \). Clearly, if \text{ExistsEnvyFree} runs in polynomial time, so does this program as \text{ExistsEnvyFree} runs at most \( [n/2] + 1 \) times. Moreover, since all schools are feasible for both groups, an envy-free allocation will exist if and only if each school has exactly equal utility, which corresponds to the scenario where we have disjoint partitions of the set \( S_1 \) and \( S_2 \) such that \( \sum_1 S_1 = \sum_1 S_2 \). Since \( S_1 \) and \( S_2 \) are symmetric, checking up to \( |A| > n/2 \) allows us to check all possible cardinalities of \( S_1 \) and \( S_2 \), completing the reduction.

\[ \square \]

\textbf{Proof}. Proof 2: In the case of an arbitrary number of demographic groups, we provide a Karp reduction from the \textit{3 − Partition} problem, which aims to separate a multiset of \( 3m, m \in \mathbb{N} \) integers into partitions of cardinality 3 such that each partition has the same sum. The reduction follows the same steps as above. Namely, we let there be \( n \) schools and \( m \) groups, each with 3 blocks of unit capacity. We again construct one school for each integer in our multiset \( S \), with each school having capacity 1 and all schools being feasible for all students. Clearly, an EF allocation exists if and only if we can assign student from each block to schools such that the sum of the utilities of those schools is equal across all \( m \) groups. This is an exact mirror of the \textit{3 − Partition} problem.

\[ \square \]

Taken together, these results show that minimizing the envy-ratio is NP-Hard.
Theorem 3.0.3. Minimizing the envy-ratio is an NP-Hard problem.

Proof. We can reduce from the envy-free decision problem directly to the envy-ratio minimization problem. Namely, if we had a polynomial time algorithm to solve the envy-ratio minimization problem, we can pass in the same inputs as the envy-free decision problem. As established above, an allocation is envy-free if and only if \( \max(\text{EnvyRatio}(x, d_i, d_j), \text{EnvyRatio}(x, d_j, d_i)) = 1 \). Since the envy-minimization problem is lower bounded by 1, an envy-free allocation exists if and only if the envy-minimization problem returns an allocation \( x \) where \( \max(\text{EnvyRatio}(x, d_i, d_j), \text{EnvyRatio}(x, d_j, d_i)) = 1 \). This completes the reduction. \( \square \)

3.0.4 Existence of Envy-Free and Proportional Allocations

We now move on to the more general case, and provide examples that show we cannot guarantee the existence of an allocation that fits with the existing definitions of fairness detailed above. Taken together, these pitfalls motivate a different approach to developing a fair algorithm which we discuss in chapter 5.

We start by showing that there exist graphs \( G \), demographic groupings \( D \) and sets of schools \( s \) such that our strongest fairness notion - envy-freeness (EF) - is not satisfiable. We then provide similar results for increasingly weak relaxations of this notion - namely envy-freeness up to one good (EF1), proportionality (Prop), and proportionality up to one good (Prop1). We then extend our up to one good results for envy-freeness and proportionality by showing that for any constant \( c \), there is some graph such that we cannot obtain envy-freeness or proportionality up to \( c \) goods.

Non-Existence of Envy-Free Allocations

First, we observe that via following proposition, it suffices to show there exists assignment problems such that no envy-free up to one good allocation exists.
Proof.

Figure 3.4: In this example, we have two groups (red and blue), 4 different blocks with unite demand, and 4 different schools with unit capacity.

Lemma 3.0.3.1. Any envy-free allocation is also an envy-free up to one good allocation.

Theorem 3.0.4. In the indivisible case, there exist graphs $G$, blocks $b$, schools $s$ and demographic groups $G$ under which no EF-1 allocation exists.

Observe figure 3.4. The only two feasible allocations are:

1. The red blocks get the value 10 schools and the blue blocks get the value 1 schools.

2. The red blocks get the value 1 schools and the blue blocks get the value 10 schools.

In both of these scenarios, even removing a single good from the bundle of items from whichever group gets assigned the good schools still yields a total utility of $10 > 1 + 1$, meaning the other group will still be envious.

The same counterexample holds for the existence of a Prop-1 allocation.

Definition 11. Let $D = \{d_1, d_2, \ldots d_n\}$ be a set of demographic groups, $b$ be a set of blocks and $s$ be a set of schools. Given an allocation $x$ let $x[d_i]$ be the spots in schools $s$ available to $|d_i| = k$ and
Let \( x_{\text{opt}} = \arg\max_x (U_{d_i}(x)) \) be an allocation where \( d_i \) is assigned the \( k \) spots that maximize its cumulative utility. We say that an allocation \( x \) satisfies proportionality up to one good if for all \( d_i \in D \), we have \( U_{d_i}(x) \geq \min_{s_k \in x_{\text{opt}}} \frac{1}{n} U_{d_i}(x_{\text{opt}}[d_i] \setminus s_k) \).

**Theorem 3.0.5.** In the indivisible case, there exist graphs \( G \), blocks \( b \), schools \( s \) and demographic groups \( G \) under which no Prop-1 allocation exists.

**Proof.** We use the same reference to figure 3.4 where there are only two feasible allocations. Without loss of generality, suppose the red group gets allocated both value 10 schools and the blue group gets allocated both value 1 schools. Then, even removing a single good from the bundle of items from the optimal allocation of red, which is both value 10 schools, we still have

\[
U_{d_{\text{red}}}(x) = 2 < \min_{s_k \in x_{\text{opt}}} \frac{1}{n} U_{d_i}(x_{\text{opt}}[d_{\text{red}}] \setminus s_k) = \frac{1}{2} (10) = 5
\]

The above example can be expanded to any number of goods and the same argument holds for proportionality. Namely, for any constant \( c \) we can construct a cycle with \( n > c \) goods in the same structure as above. By the same argument, the cycle would only have two full allocations, and we would at best be able to guarantee envy-freeness up to \( n > c \) goods. As such, we know that there is no constant \( c \) such that we can guarantee in all scenarios the existence of envy-freeness up to \( c \) goods.

**Theorem 3.0.6.** In the indivisible case, for any constant \( c \in \mathbb{N} \), there are graphs \( G \), blocks \( b \), schools \( s \) and demographic groups \( G \) under which no EF-\( c \) allocation exists.

**Theorem 3.0.7.** In the indivisible case, for any constant \( c \in \mathbb{N} \), there are graphs \( G \), blocks \( b \), schools \( s \) and demographic groups \( G \) under which no Prop-\( c \) allocation exists.
Overall it is clear that these notions from fair division literature are not directly applicable in this scenario as full-assignment of students to schools takes priority over any parity or equality of utilities. Examples like that of figure ?? where there is only one complete allocation motivate a modification of the definition of envy freeness for the context of unequal access to goods and a hard constraint on efficiency of the assignment. We propose a modified definition here based on this criteria for what we call constrained envy-freeness.

Definition 12. Given an allocation $a$ and two groups $d_1$ and $d_2$ we say that group $d_1$ envies group $d_2$ if $d_1$ can find an alternative allocation $a'$ such that:

1. $d_1$ gets assigned a subset of $d_2$'s original allocation
2. Every other group $d_i$ besides $d_1$ is fully assigned
3. $U_{d_1}(a') > U_{d_1}(a)$

The above definition incorporates the idea of hard constraints by restricting envious situations to only those allocations in which every group save the group that is envious is fully allocated. This means that one cannot be envious of an allocation where a different group has students completely left out of schools, essentially asserting that such an act provides such a large negative externally that it should not be allowed in a rational agent’s calculations. This allows us to bypass the family of examples demonstrated in 3.4 and 3.3 - examples where as the number of agents or blocks grow, the number of feasible allocations remains constant. We return to this definition after discussing the case of divisible schools.
Fairness in the Divisible Case

**EF allocations for two players** in the divisible case have long been known to exist in traditional fair division literature, with the quintessential algorithm for 2 players being the “cut and choose” protocol, independently discovered by pairs of siblings around the world [Proccacia, ]. Under this algorithm, player 1 sorts all items into 2 bundles, and then player 2 chooses which item they want with player 1 being allocated the remaining bundle. In 1992 Brams and Taylor announced
they had discovered a protocol ensuring an envy-free allocation for an arbitrary number of players in the case of divisible goods, shifting much of the conversation in the case of divisible goods to other questions like that of the computational complexity or incentive-compatibility of algorithms [Brams & Taylor, 1995].

In this chapter, we show that assignment problem, even in the divisible case it is not always possible to obtain envy-freeness or proportionality for an arbitrary number of groups. We detail a polynomial time algorithm - namely, a linear program - that aims to minimize justified envy across demographic groups, which we also provide a definition for.

### 4.0.1 Counter Example to Envy Freeness in the Divisible Case

**Theorem 4.0.1.** Suppose we treat schools $s$ as divisible goods. There are graphs $G$, blocks $b$, schools $s$ and demographic groups $G$ under which no full envy-free allocation exists.

**Proof.** Consider the graph given in Figure 4.1, where each block $b_1$, $b_2$, and $b_3$ are their own demographic group red, blue, and orange. We proceed by contradiction. Suppose there exists a full and envy-free allocation.
allocation $x_{ef}$ and let $x_{ef}[b_i]$ denote the spots assigned to $b_i$ under $x_{ef}$. First, observe that in any full allocation $x$, $b_1$ will necessarily end up with $U_{b_1}(x[b_1]) = 10$, as the utilities of all schools in its feasible allocation are 10. Thus we must have $U_{b_1}(x_{ef}[b_1]) = 10$.

Now suppose that $U_{b_2}(x_{ef}[b_2]) < 10$. Since the feasible allocation of $b_1$ is a subset of the feasible allocation of $b_2$ and we have homogeneous utility across groups, this gives us

$$U_{b_2}(x_{ef}[b_1]) = 10$$

$$U_{b_2}(x_{ef}[b_2]) < 10$$

$$\implies U_{b_2}(x_{ef}[b_2]) < U_{b_2}(x_{ef}[b_1])$$

Which violates envy-freeness. Thus we must have $U_{b_2}(x_{ef}[b_2]) \geq 10$. Now, let $n_1, n_2, n_3$ be the fractions of $b_2$ that we allocate to $s_1, s_2$ and $s_3$ respectively then we have the following system:

$$n_1 + n_2 + n_3 = 1$$

$$n_1, n_2, n_3 \geq 0$$

$$10 \cdot n_1 + 10 \cdot n_2 + n_3 \geq 10$$

The first equation comes from the fact that each block must be fully allocated. The second equation comes from the problem constraint that we cannot have negative allocations. The third equation comes from the fact that we must have $U_{b_1}(x[b_1]) = 10$. From this, we can see that $n_3$ is forced to be 0. If $n_3 > 0$ then since $n_1 + n_2 < 1$ we would have

$$10 \cdot (n_1 + n_2) + (1)n_3 = 10 \cdot (n_1 + n_2) + (1 - n_1 - n_2)(1)$$
\[= 10 * n_1 + 10 * n_2 + 1 - n_1 - n_2\]
\[= 9 * n_1 + 9 * n_2 + 1\]
\[= 9(n_1 + n_2) + 1 < 10\]

Since \(n_1 + n_2 < 1\).

Thus, we must fully allocate \(b_2\) between schools \(s_1\) and \(s_2\). That leaves space for \(b_3\) only in \(s_3\), giving us \(U_{b_3}(x_{ef}[b_3]) = 1\).

Now, let \(m_1, m_2 \geq 0\) and \(k_1, k_2 \geq 0\) be the proportion of \(b_1\) and \(b_2\) allocated to \(s_1\) and \(s_2\) respectively under \(x_{ef}\). As we established above, in any envy-free allocation, these schools must be split completely between \(b_1\) and \(b_2\). Since we enforce a full assignment for each block we have \(m_1 + m_2 = k_1 + k_2 = 1\). Moreover since in this case we have unit capacity and unit demand, we must have fully assigned schools meaning \(m_1 + k_1 = m_2 + k_2 = 1\). Observe that this thus implies \(\max(m_2, k_2) \geq 0.5\) since the two must sum to exactly 1 and both are non-negative. This means that under \(x_{ef}\) at least one of \(b_1\) or \(b_2\) must be assigned weakly more than half of \(s_2\), with half of \(s_2\) having a utility of 5. Thus we have that necessarily have that:

\[U_{b_3}(x_{ef}[b_2]_{s_2}) \geq 5 > U_{b_3}(x_{ef}[b_3]) = 1\]

\[\lor\]

\[U_{b_3}(x_{ef}[b_1]_{s_2}) \geq 5 > U_{b_3}(x_{ef}[b_3]) = 1\]

violating envy-freeness of \(x_{ef}\) and providing us with a contradiction. Thus, no envy-free allocation can exist even when the schools are divisible.

We now return to our new definition of constrained envy-freeness. By the same counter example and proof, we can see that even this restricted version of envy-freeness is not always satisfiable in
either the divisible or indivisible case.

Lemma 4.0.1.1. There are graphs $G$, blocks $b$, schools $s$ and demographic groups $G$ under which no full constrained envy-free allocation exists in either the divisible or the indivisible case.

4.0.2 Minimizing Envy

If we cannot always attain an envy-free allocation, it follows naturally to ask if we can construct an algorithm that gets us close. In the indivisible case, we used iterative algorithms that assign whole goods to players order to greedily minimize on a unitary basis the amount of envy one player has for another. In the divisible case, though, we can take advantage of the fact that we are no longer restricted to integer assignments. Such a relaxation means this problem lends itself well to optimization algorithms, and namely to linear programs. In this section, we present a linear program that minimizes the envy-ratio of the demographic groups.

We begin by defining the variables and constraints for our linear program for two demographic groups. Given an instance of the assignment problem as represented by a colored graph $G$, we introduce a variable for each edge and a constraint for each node.

1. For each edge $(b_i, s_j) \in G$, we introduce a variable $b_is_j$.

1. For every block $b_i$ we have that $\sum_{s_j \in s} b_is_j = 1$

2. For every school $s_j$ we have that $\sum_{b_i \in b} b_is_j = 1$

A satisfying solution to this program will be a set of edge weights that sum to 1 both across schools and across blocks, meaning that every block will be fully allocated and no school will be over capacity.

*Throughout this paper we assume full and tight assignments. However, in this specific case tightness is not required as in a scenario where we have more spots open than needed, the equality in the schooling constraint can simply be replaced with an inequality.
Given the above, for two demographic groups $d_1 = \{ b_1, b_2, \cdots, b_k \}$ and $d_2 = \{ b'_1, b'_2, \cdots, b'_k \}$, we want to find a suitable objective function to minimize. As detailed in 3.0.3, two standard options are to minimize the envy directly or to minimize the envy-ratio. Unfortunately, these functions do not always work as intended when we restrict which agents have access to which goods. Consider the scenario presented in 4.2.

In this situation, the red group and the blue group each have 3 high quality schools only they can be allocated to, and 4 medium to low quality schools that both groups can be assigned to. Since the other groups high quality schools are outside of each groups feasible allocation, they do not derive any utility from having spots in those schools in their allocation. As a result, no matter how we assign the 4 shared schools, we will always have

$$U_{d_{\text{red}}} (x[d_{\text{red}}]) \geq 3 \times 10$$
\[ U_{d_{red}}(x[d_{blue}]) \leq 12 \]

\[ \implies U_{d_{red}}(x[d_{red}]) > U_{d_{red}}(x[d_{blue}]) \]

The same goes for the blue group - neither red nor blue will ever envy the other. As such, allocations like that depicted on the right in 4.2 would be considered envy-free and would minimize both envy and envy-ratio as [Lipton et al., 2004] define them. Clearly, though, a fairer allocation would split the medium quality and low quality schools across the two groups.

As such we define a new objective function based on the distance to a group’s optimal allocation. First, we let \( x_{d_{1}}^{\text{best}} \) and \( x_{d_{2}}^{\text{best}} \) be the full allocations (meaning all students are allocated) such that the utility of \( d_{1} \) and \( d_{2} \) are maximized, respectively. We first note that these allocations can be found in polynomial time using an algorithm for maximum-weight matching \(^\dagger\). Now, let \( \text{Util}_{d_{1}} = \sum_{b_i \in d_{1}, s_j \in s} b_i s_j \) and \( \text{Util}_{d_{2}} = \sum_{b_i \in d_{2}, s_j \in s} b_i s_j \). We seek to minimize

\[ ||\text{Util}_{d_{1}} - U_{d_{1}}(x_{d_{1}}^{\text{best}})|| - ||\text{Util}_{d_{2}} - U_{d_{2}}(x_{d_{2}}^{\text{best}})|| \]

In other words, we want to minimize distance between how far off each group is from their optimal allocation. Intuitively, this addresses the unique scenario presented here, in that each group may have some baseline level of utility coming from schools that are always allocated to that group (either because only one group has access to that school, or because, due to the graph structure, all feasible allocations have that school being assigned completely to the same group). As a result, we can end up with scenarios like that of 4.2, where all allocations are technically envy-free, even when one group is allocated all of the spots in the best contested schools. As a result, minimizing the envy-ratio as it is traditionally defined would be ineffective here, as there would always be an envy-ratio of \(^\dagger\) See 5.0.2 for a complete proof.
Thus, we choose to minimize instead based off of how far off a group is from their best or worst possible utility. If we let \( x_{d_1}^{\text{worst}} \) and \( x_{d_2}^{\text{worst}} \) be the feasible allocations (meaning all students are allocated) such that the utility of \( d_1 \) and \( d_2 \) are minimized, then we can examine the quantities \( U_{d_1}(x_{d_1}^{\text{best}}) - U_{d_1}(x_{d_1}^{\text{worst}}) \) and \( U_{d_2}(x_{d_2}^{\text{best}}) - U_{d_2}(x_{d_2}^{\text{worst}}) \). Since we are under the constraint that all students are assigned, any utility that doesn’t go to \( d_1 \) must go to \( d_2 \) meaning these quantities are equal. We use this fact as a basis to evaluate allocation based on how good an allocation is for a group \( d_1 \) versus how good an allocation is for group \( d_2 \) in the context of all feasible allocations.

This results in the following linear program:

\[
\begin{align*}
\text{minimize} & \quad |U_{d_1}(x_{d_1}^{\text{best}}) - Util_{d_1} - U_{d_2}(x_{d_2}^{\text{best}}) + Util_{d_2}| \\
\text{subject to} & \quad \sum_{s_j \in s} b_{is} = 1 & i = 1, \ldots, n \\
& \quad \sum_{b_{ij} \in b} b_{is} = 1 & j = 1, \ldots, m
\end{align*}
\]

This program can be formulated to work with any linear program solver by replacing equality constraints with two sets of inequality, and by introducing a polynomial number of auxiliary variables and additional constraints to remove the absolute value from the objective function. As such, we know this can be computed in time polynomial in the size of the input.
Figure 4.3: In this example, there exists an allocation that is envy-free but not envy-minimizing. See appendix for explanation.
5.0.1 Existence of Allocation

The notion presented in this chapter was first proposed by Jamie Tucker-Foltz and is the result of collaboration with Professor Ariel Procaccia and Jamie.

We now return to the general case and provide a relaxation of proportionality for both the indivisible and divisible case that we call constrained proportionality.
Theorem 5.0.1. Suppose we have demographic groups $d_1$ and $d_2$, a set of blocks $b$ each of which are either part of $d_1$ or $d_2$, and a set of schools $s$. We represent the feasible schools for each block using a graph $G$. Let $x_{d_1}$ and $x_{d_1}'$ be the utility-maximizing and utility-minimizing feasible allocations for $d_1$ under the constraint that $d_2$ must also be fully allocated. Similarly, let $x_{d_2}$ and $x_{d_2}'$ be the utility-maximizing and utility-minimizing feasible allocations for $d_2$ under the constraint that $d_1$ must also be fully allocated. Then there exists an allocation $x_{fair}$ such that:

1. 
\[ U_{d_1}(x_{fair}[d_1]) \geq \frac{U_{d_1}(x_{d_1}[d_1]) + U_{d_1}(x_{d_1}'[d_1])}{2} - \max_i(U(s_i)) \]

2. 
\[ U_{d_2}(x_{fair}[d_2]) \geq \frac{U_{d_2}(x_{d_2}[d_2]) + U_{d_2}(x_{d_2}'[d_2])}{2} - \max_i(U(s_i)) \]

3. At most 1 student overall remains unallocated

Essentially, we have above a guarantee that for two groups we can always find an allocation that guarantees each group utility at least that of the average of their best and worst possible allocations minus the value of a single spot in the best available school. Moreover, such an allocation leaves at most a constant number student unallocated - namely, 1 student- independent of the size of the demographic groups or the number of schools. This is in-effect a relaxation of the notion of proportionality up to one good. However, we here allow slack not only in the utility of the agents but also in the number of students left unallocated up to a constant number of students.

In the following section, we prove that the above allocation always exist when looking for fairness across two demographic groups in the indivisible case, and by extension in the divisible case. Moreover we prove that a satisfying allocation can always be found in time polynomial in the number of students.
Proof. The goal of this proof is to construct an allocation from a generalized set of schools \( s \), blocks \( b \), demographic groups \( d_1 \) and \( d_2 \), and graph \( G \) representing feasible allocations. We will do so by constructing a matching graph \( G_{\text{fair}} \) whose nodes are our schools \( s \) and blocks \( b \), and whose edges \((b_i, s_j) \in G_{\text{fair}}\) represent an allocation of a student from block \( b_i \) to school \( s_j \).

First, we can reduce an instance of this problem with arbitrary sized blocks and arbitrary capacity schools using the following reduction. Suppose we have a graph \( G \) with arbitrarily sized blocks and schools. We construct a new graph \( G' \) in the following manner. For any blocks \( b_k \) with \( p \) students and schools \( s_j \) with capacity \( q \) in \( G \), we add to our new graph \( G' \):

1. \( p \) block nodes \( \{b_{k1}, b_{k2}, \ldots, b_{kp}\} \) with a single student

2. \( q \) school nodes \( \{s_{j1}, s_{j2}, \ldots, s_{jq}\} \) schools where \( U(s_{j1}) = U(s_{j2}) = \ldots = U(s_{jq}) \)

We let \( e' = (b_{jt}, s_{ji}) \in G' \) if and only if \( e = (b_k, s_j) \in G \). We have now reduced our graph to a unit capacity and unit demand case where any matching in \( G' \) will correspond to a matching in \( G \) with the same utility for \( d_1 \) and \( d_2 \) and vice versa. Moreover, the number of nodes added is polynomial bounded as a function of the original number of edges, original number of blocks, largest block size and largest school capacity. As such, we prove this for the unit demand and unit capacity case.

As previously established, any complete matching (meaning every student is assigned a school) is equivalent to a perfect matching on a bipartite graph where one side is all of the blocks and the other side is all of the schools. We let \( E[x_{d_1}] \) denote the set of edges of the original graph \( G \) included in \( x_{d_1} \) and \( E[x_{d_2}] \) denote the set of edges of the original graph \( G \) included in \( x_{d_2} \). Note that these edges include edges that assign spots to both the blocks in \( d_1 \) and the blocks in \( d_2 \) as we are dealing with complete allocations only. Moreover, note that \( E[x_{d_1}] \cap E[x_{d_2}] \) is not necessarily empty, as we can have assignments that are common between the two. For example, if a block only has one school it can be assigned to, then any complete allocation will include that edge.
Now, let $G'$ be a graph whose nodes are the blocks $b$, the schools $s$ and whose edges are the set $E[x_{d_1}] \cup E[x_{d_2}]$.

First, for any $e = (b_i, s_j) \in E[x_{d_1}] \cap E[x_{d_2}]$, we add $e$ to $G_{\text{fair}}$. Intuitively, we want to charge each edge in $G_{\text{fair}}$ to either the best or worst allocation for each of the two groups. If it is in the best allocation for both groups, then we should add it immediately. We remove such edges $e = (b_i, s_j)$ and their nodes from our graph $G'$ leaving us with the following lemma.

**Lemma 5.0.1.1.** The set of edges $E[x_{d_1}] \cup E[x_{d_2}] \setminus E[x_{d_1}] \cap E[x_{d_2}]$ results in $G'$ becoming a 2-regular graph.

**Proof.** First, note that in a perfect matching, each school node and each block node will have exactly one neighbor. Since our graph $G'$ was constructed using the edges $E[x_{d_1}] \cup E[x_{d_2}]$, from two perfect matchings on the same set of school and block nodes, each node in $G'$ has either one or two neighbors. Now, observe that the only nodes with one neighbor are the nodes $(b_i, s_j)$ such that $(b_i, s_j) \in E[x_{d_1}]$ and $(b_i, s_j) \in E[x_{d_2}]$. After removing these nodes and edges all remaining nodes must have exactly two neighbors. 

**Lemma 5.0.1.2.** The resulting graph which we call $G''$ decomposes into a disjoint union of cycles whose edges alternate between edges from $E[x_{d_1}]$ and edges from $E[x_{d_2}]$.

**Proof.** The fact that $G''$ decomposes into a disjoint union of cycles follows directly from the fact that it is a 2-regular graph. It remains to show that the cycles produced have alternating edges. Take any nodes $n_1, n_2, n_3 \in G''$ and edges $e_1 = (n_1, n_2), e_2 = (n_2, n_3) \in G''$ such that $e_1, e_2 \in E[x_{d_i}]$. We proceed using casework.

1. Suppose $n_1, n_3 \in b$ and $n_2 \in s$. Then if $e_1, e_2 \in E[x_{d_i}]$ this means that in $x_{d_i}$ both blocks $n_1$ and $n_3$ were assigned to the same school $n_2$. Since we are dealing with indivisible
goods, and we have unit demand and unit capacity, this violates the fact that $x_{d_1}$ is a perfect matching giving us a contradiction.

2. Suppose $n_1, n_3 \in s$ and $n_2 \in b$. Then if $e_1, e_2 \in E[x_{d_1}]$ this means that in $x_{d_1}$ a single block $n_2$ was assigned a space in two schools $n_1$ and $n_3$. Since we are dealing with indivisible goods, and we have unit demand and unit capacity, this violates the fact that $x_{d_1}$ is a perfect matching giving us a contradiction.

The same argument holds for showing that there does not exist nodes $n_1, n_2, n_3 \in G''$ and edges $e_1 = (n_1, n_2), e_2 = (n_2, n_3) \in G''$ such that $e_1, e_2 \in E[x_{d_2}]$.

We now use this graph $G''$ to construct an appropriate matching. Let $C_1, C_2, \ldots, C_k$ denote the disjoint alternating cycles of $G''$. We let $C_{d_1}^i$ be the perfect matching of $C_i$ that $d_1$ prefers, and $C_{d_2}^i$ be the perfect matching of $C_i$ that $d_2$ prefers. Note that this is equivalent to saying $e \in C_{d_1}^i$ if and only if $e \in C_i$ and $e \in E[x_{d_1}]$ and similarly for $C_{d_2}^i$.

Moreover we define a function for $0 < i < k + 1, j \in \{0, 1\}$ called $\text{diff}: C_{d_j}^i \mapsto \mathbb{R}$ such that:

$$\text{diff}(C^i) = U_{d_1}[C_{d_1}^i] - U_{d_1}[C_{d_2}^i]$$

This function takes in a cycle and returns the difference in utility for the group $d_1$ that results from using the edges from $E[x_{d_1}]$ as compared to the edges from $E[x_{d_2}]$. For example, if under $d_1$’s preferred matching they received a total utility of 10 from the schools in $C_i$ while under $d_2$’s preferred matching they received a total utility of 1 from the schools in $C_i$ then we would have $\text{diff}(C^i) = 10 - 1 = 9$. Note that since we are assuming every block is allocated and every school is filled this is a zero-sum game, meaning that to calculate the difference in utility for the group $d_2$ that results from using the edges from $E[x_{d_2}]$ as compared to the edges from $E[x_{d_1}]$ we can simply take the inverse $-\text{diff}(C^i)$. 

44
We then proceed according to the following algorithm.

**Algorithm 1 Fair Allocation**

1. Sort $C^1, C^2, \ldots, C^k$ by $|\text{diff}(C^i)|$ in increasing order and re-index as $C'^1, C'^2, \ldots, C'^k$
2. $\text{Sum}d_1 \leftarrow 0$
3. $\text{Sum}d_2 \leftarrow 0$
4. Allocation $G_{\text{fair}} \leftarrow \text{empty}$
5. for $C'^w$ in $C'^1, C'^2, \ldots, C'^{k-1}$:
6. \hfill if $\text{Sum}d_1 < \text{Sum}d_2$ then
7. \hfill Add $C'^w$ to $G_{\text{fair}}$
8. \hfill $\text{Sum}d_1 \leftarrow \text{Sum}d_1 + U_{d_1}[C'^w]$
9. \hfill $\text{Sum}d_2 \leftarrow \text{Sum}d_2 + U_{d_2}[C'^w]$
10. \hfill else
11. \hfill Add $C'^w$ to $G_{\text{fair}}$
12. \hfill $\text{Sum}d_1 \leftarrow \text{Sum}d_1 + U_{d_1}[C'^w]$
13. \hfill $\text{Sum}d_2 \leftarrow \text{Sum}d_2 + U_{d_2}[C'^w]$
14. Return $\text{SplitGraph}(G_{\text{fair}}, C^k, \text{Sum}d_1, \text{Sum}d_2)$

Essentially, what we are doing is greedily choosing which matching to use on each cycle in order to keep the utilities of both groups as balanced as possible up to the $k-1$th cycle. Since the cycles are sorted in increasing order of $|\text{diff}|$, we know that after the first $k-1$ cycles have been assigned, the difference between $\text{Sum}d_1$ and $\text{Sum}d_1$ is weakly less than $|\text{diff}(C^k)|$. This means that if we could arbitrarily split $U(C^k)$ between $d_1$ and $d_2$ then we could obtain exactly equal utilities across both groups. Since this is a zero sum game, this guarantee would be equivalent to guaranteeing utility greater than or equal to $\frac{U(x_1[d_1]) + U(x'_1[d_1])}{2}$ and $\frac{U(x_2[d_2]) + U(x'_2[d_2])}{2}$.

We now prove that there exists some procedure which we call $\text{SplitGraph}$ that can return an approximation to such a matching. Namely, suppose that if after assigning the first $k-1$ cycles we have $\text{Sum}d_1 - \text{Sum}d_2 = \beta$ and $\text{diff}(C^k) = \alpha$. As a result we are targeting a matching $M^{C^k}$ such that $U_{d_1}' = U_{d_1}(M^{C^k}) - U_{d_1}(M^{C^k}) = \gamma - \beta$ and $U_{d_2}' = U_{d_2}(M^{C^k}) - U_{d_2}(M^{C^k}) = \gamma$ where $2\gamma - \beta = \alpha$ for some $\gamma$. Essentially, we are looking for a matching where of the utility that is
available to swing between $d_1$ and $d_2$, which we are calling \( \text{diff}(C^k) = \alpha \), we want $d_2$ to get $\beta$ more of it. We now show that there exists some procedure \textit{SplitGraph} that returns a matching $M^{C_k}$ such that $|U'_{d_1} - (\gamma - \beta)| \leq \max_{s_i \in \mathcal{K}(s_i)}(U(s_i))$ and $|U'_{d_2} - \gamma| \leq \max_{s_i \in \mathcal{K}(s_i)}(U(s_i))$.

\textbf{Lemma 5.0.1.3.} There exists some procedure \textit{SplitGraph} that returns a matching $M^{C_k}$ such that $|U'_{d_1} - (\gamma - \beta)| \leq \max_{s_i \in \mathcal{K}(s_i)}(U(s_i))$ and $|U'_{d_2} - \gamma| \leq \max_{s_i \in \mathcal{K}(s_i)}(U(s_i))$.

\textit{Proof.} Without loss of generality, suppose that given $C^k$ we have $U'_{d_2}[M^{C_k}_{d_2}]$ and $U'_{d_2}[M^{C_k}_{d_1}]$ where $M^{C_k}_{d_2}$ is the matching of $C^k$ preferred by $d_2$ and $M^{C_k}_{d_1}$ is the matching of $C^k$ preferred by $d_1$.

First, note that starting with the matching $M^{C_k}_{d_1}$ and switching one student $b_r$ at a time from the school it is assigned to in $M^{C_k}_{d_1}$, call it $s_r$, to the school it is assigned to in $M^{C_k}_{d_2}$, $s'_r$, changes the total utility of $d_1$ and $d_2$ by at most $\max_{s_i \in \mathcal{K}(s_i)}(U(s_i))$. Since we are starting from a perfect matching, all $d_1$ students in $M^{C_k}_{d_1}$ are assigned, meaning by switching one student the utility increase or decrease to $d_1$ is at most $\max(s_i) - \min(s_i) \leq \max(s_i)$. Similarly, by switching one student we are displacing at most one $d_2$ student, leaving them without a school and thus decreasing the total utility of $d_2$ by at most $\max(s_i)$. Note also that in this first step, as mentioned above, exactly one student is left unallocated as we start with a full allocation and are displacing a maximum of 1 student.

If no student was left unallocated, then we choose another random block and repeat the above process. If there is such a student left unallocated, call that student $b_{(r + 1) \mod n}$ since we know such a node will be adjacent to our starting node $b_r$.

Now, we repeat and assign $b_{(r + 1) \mod n}$ to the school it is assigned to in $M^{C_k}_{d_1}$, namely $s_{(r + 1) \mod n}$. Note that regardless of whether or not we start with an existing unallocated student or with a random student, we still after the new assignment still have at most 1 student left unassigned. If we start with a new random student that means that there were no unassigned students in the first place and in this step we are displacing at most 1 student. If we are starting with an unassigned student $b_{(r + 1) \mod n}$, then we are again displacing at most one additional student but the original unass-
signed student $b_{(r \pm 1)} \mod n$ becomes assigned, again leaving us with at most 1 unsassigned student.

We can repeat this process until all students have been swapped from their allocations in $M_{d_1}^{C_k}$ to their allocations in $M_{d_2}^{C_k}$. At each step in the process, we are changing the utility of both $d_1$ and $d_2$ by at most $\text{MAX}(s_i)$. These act as discrete steps towards our goal utility. This means that for any $y \in \mathbb{R}, -\alpha \leq y \leq \alpha$ we can achieve an intermediary matching $M_{\text{opt}}^{C_k}$ such that $|U_{d_1}^r[M_{\text{opt}}^{C_k}] - y| \leq \text{MAX}s_i$. Since at each intermediary matching we have at most one student unallocated we have $\alpha \geq U_{d_1}^r[M_{\text{opt}}^{C_k}] + U_{d_2}^r[M_{\text{opt}}^{C_k}] \geq \alpha - \text{MAX}s_i$. Taken together means that we also have $|U_{d_2}^r[M_{\text{opt}}^{C_k}] - (y + \beta)| \leq \text{MAX}s_i$. Letting $y = \gamma - \beta$ since $\alpha \leq \gamma - \beta \leq \alpha$ thus completes the proof of the lemma.

Taken together with the rest of Algorithm 1, this lemma completes the proof of the theorem.

Namely, we are targeting an allocation $x_{\text{fair}}$ such that

$$U(x_{\text{fair}}[d_1]) = \frac{U(x_1[d_1]) + U(x'_1[d_1])}{2}$$
Figure 5.2: An example of a secondary student swap from one optimal allocation (blue) to a second optimal allocation (red). Observe that the switch still only leaves net one student without a school and one school without a student as circled in red.

\[ U(x_{\text{fair}}[d_2]) = \frac{U(x_2[d_2]) + U(x_1'[d_2])}{2} \]

Note \( U(x_1[d_1]) - U(x_1'[d_1]) = U(x_2[d_2]) - U(x_1'[d_2]) = \beta \) since in a perfect matching with homogenous utilities, any utility that \( d_1 \) looses goes to \( d_2 \). This means we are essentially trying to divide \( \beta \) utility between \( d_1 \) and \( d_2 \) evenly. Greedily assigning the schools associated with the first \( k - 1 \) cycles as done in Algorithm 1 gets us to an allocation \( x_{\text{fair'}} \) where \( \left| U(x_{\text{fair'}}[d_1]) - \frac{U(x_1[d_1]) + U(x_1'[d_2])}{2} \right| \leq \text{diff}(C^k) \) and \( \left| U(x_{\text{fair'}}[d_2]) - \frac{U(x_2[d_2]) + U(x_1'[d_2])}{2} \right| \leq \text{diff}(C^k) \). Finally, running \( \text{SplitGraph} \) divides \( C^k \) in such a way as to give us an allocation \( x_{\text{fair}} \) that satisfies the conditions given in the theorem statement - namely, they allow us to target the last bit of utility within the value of a spot in the highest utility school.
5.0.2 Runtime of Finding a Constrained Prop1 Allocation

We now show that we can compute the above in time polynomial in the number of nodes on our graph. Recall that our algorithm FairAllocation is split into three parts:

1. Step 1: Identify best and worst possible allocations for $d_1$ and $d_2$ and use them to construct a bipartite feasibility graph $G'$.

2. Step 2: Decompose $G'$ into a set of disjoint cycles $C_1 \cdots C_k$.

3. Step 3: Run FairAllocation

We begin by showing that we can identify the best and worst possible allocations for $d_1$ and $d_2$ in polynomial time.

**Lemma 5.0.1.4.** If at least one perfect matching exists, we can find an optimal perfect allocation for a given demographic group in polynomial time.

*Proof.* We begin with a bipartite graph $G$ with $|G| = 2n \cdot n$ school nodes and $n$ block nodes - which includes all feasible allocations. Without loss of generality, suppose we are optimizing for the demographic group $d_1$. We first take all $e = (b_w, s_j) \in G$ such that $b_w \in d_1$ and assign them weight $U(s_j) + (n) \max_i (U(s_i))$. For all other edges, namely the edges connected to blocks from $d_2$, we assign them weight $n \max_i (U(s_i))$. We call this new weighted graph $G_w$.

We then run any maximum-weight matching algorithm such as the Hungarian algorithm.

Correctness: To check for correctness we need to show two things. Assuming at least one perfect matching exists, we need to show that:

1. The above procedure admits a perfect matching

2. Among all perfect matchings, the above procedure admits the matching that maximizes the value $\sum_{(b, s_j) \in G, b \in d_1} w(e)$
Lemma 5.0.1.5. If a perfect matching exists in $G_w$, then any maximal-weight matching must also be a perfect matching.

Proof. Suppose towards contradiction that there was a maximal-weight matching $M$ such that for some $b_i \in b$ and $\forall s_j \in s$, we have that $e = (b_i, s_j) \not\in M$ (meaning $b_i$ is unmatched). The total sum of edge weights of this matching is upper bounded by:

\[
\sum_{e \in M} w(e) = \sum_{s_j \in M} U(s_j) + \sum_{i} \text{MAX}_i(U(s_i)) \leq (n - 1)(\text{MAX}_i(U(s_i)) + (n)\text{MAX}_i(U(s_i))) \leq (n - 1)(n + 1)\text{MAX}_i(U(s_i))) \leq (n^2 - 1)\text{MAX}_i(U(s_i)))
\]

Now, observe that in any perfect matching $M_p$, the total sum of the edge weights is lower bounded as follows:

\[
\sum_{e \in M} w(e) = \sum_{s_j \in M_p} U(s_j) + (n)\text{MAX}_i(U(s_i))
\]
\[ \geq (n)(n)\text{MAX}_i(U(s_i)) \]
\[ = n^2(\text{MAX}_i(U(s_i))) \]

Since we know a perfect matching will have exactly \( n \) edges. These two facts combined show that any maximal-weight matching must be a perfect matching if one exists in this reduction. \( \square \)

It remains to show that this procedure admits the matching that maximizes the value \( \sum_{b,s_j \in G, b \in d_1} w(e) \). This comes directly from the correctness of the maximum matching. Namely, if the above returned any perfect matching \( M_p \) and there existed some perfect matching \( M'_p \) such that \( \sum_{b,s_j \in M'_p, b \in d_1} w(e) > \sum_{b,s_j \in M_p, b \in d_1} w(e) \) then we would necessarily have that:

\[
\sum_{e \in M'_p} w(e) = \sum_{(b,s_j) \in M'_p, b \in d_1} w(e) + \sum_{(b,s_j) \in M'_p, b \in d_2} w(e)
\]
\[ = \sum_{(b,s_j) \in M_p, b \in d_1} w(e) + n(\text{MAX}_i(U(s_i))) \]
\[ > \sum_{(b,s_j) \in M_p, b \in d_1} w(e) + n(\text{MAX}_i(U(s_i))) \]
\[ = \sum_{(b,s_j) \in M_p, b \in d_1} w(e) + \sum_{(b,s_j) \in M_p, b \in d_2} w(e) + \sum_{e \in M_p} w(e) \]

Meaning that \( M_p \) was not a maximal weight matching, giving us a contradiction. Thus we know that the above process correctly returns the optimal allocation for a given demographic group.

Runtime: Finding the school with the maximal utility simply requires interating over the list of schools which requires \(|S|\), linear time in the number of nodes. Assigning weights to each edge requires iterating over all the edges which requires \( c|E| \) where \( c \) is the maximum number of neigh-
bors of any node, which is linear time in the number of nodes. Finally, finding a maximum-weight matching on a bipartite graph is known to be a strongly polynomial time problem.

Corollary 5.0.1.1. If at least one perfect matching exists, we can find a worst-case perfect allocation for a given demographic group in polynomial time.

It remains to show that steps 2 and 3 can be accomplished in polynomial time.

Lemma 5.0.1.6. We can decompose a 2-regular graph into a list of disjoint cycles in polynomial time.

Proof. This comes directly from the fact that a 2-regular graph is a composition of disjoint cycles. As such, we can simply use depth-first search which takes time $O(n)$ in this case since we have exactly $n$ nodes and $n$ edges.

Lemma 5.0.1.7. SplitGraph runs in polynomial time.

Proof. In order to prove we can perform SplitGraph in time polynomial in the number of nodes of our original graph, we provide a specific procedure. Without loss of generality suppose that $\text{Sum}d_1 < \text{Sum}d_2$.

Finally, taken together we have that:

Lemma 5.0.1.8. FairAllocation runs in polynomial time.

Proof. In the first part of our algorithm, the iterate through each cycle exactly once. Calculating the difference in utilities between the $d_1$-preferred matching and the $d_2$ requires iterating over each block node and each school node at most twice (once for each matching) and performing the constant time action of adding the utility of the school to a stored variable, giving us $O(n)$ operations.
Algorithm 2 Split Graph

Require: $\text{Sum}_{d_1}, \text{Sum}_{d_2}, G_{\text{fair}}, C^k$

1: NewMatching $\leftarrow M_{d_2}^{C^k}$
2: $e = (b_i, s_j) \leftarrow \max_{e \in E(C^k \cdot d_1)} U_{d_1}[e]$
3: NewMatching $\leftarrow \text{ReplaceEdge}(e, \text{NewMatching})$
4: $M_{d_2}^{C^k} \leftarrow M_{d_2}^{C^k} - e$
5: while $\text{Sum}_{d_1} < \text{Sum}_{d_2}$ do
6: $b \leftarrow \text{FindUnallocated}($NewMatching$)$
7: $e = (b, s_k) \leftarrow \text{ReturnEdge}(b, M_{d_1}^{C^k})$
8: NewMatching $\leftarrow \text{ReplaceEdge}(e, \text{NewMatching})$
9: $\text{Sum}_{d_1} \leftarrow \text{Sum}_{d_1} + U_{d_1}[e]$
10: $\text{Sum}_{d_2} \leftarrow \text{Sum}_{d_2} + U_{d_2}[e]$
11: Return NewMatching

For each cycle we do three constant time operations, adding whichever group’s matching to our store $G_{\text{fair}}$ and updating our sums. All of this together takes $O(n)$ operations.

It remains to show that the second part of our algorithm runs in time polynomial in the number of graph nodes.

We know this algorithm terminates as we know a priori that $\text{Sum}_{d_2} - \text{Sum}_{d_1} \leq |\text{diff}(C^k)|$.

Moreover, since at each iteration we are switching one edge from $M_{d_2}^{C^k}$ to $M_{d_1}^{C^k}$, eventually we will have the complete matching $M_{d_1}^{C^k}$. This means we will eventually reach $\text{Sum}_{d_1} + \text{diff}[C^k] \geq \text{Sum}_{d_2}$. Moreover, this will happen after at most $n$ iterations, one for each edge in the matching.

ReplaceEdge takes a matching, adds in the new edge $e$ and removes whatever edges it conflicts with, if any. This requires at worst linear time if using an adjacency list or constant time if using an adjacency matrix to store the graph.

FindUnallocated Returns an unallocated student if one exists, otherwise it returns a random student whose matching is still from $M_{d_1}^{C^k}$. Namely, on the $t^{th}$ iteration indexed starting at 1, FindUnallocated will return the node $b_{(i+t) \text{ mod } n}$. This is a constant time operation.

ReturnEdge($b$, $M$) returns the edge associated with $b$ in the matching $M$. Note that by definition
of a matching, this edge will be unique, and can be found in constant time if using an adjacency matrix.

Thus, since we are executing three polynomial time functions at most $n$ times, we have this algorithm runs in time polynomial in the number of nodes $n$. Namely, if given an already-constructed adjacency matrix this runs in $O(n)$ time.

Together, this shows that our algorithm runs in polynomial time in the number of nodes $n$. \hfill \Box
In this work, we address the problem of fairly assigning students to schools. We establish a setting where, instead of a simple utility penalty for undesirable goods, agents are limited in what goods are visible to them. We further constrain our setting by imposing a principle any student left unallocated should incur a high cost, and that any reasonable methodology should leave at most some constant number of students without schools.

Given this setting, we consider fairness across groups of these agents, and show through a series of
results and examples that combined these constraints cause the classical algorithms, methodologies, and intuitions that hold true throughout most of fair division literature to fail.

This includes specific failures like that of the Round Robin procedure, and broader truths such as the fact that in this context and despite homogenous valuations, envy-free allocations are not guaranteed to exist even in the divisible case. Many of our results center around two major concepts. The first is that due to our constraints, it is easy to construct examples where the number of feasible allocations is small even as the number of agents and items grow arbitrarily large. This paradigm allows us to construct examples where a single decision or a small number of decisions can already cause the rest of the allocation to crystallize, leaving little flexibility. The second concept is something likely faces in the real world of school allocation. Namely, that due to the fact that we want all agents to be allocated, agents who have only good options around them end up with a larger than ideal share of the total utility than would be initially desirable in a fair allocation.

These complexities motivate new derivative definitions of fairness that can be attained in this setting specifically for cases where only two groups are involved, and one of which we conjecture can be extended to groups of agents larger than 2. We take a pragmatic approach that attempts to answer not what should we guarantee in an ideal world but what can we guarantee given the current state of the issue. In the context of schooling, these results already provide a basic methodology for democratizing access to schools that lands between the naive model of residential-based assignment - the methodology used by the majority of small cities and towns in the United States - and the complex field that has become school choice. We hope to extend and strengthen these notions in future works, proving tight bounds and considering intersections of demographic groups. Future directions could include the relaxation of the assumption that groups are centrally organized to a scenario where even within a group’s assignment, students compete greedily for the best spots. Alternatively, future work could look into the problem’s structure and classify rigorously in which situations stronger notions of fairness can be attained. Such results could complement the literature.
on envy-freeness with subsidy, and could lead to algorithms or theorems that allow us to compute in which schools should we invest resources in order to turn a graph from one where no fair solutions exists to one where there does exist some fair allocation.

Broadly, the idea of envy-freeness and fairness across groups of agents who have access to different goods appears rich, and could be used in other scenarios where a large utility penalty is insufficient and a hard constraint is necessary.
A.0.1 Fairness when all schools are feasible for all students

Lemma A.0.0.1. If every school is feasible for every student, then we can always find an EF-1 allocation of students to schools for an arbitrary number of demographic groups.

Proof. We apply the round-robin protocol, which is known to generate EF-1 allocations in cases where all goods are available to all players.


[Proccacia, ] Proccacia, A. Lecture notes. *CS238*.

