JMID

Journal of Mechanism and Institution Design

CONTENTS

Volume 8, Issue 1, December 2023
ISSN: 2399-844X (Print), 2399-8458 (Online)

Editorial board

Published by
The Society for the Promotion of Mechanism and Institution Design
Editorial office, Centre for Mechanism and Institution Design
University of York, Heslington, York YO10 5DD
United Kingdom
http://www.mechanism-design.org

ISSN: 2399-844X (Print), 2399-8458 (Online), DOI: 10.22574/jmid

The founding institutional members are
University of York, UK
Alpen-Adria-Universität Klagenfurt, Austria
Southwestern University of Economics and Finance, China.

Cover & Logo Artwork @ Jasmine Yang
LaTeX Editor & Journal Design @ Paul Schweinzer (using ‘confproc’)
LaTeX Editorial Assistant @ Daniel Rehsmann

Printed in York and Klagenfurt by the Centre for Mechanism and Institution Design

A Letter from the Editors

When the World Health Organization declared the COVID-19 pandemic over in May this year, the scientific community had an opportunity to reflect on the events since the first reported outbreaks in December 2019. Within weeks, the genetic information of the new pathogen was sequenced, internationally shared, and work on novel prototype mRNA-based vaccines began at research institutes and international pharmaceutical corporations. In under a year, the developed vaccines were tested, approved, and became available: An unprecedented success in international, scientific cooperation saving millions of lives.

On the downside, the relative political stability of the last decades is dissolving: Inhumanly cruel terror, wars of aggression, radicalisation, and climate-change-triggered mass migration stand against the merely economic problems of the rich countries. Nationalism, populism, and authoritarianism are rising in various regions and countries. Many international as well as domestic mechanisms and institutions are malfunctioning, leading to severe conflicts and deep divisions—unfortunately resulting in ample research opportunities and obligations for our community.

In the coming year, the Conference on Mechanism and Institution Design 2024 will take place July 8–12 at the Corvinus University of Budapest, Hungary. We are grateful to Péter Biró for organizing the event. At the conference, we will celebrate Vincent Crawford's 75th birthday and his fundamental contributions to economic theory, game theory, and our society. The confirmed keynote speakers include Paul Milgrom (Stanford University), Roger Myerson (University of Chicago), Al Roth (Stanford University), and Eva Tardos (Cornell University). We are looking forward to the conference and celebrations!

As with each new issue, we wish to thank all associate editors, referees, and supporters for their contributions to the Journal. Everyone involved in the scientific selection and technical production of the Journal is volunteering their capabilities, time, and effort. As a result, our Journal of Mechanism and Institution Design can publish high-quality research free of charge to both authors and readers, allowing for free and open access to the public. We hope the community sharing this vision will grow and become stronger and more successful in tackling the most pressing design issues of our time.

Paul Schweinzer & Zaifu Yang, Klagenfurt & York, December 20th, 2023.

Journal of Mechanism and Institution Design
ISSN: 2399-844X (Print), 2399-8458 (Online)
DOI:10.22574/jmid.2023.12.001

STABLE AND ENVY-FREE LOTTERY ALLOCATIONS FOR AFFORDABLE HOUSING

Zhan Wang

Southwestern University of Finance and Economics and Sichuan University, China
wangzhan@swufe.edu.cn

Jinpeng Ma

Rutgers University-Camden, USA
jinpeng@camden.rutgers.edu

Hongwei Zhang

Sichuan University, China
zhanghongwei@scu.edu.cn

ABSTRACT

Affordable housing lotteries often enforce a rule preventing duplicate lottery entries that makes the model in Hylland & Zeckhauser (1979) (HZ) inapplicable. We revisit HZ and propose a new individually stable (IS) allocation that can be achieved by a Tickets algorithm and accommodate the rule. A strictly envy-free (SEF) allocation is shown to be the unique IS and Pareto-optimal allocation, the outcome of the unique strong Nash equilibrium of a congestion game, and the unique Pseudo market equilibrium allocation in HZ. The algorithm always obtains the unique SEF allocation (if any) and fixes a designed flaw of existing lotteries.

Keywords: Affordable housing, lottery allocations, Tickets algorithm

JEL Classification Numbers: H42, C78, C72, D63

1. INTRODUCTION

Affordable housing is a major source of housing for low income families, the elderly, and disabled households–not only in the United States (U.S.)

All authors declare there are no conflicts of interest. Any errors are our own.

but also in countries like China and India. It may even be the only source of housing for these households in major cities like Hong Kong, Manhattan, and Boston. As such, it is both legally[1] and ethically necessary that the allocation of these units be both fair and equitable.

Lotteries are often, but not always, used as the method of allocation. In a typical lottery for affordable housing, a developer (e.g. New York City (NYC)) advertises to the public the units for rent or sale and the eligibility requirements in income and family sizes. Eligible applicants file their applications before a deadline. Applicants are then pooled and randomly selected by a lottery. In such a practice, the Housing Authority or developer often prohibits duplicate lottery entries in order to give each eligible applicant an equal access and an equal opportunity to access affordable housing. This rule has been widely used in practice. An example attached in the supplemental materials is Sachem's Path at Cape Cod in 2015, which used a lottery to sell 36 newly built single affordable houses with three different sizes of 1 BR, 2 BR and 3 BR. An applicant may qualify for different sizes but he can choose one and only one size in his application, i.e., an applicant holds one and only one lottery ticket in one pool. This restriction provides a constraint on feasible lottery allocations.

A lottery is an allocation mechanism that assigns each profile of preferences an ex ante allocation of lottery tickets. Whether a lottery is fair or not depends on the ex ante allocation it produces for each given profile. An agent is said to envy the other if he prefers that agent's lottery tickets to his own. An ex ante allocation is thus envy-free (EF) if no agent envies any others. A stronger version of this is that an ex ante allocation is strictly EF (SEF) if every agent strictly prefers his own ticket to any different ticket held by the other. The EF notion has been extensively studied in the economic literature for economies with divisible goods (Foley, 1967; Pazner & Schmeidler, 1974; Varian, 1974; Crawford, 1977), however the closely related SEF notion plays a new and important role for the economy in this paper.

Affordable housing in the U.S.[2] has deeply discounted fixed rents or selling prices.[3] There are often far more applicants than there are units available. The

[1] As stated in the Fair Housing Act of 1968 and the Affirmatively Furthering Fair Housing Rule of 2015, see http://portal.hud.gov/hudportal/HUD?src=/program_offices/fair_housing_equal_opp/progdesc/title8.

[2] Affordable housing in the U.S. includes Section 8 Housing, Income Based Housing, Low Income Rentals, Public Housing, Subsidized Housing, Low Income Apartments and so forth. More than 1.2 million households in the U.S. are living in affordable housing.

[3] A number of papers have studied economies where rents or selling prices are flexible. Yet they

question now raised is whether or not these lotteries indeed produce an optimal and fair allocation among all the households that participate. The following summarizes our common understanding of an affordable housing lottery:

> Affordable housing lotteries exist to ensure fair and equitable distribution of housing to eligible applicants. [...]. The workings of a lottery are very simple. All qualified applicants are pooled and chosen randomly.[4]

We find, on the contrary, that the above noted restriction used in common practice causes serious issues with fairness not noted before in both theory and practice. For example, the affordable housing lotteries in NYC and the City of Chongqing in China may generate allocations that are not EF and Pareto optimal (PO), even if such an allocation exists. The same design flaw exists for the school choice lotteries that assign students of all grades to the KIPP Houston Public Charter Schools in the State of Texas. Thus, the flaw is not limited to one specific lottery. See Section II for detail. In fact, it is not that hard to understand why a rule imposed on a mechanism may affect its outcome (Crawford, 1979, 1980)

This motivates our study in this paper about how to design an ex ante fair and efficient lottery for affordable housing that can accommodate this rule and resolve the design flaw in existing lotteries. To achieve that goal, we revisit the seminal model in Hylland & Zeckhauser (1979) (HZ) by adding a restriction such that each agent only holds one lottery ticket in one lottery pool in an ex ante lottery allocation. Because an EF allocation may not exist due to the rule, we define a weaker notion of envy-freeness called individual stability.

A lottery allocation is individually stable (IS) if no agent has any incentive to give up his existing lottery ticket in exchange for a better lottery ticket while all other agents still keep theirs. The IS notion is quite natural, slightly different from the EF notion in the sense that an agent will act to change her lot if she does envy another agent and such an action will make her better off. This IS notion makes sense because fairness of a lottery is based on an ex ante allocation, in

---

are still controlled, falling into an upper and a lower bound. How to determine these controlled rents or prices and ex post allocations are the major issues addressed therein, see, e.g., Talman & Yang (2008); Zhu & Zhang (2011); Andersson & Svensson (2014); Andersson et al. (2015, 2016).

4 Housing Assistance Corporation at Cape Cod in MA. See http://www.haconcapecod.org/programs-and-services/homeownership-lottery

which eligible agents voluntarily file or withdraw their applications before a deadline. In fact, the stability thus defined follows that of a stable matching in Gale & Shapley (1962), where their noted deferred acceptance algorithm is a procedure to achieve a stable matching. Their algorithm has an optimization process that allows those who receive proposals to improving their positions, which are only finalized at the moment when there are no more new proposals to improve their positions. We use this idea in the design of our "Tickets" algorithm with a rejection-acceptance optimization process.

Our algorithm works as follows. A builder provides a finite number of pools of lotteries in a room with a single door for entrance. Each application is seen as a lottery ticket from one pool. Self-interested applicants form a queue to enter the room one at a time to pick up the best lottery ticket from one pool–the best for him at the moment he enters the market. The first agent in the queue enters the room and picks up the lottery he likes the most. Clearly, the market reaches a stability for a single agent economy. The stability in the market may be broken after a new entry enters the room to pick up the best ticket she likes. If that is the case, the Tickets algorithm lets an applicant, who has been made worse off due to the new entry, return his existing ticket, go back to the door and then reenter the market as if he is a new entry; He reenters to pick up the ticket he like the best–a ticket he likes better than his returned ticket. This chain-reaction process can continue until the stability in the room is restored. After that, the next new applicant in the queue enters the room and the process continues until there is no new agent in the queue. We show that this process will end within finite steps and the finalized lottery is stable.

The proposed Tickets algorithm is similar to a dynamic rematching process that has been used by Roth & Vate (1990) and Ma (1994) to reach a stable matching for the two-sided marriage market. The rematching process there goes from one stability to the next with one agent entering the market to make a proposal to his or her best mate of the opposite sex. Such an algorithm is called procedurally fair in Klaus & Klijn (2006). In this paper, the IS notion is justified by the Nash equilibrium (NE) of a lottery congestion game along the line of Milchtaich (1996).

A lottery allocation is (ex ante) PO if there is no other allocation that makes at least one applicant better off and no applicant worse off. A lottery allocation is (ex ante) weakly PO if there is no other allocation that makes all applicants better off. We find that an EF lottery allocation may not be PO, as in Pazner & Schmeidler (1974); Varian (1974); Crawford (1977) for divisible goods.

However, there always exists an IS allocation that is weakly PO in our economy. Moreover, an SEF allocation is the unique one that is both IS and PO.

The uniqueness of the SEF allocation is somewhat a surprise. An even more interesting result is that this SEF allocation is also the unique Pseudo market equilibrium in HZ, with the equilibrium price of a house being equal to the reciprocal of the lottery drawing probability of that house under the SEF allocation. Moreover, the unique SEF allocation is also the unique strong NE (Aumann, 1959) outcome of the lottery congestion game. The existence of an SNE is a complicated matter. Even a potential game (e.g., a Prisoner's dilemma game) may not have an SNE. The existence of an SNE for an unweighted congestion game in Milchtaich (1996) is shown by Konishi et al. (1997), who provided a different framework than that in Milchtaich. It is known, however, from Voorneveld et al. (1999) that the two classes of games in Milchtaich (1996); Konishi et al. (1997) are equivalent. Milchtaich (1996) showed that his game may not have the finite improvement property (Monderer & Shapley, 1996) and even a best-reply improvement path or dynamics may also last infinitely; He showed the existence theorem by establishing the existence of a finite best-reply improvement path that ends up with a pure NE. Konishi et al. (1997) used a different strategy by showing that any NE that is not SNE is Pareto-dominated by another NE. Our paper provides a condition under which there is an unique SNE for their games in Milchtaich and Konishi et al. Our Tickets-TTC algorithm incorporates Gale's Top Trading Cycles (TTC) (Shapley & Scarf, 1974) in order to reach a stable allocation that is immune to coalitional deviations. The Tickets-TTC algorithm always ends with an SNE outcome.

Given so many appealing properties associated with the unique SEF allocation, one may wonder whether the NYC affordable housing lottery achieves it or not. We find out that even if an economy has the SEF allocation, the NYC lottery, like many others, fails to achieve it. This causes a major concern because it implies that the NYC lottery is neither EF nor PO and thus suffers from a design flaw.

2. RELATED LITERATURE

To the best of our knowledge, there is no paper in the literature that has addressed the question of whether an affordable housing lottery is fair or not from a theoretical aspect. There is a growing interest in theory about random allocations for the housing allocation model in Hylland & Zeckhauser

(1979). They studied a housing allocation economy with a finite number of indivisible houses and unit demand and showed that there always exists an ex ante EF and Pareto-optimal (PO) random allocation when agents have the von Neumann-Morgenstern (vNM) utility over houses. By means of a pseudo market where each agent is artificially assigned an equal budget, they showed that their EF and PO random allocation can be achieved by a market mechanism at a competitive equilibrium. The equilibrium allocation is a doubly stochastic matrix so that a centralized lottery can be found to implement it with a collection of deterministic allocations via the celebrated Birkhoff-von Neumann decomposition theorem. Bogomolnaia & Moulin (2001) (BM) studied a similar economy but with ordinal preferences and provided an existence theorem of EF and ordinal PO random allocations by designing a novel probabilistic serial mechanism using an "eating" algorithm, in which all agents start to eat houses at the same time at an equal rate in the order of their first, their second, etc., ranked house (of unit supply) until an agent's total eating shares sum up to at most 1. Katta & Sethuraman (2006) extended EF and ordinal PO results in BM to the case with weak preferences and found a conflict between EF and ordinal PO properties and (weak) strategy-proofness. Budish et al. (2013) (BCKM) offered several extensions of those results in HZ and BM to economies that allow multiple unit demands and two-sided matching. In particular, BCKM provided an extension of the Birkhoff-von Neumann theorem and found an important necessary and sufficient condition for the existence of a centralized lottery to implement a random allocation with a collection of deterministic ones. A salient feature of these random allocations in the literature is the fact that they all require multiple lottery entries and a broad domain of feasible deterministic allocations for their implementation. Thus, the results in HZ, BM and BCKM that apply to a broader domain are not applicable to affordable housing lotteries studied in this paper because an identified equilibrium in these papers may use allocations that are not feasible.

The Tickets algorithm is related to the serial dictatorship in Svensson (1994) and the random serial dictatorship (RSD) in Abdulkadiroğlu & Sönmez (1998). A main difference is the dynamic optimization process in the Tickets algorithm that RSD does not have. The main issue with RSD for affordable housing is that it is not ex ante efficient and does not produce an ex ante stable allocation, known from Zhou's impossibility theorem Zhou (1990). RSD is also different from the NYC lottery and the Tickets algorithm by the fact that a random allocation under RSD is due to a random drawing of orders of applicants, not a lottery drawing as in the NYC lottery or the Tickets algorithm.

There is a surging interest in the study of ex post efficient allocations for affordable housing. For example, Talman & Yang (2008) investigated a housing market with rents that are either fixed or limited to an interval. They proposed a dynamic auction that finds an allocation of houses and its supporting rents such that the allocation is ex post PO and lies in the core Andersson et al. (2015). Andersson & Svensson (2014, 2016) investigated a more general but closely related market and demonstrated the existence of an ex post efficient allocation achieved by a strategy-proof mechanism. For an additional study on an ex post allocation mechanism for the housing market, see Andersson et al. (2016). In our lottery model, rents or selling prices are fixed by the housing authority or the developer in the advertisement and are not subject to any change. Our aim and approach in this paper are totally different from theirs in that we focus on the ex ante allocations and investigate the issue whether an affordable housing lottery is ex ante fair and PO.

The rest of the paper is organized as follows. Section II discusses the NYC lottery in detail and its design flaw. We also include a brief introduction to some other lotteries. Section III presents the model and definitions. Section IV presents the main results. Section V discusses the Pseudo market and the random assignments in HZ. Section VI presents a strategic form lottery congestion game and makes a connection of an IS (SEF) allocation to a pure (strong) Nash equilibrium. Section VII concludes.

3. LOTTERIES IN PRACTICE

3.1. The NYC Lottery and Its Design Flaw

We use NYC as example to show how a lottery for affordable housing has been used to create a waitlist. Our focus is on the design of a lottery as a mechanism, not on a particular waitlist generated. Once there is a development available in NYC, the developer/market agent starts to advertise the availability of the development and the eligibility requirements of household income and family sizes. Rents for these apartments are fixed and heavily discounted. Discounted rents create excess demand. In NYC, there are about 1,000 applicants for every 1 available affordable rental apartment. A developer or market agent runs the lottery for that development, independently of the other, and the developer manages the waitlist after the lottery. After a waitlist is finalized, the developer allocates the apartments according to the list, a list of log numbers ranked from the low to the high. Because many applications that are opened in a lottery do not qualify and the developer has no obligation to fill all positions in the list ex post, the applications that are open in a lottery are at least 20 times the total number of available units. Such a long waitlist is helpful in protecting the developer's interest in an ex post allocation of the units and for the future to fill vacancies. This practice is not without a cost. Applicants who get a very high log number may easily get depressed. A high log number gives an applicant a hope but it may remain a hope indefinite time. See the following blog posted by **q41apartments** on [5]:

> I was recently awarded a middle income lottery and i wanted to share my entire experience form A to Z with everyone. I constantly see someone asking the same question everyday on this forum and i hope my post helps. This is to give anyone applying an insight to the process and what to expect. Now, I am going to be up front and let everyone know that it was a very hard and painful process. I cried many nights.

To model an ex ante lottery for a development in theory, we need to assume that all applications in the lottery are eligible for the units to which they have been assigned[6] and each application in the same pool has an equal right to be drawn.

A development in NYC may offer different unit sizes, called "communities" in this paper. Each size has a quota, the maximum number of available units of the same size. For example, a development may offer 50 1-bed and 30 2-bed apartments. Thus, we say that this developer offers two types of communities with 50 of the first size and 30 of the second size. Some applicant may qualify for both sizes but applies to one single desired size for which she is qualified.

---

5 http://www.city-data.com/forum/new-york-city/2241280-how-housing-lottery-process-works-z.html

6 In NYC, it is the officer who makes a single choice of the size of a single rental unit on behalf of an applicant by the end of each application form. Here we have to assume that the choice made by the officer is indeed in the best interest of the applicant. In particular, the officer knows the number of applications or applicants that have been filed for a unit. Application forms, online or by mail, are the same across all developments. For the sake of the reader, one application form is attached in our supplemental materials. Because no duplicate applications are allowed by the developer and the officer makes a single choice about the size of a rental unit, duplicate lottery entries are prohibited in the NYC lottery.

Some applicant may be just qualified for one size only. Under such a case, we assign zero utility to an applicant for those apartments for which she does not qualify.

Applicants are encouraged to register and apply for lotteries through the NYC Housing Connect[7]. An advantage of such a centralized application system is that an applicant can apply for multiple developments at the same time but each applicant must submit one application for each lottery, because each development must conduct its lottery independently. Once an applicant is awarded an apartment, his applications for others must be voided. Thus, the applications for lotteries in NYC have been centralized while each lottery is run by a market agent or developer in a decentralized manner. There are several important rules for the application of a lottery in NYC[8]:

> R1. Submit only one entry per household. You will be disqualified if more than one entry is received per lottery for your household. If you submit an entry online, you may NOT submit an entry via mail. If you submit an entry via mail, you may NOT submit an entry online.
> R2. Entries are selected randomly through a lottery. Depending on the volume received, it may not be possible for all entries to be included in the waiting list [in an ex post allocation process].
> R3. The entry must be submitted no later than the deadline indicated for each development.
> R4. Once the application has been submitted, it cannot be withdrawn. The process must be the same for both paper and online applications, and people are not able to withdraw a paper application once it has been mailed in.

The lottery is randomly drawn against the total pool of applicants for the available units of the development. Because an applicant has been assigned to a single unit by the officer, the probability of an applicant that has been drawn is the same as if each unit size has a pool of lottery tickets of a color. Assume that all drawn applicants are eligible for the units they were drawn and

[7] `https://a806-housingconnect.nyc.gov/nyclottery/lottery.html`. Paper applications in mail are still allowed, with one application for one lottery.

[8] See the application instructions from `http://www1.nyc.gov/site/hpd/renters/housing-connect.page`. We do not provide the instructions on the eligibility requirements, which are not relevant to our study.

they will choose ex post those units that were ex ante assigned to them. Under these cases, the extra number of drawn applicants beyond the quota in practice becomes irrelavent in theory.

The reason we need to design the Tickets algorithm is that the NYC lottery as a mechanism can generate an ex ante lottery that is not stable. Consider such an example in which there are two communities, 1 and 2, each of which has one single rental unit. There are six agents, $1, 2, \cdots, 6$. Agents 1 and 2 have the same vNM utilities $(2, 3)$. That is, both agents prefer an apartment in community two to an apartment in community one under ordinal preferences. Agents 3 to 6 have the same vNM utility $(0, 5)$. Consider the order $\sigma = (1, 2, 3, 4, 5, 6)$ of applications, where applicant 1 applies the first, applicant 2 applies the second, etc. Under the NYC lottery, with self-interested and myopically sincere agents[9], applicant 1 enters the market and chooses community two. When applicant 2 files his application, he does not choose community two because his expected utility by choosing two equals 1.5, smaller than 2. Note that by R4, applicants 1 and 2 will never change their choices no matter who comes later in the order. The ex ante lottery (ticket) allocation is given under the NYC lottery by $\mathcal{L}^\sigma$ that induces a random lottery allocation or probability distribution matrix $P(\mathcal{L}^\sigma)$ (the last column is the "null" object):

$$\mathcal{L}^\sigma = \begin{pmatrix} \mathcal{L}_1^\sigma \\ \mathcal{L}_2^\sigma \\ \mathcal{L}_3^\sigma \\ \mathcal{L}_4^\sigma \\ \mathcal{L}_5^\sigma \\ \mathcal{L}_6^\sigma \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \end{pmatrix} \quad \textit{and} \quad P(\mathcal{L}^\sigma) = \begin{pmatrix} 0 & \frac{1}{5} & \frac{4}{5} \\ 1 & 0 & 0 \\ 0 & \frac{1}{5} & \frac{4}{5} \\ 0 & \frac{1}{5} & \frac{4}{5} \\ 0 & \frac{1}{5} & \frac{4}{5} \\ 0 & \frac{1}{5} & \frac{4}{5} \end{pmatrix}.$$

Thus, the total expected utility under $\mathcal{L}^\sigma$ equals $6\frac{3}{5}$. Applicant 1 has an incentive to withdraw his lottery in community 2 in exchange for a lottery in community 1 under $\mathcal{L}^\sigma$. It does not satisfy R4 voluntarily because it gives an agent the incentive to withdraw his lottery. In contrast, our Tickets algorithm

[9] This is the same assumption used in the noted deferred acceptance algorithm in Gale & Shapley (1962). We thank Vince Crawford for pointing out the importance of this assumption under the NYC lottery and the Tickets algorithm.

will provide a stable lottery allocation given by

$$\mathcal{L}' = \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad P(\mathcal{L}') = \begin{pmatrix} \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & \frac{1}{4} & \frac{3}{4} \\ 0 & \frac{1}{4} & \frac{3}{4} \\ 0 & \frac{1}{4} & \frac{3}{4} \\ 0 & \frac{1}{4} & \frac{3}{4} \end{pmatrix},$$

which yields total expected utility of 7, $\frac{2}{5}$ higher than the lottery $\mathcal{L}^\sigma$. Under different orders σ, the NYC lottery may provide different allocations. On the other hand, our Tickets algorithm yields the same lottery $\mathcal{L}'$ for any order σ, out of 720 in total, and it is not Pareto dominated by any other lottery allocation because $\mathcal{L}'$ is SEF.

The fact that the NYC lottery fails to achieve the unique SEF allocation has a serious consequence. It means that the NYC lottery is neither EF nor PO. This problem has been caused by the rule R4, which is designed for convenience in the administration of the lottery with the mail-in applications. In our design of the Tickets algorithm, we remove R4 from the NYC lottery temporarily by adding an iterated optimization process to it so that the finalized lottery satisfies both R1 and R4 voluntarily. The main idea is to allow an applicant to reject herself in her existing pool and then accept herself to the other better pool temporarily, similar to those processes in Roth & Vate (1990) and Ma (1996).

Consider the above example without agents 3 to 6. Using the noted eating algorithm in BM and BCKM, we obtain the random assignment given by $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{pmatrix}$, under which each agent's assignment (row vector) involves two lottery entries. Because x is also the random assignment in HZ for prices $(0, 2)$ when each agent is assigned a unit budget, a lottery outcome identified in HZ for cardinal utilities may also need multiple lottery entries, a clear violation of R1.

Let us go back to the above example with the six agents. Now applying the eating algorithm to it again, assuming agents eat nothing for those houses with zero utility, we get the random assignment under the BM algorithm as follows:

$$P_{BM} = \begin{pmatrix} \frac{1}{2} & \frac{1}{6} & \frac{1}{3} \\ \frac{1}{2} & \frac{1}{6} & \frac{1}{3} \\ 0 & \frac{1}{6} & \frac{5}{6} \\ 0 & \frac{1}{6} & \frac{5}{6} \\ 0 & \frac{1}{6} & \frac{5}{6} \\ 0 & \frac{1}{6} & \frac{5}{6} \end{pmatrix},$$

which can be implemented by a centralized lottery over six deterministic assignments (BCKM):

$$P_{BM} = \frac{1}{6}\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix} + \frac{1}{6}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix} + \frac{1}{6}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix} + \frac{1}{6}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix} + \frac{1}{6}\begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + \frac{1}{6}\begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}.$$

Total expected utility under the random assignment P_{BM} equals $4\frac{5}{6}$, substantially less than under the random assignments $\mathcal{L}^{\sigma}$ and $\mathcal{L}'$. More importantly, any pure assignment in the above implementation that assigns an agent to the null object will give her the incentive to unilaterally deviate from such an assigned choice. Thus, such an assignment is not IS and cannot be at a Nash equilibrium of our lottery game defined in Section VI. Our Tickets algorithm successfully resolves this issue because its outcome is always at a Nash equilibrium for any order of applicants.

3.2. The Chongqing Lottery

In the City of Chongqing in China, public rental apartments are assigned to qualified applicants with a lottery managed by the City Public Rental Housing Authority[10]. Unlike NYC, the City of Chongqing asks each applicant to choose one and only one size of rental units in her application form. The lottery is randomly drawn against each pool of applicants of the same size. Because this lottery system, as in NYC, lacks the rejection-acceptance optimization process in the Tickets algorithm, it has the same design flaw of generating an assignment that is not EF and PO.

[10] See http://www.cqgzfglj.gov.cn/flfg/201102/t20110211_164260.html, written in Chinese.

3.3. The KIPP Houston School Lottery

The KIPP Houston Public Schools are public charter schools that are free and open to students of all grades in the State of Texas. Students are assigned to the KIPP Houston Public Charter Schools by a lottery system.[11] Each student fills a single application for one and only one school. If the number of applications for a school is less than its capacity, all applicants are admitted to that school. If the number of applications is more than the capacity of a school, a lottery is drawn to assign applicants to that school. Interestingly, the lottery system leaves room for an applicant to file a duplicate application, say, by a mistake. In such a case, the new application becomes effective and the old one will be removed. Thus, the KIPP Houston School Lottery allows an applicant to make a mistake by rejecting an old application and accepting a new one, a key feature in our Tickets algorithm. The major difference is that applicants in the Kipp Houston School Lottery do not know the number of applicants that have been applying for a school when they file a new application by an "error". Thus, under the KIPP Houston School Lottery, a new application is filed based on the schools not on the schools and the number of filed applications, as in the Tickets algorithm. Thus, the Kipp Houston School Lottery has the same design defect as the NYC lottery. Nonetheless, the Kipp Houston School Lottery can be implemented similarly to our Tickets algorithm if it encourages applicants to file new applications, one at a time, to replace their old application, and the number of applicants for a school at any moment of time in the process is provided to the public.

4. MODEL AND DEFINITIONS

Because each development runs a lottery independently of the others, we model each development as an economy with indivisible objects and unit demand, as in HZ. Let $C = \{1, 2, \cdots, m\}$ denote the set of communities (i.e., objects or sizes) and community $j \in C$ has q_j number of idential apartments, where $q_j \geq 1$ is a finite integer. Thus, community j is indexed with the number $j = 1, 2, \cdots, m$. Let $q = (q_1, q_2, \cdots, q_m)$. For example, if a development provides 3 units of 1-bedroom and 5 units of 2-bedroom rental apartments, then $C = \{1, 2\}$ and $q = (3, 5)$. Let $A = \{1, 2, \cdots, n\}$ denote the set of agents

[11] See http://kipphouston.org/lottery-faqs for a detailed explanation of the lottery system.

or applicants and u_{ij} denote the vNM utility of agent $i \in A$ for renting an apartment in a community $j \in C$. Let $u_i = (u_{ij})$ be a given vNM utility vector for agent $i \in A$ such that $u_{ij} \geq 0$ for all $j \in C$ and $u_{ij} > 0$ for at least one $j \in C$. Let $u = (u_1, u_2, \cdots, u_n)'$ be the utility matrix. A lottery allocation problem for affordable housing is denoted (A, C, q, u).

The total capacity in the economy equals $\sum_{j=1}^{m} q_j$, which can vary from a few to several hundreds in NYC. We consider an economy where n is much larger than the total capacity. For examples, in 2014, there were 58,832 lottery applications for 105 affordable units in a new building in Greenpoint, Brooklyn; 98 rental apartments in the Sugar Hill development in Upper Manhattan attracted more than 48,000 applicants.[12] Note that it will not be easy to use a Pseudo market in HZ to find an equilibrium outcome for such a large scale economy.

A lottery assignment or allocation is an $n \times m$ matrix x such that $x_{ij} = 1$ if agent i is assigned an apartment in community j and $x_{ij} = 0$ otherwise. The set of feasible lottery assignments is defined by $X = \{x | \sum_{j \in C} x_{ij} = 1, \forall i \in A\}$, because each agent can only rent one apartment ex post and no agent can submit duplicate lottery entries ex ante by the rule. Thus, x_i is a vector with one j such that $x_{ij} = 1$ for $x \in X$. That is, each community has a pool of lottery tickets and an agent holds one lottery ticket in one single pool, since each qualified or eligible agent must have an equal access to affordable housing. By the equal opportunity, we assume that each lottery ticket has an equal chance to be drawn against the quota each community has. We also use x_i for the component j such that $x_{ij} = 1$ for notational convenience. This will not cause any confusion from the context. It is this feasibility constraint on X that makes the Pseudo market in HZ in the general domain and the eating algorithm in BM and BCKM inapplicable to the problem (A, C, q, u). Note that x itself is not a random allocation. There is no restriction on the number of applicants for a community $j \in C$, i.e., $0 \leq \sum_{i \in A} x_{ij} \leq n$. Let

$$n_j(x) = |\{i \in A | x_{ij} = 1\}| \tag{4.1}$$

denote the number of agents who have been assigned to community j under $x \in X$. If we consider each community to have a lottery pool, $n_j(x)$ is the number of lottery tickets in the pool of community j under allocation x. The

[12] Mireya Navarrojan: "Long Lines, and Odds, for New York's Subsidized Housing Lotteries," New York Times, Jan. 29, 2015.

probability that an agent i is allocated with an apartment in community j under a feasible lottery assignment $x \in X$ is given by $z_j(x)$, where

$$z_j(x) = min\{1, \frac{q_j}{n_j(x)}\}. \tag{4.2}$$

An agent i's expected utility $EU_i(x)$ under $x \in X$ depends on the apartment in community j assigned to i and the total number of agents who have been assigned to community j, where $EU_i(x)$ is defined by

$$EU_i(x) = \sum_{j \in C} x_{ij} u_{ij} z_j(x). \tag{4.3}$$

Equivalently, $EU_i(x) = u_{ij} z_j(x)$ for j such that $x_{ij} = 1$ or $x_i = j$.

Remark. In HZ, the Pseudo market operates with vNM utility functions. This assumption of vNM utility functions in this paper is for expository convenience. Our main results still hold even if the expected utility EUs are not the vNM utility functions. For example, the utility functions may be defined by $EU_i(x) = u_{ij} f_i(z_j(x))$ for $x_{ij} = 1$ and $EU_i(x) = 0$ otherwise, where f_i is monotonically nonincreasing or nondecreasing in $z_j(x)$ for j such that $x_{ij} = 1$.

Definition 1. A feasible allocation $x \in X$ is envy-free (EF) if for any agent i, community $l \in C$, it holds that $EU_i(x) \geq u_{il} z_l(x)$.

A feasible allocation is EF if no agent prefers others' to his own lottery assignment (Foley, 1967; Pazner & Schmeidler, 1974; Varian, 1974; Crawford, 1977). Note that others' assignments may include those apartments that are not assigned to any agent because public rental units are initially owned by the public, not by the agents or applicants.

Example 1. An EF allocation may not exist. Let $A = \{1, 2\}$, $C = \{1, 2\}$, $(q_1, q_2) = (1, 1)$, and $u_1 = u_2 = (2, 3)$. Consider all four feasible allocations in X:

$$x1 = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}, \; x2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \; x3 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \; \text{and } x4 = \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}.$$

Under allocation $x1$, agents 1 and 2 each hold a ticket in community one and no one holds a ticket in community two. Thus, allocation $x1$ induces a probability matrix $P(x1) = \begin{pmatrix} \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} \end{pmatrix}$, where the last column is for the "null" object.

Thus, the expected utility under $x1$ equals $EU(x1) = (EU_1, EU_2)(x1) = (1, 1)$. Similarly, the expected utilities under allocations $x2$ and $x3$ are given by $EU(x2) = (2, 3)$ and $EU(x3) = (3, 2)$, respectively. The expected utility under $x4$ is given by $EU(x4) = (1.5, 1.5)$. Clearly, $x2$ and $x3$ are not EF. Under $x1$, agent 1 or 2 prefers the unassigned apartment in C_2 to the assigned lottery over the apartment in C_1. Under $x4$, agent 1 or 2 prefers the unassigned apartment in C_1 to the assigned lottery for the apartment in C_2. So, no allocation in X is EF.

5. MAIN RESULTS

Example 1 has shown that an EF allocation may not exist for (A,C,q,u). Our resolution to this problem is the definition of an individually stable (IS) allocation. We provide an algorithm whose outcome is an IS allocation. However, an EF or IS allocation may not be PO. To find a PO allocation, we define a strictly envy-free (SEF) allocation and discuss a number of promising properties associated with it.

5.1. Tickets Algorithm and Stable Lottery Allocations

Let (x_{-i}, k) denote a feasible allocation in X such that only agent i's assignment j under x, i.e., $x_{ij} = 1$, is replaced by $k \in C$, $k \neq j$, and everyone else's assignment remains the same.

Definition 2. A feasible allocation $x \in X$ is an individually stable (IS) allocation if $EU_i(x) \geq EU_i((x_{-i}, k))$ for all $i \in A$ and $k \in C$.

The IS notion is quite natural, slightly different from the EF notion in the sense that an agent will act if they envy another agent, if such an action makes them better off. Thus, a lottery allocation is IS if no agent has any such incentive to move to another lottery lot. It is a notion that follows that of a stable matching in Gale and Shapley. It is also related to the notion of a Nash equilibrium, but the two are not the same because there are no strategies or games defined with the IS notion.

Theorem 1. *For any affordable housing economy (A, C, q, u), there always exists an IS lottery allocation.*

The proof is in Appendix IX.B. We prove the theorem by designing an algorithm. Now imagine that each community has a unique ticket machine j. Each machine j, $j = 1, 2, \cdots, m$, prints n identical lottery tickets. Each ticket gives its holder an equal right to be drawn to rent an apartment in that community. Different machines print different tickets so that no two machines print the same tickets, using distinctive colors or other means. We can put all the ticket machines in a room with a single entrance. Agents or applicants form an arbitrary random queue or order to enter the room to get tickets, one ticket for one agent.

Tickets Algorithm. Without loss of generality, let $(1,\ 2,\ \cdots,\ n)$ be the order of agents. That is, agent 1 is the first in the queue to enter the room to get a ticket. We now specify in steps where each agent gets his ticket. The algorithm operates similarly for other orders of agents. Note that agents are self-interested and myopically sincere, as in the deferred acceptance algorithm. The lottery allocation part of the algorithm consists of Step 1 to Step n:

Step 1. Let agent 1 enter the room first and choose a lottery ticket she likes the most. In case there are ties, agent 1 may use any tie-breaker. Let x^1 denote the allocation by the end of this step.

Step 2. Let agent 2 enter the room to choose the lottery ticket he likes the most. If agent 2 chooses a different lottery ticket from agent 1's, then the process moves to the next step. If 2's choice is the same as agent 1's, his choice reduces 1's expected utility. Thus, agent 1 may like some other lottery tickets better than what she held in Step 1. If that is the case, let agent 1 return her ticket and go out of the door, and then reenter, before any other agent in the order, to re-choose a lottery ticket mostly preferred by agent 1. The process moves to the next step. Agents 1 and 2 may use any tie-breaker in case there are ties. Let x^2 denote the allocation by the end of this step.

Step i. Let agent i enter the room to choose the lottery ticket he likes the most. The choice made by agent i reduces the expected utility for those who have been in the room and held the same lottery tickets as agent i does. Some of those agents may want to return their tickets in exchange for some other tickets they like better. If there are such agents, then let one and only one agent h among them return his lottery ticket and go out of the door, and then reenter the room, to re-choose a lottery ticket mostly preferred by agent h. Repeat this process by treating agent h as agent i, until no agent in the room wants to return his ticket. Let x^i denote the allocation by the end of this step.

Repeat **Step i** above until all agents in the order are exhausted.

This procedure generates a sequence of allocations $x^1, x^2, \cdots, x^i, \cdots, x^n$, each of which is IS with respect to the agents in the room. Note this sequence of allocations is order-dependent because different orders may have different sequences.

NYC Lottery. In the NYC lottery, each applicant enters the room to pick up a lottery ticket that she likes the most and holds her lottery ticket to the close. That is, the NYC lottery is equivalent to the serial dictatorship for lottery tickets, not for the actual apartments. Then these lottery tickets are drawn to generate a waitlist in an ex post allocation. Because different drawings create different waitlists, the NYC lottery is not quite the same as RSD.

The Tickets algorithm incorporates a rejection-acceptance optimization process into the NYC lottery. A main change comes from those steps after Step 1. We have to make the change because the NYC lottery is not stable so that R4 cannot hold voluntarily, as shown in Section II.A.

A key observation in our algorithm is that any agent in the room who rejects and accepts herself due to the new entry only does that once in order to restore the stability in the market. Thus, the process of the adjustment must be finite. It is critical to let just one agent from the order enter the market each time. If one lets two or more agents enter the market at the same time, there is no guarantee the process will converge. In fact, if two agents are allowed to exchange their lots at the same time, it is easy to construct an example where the two agents can form a cycle by moving back and forth between two communities, no convergence to stability.

Next we use an example to illustrate how the Tickets algorithm reaches a stable lottery allocation for every order of applicants and point out some general properties that will be studied in the next sections.

Example 2. Let $A = (1, 2, 3, 4),\ C = (1, 2, 3)$ and $q = (q_1, q_2, q_3) = (1, 1, 1)$.

Table 1: Tickets Algorithm for Example 2

Step \ Order	(1,2,3,4)	(1,2,4,3)	(1,3,2,4)
1	$(C_2, , ,)$	$(C_2, , ,)$	$(C_2, , ,)$
2	$(C_2, C_2, ,)$	$(C_2, C_2, ,)$	$(C_2, , C_2,)$
Agent 1 re-assigned	$(C_3, C_2, ,)$	$(C_3, C_2, ,)$	$(C_3, , C_2,)$
3	$(C_3, C_2, C_2,)$	$(C_3, C_2, , C_1)$	$(C_3, C_2, C_2,)$
4	(C_3, C_2, C_2, C_1)	(C_3, C_2, C_2, C_1)	(C_3, C_2, C_2, C_1)

The utility matrix is given by

$$\begin{pmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{pmatrix} = \begin{pmatrix} 2 & 4 & 3 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 3 & 0 & 2 \end{pmatrix}.$$

There are 24 orders in total under the Tickets algorithm. We apply the Tickets algorithm to three orders to illustrate the algorithm here. We use C_j for community j, to avoid confusion.

Consider three orders $(1,2,3,4)$, $(1,2,4,3)$ and $(1,3,2,4)$, where agent 1 enters the room first (see Table 1) and chooses C_2 because C_2 is agent 1's best choice in **Step 1**. The allocation is simply $x^1 = (C_2, , ,)$, where an empty component indicates an agent is assigned with no apartment.

Step 2. Now C_2 is the community that agents 2 and 3 prefer the most. Therefore, they choose C_2 after they enter. After agents 2 and 3 enter, their choices reduce the expected utility for agent 1, who can find a better lottery by returning the existing one. Therefore, agent 1 returns her ticket and goes out of the door and reenters to choose C_3.[13] $x^2 = (C_3, C_2, ,)$ is the allocation in Table 2 with the order $(1,2,3,4)$ because no agent in $\{1,2\}$ would like to return her ticket. Similarly, we get x^2 for the other two orders.

Step 3. Now a new agent from the order enters the room. With the order $(1,2,3,4)$, agent 3 enters the room and chooses C_2. Because no agent in $\{1,2,3\}$ would like to return her ticket after agent 3 took a ticket from C_2, the algorithm at Step 3 ends with $x^3 = (C_3, C_2, C_2,)$. Similarly, we get x^3 for the other two orders.

[13] This is the key difference between the Tickets algorithm and the serial dictatorship or the NYC lottery.

Step 4. When agent 4 enters the room under the order $(1, 2, 3, 4)$, she chooses C_1. Because no agent in the room would like to return her ticket after agent 4 took a ticket from C_1, the algorithm at Step 4 ends with $x^4 = (C_3, C_2, C_2, C_1)$. Because there are no more agents in the order, this ends the algorithm with x^4.

There are several interesting properties that should be noted in this example. It has two IS allocations:

$$x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad and \quad x^* = \begin{pmatrix} x_1^* \\ x_2^* \\ x_3^* \\ x_4^* \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}.$$

The expected utilities are given by $EU(x) = (2, 1, 1, 2)$ and $EU(x^*) = (3, 1, 1, 3)$. Thus, x^* Pareto dominates x. Under all three orders, the final allocation achieved by the Tickets algorithm is the allocation x^*, not the less efficient x. We show in Theorem 6 that this is a general result. The Tickets algorithm will always reach this more efficient allocation x^*, if any, over all other less efficient IS allocations. Moreover, this conclusion does not depend on the orders (see Appendix IX.F for the rest orders).

Under the NYC lottery with the application order $(1, 2, 3, 4)$, applicants 1, 2 and 3 get their tickets in C_2 and applicant 4 gets his from C_1. Clearly, such an allocation leaves C_3 unfilled and is not stable.

5.2. IS and Weakly PO Allocations

We start with the definition of coalitionally stable lottery allocation.

Definition 3. An allocation $x \in X$ is coalitionally stable (CS) if there is no other allocation $x' \in X$ that satisfies for any agent $i \in A' = \{a | x_a \neq x'_a\}$, inequality $EU_i(x') > EU_i(x)$ holds.

An allocation that is CS must be IS but the converse is not right. If an allocation is not coalitionally stable, then a coalition of agents may jointly withdraw their lottery tickets in exchange for other different lottery tickets so that everyone in the coalition prefers their new lottery to the existing one under the allocation.

Definition 4. A feasible allocation $x \in X$ is weakly Pareto optimal if there is no other feasible allocation $y \in X$ such that $EU_i(y) > EU_i(x)$ for all agents $i \in A$.

A feasible allocation is weakly PO if there is no other feasible allocation that gives each agent a higher expected utility. Note that an IS allocation x may not be weakly PO but every CS allocation is weakly PO.

Thus, an interesting question is whether there always exists an IS and weakly PO allocation in the economy (A, C, q, u). The answer is not obvious. We need to construct a lottery housing market to trade lottery "houses". For example, it is possible to use TTC to achieve x^* from x in Example 2 (see Theorem 5 for a general result associated with SEF). Starting with the IS allocation x, we can construct a lottery housing market in which a top trading cycle between agents 1 and 4 is formed to obtain the unique SEF x^*. What happens if an economy does not have an SEF allocation? The next result answers this question.

Since $u_{ij} > 0$ for at least one j, this implies that each agent must have a lottery ticket $j \in C$ under an IS allocation x. Given an IS allocation x, we now construct a housing market to find the strict core allocation as follows. Each lottery ticket is seen as a "house". Let the lottery allocation x be the initial endowment in the lottery housing market. Because there are m communities, there are just m different houses. Some ties may exist.[14] An agent ranks n lottery houses according to his expected utilities $u_{ij}z_j(x)$ of those lottery tickets of j type, $j = 1, 2, \cdots, m$. In case there is a tie between his initial endowed lottery house and any other lottery houses of the same kind, he always prefers his own lottery house to others' lottery houses and uses an arbitrary tie-breaker to break other ties, if any. In case there is a tie between his initial endowed lottery house and any other lottery houses that are not the same type, he prefers others' lottery houses to his initial endowed lottery house and uses an arbitrary tie-breaker to break other ties, if any. This completes our construction of a lottery housing market. Then we can apply TTC to this lottery housing market to get the final lottery allocation, denoted x^*. We can use the TTC algorithm to find the allocation that is IS and in the strict core for the lottery housing market with the initial endowment x. The algorithm that uses the Tickets algorithm first and then the TTC algorithm is called the Tickets-TTC algorithm.

Theorem 2. *For any lottery allocation problem (A,C,q,u), the allocation* x^*

[14] How to break ties in a housing market turns out to be critical for TTC, see, e.g., Saban & Sethuraman (2013) and references therein.

that is obtained from the Tickets-TTC algorithm is CS, i.e., an IS and weakly PO allocation always exists.

Proof. Assume, on the contrary, that x^* is not coalitionally stable. That is, there is an allocation, denoted x', such that there is a subset of agents $A' = \{a | x_a \neq x'_a\}$, which is clearly nonempty, such that $x_i \neq x'_i$ and the inequality $EU_i(x') > EU_i(x^*)$ holds for all $i \in A'$.

We consider two cases: (a) for all $j \in C$, $n_j(x^*) = n_j(x')$; (b) there is a community $j \in C$ such that $n_j(x') > n_j(x^*)$.

We prove Case (a) first. By the assumption $n_j(x^*) = n_j(x')$ for all $j \in C$, this implies that x' is obtained from x^* by forming some trading cycles. But the fact that all agents in A' strictly prefer x' to x^* contradicts to the use of the TTC algorithm. This completes the proof of Case (a).

Next, we prove Case (b). Because there is a community j such that $n_j(x') > n_j(x^*)$, there must be at least one agent, denoted i', satisfying 1). $i' \in A'$ and 2). $x'_{i'j} = 1$.

By the assumption $n_j(x') > n_j(x^*)$, we have

$$n_j(x') \geq n_j(x^*_{-i'}, j), \tag{5.1}$$

which implies

$$EU_{i'}((x^*_{-i'}, j)) \geq EU_{i'}(x') \tag{5.2}$$

By Theorem 1, we know that x^* is IS. Thus,

$$EU_{i'}(x^*) \geq EU_{i'}((x^*_{-i'}, j)) \tag{5.3}$$

Thus, by inequalities (5.2) and (5.3), we have

$$EU_{i'}(x^*) \geq EU_{i'}(x'), \tag{5.4}$$

which is a contradiction to the assumption agent i' is in A'. This completes the proof. □

Our next example answers the question whether all CS allocations can be obtained by the Tickets-TTC algorithm. The answer is negative. Surprisingly the Tickets-TTC algorithm is selective among all CS allocations by choosing the ones that are the most efficient.

Example 3. A CS allocation may not be obtained by the Tickets-TTC algorithm. Let an affordable housing economy be given by $A = \{1, 2, 3, 4\}$, $C = \{1, 2\}$, $u_1 = u_2 = u_3 = (12, 7)$, and $u_4 = (5, 2)$. Each community has a unit capacity, i.e., $q = (1, 1)$.

There are four CS allocations:

$$x1 = \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 1 & 0 \end{pmatrix}, x2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 1 & 0 \end{pmatrix}, x3 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \end{pmatrix} \text{ and } x4 = \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

The expected utility vectors for allocations from $x1$ to $x3$ are, respectively, $(4, 4, 7, \frac{5}{3})$, $(4, 7, 4, \frac{5}{3})$ and $(7, 4, 4, \frac{5}{3})$, with the same total expected utility of $14 + \frac{5}{3}$. The expected utility vector for allocation $x4$ is $(4, 4, 4, 2)$, with the total expected utility of 14, less efficient than others. To achieve $x4$, we need agent 4 in community C_2. But we find out that the Tickets-TTC algorithm can only achieve the three more efficient CS allocations from $x1$ to $x3$ but cannot achieve $x4$, even though $x4$ is CS.

Table 2: Outcomes of the Tickets-TTC Algorithm with All Orders

Agents′ Orders	*Agents in Community C_1*	*Agents in Community C_2*
$(1,2,3,4)$; $(1,2,4,3)$; $(1,4,2,3)$; $(3,2,1,4)$ $(3,2,4,1)$; $(3,4,2,1)$; $(4,2,3,1)$; $(4,2,1,3)$	1, 3, 4	2
$(1,3,2,4)$; $(1,3,4,2)$; $(1,4,3,2)$; $(2,3,1,4)$ $(2,4,3,1)$; $(3,2,4,1)$; $(4,3,2,1)$; $(4,3,1,1)$	1, 2, 4	3
$(2,1,3,4)$; $(2,1,4,3)$; $(2,4,1,3)$; $(3,1,2,4)$ $(3,1,4,2)$; $(3,4,1,2)$; $(4,1,3,2)$; $(4,1,2,3)$	2, 3, 4	1

Table 2 provides the outcomes of the Tickets-TTC algorithm for all possible orders. As one can see from Table 2, agent 4 is not allocated to community C_2, no matter which order has been used. It is worthwhile noting that under the RSD algorithm, agent 4 will be assigned to community C_2 as long as she is ranked the second in an order. This example shows that there is a substantial difference between RSD and the Tickets-TTC algorithm.

5.3. Strictly Envy-free Allocations

Definition 5. A feasible allocation $x \in X$ is (ex ante) Pareto optimal (PO) if there exists no other feasible allocation $x' \in X$ such that $EU_i(x') \geq EU_i(x)$ for all $i \in A$ and the strict inequality holds for at least one $i \in A$.

If a feasible allocation x is PO, then there does not exist any feasible allocation y that makes at least one agent better off without making the rest worse off. It is easy to have an example such that an EF allocation is not PO.

The following definition of a strictly envy-free (SEF) allocation is just slightly stronger than the EF allocation, i.e., inequality is strict for every agent in A and any community $l \neq j$ such that $x_{ij} = 1$ in the EF definition. This definition is equivalent to an EF allocation in an economy with strict ordinal preferences. As in the housing market in Shapley & Scarf (1974) where strict preferences are quite important to the model, the strict inequality in the definition of an SEF allocation has an enormous consequence for the problem (A, C, q, u). Recall our notation of x_i that also denotes the community j such that $x_{ij} = 1$.

Definition 6. A feasible allocation $x \in X$ is SEF if $EU_i(x) > u_{il} z_l(x)$ for all agents $i \in A$ and communities $l \in C$ such that $l \neq x_i$.

Theorem 3. *Every SEF allocation $x^* \in X$ is Pareto optimal.*

Proof. Let x^* be an SEF allocation. Assume, on the contrary, that x^* is not Pareto optimal. Thus, there must be a feasible allocation $x' \in X$ such that x' Pareto dominats x^*, denoted by $x'Px^*$. We consider two cases: (a) for all $j \in C$, $n_j(x^*) = n_j(x')$; and (b) there is a community j such that $n_j(x') > n_j(x^*)$.

We prove Case (a) first. Because x^* is SEF, we have $EU_i(x^*) > u_{ij} z_j(x^*)$ for all i such that $j \neq x_i^*$ and $EU_i(x^*) \geq u_{ij} z_j(x^*)$ for all i and j. By the assumption that $n_j(x^*) = n_j(x')$ for all j, we have that for all j, $z_j(x^*) = z_j(x')$. Thus, we obtain

$$EU_i(x') = u_{ix_i'} z_{x_i'}(x') = u_{ix_i'} z_{x_i'}(x^*) \leq EU_i(x^*)$$

for all i, which is a contradiction to the assumption $x'Px^*$. This completes the proof of Case (a).

Next we prove Case (b): $\exists j$ such that $n_j(x') > n_j(x^*)$. This implies that there must be at least one agent, denoted i^*, such that $x'_{i^*} = j$ and $x^*_{i^*} \neq j$. Moreover, the fact that $n_j(x') > n_j(x^*)$ implies $z_j(x') \leq z_j(x^*)$. Because x^*

is SEF, we have that for all i such that $j \neq x_i^*$, $EU_i(x^*) > u_{ij}z_j(x^*)$. Thus, it follows $EU_{i^*}(x^*) > u_{i^*j}z_j(x^*)$ and then

$$EU_{i^*}(x^*) > u_{i^*j}z_j(x^*) \geq u_{i^*j}z_j(x') = EU_{i^*}(x'),$$

which is a contradiction to the assumption $x'Px^*$. This completes the proof of Case (b) and the proof of the theorem. □

It is clear that an SEF allocation is also IS. The following result shows that an SEF allocation is the unique allocation that is both IS and PO. The uniqueness is somehow a surprise, given the fact that there are so many PO allocations. The proof of this result is given in Appendix IX.C.

Theorem 4. *An SEF allocation $x^* \in X$ is the unique allocation that is both individually stable and Pareto optimal.*

Example 4. An IS allocation may not be SEF or PO. Consider Example 1 with $u_1 = (2, 3)$ and $u_2 = (3, 2)$. Consider two allocations $x' = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ and $x^* = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$. The expected utilities of the two allocations are given by $EU(x') = (2, 2)$ and $EU(x^*) = (3, 3)$. Note that both x' and x^* are IS. But x^* is the only SEF allocation. Also note that x^*Px'. So an IS allocation may be neither PO nor weakly PO.

Even though an IS allocation may not be SEF, starting with an IS allocation, we can always achieve the unique SEF allocation, if any exists, using the noted TTC.[15]

Theorem 5. *Starting with an IS allocation x that is not SEF, one can get the SEF allocation x^*, if any, by using the Tickets-TTC algorithm.*

Proof. Follow from Theorems 2 and 4. If the allocation x^* is not SEF, then x^* is not in the strict core of the housing market constructed, a contraction to the result in Roth & Postlewaite (1977) that shows that the strict core is a singleton consisting of x^*. □

[15] For another use of TTC, see Pápai (2000). Sen (2008) also provided a detailed discussion of the TTC algorithm.

A question of interest is how to know when the SEF allocation exists or does not exist for a given problem (A, C, q, u). We find that our Tickets algorithm can answer this question. There may exist many IS allocations, but there exists at most one SEF allocation. If an SEF allocation does not exist, our algorithm reaches an IS allocation. If an SEF allocation exists, then our algorithm only achieves that SEF allocation, no matter how many IS allocations are there, without using TTC. This result in Theorem 6 does not depend on the random orders under which agents enter the room. Thus, Theorem 6 is important because we can apply the Tickets algorithm to any single order. In practice such an order is naturally formed in an application process. Recall x^n denotes the allocation at the close of the Tickets algorithm.

Theorem 6. *For any lottery allocation problem (A, C, q, u), an SEF allocation exists if and only if x^n is SEF.*

The proof is in Appendix IX.D. We have used Example 2 to illustrate how the Tickets algorithm has reached this more efficient SEF lottery allocation, if any, over those that are less efficient (see Appendix IX.F). Let us use a metaphor to illustrate the "peculiar" feature in Theorem 6. Imagine there are millions of fish in red (IS) or blue (PO) colors. There is one and only one fish in a purple color (SEF). For some reason we don't know, our Tickets algorithm, as a kind of "net", can catch the only purple fish, no matter which order has been used and how many fish are there in red or blue colors.

6. HYLLAND AND ZECKHAUSER PSEUDO MARKET

Hylland & Zeckhauser (1979)(HZ) studied an economy with a finite number of indivisible houses (or jobs) and agents with each agent consuming one house (or assigned to one job). In their original model each house may have multiple copies of the same kind with the total number of houses in the economy being equal to the number of agents so that each agent is assigned to at least one house (one job). Specifically, HZ studied a housing allocation problem (A, C, q, u), where $A = \{1, 2, \cdots, n\}$ is the set of agents, $C = \{1, 2, \cdots, m\}$ is the set of houses with a capacity $q = (q_1, q_2, \cdots, q_m)$ such that $q_j \geq 1$ for all $j \in C$ and $\sum_{j=1}^{m} q_j = n$. $u = (u_1, u_2, \cdots, u_n)'$ denotes the vNM utility matrix with $u_i = (u_{ij})$ as agent i's utility vector over house $j \in C$ for $i \in A$. A pure assignment μ for the economy is a matching $\mu : A \to C$ that assigns an agent a house, each copy of which is assigned to at most one agent. Each matching μ

can be represented by a matrix $[p_{ij}]$, each of whose components equals either 0 or 1, satisfying $\sum_{j=1}^{m} p_{ij} = 1$ and $\sum_{i=1}^{n} p_{ij} = q_j$ for all $i \in A$ and $j \in C$. If $q_j = 1$ for all $j \in C$, $[p_{ij}]$ becomes a permutation matrix.

HZ introduced a Pseudo market with a rent system $(h_1, h_2, \cdots, h_m)$ over the m houses in C. In their market each agent $i \in A$ is assigned an equal budget $B_i > 0$ of an artificial money and chooses a probability distribution or lottery over all houses $(p_{i1}, p_{i2}, \cdots, p_{im})$ to maximize his expected utility $\sum_{j=1}^{m} p_{ij} u_{ij}$ subject to his budget constraint $\sum_{j=1}^{m} p_{ij} h_j \leq B_i$. They showed that there always exists a rent system $(h_1^*, h_2^*, \cdots, h_m^*)$ and a lottery allocation $[p_{ij}^*]$ that form a competitive equilibrium. That is, the lottery $(p_{i1}^*, p_{i2}^*, \cdots, p_{im}^*)$ maximizes the expected utility for each agent i among all lotteries $(p_{i1}, p_{i2}, \cdots, p_{im})$ that satisfy agent i's budget constraint and the market clearing condition (by the property of probability matrix $[p_{ij}^*]$). Note that entries in $[p_{ij}^*]$ are shares, not necessarily zero or 1. The random allocation or assignment $[p_{ij}^*]$ is an ex-ante EF and PO. Most importantly, the competitive equilibrium random assignment $[p_{ij}^*]$ can be implemented by a centralized draw of a lottery over all pure assignments. To see how this may be done, we may assume that $q_j = 1$ for all $j \in C$ so that $n = m$. A competitive equilibrium random assignment $[p_{ij}^*]$ is doubly stochastic, i.e., a matrix each of whose rows and columns sums to 1. The noted Birkhoff-von Neumann decomposition theorem shows that a doubly stochastic matrix can be represented by a convex combination (i.e., a centralized lottery) of permutation matrices. This result can be extended to the more general case where each house has multiple identical copies because each identical copy may be considered as a house with a unit capacity. Note that identical houses must have identical prices at a competitive equilibrium.

HZ's model does not apply to affordable housing directly. There are three obstacles for applying it to affordable housing. First, the capacity in the economy will be much smaller than the number of agents in the economy for affordable housing. Second, the random assignment at an equilibrium in HZ may not satisfy the rule–no duplicate lottery entries, which means that not all IS allocations in our model can be realized by a competitive equilibrium in a Pseudo market in HZ and a competitive equilibrium in HZ may not be feasible for our model. Third, more importantly, the Pseudo market in HZ is quite complicated and may not be applicable to affordable housing. This is because the competitive rents $(h_1^*, h_2^*, \cdots, h_m^*)$ are unknown, an auctioneer must find an algorithm to integrate the demand information of an individual random assignment from each agent so that an adjustment of the rents, which are out of equilibrium, can be made to approach $(h_1^*, h_2^*, \cdots, h_m^*)$. It is unclear if there is an effective algorithm to do the job. If there is one, it will not be a simple task for an economy with a large scale of, say, 48,000 agents. In fact, the same question regarding applicability of the HZ Pseudo market approach has been raised and discussed before in BCKM.

Under one specific case, we are able to relate our IS allocation to the result in Hylland & Zeckhauser (1979), using a much simpler way to compute the competitive equilibrium of their Pseudo market. Let us add a "null" community $m+1$ with $n-m$ number of the same copies into our model and that in HZ. For all agents i, we set $u_{i(m+1)} = 0$ for both models. A lottery allocation $[p_{ij}]$ in HZ for the null community satisfies $\sum_{i=1}^{n} p_{i(m+1)} = n-m$. A lottery allocation in our model has a specific restriction such that for all agent $i = 1, 2, \cdots, n$ and communities $j = 1, 2, ..., m$, there is one and only one p_{ij} that is greater than zero. That is, each agent can only be assigned with a positive probability to one house that is not null in order to satisfy the rule. By the equal access principle, each eligible agent has applied for one unit in a community $j \in C$.

Example 5. Let $C = \{1, 2\}$ and $q = (1, 1)$. There are three agents whose vNM utilities are given by $u_1 = (1, 0)$ and $u_2 = u_3 = (3, 2)$.

Now we set $B_i = 1$ for all $i \in A$. Then under the Pseudo market approach, we get the market-clearing rent vector $h^* = (\frac{9}{5}, \frac{6}{5}, 0)$ and the market-clearing probability matrix

$$P_{HZ} = [p_{ij}^*] = \begin{pmatrix} \frac{5}{9} & 0 & \frac{4}{9} \\ \frac{2}{9} & \frac{1}{2} & \frac{5}{18} \\ \frac{2}{9} & \frac{1}{2} & \frac{5}{18} \end{pmatrix},$$

which yields total expected utility of $3\frac{8}{9}$. The expected utilities for the three agents are given by $(\frac{5}{9}, 1\frac{2}{3}, 1\frac{2}{3})$. Under our Tickets algorithm, with the agent order $\sigma = (1, 2, 3)$, we get

$$x^n(\sigma) = \begin{pmatrix} x_1^n \\ x_2^n \\ x_3^n \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \end{pmatrix} \quad and \quad P(x^n(\sigma)) = \begin{pmatrix} \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 1 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} \end{pmatrix},$$

with total expected utility of 4. The expected utilities for the three agents are given by $(\frac{1}{2}, 2, 1\frac{1}{2})$. Clearly, $P(x^n(\sigma)) \neq P_{HZ}$. There are five more orders under the Tickets algorithm, each of which gives rise to an IS allocation with the same total expected utility of 4. So, the Tickets algorithm achieves a better

PO allocation than the competitive equilibrium in HZ. An interesting aspect of this observation is the fact that these IS allocations are all at pure Nash equilibria (see Section VI), which outperform the Pseudo market at competitive equilibrium. Nonetheless, this is not a general result because there are examples in which the Pseudo market at competitive equilibrium outperforms the Tickets algorithm or the NE.

Theorem 7. *Let x^* be the SEF allocation and $P(x^*)$ be the induced probability distribution over communities C and the null community. Then $P(x^*) = P_{HZ}$, where P_{HZ} is the competitive equilibrium random assignment under the Pseudo market in HZ with each agent a budge of 1.*

An important observation in the proof of Theorem 7 is that the equilibrium rent of a house h that is not null equals the reciprocal of the lottery drawing probability in $P(x^*)$ for the house h under the SEF allocation. That is, the lottery drawing probability matrix uniquely determines the equilibrium prices of the houses. A house's price at equilibrium is higher because it is more overly demanded. Thus, for an economy with an SEF allocation, the Tickets algorithm provides a simple way to implement the unique desired EF and PO competitive equilibrium in HZ. This overcomes the obstacle using a centralized auctioneer to integrate individual demand information into the prices in the Pseudo market.

7. LOTTERY CONGESTION GAMES

Rosenthal (1973) originated the study of a class of symmetric congestion games. Milchtaich (1996) studied a class of unweighted (singleton) congestion games with player-specific payoff functions. A congestion game in the class of Rosenthal has a potential function and the Finite Improvement property (FIP) (Monderer & Shapley, 1996). Milchtaich showed that a game in his class may not have the FIP or a potential function. He found out that a best-reply improvement path or dynamics may also last infinitely by forming a cycle. Nevertheless, he found a finite best-reply improvement path that ends up with a pure Nash equilibrium (NE). Konishi et al. (1997) provided a general class of congestion games that are defined differently from that in Milchtaich. But their assumptions on the payoff functions, namely, the *independence of irrelevant choices*, *anonymity*, and *partial rivalry*, make their game equivalent to that in Milchtaich (Voorneveld et al., 1999). Konishi et al. showed that a pure strategy

SNE (Aumann, 1959) always exists. The existence theorem about SNE for the class of congestion games in Rosenthal was proved in Holzman & Law-Yone (1997). Note that not all potential games have an SNE. In this paper, we go one step further to provide a condition such that there is a unique SNE, by means of the SEF notion. If the condition is not satisfied, then the Tickets-TTC algorithm always achieves an SNE, not necessarily unique.

Our Tickets algorithm is initially designed by following the two-sided matching literature. We recognize later that our algorithm is in fact an explicit way to identify a finite best-reply improvement path implicitly stated in Milchtaich once we transfer our affordable housing lottery into a lottery congestion game along his framework.[16] Our recognition that an unweighted congestion game in Milchtaich may be transferred into an affordable housing lottery, and vice versa, provides a new link between the housing allocation model in HZ and the congestion models in Rosenthal, Konishi et al., and Milchtaich (see Tables 3 and 4) or the potential games in Monderer and Shapley. In particular, the unique SNE is equivalent to the unique Pseudo market equilibrium outcome in HZ, a result that justifies the artificial Pseudo market economy in HZ. Moreover, the Tickets algorithm always reaches the unique SNE, if any SEF exists, for any order of applicants, a novel contribution to the literature of congestion or potential games. The reason the Tickets algorithm can always reach the unique SEF or SNE is that it achieves an NE that is not Pareto dominated by others (via Konishi et al.).

Beyond the unique SNE, our main contribution in this section is to show how to transfer an affordable housing lottery into a congestion game in Milchtaich. Moreover, an unweighted congestion game in Milchtaich that may be transferred into an affordable housing lottery along the framework of HZ provides a new link between the housing allocation model in HZ and the congestion models. Furthermore, the unique SNE is equivalent to the unique Pseudo market equilibrium outcome in HZ, a result that contributes to a new view of the Pseudo market of HZ.

Next we construct a lottery congestion game that belongs to the class studied by Milchtaich from an affordable housing lottery. Our first result of this section is that every pure Nash equilibrium outcome of the lottery congestion game

[16] We thank Abraham Neyman for communications and comments that brought our attention to the congestion game in Milchtaich during the 2015 International Conferences in honor of Abraham Neyman and Sergiu Hart in which our main results about the Tickets algorithm were presented.

induces an IS allocation, and vice versa. First, we introduce the congestion game in Milchtaich:

Definition 7. (Milchtaich, 1996). An unweighted congestion game with player-specific payoff functions $\Gamma_{(A,C,U)}$ is defined as follows: Let A be the set of players and C be the common set of strategies with $q_j = 1$ for all $j \in C$. Let $s = (s_1, s_2, \cdots, s_n) \in C^n$ be a profile of strategies. The payoff player i receives for playing strategy j, i.e., $s_i = j$, is a monotonically nonincreasing function U_{ij} of the total number $n_j(s)$ of players playing the jth strategy, i.e., $n_j(s) = |\{i \in A | s_i = j\}|$ for all $j \in C$.

An unweighted congestion game with player-specific payoff functions becomes a symmetric (singleton) congestion game studied by Rosenthal if there is a payoff function U_j such that $U_{ij} = U_j$ for all $i \in A$ and $j \in C$. It is easy to see how a game $\Gamma_{(A,C,U)}$ may be transferred into an affordable housing lottery problem. Each $j \in C$ may be seen as a house and the vNM utility for a house is derived by limiting the strategy profiles to those s such that $n_j(s) = 1$ for all $j \in C$. If the utility does not satisfy the vNM condition, then we can derive the ordinal preferences from U_{ij} by varying all $s \in C^n$.

Definition 8. A strategy profile s is a Nash equilibrium (NE) of an unweighted congestion game with player-specific payoff functions if and only if each s_i is a best-reply strategy:

$$U_{is_i}(n_{s_i}(s)) \geq U_{ij}(n_j(s) + 1), \quad for\ all\ i \in A\ and\ j \in C. \tag{7.1}$$

Monderer and Shapley showed that an ordinal potential game is equivalent to a game that has FIP. They defined a path by a sequence $\gamma = (s(0), s(1), s(2), \cdots)$ such that for every $k \geq 1$, there exists a unique deviator i, say, such that $s(k) = (s_{-i}(k-1), s_i(k))$ with $s_i(k) \neq s_i(k-1)$. If the path is finite, $s(0)$ is the initial point of γ and the last element is called the terminal point. It is a finite improvement path if for all $k \geq 1$, $U_{is_i(k)}(n_{s_i(k)}(s(k))) > U_{is_i(k-1)}(n_{s_i(k-1)}(s(k-1)))$ for the unique deviator i at step k. The game has FIP if every improvement path is finite. It is a best-reply improvement path (BRIP) if for all $k \geq 1$, $U_{is_i(k)}(n_{s_i(k)}(s(k))) > U_{is_i(k-1)}(n_{s_i(k-1)}(s(k-1)))$ for the unique deviator i at step k and $s_i(k)$ is a best reply strategy against $s_{-i}(k-1)$. A game that has FIP must have a finite BRIP, but the converse is not true. Even worse, a best-reply improvement path may last infinite for a game in Milchtaich.

Example 6. (Milchtaich, 1996). Let $A = \{1, 2, 3\}$ and $C = \{C_1, C_2, C_3\}$. The ordinal preferences of the three players over (C_j, n_{C_j}) of communities $C_j \in C$ and its congestion n_{C_j} are given in Table 3:

Table 3: Players' preferences

Agents′ Orders	*Agents in Community* C_1	*Agents in Community* C_2
$(1,2,3,4)$; $(1,2,4,3)$; $(1,4,2,3)$; $(3,2,1,4)$ $(3,2,4,1)$; $(3,4,2,1)$; $(4,2,3,1)$; $(4,2,1,3)$	1, 3, 4	2
$(1,3,2,4)$; $(1,3,4,2)$; $(1,4,3,2)$; $(2,3,1,4)$ $(2,4,3,1)$; $(3,2,4,1)$; $(4,3,2,1)$; $(4,3,1,1)$	1, 2, 4	3
$(2,1,3,4)$; $(2,1,4,3)$; $(2,4,1,3)$; $(3,1,2,4)$ $(3,1,4,2)$; $(3,4,1,2)$; $(4,1,3,2)$; $(4,1,2,3)$	2, 3, 4	1

The numerical game is given in Table 4, in which we assign numerical numbers 5, 4, 3, 2, and 1, respectively, for the first, the second, $\cdots$, and the fifth ranked choice in a player's preferences and zero for the rest of the choices.

Let $\gamma = (s(0), s(1), s(2), \cdots, s(5), \cdots)$ be an improvement path given by $s(0) = (C_2, C_1, C_1)$, $s(1) = (C_3, C_1, C_1)$, $s(2) = (C_3, C_3, C_1)$, $s(3) = (C_3, C_3, C_2)$, $s(4) = (C_2, C_3, C_2)$, $s(5) = (C_2, C_1, C_2)$, with a cycle

$$s(0) \to s(1) \to s(2) \to s(3) \to s(4) \to s(5) \to s(0)$$

with corresponding deviator sequences of $(1 : C_2 \to C_3)$, $(2 : C_1 \to C_3)$, $(3 : C_1 \to C_2)$, $(1 : C_3 \to C_2)$, $(2 : C_3 \to C_1)$, $(3 : C_2 \to C_1)$, where $(i : C_j \to C_{j'})$ denotes player i deviates from C_j to $C_{j'}$. One can check that each deviator uses a best-reply strategy at each $k \geq 1$. Thus, γ is a best-reply improvement path that lasts infinitely. Nonetheless, the game has two pure NE: $s^* = (C_3, C_1, C_2)$ and $s^0 = (C_2, C_3, C_1)$; it is the former that Pareto dominates the later.

Note that a best-reply improvement path that starts with an NE must stop there. Thus, a best-reply improvement path approach can reach both NE, including the less efficient one with equilibrium payoffs $(4, 4, 4)$. Under the Tickets algorithm, the more efficient NE (C_3, C_1, C_2) is the only one that is obtained.

Theorem 8. (Milchtaich, 1996). *Every unweighted congestion game with player-specific payoff functions has a Nash equilibrium in pure strategies.*

Table 4: A Lottery Game with Infinite BRIP

A_3: C_1

A_1 \ A_2	C_1	C_2	C_3
C_1	(0,0,0)	(0,1,3)	(0,4,3)
C_2	(4,2,3)	(2,0,4)	**(4,4,4)**
C_3	(5,2,3)	(3,0,4)	(2,3,4)

A_3: C_2

A_1 \ A_2	C_1	C_2	C_3
C_1	(0,2,5)	(1,0,2)	(1,4,5)
C_2	(3,5,2)	(0,0,0)	(3,4,2)
C_3	**(5,5,5)**	(1,3,1)	(2,3,5)

A_3: C_3

A_1 \ A_2	C_1	C_2	C_3
C_1	(0,2,1)	(1,1,1)	(1,3,0)
C_2	(4,5,1)	(3,0,1)	(4,3,0)
C_3	(2,5,0)	(2,1,0)	(0,0,0)

The strategy in the proof of Theorem 8 is to find a best-reply improvement path that is finite. Milchtaich started with a one player game $\Gamma_{(A,C,U)}$ with $\mid A \mid= 1$, which has a pure NE. Then let $\Gamma_{(A,C,U)}$ be a game with $\mid A \mid= i$ and assume that it has a pure NE s^i. Starting with the equilibrium profile s^i, Milchtaich found a finite best-reply improvement path that ends up with a pure NE s^{i+1} in the game $\Gamma_{(A,C,U)}$, where $\mid A \mid= i + 1$. His induction approach is motivated by the definitions of FIP and the best-reply improvement path, each of which allows a single deviator at each step k. The Tickets algorithm follows the random order rematching process in Ma (1996) for the marriage problem that allows a single player to enter the market to make a proposal. The induction approach consists of a sequence of games such as the game defined in Table 4. The Tickets algorithm consists of a sequence of applicants working directly on preferences given in Table 3, with each applicant knowing her private preference but without knowing the private preferences of all others.

In fact we can transfer a lottery allocation problem (A,C,q,u) into a lottery congestion game $\mathcal{G}$ as follows. A is the set of players and C is the common set of strategies. Each strategy $j \in C$ has an integer $q_j \geq 1$. A profile of strategies $s = (s_1, s_2, \cdots, s_n) \in \Pi_{i=1}^{n} C$ induces a lottery allocation $x(s)$ such that $x_{ij}(s) = 1$ if $s_i = j$ and $x_{ij}(s) = 0$ otherwise. Note that $x(s) \in X$. Then we define a vector of congestion by $n(s) = (n_1(x(s)), n_2(x(s)), \cdots, n_m(x(s)))$, where $n_j(x(s))$ is defined by (4.1) by replacing x with $x(s)$. Given a strategy profile s, the payoff vector $U(s) = (U_1(s), U_2(s), \cdots, U_n(s))$ is defined by $U_i(s) = EU_i(x(s))$ for all $i \in A$ via (4.3). Equivalently, using notation in Milchtaich, we have $U_{is_i}(n_{s_i}(s)) = EU_i(x(s))$.

In $\mathcal{G}$, $n_j(x(s))$ is a measure for congestion of strategy j under strategy profile s. It follows from (4.2) that $z_j(x(s))$ is a monotonically nonincreasing function of $n_j(x(s))$. Thus the payoff function $EU_i(s)$ is a monotonically nonincreasing function of $n_j(x(s))$. So our lottery congestion game $\mathcal{G}$ belongs to the class of unweighted congestion games with player-specific payoff functions in Milchtaich. As in the construction of our lottery game, a pure strategy profile s induces a feasible lottery allocation $x(s)$ in X. Moreover, a lottery allocation x naturally induces a strategy profile $s(x)$ for the game $\mathcal{G}$ by setting $s_i(x) = j$ for $x_{ij} = 1$ for all $i \in A$. It is easy to see that a pure Nash equilibrium s induces a lottery allocation that is IS. An IS allocation x induces a Nash equilibrium in pure strategies $s(x)$ for the game $\mathcal{G}$. Thus, Theorem 1 proves Theorem 8, and vice versa.

Definition 9. (Envy-freeness). A strategy profile s is envy-free (EF) in an unweighted congestion game with player-specific payoff functions if and only if no player i wants to change his strategy with any other's, including those strategies that may not be played:

$$U_{is_i}(n_{s_i}(s)) \geq U_{ij}(n_j(s)) \quad for\ all\ i \in A\ and\ j \in C. \tag{7.2}$$

Thus, a strategy profile s is strictly envy-free (SEF) if and only if the inequality above strictly holds for all $j \neq s_i$. For example, the NE s^* in Example 6 is SEF but the NE s^0 is not. Clearly, an EF or SEF strategy profile s must be an NE; the converse is not true. Next we discuss how an SNE is related to the SEF lottery allocation.

Definition 10. A profile s^* is a strong (pure) Nash equilibrium (SNE) of the game $\mathcal{G}$ if there is no coalition $T \subset A$ and strategies s_T for the coalition T such that

$$EU_i(s^*_{-T}, s_T) > EU_i(s^*) \quad \forall i \in T.$$

An SNE is a Nash equilibrium whose outcome is weakly PO. Because a symmetric congestion game in Rosenthal is an exact potential game (Monderer and Shapley), SNE, if any exists, coincides with NE, which is in the Argmax set of the potential function, as shown by Holzman and Law-Yone. This is no longer true for the game $\mathcal{G}$. In Example 6, s^0 is Pareto dominated by s^* (Konishi et al.).

Theorem 9. *(Konishi et al., 1997) Every unweighted congestion game with player-specific payoff functions has a pure strategy strong Nash equilibrium.*

Proof. A CS allocation induces an SNE. The existence follows from Theorem 2. □

With the SEF notion for the game $\mathcal{G}$, we prove a result that is much stronger than Theorem 9. The following directly follows from Theorem 6.

Theorem 10. *An unweighted congestion game with player-specific payoff functions $\mathcal{G}$ has an NE that is SEF if and only if the outcome of the Tickets algorithm is SEF in the related affordable housing lottery.*

Theorem 11. *Let s^* be an NE in the game $\mathcal{G}$ that is also SEF. Then the NE s^* is the unique SNE in the game $\mathcal{G}$. Moreover, such an NE s^* induces a unique probability distribution matrix at a competitive equilibrium under the Pseudo market in HZ with each agent being assigned an equal budget of* 1.

Proof. An SEF allocation is the unique CS allocation which induces a unique SNE; Moreover, it can be achieved by the Tickets algorithm for any given order (Theorems 5 and 6). The connection between SNE and the competitive equilibrium in HZ follows from Theorem 7. □

8. CONCLUSIONS

There are so many affordable housing lotteries using a simple rule–preventing duplicate lottery entries on ex ante lottery allocations, a rule that is used in the design that guarantees an equal access and opportunity for eligible applicants to affordable housing. The same type rule is used as well in many school choice lotteries. This rule is incompatible with, in the general domain, the competitive equilibrium of a Pseudo market in HZ and the EF and PO allocations generated by the algorithms in BM and BCKM. We study an affordable housing lottery with a model in HZ with a more restrictive domain, which fits well those affordable housing or school choice lotteries implemented in practice. These lotteries have a flaw in their designs because they generate outcomes that are not EF and PO, even if they exist. We provide a fix by a Tickets algorithm that incorporates an optimization process with the rejection-acceptance idea in the noted deferred acceptance algorithm. We have proposed a new fairness notion of individual stability along the lines of a stable matching in Gale and Shapley. Such a notion is justified with the Nash equilibrium in a lottery congestion game in Milchtaich. In fact, we provide a way to transfer an affordable housing lottery that is modeled by the housing allocation problem in HZ into a congestion game in Milchtaich, and vice versa, to establish several equivalence results between IS and NE, between SEF and the competitive equilibrium in HZ, and between the CS allocations and the SNE outcomes. In particular, we provide conditions for the existence of a unique SNE for the lottery congestion game and show that the Tickets algorithm always reaches the unique SNE for any order of applicants, a result contributing to two different strands of the literature of the housing allocation problem in HZ and the congestion or potential games. This link paves a new way to address other important issues in affordable housing lotteries by means of congestion or potential games, which have been more extensively studied.

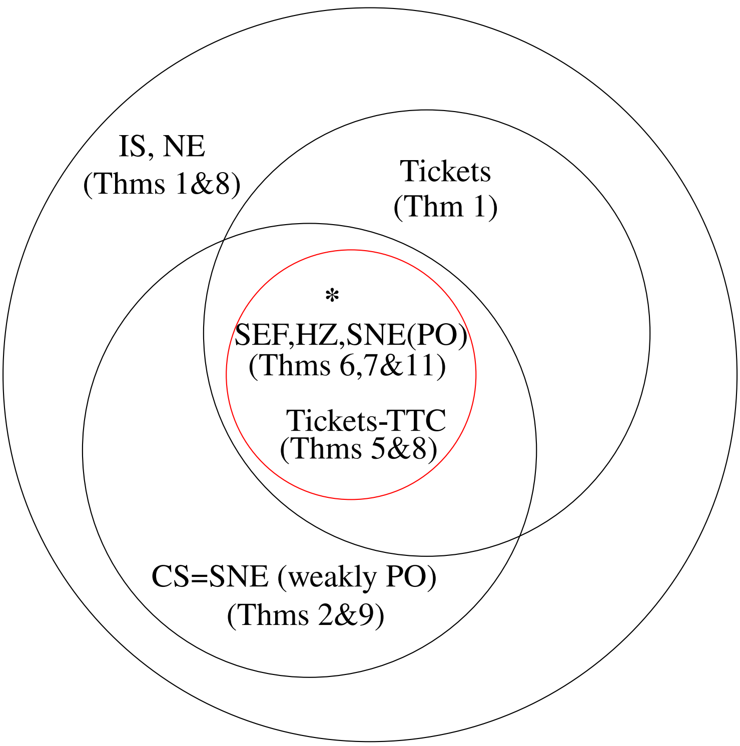
No issue with affordable housing lotteries is small. Households in the millions have been living in affordable housing in the U.S. alone and there are more to come as time goes by. There are also many different lottery protocols for affordable housing in the U.S. and across the world, each of which may be considered fair. But fairness should not be taken for granted as with the NYC lottery. Any slight change in a lottery protocol can affect the lottery allocations ex ante or ex post. A change in the rule also changes the outcome completely because agents play very different games, as noted in the celebrated papers by Crawford (1979, 1980). It remains open if our main results in this paper hold for other lottery mechanisms with different rules.

The central goal of this research is to explore the HZ economy and provide a model for achieving a fair lottery allocation to practical problems like affordable housing or school choice lotteries. The HZ model and our Tickets algorithm may be useful for other applications involving lotteries. Subsequent works should target other rules over allocation–like rental control–so that the unique SEF lottery allocation always exists. The rental control literature in Talman & Yang (2008); Andersson & Svensson (2014, 2016); Andersson et al. (2015) may be an excellent start. Since ex post allocations can affect the ways agents behave ex ante, another interesting extension of this paper would be a combination of this paper's model with the ex post allocation proposed in Andersson et al. (2016).

9. APPENDICES: SUMMARY OF MAIN RESULTS AND PROOFS

9.1. Summary of Main Results

Figure 1. The Relations of the Main Results



Note: Figure 1 summarizes our main results in this paper. Note that Figure 1 is drawn for an economy with and without an SEF lottery allocation in the same graph. For an economy with an SEF allocation, the three inner circles in Figure 1 will all shrink to a single point $*$: the SEF allocation is a singleton that is achieved by the Tickets algorithm without using TTC; it is also the unique SNE outcome and the unique Pseudo market competitive equilibrium in HZ. For an economy that does not have an SEF allocation, the allocations that can be achieved by the Tickets-TTC algorithm are a subset of all CS lottery allocations or SNE outcomes. The Tickets-TTC algorithm is selective because it only achieves those NE outcomes that are not Pareto dominated. The Tickets algorithm achieves an IS allocation or an NE outcome that may not be CS or SNE. The IS allocations or NE outcomes achieved by the Tickets algorithm are a subset of all IS allocations or NE outcomes.

Proof of Theorem 1. The sequence $\{x^i\}, i = 1, 2, 3, ..., n$, consists of lottery allocations such that each x^i is an IS allocation with respect to agents $A^i = \{1, 2, \cdots, i\}$ in the room at step i and the communities C, with the same q. We have to show that the algorithm ends in a finite number of steps and the final allocation x^n is an IS allocation for the problem (A, C, q, u). We can compare the two lottery allocations x^{i-1} and x^i. After i enters the room, the expected utilities for agents from 1 to $(i-1)$ under x^i are non-increasing and there exists at least one agent from 1 to $(i-1)$ whose expected utility has been decreased if there is anyone from 1 to $(i-1)$ who returns his ticket, due to agent i's entry. This is because only one agent enters the door each time during the process. The number of agents who hold a lottery ticket from a machine j, for all j, under x^i is also nondecreasing under x^{i-1} after agent i enters the room. Moreover, there exists one and only one machine such that the number of agents who hold a lottery ticket from that machine increases by one from x^{i-1} to x^i. Note that u_{aj} is positive for at least one $j \in C$ and for any $a \in A$.

It is clear that x^1 is IS with respect to agent 1, when he is the only agent in the room and he chooses the lottery ticket that he likes the most. Assume x^{i-1} is IS with respect to agents $\{1, 2, \cdots, i-1\}$ and C. After agent i enters the room, if there is no agent h, $h = 1, 2, \cdots, i-1$, who would like to return his lottery ticket and goes back to the door, then we obtain x^i, which is IS with respect to $\{1, 2, \cdots, i\}$ and C, by the algorithm and the fact that x^{i-1} is IS with respect to $\{1, 2, \cdots, i-1\}$ and C. If there is an agent h who would like to return his lottery ticket and goes out of the door, then agent h is seen as i. Note that the lottery allocation in the room before h reenters the room is IS with respect to $\{1, 2, \cdots, i\} \setminus \{h\}$ and C. After h reenters, he chooses the lottery ticket that he likes the most. Such a lottery ticket cannot be the one he just returned.

Since h has chosen the lottery ticket he likes the most when he reenters, he will not return his ticket again and go back to the door within the process from x^{i-1} to x^i. Note that as long as there is an agent by the door during the process, the number of lottery tickets issued by each machine remains the same. This implies that the process from x^{i-1} to x^i will not form a cycle because the number of agents in the room is finite at $(i-1)$ and no agent will visit the door twice in the process. Therefore, the process will end with a lottery allocation x^i in a finite number of steps. During the process, starting with i, x^i is obtained such that no agent from $\{1, \cdots, i\}$ would like to return his lottery ticket anymore. Thus, x^i must be IS with respect to $\{1, 2, \cdots, i\}$ and C. This shows that x^n is an IS allocation for (A, C, q, u). Note that the total number of steps to achieve x^n is bounded by $(n(n+1))/2$. Thus, this completes the proof of Theorem 1. $\square$

Proof of Theorem 4. Let x^* be an SEF allocation. We know from Theorem 3 that x^* is PO. Assume, on the contrary, that there is another allocation $x' \in X$ which is also individually stable and Pareto optimal. As in the proof of Theorem 3, we consider two cases: (a) $n_j(x^*) = n_j(x')$ for all $j \in C$; and (b) there is a community j_0 such that $n_{j_0}(x^*) < n_{j_0}(x')$.

We prove Case (a) first. Because $x^* \neq x'$, there exists at least an agent i' such that $x^*_{i'} \neq x'_{i'}$. By the assumption x^* is SEF, it follows that $EU_{i'}(x^*) > EU_{i'}(x')$. Using the fact that x^* is SEF and $n_j(x^*) = n_j(x')$, we also have that $EU_i(x^*) \geq EU_i(x')$ for all agents i. But, this implies that $x^* P x'$, which contradicts the assumption that x' is Pareto optimal. This completes the proof of Case (a).

Now we prove Case (b): $\exists j_0$ such that $n_{j_0}(x') > n_{j_0}(x^*)$. In this case, we show that we can find an infinite chain of communities, $j_0, j_1, j_2, \cdots$, with distinct vertexes, such that

$$n_{j_0}(x^*) < n_{j_0}(x'),$$

$$n_{j_1}(x^*) + n_{j_0}(x^*) < n_{j_1}(x') + n_{j_0}(x'),$$

$$n_{j_2}(x^*) + n_{j_1}(x^*) + n_{j_0}(x^*) < n_{j_2}(x') + n_{j_1}(x') + n_{j_0}(x'),$$

$$\cdots\cdots$$

But this is impossible because the number of communities is finite.

To complete the proof, we start with a chain, $j_0, j_1, \cdots, j_k$, with distinct vertexes, such that

$$n_{j_k}(x^*) + \cdots + n_{j_1}(x^*) + n_{j_0}(x^*) < n_{j_k}(x') + \cdots + n_{j_1}(x') + n_{j_0}(x'). \quad (9.1)$$

Then we claim that there exists j_{k+1} such that $j_{k+1} \notin \{j_0, j_1, \cdots, j_k\}$ and

$$n_{j_{k+1}}(x^*)+n_{j_k}(x^*)+\cdots+n_{j_1}(x^*)+n_{j_0}(x^*) < n_{j_{k+1}}(x')+n_{j_k}(x^*)+\cdots+n_{j_1}(x')+n_{j_0}(x'). \quad (9.2)$$

By assumption, (9.1) holds for $k = 0$, i.e., $n_{j_0}(x^*) < n_{j_0}(x')$, there exists at least one agent i_0 such that $x'_{i_0} = j_0$ and $x^*_{i_0} \neq j_0$. For notational convenience, let $j_1 = x^*_{i_0}$. Clearly, $j_0 \neq j_1$.

Because x' is IS, we have, by the IS definition,

$$EU_{i_0}(x') \geq EU_{i_0}(x'_{-i_0}, j_1),$$

which leads to

$$u_{i_0 j_0} z_{j_0}(x') \geq u_{i_0 j_1} z_{j_1}(x'_{-i_0}, j_1). \quad (9.3)$$

By the assumption $n_{j_0}(x^*) < n_{j_0}(x')$, we have $z_{j_0}(x^*) \geq z_{j_0}(x')$. Then

$$u_{i_0 j_0} z_{j_0}(x') \leq u_{i_0 j_0} z_{j_0}(x^*) \quad (9.4)$$

must hold. Because x^* is SEF, it follows from the SEF definition that

$$u_{i_0 j_0} z_{j_0}(x^*) < u_{i_0 j_1} z_{j_1}(x^*). \quad (9.5)$$

Now it follows from inequalities (9.3)-(9.5) that

$$u_{i_0 j_1} z_{j_1}(x^*) > u_{i_0 j_1} z_{j_1}(x'_{-i_0}, j_1),$$

which gives us $z_{j_1}(x^*) > z_{j_1}(x'_{-i_0}, j_1)$ and then

$$n_{j_1}(x^*) \leq n_{j_1}(x'). \quad (9.6)$$

It follows from (9.6) and $n_{j_0}(x^*) < n_{j_0}(x')$ that

$$n_{j_1}(x^*) + n_{j_0}(x^*) < n_{j_1}(x') + n_{j_0}(x'). \quad (9.7)$$

Now, we can follow the procedure to find j_2, etc, up to k for the chain $j_0, j_1, \cdots, j_k$. We want to show that we can find the next j_{k+1}. Once again, it follows from (9.1) that there must exist at least one agent i_k such that $x'_{i_k} \in \{j_0, j_1, \cdots, j_k\}$ and $x^*_{i_k} \notin \{j_0, j_1, \cdots, j_k\}$. Let $x'_{i_k} = j_\alpha \in \{j_0, j_1, \cdots, j_k\}$. For convenience, let $j_{k+1} = x^*_{i_k}$. Note that $j_{k+1} \notin \{j_0, j_1, \cdots, j_k\}$.

When $j_\alpha = j_0$, the proof (9.3)-(9.5) yields $n_{j_1}(x^*) \leq n_{j_1}(x')$, which gives us

$$z_{j_1}(x^*) \geq z_{j_1}(x'). \tag{9.8}$$

Now we show the case when $j_\alpha = j_1$. Once again, we can follow the proof (9.3)-(9.5) (in which (9.8) has been used) to get $n_{j_2}(x^*) \leq n_{j_2}(x')$, which gives us the desired inequality

$$z_{j_2}(x^*) \geq z_{j_2}(x'). \tag{9.9}$$

Thus, by duplicating the proof of (9.3)-(9.5) for $\alpha = 0, 1, 2, \cdots, k$, we get the following sequence

$$\begin{aligned} n_{j_0}(x^*) &< n_{j_0}(x') \Rightarrow z_{j_0}(x^*) \geq z_{j_0}(x') \\ \Rightarrow n_{j_1}(x^*) &\leq n_{j_1}(x') \Rightarrow z_{j_1}(x^*) \geq z_{j_1}(x') \\ \cdots & \quad \cdots \\ \Rightarrow n_{j_k}(x^*) &\leq n_{j_k}(x') \Rightarrow z_{j_k}(x^*) \geq z_{j_k}(x') \end{aligned} \tag{9.10}$$

Note that $j_{k+1} = x^*_{i_k}$ and $j_{k+1} \notin \{j_0, j_1, \cdots, j_k\}$. Once again, due to $z_{j_k}(x^*) \geq z_{j_k}(x')$, we can follow the proof of (9.3)-(9.5) to get $n_{j_{k+1}}(x^*) \leq n_{j_{k+1}}(x')$, which, together with (9.1), yields (9.2). This completes the proof. $\square$

Discussion of the proof of Theorem 6

In the proof of Theorem 6, we introduce a special door clock to track agents and their allocated tickets. We set a clock on the door and let it tick once when an agent is entering the room. The clock does not tick when an agent inside goes out of the door. When that agent reenters the room, the clock ticks one more time. For example, if 1 and 2 are the first two agents in an order to enter the room, then, when agent 1 is entering the room, the clock ticks once and records a time $z = 1$. The clock ticks one more time and records a time $z = 2$ when agent 2 is entering the room. Things now become a little bit more complicated because agent 1 may want to return his ticket after agent 2 enters and takes a ticket from the same pool as agent 1's. If that is the case, we allow agent 1 to go out of the room but the clock does not tick when agent 1 goes out. The clock ticks one more time and records a time $z = 3$ when agent 1 is reentering. If agent 1 does not want to return his ticket after agent 2 enters, agent 1 stays in the room and the clock stops at $z = 2$. After that, the next agent

Table 5: Door time for Example 3

agent at the door	door time	(1,2,3,4)
1	1	(, , ,)
2	2	$(C_2, , ,)$
1	3	$(, C_2, ,)$
3	4	$(C_3, C_2, ,)$
4	5	$(C_3, C_2, C_2,)$
null	6	(C_3, C_2, C_2, C_1)

in the order can get in and the clock ticks accordingly. When all agents in A form any fixed or random order to enter the room, the clock records a sequence of door times $\{z\} = 1, 2, \cdots, T$, where T is the door time when the algorithm ends. Note T is at least as large as n, and at most as large as $\frac{n(n+1)}{2}$.

An agent's door time is a dynamic process, depending on the order of agents entering the room and the optimization process taking place in the room. But, we can assign a unique door time $\{t_1, t_2, \cdots, t_n\}$ for agents in A, where t_i is defined as the door time when agent i enters the door the last time. That is, agent i enters the room and gets his ticket at door time t_i, and then keeps the same ticket from the door time t_i to the door time T. Clearly, each t_i is a time in the sequence of door times $1, 2, \cdots, T$.

For each community, there is a unique door time t_j, where t_j is the first door time such that community j has a number of outside tickets that equals $n_j(x^n)$. That is, the total number of outside tickets held by agents remains the same for j from time t_j to time T.

Table 7 presents the door time in Example 2 for the order (1,2,3,4) under the Tickets algorithm. According to our definition, the unique door times of the four agents are given by $t_1 = 3,\ t_2 = 2,\ t_3 = 4,\ t_4 = 5$, while the unique door times of the three communities are given by $t_{C_1} = 6,\ t_{C_2} = 5,\ t_{C_3} = 4$. Agent 1 reenters the room and the clock ticks at $z = 3$ and he chooses C_3 and then holds C_3 to the end. Thus, $t_1 = 3$. At the door time $z = 5$, the number of agents who hold C_2 equals 2, that remains the same from $z = 5$ to $z = 6$, which is the end. Thus, $t_{C_2} = 5$.

We now define a sequence of allocations $(y_1, y_2, \cdots, y_T)$, where each y_z is the allocation in the room at the door time z for those agents who have been inside the room, $1 \leq z \leq T$, and the agent at time z, if any, who is willing to return his lottery, has gone out of the door.

It should be noted that the allocation $(y_1, y_2, \cdots, y_T)$ may be different from the allocation $(x^1, x^2, \cdots, x^n)$ in the algorithm because T may be different from n. But, note that $y_2 = x^1$ and $y_T = x^n$. Let $n_j(y_z)$ denote the number of agents who hold tickets in community j under y_z, $z = 1, 2, \cdots, T$. Let t_1 and t_2 be two door times such that $y_{t_1} = x^i$ and $y_{t_2} = x^{i+1}$. Then, for any door time t such that $t_1 \leq t < t_2$ and any community $j \in C$, we have $n_j(y_t) = n_j(y_{t_2})$. Thus, $n_j(y_z)$ must be monotonically increasing with respect to door time $z = 1, 2, \cdots, T$.

For each community, there is a unique door time t_j, where t_j is the first door time such that community j has a number of outside tickets that equals $n_j(x^n)$. That is, the total number of outside tickets held by agents remains the same for j from time t_j to time T. This definition makes sense because $n_j(y_z) \leq n_j(y_{z+1})$ for $z = 1, 2, \cdots, T-1$. Thus, each community j has a unique t_j.

There are useful properties associated with these definitions of door times.

(A). $t_{i_1} \neq t_{i_2}$ if and only if $i_1 \neq i_2$ for any two agents i_1 and i_2 in A.

(B). $t_{j_1} \neq t_{j_2}$ if and only if $j_1 \neq j_2$ for any two communities j_1 and j_2 in C.

(C). If $t_{i_1} > t_{i_2}$ for two agents i_1 and i_2, then either i_1 is an agent from the order who enters the room at time t_{i_1} and keeps his choice from t_{i_1} to T, or agent i_1 must have gone out of the room and reentered at least once after door time t_{i_2}.

(D). $n_j(y_k)$ is nondecreasing in door time k for all $j \in C$. That is, $n_j(y_1) \leq n_j(y_2) \leq \cdots \leq n_j(y_T)$ for all $j \in C$. That is, no community becomes less congested as more and more agents enter the room.

(E). It follows from the definition of t_j, $n_j(y_t) < n_j(y_T)$ for all $t = 1, 2, \cdots, t_{j-1}$ and $n_j(y_t) = n_j(y_T)$ for all $t = t_j, t_{j+1}, \cdots, T$ for all $j \in C$.

(F). Allocation y_t, $t = 1, 2, \cdots, T$, is always IS with respect to agents in the room. Under our assumption $u_{ij} > 0$ for at least one $j \in C$, each agent in the room under y_t holds one and only one lottery ticket.

(G). The agent who enters the room as the clock ticks and records door time $t_j - 1$ must have chosen community j. For any $t > t_j$, if there exists an agent who chooses j, there must exist the other agent who returns his j ticket and gets out of the door and reenters to choose a ticket that is not j.

Proof of Theorem 6. Assume, on the contrary, there exists an SEF allocation x^* and $x^n \neq x^*$. We consider two cases.

Case 1. There is a community j_0 such that $n_{j_0}(x^*) < n_{j_0}(x^n)$. This leads to

Case (b) in the proof of theorem 4. Note that in the proof of theorem 4 in Case (b), we only use the IS, not PO assumption. From that proof, we can conclude that x^n is not IS, which is a desired contradiction.

Case 2. Assume that $n_j(x^*) = n_j(x^n)$ for all $j \in C$. Because $x^* \neq x^n$, x^* is SEF, x^n is IS, it follows from Theorem 5 that we can get x^* by using the TTC algorithm from x^n. This is due to the fact that x^* is unique and x^n is IS. Let $\Omega_1 = \{i_1, i_2, \cdots, i_\omega\}$ be any TTC cycle from the TTC algorithm such that $\omega \geq 2$. Because Ω_1 is a TTC cycle, without loss of generality, we may suppose the cycle Ω_1 has been formed as follows: Agent i_θ receives agent $i_{(\theta+1)}$'s ticket for all $\theta = 1, 2, \cdots, \omega - 1$ and agent i_ω receives agent i_1's ticket.

For notation convenience, we use θ for i_θ, $\theta = 1, 2, \cdots, \omega$. Let $\mu_\theta = x^*_\theta$ and $\nu_\theta = x^n_\theta$. μ_θ and ν_θ are the tickets held by agent θ under x^* and x^n, respectively. Note that agents θ, $\theta = 1, 2, \cdots, \omega$, trade their tickets under x^n to get their tickets under x^*. We need to show that there exists no Ω_1 cycle such that $|\Omega_1| \geq 2$, by a contradiction.

Because x^* is strictly envy-free, all agents θ become strictly better off by trading their tickets in the cycle Ω_1. The following must hold for all agents θ, $\theta = 1, 2, \cdots, \omega$,

$$EU_\theta(x^n) < EU_\theta(x^*), \tag{9.11}$$

which implies that

$$u_{\theta\mu_\theta} z_{\mu_\theta}(x^*) > u_{\theta\nu_\theta} z_{\nu_\theta}(x^n), \tag{9.12}$$

since $n_j(x^*) = n_j(x_n)$ for all $j \in C$, by assumption.

Now, we state three lemmas, which only apply to Case 2.

Lemma 1. *There exists at least one agent i_θ in Ω_1 such that $t_{\nu_\theta} < t_\theta$.*

Recall that ν_θ is agent θ's lottery ticket under x^n, the outcome from the algorithm. From (E), the total number of agents who hold a lottery in ν_θ remains the same for $t \geq t_{\nu_\theta}$. By the definition t_θ, agent θ will keep the same lottery for all $t \geq t_\theta$. Due to (F), Lemma 1 states that there exists one agent θ in the cycle Ω_1 such that agent θ enters the room and chooses the lottery ν_θ and keeps the same lottery to the close of the algorithm. Moreover, the claim $t_{\nu_\theta} < t_\theta$ implies that after agent θ enters at t_θ to choose ν_θ, there exists another agent who must return his lottery ticket in community ν_θ and reenters to choose a lottery ticket that is not ν_θ. We can even make a stronger statement. Define η to be the agent such that $t_\eta = \max\{t_\theta | \theta = 1, 2, \cdots, \omega\}$. That is, all agents in

Ω_1, except η, keep their lottery tickets at least from door time $(t_\eta - 1)$ to the end of the algorithm. η is the last agent in Ω_1 who will keep the same lottery from the door t_η to the end of the algorithm. That is, all agents in Ω_1 will keep the same lottery tickets from door time t_η to the end of the algorithm. The next lemma shows that Lemma 1 also applies to this agent η.

Lemma 2. *The inequality $t_{\nu_\eta} < t_\eta$ also holds.*

We now continue our proof of Theorem 6. By Lemma 2, we have $t_{\nu_\eta} < t_\eta$. It follows from the definition of t_η and Lemma 2 that there must exist at least one agent, denoted γ_1, who holds a ticket in community ν_η under the allocation y_{t_η} at the door time t_η but a ticket in a different community, ν_{γ_1}, say, under the lottery allocation x^n. Because agent γ_1 must go back to the door after t_η at least once, it follows that $t_{\gamma_1} > t_\eta$. The next lemma shows that if that is the case, then we can find another agent γ_2 such that $t_{\gamma_2} > t_{\gamma_1}$.

Let Ω_2 be the cycle where agent γ_1 belongs and $t_{\eta_2} = \max\{t_\delta \mid \delta \in \Omega_2\}$. Note that $\gamma_1 \notin \Omega_1$ since $t_\eta < t_{\gamma_1}$. Thus, Ω_1 and Ω_2 are two different cycles. After we find this γ_1, we show that this γ_2 must be in the third cycle Ω_3.

Lemma 3. *There exists an agent who is in the third cycle Ω_3.*

The remaining proof of Theorem 6 directly follows Lemma 3 because Ω_3 is either a singleton set or consists of at least two agents. In either case, we can find the fourth cycle Ω_4 that is nonempty (see the proof of Lemma 3 in the Appendix). This process leads to an infinite sequence of cycles, $\Omega_1, \Omega_2, \cdots$, a desired contraction because the number of cycles can only be finite. This completes the proof of Theorem 6. □

Proof of Lemma 1. Recall our definitions of the door times t_θ for agent θ and t_{ν_θ} for his lottery in community ν_θ: t_{ν_θ} is the door time such that $n_{\nu_\theta}(t) = n_{\nu_\theta}(x^n)$ for all door time $t = t_{\nu_\theta}, (t_{\nu_\theta} + 1), \cdots, T$. Thus, one agent must give up his lottery in community ν_θ and goes out of the door if there is any other agent who enters and chooses a lottery ticket from ν_θ after time t_{ν_θ}. t_θ is the door time such that agent θ keeps his lottery ticket ν_θ to the end of the algorithm, the door time T. The inequality $t_{\nu_\theta} < t_\theta$ implies that either agent θ returns his ticket at the door time $(t_\theta - 1)$ and reenters to choose ν_θ at the door time t_{ν_θ}; or he is an agent from the original order who just enters at the door time t_θ and chooses ν_θ. By $t_{\nu_\theta} < t_\theta$, Lemma 1 claims that after θ chooses his ticket ν_θ, by property (G), there must exist one other agent in the room who will return his lottery in community ν_θ and gets out of the door. This agent reenters and chooses a ticket that is not in community ν_θ.

Define $t' = \max_{\theta=1,2,\cdots,\omega} t_{\nu_\theta}$. Then, for all t such that $t = (t'+1), (t'+2), \cdots, T$, for all agents θ in Ω_1, the following holds

$$n_{\nu_\theta}(y_t) = n_{\nu_\theta}(y_T).$$

Moreover, for all t such that $t \le t'$, there is at least one $\theta \in \Omega_1$ such that

$$n_{\nu_\theta}(y_t) < n_{\nu_\theta}(y_T),$$

by property (D). Therefore, there is one and only one community ν^* in $\{\nu | \nu_\theta = \nu, \theta \in \Omega_1\}$ such that

$$n_{\nu^*}(y_{t'}) < n_{\nu^*}(y_T)$$

at the door time t'. Thus, $n_{\nu^*}(y_t) < n_{\nu^*}(y_T)$ for all door time $t < t'$ and $n_{\nu^*}(y_t) + 1 = n_{\nu^*}(y_T)$ for $t = t'$. Moreover, $n_{\nu^*}(y_t) = n_{\nu^*}(y_T)$ for all door time $t > t'$.

Now, assume, on the contrary, that $t_\theta \le t_{\nu_\theta}$ for all $\theta \in \Omega_1$. This means that all agents $\theta \in \Omega_1$ will keep their lotteries to the door time T. That is, there is no agent $\theta \in \Omega_1$ who will go outside of the door after the door time t'. Let θ^* be the agent in Ω_1 such that $\mu_{\theta^*} = \nu^*$, where μ_{θ^*} is the lottery ticket in community ν^* agent θ^* obtains after forming the cycle Ω_1. Then, at the door time t', the following holds

$$n_{\nu_{\theta^*}}(y_{t'}) = n_{\nu_{\theta^*}}(y_T) \ \text{ and } \ n_{\mu_{\theta^*}}(y_{t'}) + 1 = n_{\mu_{\theta^*}}(y_T).$$

Thus, at the door time t', agent θ^* holds the ticket in community ν_{θ^*}, but he strictly prefers a lottery in community ν^*, by (9.11). This implies that the lottery allocation at the door time t' cannot be IS (with respect to the set of agents in the room and communities C) because agent θ^* will return his lottery ticket ν_{θ^*} for the better lottery ticket ν^*. However, this is a contradiction to the property (F). This completes the proof of Lemma 1. □

Proof of Lemma 2. We need to prove that, at the door time t_η when agent η enters the room and chooses a lottery ticket from community ν_η, there exists another agent who will return his lottery ticket in community ν_η and go back to the door. Here we have used property (D).

By Lemma 1, there is an agent $\theta^* \in \Omega_1$ such that $t_{\nu_{\theta^*}} < t_{\theta^*}$. If $\theta^* = \eta$, we are done. If not, then $\theta^* \neq \eta$ and

$$t_\eta > t_{\theta^*} > t_{\nu_{\theta^*}}.$$

Now consider agent $(\theta^*+1) \in \Omega_1$ and his lottery ticket $\nu_{(\theta^*+1)}$. By the definition of Ω_1, it follows from (9.11) that

$$EU_{\theta^*}(x^n) < EU_{\theta^*}(x^*) = u_{\theta^*\mu_{\theta^*}} z_{\mu_{\theta^*}}(x^*). \tag{9.13}$$

The equality "=" above follows from the definition of expected utility and the lottery 'house' μ_{θ^*} is obtained by θ^* in the cycle Ω_1. For all $j \in C$, $z_j(x^n) = z_j(x^*)$ and $\mu_{\theta^*} = \nu_{(\theta^*+1)}$, we have

$$u_{\theta^*\mu_{\theta^*}} z_{\mu_{\theta^*}}(x^*) = u_{\theta^*\nu_{(\theta^*+1)}} z_{\nu_{(\theta^*+1)}}(x^n). \tag{9.14}$$

For notational convenience, let $y_{t_\eta} = O$. Because O is IS by the property (F), we have

$$EU_{\theta^*}(O) \geq EU_{\theta^*}(O_{-\theta^*}, \nu_{(\theta^*+1)}) = u_{\theta^*\nu_{(\theta^*+1)}} z_{\nu_{(\theta^*+1)}}(O_{-\theta^*}, \nu_{(\theta^*+1)}). \tag{9.15}$$

Because $t_{\nu_{\theta^*}} < t_\eta$, we have $n_{\nu_{\theta^*}}(O) = n_{\nu_{\theta^*}}(x^n)$. Thus,

$$EU_{\theta^*}(x^n) = EU_{\theta^*}(O). \tag{9.16}$$

By equations (9.13)-(9.16), we have

$$u_{\theta^*\nu_{(\theta^*+1)}} z_{\nu_{(\theta^*+1)}}(x^n) > u_{\theta^*\nu_{(\theta^*+1)}} z_{\nu_{(\theta^*+1)}}(O_{-\theta^*}, \nu_{(\theta^*+1)}). \tag{9.17}$$

Thus, we must have

$$n_{\nu_{y_{(\theta^*+1)}}}(t_\eta) + 1 > n_{\nu_{(\theta^*+1)}}(x^n), \tag{9.18}$$

which yields

$$n_{\nu_{y_{(\theta^*+1)}}}(t_\eta) \geq n_{\nu_{(\theta^*+1)}}(x^n), \tag{9.19}$$

and then,

$$t_{\nu_{(\theta^*+1)}} < t_\eta. \tag{9.20}$$

Now, due to (9.20), we can apply the same idea in the proof with agent θ^* above to agent (θ^*+1) to find agent (θ^*+2) and $\nu_{(\theta^*+2)}$, and then the following inequality also holds

$$t_{\nu_{(\theta^*+2)}} < t_\eta.$$

Continuing with this process, we must eventually reach the agent η since he is also an agent in Ω_1. Thus, the inequality

$$t_{\nu_\eta} < t_\eta$$

must hold. This completes the proof of Lemma 2. □

Proof of Lemma 3. Now we consider two cases for γ_1. Case (α). $|\Omega_2| \geq 2$; Case (β). $\Omega_2 = \{\gamma_1\}$.

First, we prove Case (α). By the assumption, Ω_2 consists of at least two agents. Let $\gamma_1 \in \Omega_2 = \{i'_1, i'_2, \cdots, i'_h\}$. But, for notation convenience, we use δ for i'_δ, $\delta = 1, 2, \cdots, h$. Let $\eta_2 \in \Omega_2$ such that $t_{\eta_2} = \max\{t_\delta | \delta \in \Omega_2\}$. By Lemma 2, we have $t_{\eta_2} > t_{\nu_{\eta_2}}$. Now, property (G) implies that when η_2 enters and chooses the lottery ν_{η_2} in community ν_{η_2}, there exists an agent who returns his lottery in community ν_{η_2} and goes out of the door. This implies that there is an agent γ_2, say, such that $t_{\gamma_2} > t_{\eta_2}$, which implies that γ_2 must be in the third cycle Ω_3. This completes the proof of Case (α).

Next, we prove Case (β). Under this case,

$$EU_{\gamma_1}(x^n) = EU_{\gamma_1}(x^*). \tag{9.21}$$

Recall that $\eta \in \Omega_1$ and γ_1 does not hold a lottery in the community where η is, under x^n. Because x^* is SEF, and η and γ_1 are not in the same communities under x^n, we have

$$EU_{\gamma_1}(x^*) > u_{\gamma_1 \nu_\eta} z_{\nu_\eta}(x^*). \tag{9.22}$$

By our assumption in Case (b), $n_j(x^*) = n_j(x^n)$ for all $j \in C$, which implies that $z_j(x^*) = z_j(x^n)$ for all $j \in C$. Thus, $u_{\gamma_1 \nu_\eta} z_{\nu_\eta}(x^*) = u_{\gamma_1 \nu_\eta} z_{\nu_\eta}(x^n)$. And then (9.22) gives us

$$EU_{\gamma_1}(x^*) > u_{\gamma_1 \nu_\eta} z_{\nu_\eta}(x^n). \tag{9.23}$$

Since $EU_{\gamma_1}(x^n) = u_{\gamma_1 \nu_{\gamma_1}} z_{\nu_{\gamma_1}}(x^n)$, it follows from (9.21) and (9.23) that

$$u_{\gamma_1 \nu_{\gamma_1}} z_{\nu_{\gamma_1}}(x^n) > u_{\gamma_1 \nu_\eta} z_{\nu_\eta}(x^n). \tag{9.24}$$

Because y_{t_η} is IS (with respect to agents in the room at time t_η) and γ_1 holds a lottery ticket in community ν_η, which is not equal to ν_{γ_1}, we obtain, by the IS definition,

$$EU_{\gamma_1}(O) \geq EU_{\gamma_1}((O_{-\gamma_1}, \nu_{\gamma_1})),$$

where $O = y_{t_\eta}$, and then

$$u_{\gamma_1 \nu_\eta} z_{\nu_\eta}(O) \geq u_{\gamma_1 \nu_{\gamma_1}} z_{\nu_{\gamma_1}}((O_{-\gamma_1}, \nu_{\gamma_1})). \tag{9.25}$$

Now, by Lemma 2, at the door time t_η, the inequality $t_{\nu_\eta} < t_\eta$ holds. Thus, by property (E), we get

$$z_{\nu_\eta}(O) = z_{\nu_\eta}(x^n). \tag{9.26}$$

By (9.26), $u_{\gamma_1 \nu_\eta} z_{\nu_\eta}(x^n) = u_{\gamma_1 \nu_\eta} z_{\nu_\eta}(O)$. Then, by (9.24) and (9.25), we get

$$u_{\gamma_1 \nu_{\gamma_1}} z_{\nu_{\gamma_1}}(x^n) > u_{\gamma_1 \nu_\eta} z_{\nu_\eta}(O) \geq u_{\gamma_1 \nu_{\gamma_1}} z_{\nu_{\gamma_1}}((O_{-\gamma_1}, \nu_{\gamma_1})), \tag{9.27}$$

which yields

$$n_{\nu_{\gamma_1}}(x^n) < n_{\nu_{\gamma_1}}(O) + 1,$$

and then

$$n_{\nu_{\gamma_1}}(x^n) \leq n_{\nu_{\gamma_1}}(O). \tag{9.28}$$

But, by property (D), $n_j(t)$ is nondecreasing function in t for all $j \in C$. Thus, it must hold that

$$n_{\nu_{\gamma_1}}(x^n) \geq n_{\nu_{\gamma_1}}(O),$$

which, together with (9.28), gives us

$$n_{\nu_{\gamma_1}}(x^n) = n_{\nu_{\gamma_1}}(O), \tag{9.29}$$

which implies that $t_{\nu_{\gamma_1}} < t_\eta$, and then the following must hold

$$t_{\nu_{\gamma_1}} < t_\eta < t_{\gamma_1}.$$

Now, by property (G), there must exist an agent, denoted γ_2, who holds the ticket of ν_{γ_1} at the door time t_{γ_1}, but not a ticket of a different community, ν_{γ_2}, say, in the lottery allocation x^n. Thus, we have $t_{\gamma_2} > t_{\gamma_1}$, which implies that γ_2 is in the third cycle Ω_3 because Ω_2 is singleton. This completes the proof of Case (β).

□

Proof of Theorem 7. Let $B_i = 1$ be the budgets for all $i \in A$ and $h_j = \frac{1}{z_j(x^*)}$ for $j \in C$ be the prices. For the null commodity, let its price h_{m+1} be zero. Now consider agent i's optimization problem in the choice of random allocations, as if all goods are divisible, under his budget constraint:

$$\max_{(p_{i1},p_{i2},\cdots,p_{im},p_{i(m+1)})} \sum_{j=1}^{m+1} p_{ij}u_{ij}$$

$$\text{subject to} \sum_{j=1}^{m+1} p_{ij}h_j = 1.$$

For all $i \in A$ and $j \in C \cup \{m+1\}$, let

$$p^*_{ij} = \begin{cases} z_j(x^*) & \text{if } x^*_{ij} = 1, j \neq m+1; \\ 1 - z_j(x^*) & \text{if } x^*_{ij} = 1, j = m+1; \\ 0 & \text{if } x^*_{ij} = 0. \end{cases}$$

We claim that $[p^*_{ij}]$ satisfies

a). the budget constraints for all agents i: For all $i \in A$,

$$\sum_{j=1}^{m+1} p^*_{ij} \times h_j = \frac{1}{z_j(x^*)}(z_j(x^*)) + 0 = 1;$$

b). the market clearing condition:

$$\sum_{i=1}^{n} p^*_{ij} = q_j;$$

c). the optimization condition where each agent maximizes his expected utility subject to his budget constraint: Because x^* is SEF, we have, for all i, j such that $x^*_{ij} \neq 1$ and $x^*_{ij^*} = 1$,

$$u_{ij^*}z_{j^*}(x^*) > u_{ij}z_j(x^*). \tag{9.30}$$

Let $P' = [p'_{ij}]$ be a random assignment that satisfies the budget constraints

$$\sum_{j \in C} p'_{ij}h_j = 1, \quad \forall i \in A.$$

Denote $p'_{ij}h_j = o_{ij}$. Then we have

$$\sum_{j\in C} o_{ij} = 1, \quad \forall i \in A. \tag{9.31}$$

Thus, the expected utility of agent i under the random assignment P' is given by

$$EU_i(P') = \sum_{j\in C} u_{ij}p'_{ij} = \sum_{j\in C} u_{ij}\frac{1}{h_j}o_{ij} = \sum_{j\in C} u_{ij}z_j(x^*)o_{ij}.$$

Replacing $u_{ij}z_j(x^*)$ with $u_{ij^*}z_{j^*}(x^*)$ in the above, and then using inequality (9.30) and equality (9.31), we have, $\forall i \in A$,

$$\sum_{j\in C} u_{ij}z_j(x^*)o_{ij} \leq u_{ij^*}z_{j^*}(x^*).$$

Thus, we have $EU_i(P^*) \geq EU_i(P')$ for all $i \in A$. Moreover, the equality holds only if $P' = P^*$. The uniqueness follows. Therefore, $P(x^*) = P_{HZ}$. This completes the proof. □

9.2. Tickets Algorithm versus Lottery Game in Example 2

Example 2 (Continued). In Table 6, we present the Tickets algorithm for the remaining 21 orders in Example 2. In Table 7, in the game of Example 2, we denote agent i by A_i, to avoid confusion. Agent A_1 is the row player, agent A_2 is the column player, agent A_4 is the matrix player, and agent A_3 is the matrix group player. There are two pure Nash equilibria $s = (C_1, C_2, C_2, C_3)$ and $s^* = (C_3, C_2, C_2, C_1)$, with equilibrium payoffs $(2, \frac{1}{2}, \frac{1}{2}, 2)$ and $(3, \frac{1}{2}, \frac{1}{2}, 3)$, respectively. s^* is SNE while s is NE but not SNE (Konishi et al., 1997). We find out that our Tickets algorithm can reach the only SNE s^* no matter which order has been used, as shown in the table.

Table 6: Tickets Algorithm for Example 2

Order \ steps	(1,3,4,2)	(1,4,2,3)	(1,4,3,2)
1	$(C_2, , ,)$	$(C_2, , ,)$	$(C_2, , ,)$
2	$(C_2, ,C_2,)$	$(C_2, , ,C_1)$	$(C_2, , ,C_1)$
	$(C_3, ,C_2,)$		
3	$(C_3, ,C_2,C_1)$	$(C_2,C_2, ,C_1)$	$(C_2, ,C_2,C_1)$
		$(C_3,C_2, ,C_1)$	$(C_3, ,C_2,C_1)$
4	(C_3,C_2,C_2,C_1)	(C_3,C_2,C_2,C_1)	(C_3,C_2,C_2,C_1)
Order \ steps	(2,1,3,4)	(2,1,4,3)	(2,3,1,4)
1	$(,C_2, ,)$	$(,C_2, ,)$	$(,C_2, ,)$
2	$(C_3,C_2, ,)$	$(C_3,C_2, ,)$	$(,C_2,C_2,)$
3	$(C_3,C_2,C_2,)$	$(C_3,C_2, ,C_1)$	$(C_3,C_2,C_2,)$
4	(C_3,C_2,C_2,C_1)	(C_3,C_2,C_2,C_1)	(C_3,C_2,C_2,C_1)
Order \ steps	(2,3,4,1)	(2,4,1,3)	(2,4,3,1)
1	$(,C_2, ,)$	$(,C_2, ,)$	$(,C_2, ,)$
2	$(,C_2,C_2,)$	$(,C_2, ,C_1)$	$(,C_2, ,C_1)$
3	$(,C_2,C_2,C_1)$	$(C_3,C_2, ,C_1)$	$(,C_2,C_2,C_1)$
4	(C_3,C_2,C_2,C_1)	(C_3,C_2,C_2,C_1)	(C_3,C_2,C_2,C_1)
Order \ steps	(3,1,2,4)	(3,1,4,2)	(3,2,1,4)
1	$(, ,C_2,)$	$(, ,C_2,)$	$(, ,C_2,)$
2	$(C_3, ,C_2,)$	$(C_3, ,C_2,)$	$(,C_2,C_2,)$
3	$(C_3,C_2,C_2,)$	$(C_3, ,C_2,C_1)$	$(C_3,C_2,C_2,)$
4	(C_3,C_2,C_2,C_1)	(C_3,C_2,C_2,C_1)	(C_3,C_2,C_2,C_1)
Order \ steps	(3,2,4,1)	(3,4,1,2)	(3,4,2,1)
1	$(, ,C_2,)$	$(, ,C_2,)$	$(, ,C_2,)$
2	$(,C_2,C_2,)$	$(, ,C_2,C_1)$	$(, ,C_2,C_1)$
3	$(,C_2,C_2,C_1)$	$(C_3, ,C_2,C_1)$	$(,C_2,C_2,C_1)$
4	(C_3,C_2,C_2,C_1)	(C_3,C_2,C_2,C_1)	(C_3,C_2,C_2,C_1)
Order \ steps	(4,1,2,3)	(4,1,3,2)	(4,2,1,3)
1	$(, , ,C_1)$	$(, , ,C_1)$	$(, , ,C_1)$
2	$(C_2, , ,C_1)$	$(C_2, , ,C_1)$	$(,C_2, ,C_1)$
3	$(C_2,C_2, ,C_1)$	$(C_2, ,C_2,C_1)$	$(C_3,C_2, ,C_1)$
	$(C_3,C_2, ,C_1)$	$(C_3, ,C_2,C_1)$	
4	(C_3,C_2,C_2,C_1)	(C_3,C_2,C_2,C_1)	(C_3,C_2,C_2,C_1)
Order \ steps	(4,2,3,1)	(4,3,1,2)	(4,3,2,1)
1	$(, , ,C_1)$	$(, , ,C_1)$	$(, , ,C_1)$
2	$(,C_2, ,C_1)$	$(, ,C_2,C_1)$	$(, ,C_2,C_1)$
3	$(,C_2,C_2,C_1)$	$(C_3, ,C_2,C_1)$	$(,C_2,C_2,C_1)$
4	(C_3,C_2,C_2,C_1)	(C_3,C_2,C_2,C_1)	(C_3,C_2,C_2,C_1)

Table 7: Tickets Algorithm for Example 2

A_3: C_1

A_1 \ A_2	C_1	C_2	C_3
C_1	$(\frac{1}{2},0,0,\frac{3}{4})$	$(\frac{4}{3},1,0,1)$	$(\frac{4}{3},0,0,1)$
C_2	$(4,0,0,1)$	$(2,\frac{1}{2},0,\frac{3}{2})$	$(4,0,0,\frac{3}{2})$
C_3	$(3,0,0,1)$	$(3,1,0,\frac{3}{2})$	$(\frac{3}{2},0,0,\frac{3}{2})$

A_4: C_1

A_1 \ A_2	C_1	C_2	C_3
C_1	$(\frac{2}{3},0,0,0)$	$(1,\frac{1}{2},0,0)$	$(1,0,0,0)$
C_2	$(2,0,0,0)$	$(\frac{4}{3},\frac{1}{3},0,0)$	$(2,0,0,0)$
C_3	$(3,0,0,0)$	$(3,\frac{1}{2},0,0)$	$(\frac{3}{2},0,0,0)$

A_4: C_2

A_1 \ A_2	C_1	C_2	C_3
C_1	$(\frac{2}{3},0,0,2)$	$(1,1,0,2)$	$(1,0,0,1)$
C_2	$(4,0,0,2)$	$(2,\frac{1}{2},0,2)$	$(4,0,0,1)$
C_3	$(\frac{3}{2},0,0,1)$	$(\frac{3}{2},1,0,2)$	$(1,0,0,\frac{2}{3})$

A_4: C_3

A_3: C_2

A_1 \ A_2	C_1	C_2	C_3
C_1	$(\frac{2}{3},0,1,1)$	$(1,\frac{1}{2},\frac{1}{2},\frac{3}{2})$	$(1,0,1,\frac{3}{2})$
C_2	$(2,0,\frac{1}{2},\frac{3}{2})$	$(\frac{4}{3},\frac{1}{3},\frac{1}{3},3)$	$(2,0,\frac{1}{2},3)$
C_3	$(3,0,1,\frac{3}{2})$	$\mathbf{(3,\frac{1}{2},\frac{1}{2},3)}$	$(\frac{3}{2},0,1,3)$

A_4: C_1

A_1 \ A_2	C_1	C_2	C_3
C_1	$(1,0,\frac{1}{2},0)$	$(2,\frac{1}{3},\frac{1}{3},0)$	$(2,0,\frac{1}{2},\frac{1}{2})$
C_2	$(\frac{4}{3},0,\frac{1}{3},0)$	$(1,\frac{1}{4},\frac{1}{4},0)$	$(\frac{4}{3},0,\frac{1}{3},0)$
C_3	$(3,0,\frac{1}{2},0)$	$(3,\frac{1}{3},\frac{1}{3},0)$	$(\frac{3}{2},0,\frac{1}{2},0)$

A_4: C_2

A_1 \ A_2	C_1	C_2	C_3
C_1	$(1,0,1,2)$	$\mathbf{(2,\frac{1}{2},\frac{1}{2},2)}$	$(2,0,1,1)$
C_2	$(2,0,\frac{1}{2},2)$	$(\frac{4}{3},\frac{1}{3},\frac{1}{3},2)$	$(2,0,\frac{1}{2},1)$
C_3	$(\frac{3}{2},0,1,1)$	$(\frac{3}{2},\frac{1}{2},\frac{1}{2},2)$	$(1,0,1,\frac{2}{3})$

A_4: C_3

A_3: C_3

A_1 \ A_2	C_1	C_2	C_3
C_1	$(\frac{2}{3},0,0,1)$	$(1,1,0,\frac{3}{2})$	$(1,0,0,\frac{3}{2})$
C_2	$(4,0,0,\frac{3}{2})$	$(2,\frac{1}{2},0,3)$	$(4,0,0,3)$
C_3	$(\frac{3}{2},0,0,\frac{3}{2})$	$(\frac{3}{2},1,0,3)$	$(1,0,0,3)$

A_4: C_1

A_1 \ A_2	C_1	C_2	C_3
C_1	$(1,0,0,0)$	$(2,\frac{1}{2},0,0)$	$(2,0,0,0)$
C_2	$(2,0,0,0)$	$(\frac{4}{3},\frac{1}{3},0,0)$	$(2,0,0,0)$
C_3	$(\frac{3}{2},0,0,0)$	$(\frac{3}{2},\frac{1}{2},0,0)$	$(1,0,0,0)$

A_4: C_2

A_1 \ A_2	C_1	C_2	C_3
C_1	$(1,0,0,1)$	$(2,1,0,1)$	$(2,0,0,\frac{2}{3})$
C_2	$(4,0,0,1)$	$(2,\frac{1}{2},0,1)$	$(4,0,0,\frac{2}{3})$
C_3	$(1,0,0,\frac{1}{3})$	$(1,1,0,\frac{1}{3})$	$(\frac{3}{4},0,0,\frac{1}{2})$

A_4: C_3

References

Abdulkadiroğlu, A., & Sönmez, T. (1998). Random serial dictatorship and the core from random endowments in house allocation problems. *Econometrica*, *66*(3), 689–701.

Andersson, T., Ehlers, L., & Svensson, L.-G. (2016). Transferring ownership of public housing to existing tenants: A market design approach. *Journal of Economic Theory*, *165*, 643-671.

Andersson, T., & Svensson, L.-G. (2014). Non-manipulable house allocation with rent control. *Econometrica*, *82*(2), 507-539.

Andersson, T., & Svensson, L.-G. (2016). Strategy-proof house allocation with price restrictions. *Journal of Economic Theory*, *163*, 167-177.

Andersson, T., Yang, Z., & Zhang, D. (2015). How to efficiently allocate houses under price controls? *Economics Letters*, *130*, 97-99.

Aumann, R. J. (1959). Acceptable points in general cooperative n-person games. *Contributions to the Theory of Games*, *4*(40), 287–324.

Bogomolnaia, A., & Moulin, H. (2001). A new solution to the random assignment problem. *Journal of Economic Theory*, *100*(2), 295–328.

Budish, E., Che, Y.-K., Kojima, F., & Milgrom, P. (2013). Designing random allocation mechanisms: Theory and applications. *American Economic Review*, *103*(2), 585–623.

Crawford, V. P. (1977). A game of fair division. *Review of Economic Studies*, *44*(2), 235–247.

Crawford, V. P. (1979). A procedure for generating pareto-efficient egalitarian-equivalent allocations. *Econometrica*, 49–60.

Crawford, V. P. (1980). A self-administered solution of the bargaining problem. *Review of Economic Studies*, *47*(2), 385–392.

Foley, D. (1967). Resource allocation and the public sector. *Yale Economic Essays*, *7*, 45–98.

Gale, D., & Shapley, L. S. (1962). College admissions and the stability of marriage. *American Mathematical Monthly*, *69*(1), 9–15.

Holzman, R., & Law-Yone, N. (1997). Strong equilibrium in congestion games. *Games and Economic Behavior*, *21*(1-2), 85–101.

Hylland, A., & Zeckhauser, R. (1979). The efficient allocation of individuals to positions. *Journal of Political Economy*, *87*(2), 293–314.

Katta, A.-K., & Sethuraman, J. (2006). A solution to the random assignment problem on the full preference domain. *Journal of Economic Theory*, *131*(1), 231–250.

Klaus, B., & Klijn, F. (2006). Procedurally fair and stable matching. *Economic Theory*, *27*, 431–447.

Konishi, H., Le Breton, M., & Weber, S. (1997). Equilibria in a model with partial rivalry. *Journal of Economic Theory*, *72*(1), 225–237.

Ma, J. (1994). Strategy-proofness and the strict core in a market with indivisibilities. *International Journal of Game Theory*, *23*, 75–83.

Ma, J. (1996). On randomized matching mechanisms. *Economic Theory*, 377–381.

Milchtaich, I. (1996). Congestion games with player-specific payoff functions. *Games and Economic Behavior*, *13*(1), 111–124.

Monderer, D., & Shapley, L. S. (1996). Potential games. *Games and Economic Behavior*, *14*(1), 124–143.

Pápai, S. (2000). Strategyproof assignment by hierarchical exchange. *Econometrica*, *68*(6), 1403–1433.

Pazner, E. A., & Schmeidler, D. (1974). A difficulty in the concept of fairness. *Review of Economic Studies*, *41*(3), 441–443.

Rosenthal, R. W. (1973). A class of games possessing pure-strategy Nash equilibria. *International Journal of Game Theory*, *2*, 65–67.

Roth, A. E., & Postlewaite, A. (1977). Weak versus strong domination in a market with indivisible goods. *Journal of Mathematical Economics*, *4*(2), 131–137.

Roth, A. E., & Vate, J. H. V. (1990). Random paths to stability in two-sided matching. *Econometrica*, 1475–1480.

Saban, D., & Sethuraman, J. (2013). House allocation with indifferences: a generalization and a unified view. In *Proceedings of the Fourteenth ACM Conference on Electronic Commerce* (pp. 803–820).

Sen, A. (2008). 'fair and lovely': Some theoretical considerations in the equitable allocation of resources. *Indian Statistical Institute*.

Shapley, L., & Scarf, H. (1974). On cores and indivisibility. *Journal of Mathematical Economics*, *1*(1), 23–37.

Svensson, L.-G. (1994). Queue allocation of indivisible goods. *Social Choice and Welfare*, *11*, 323–330.

Talman, D., & Yang, Z. (2008). A dynamic auction for differentiated items under price rigidities. *Economics Letters*, *99*(2), 278–281.

Varian, H. R. (1974). Equity, envy, and efficiency. *Journal of Economic Theory*, *9*, 63–91.

Voorneveld, M., Borm, P., Van Megen, F., Tijs, S., & Facchini, G. (1999). Congestion games and potentials reconsidered. *International Game Theory Review*, *1*, 283–299.

Zhou, L. (1990). On a conjecture by gale about one-sided matching problems. *Journal of Economic Theory*, *52*(1), 123–135.

Zhu, J., & Zhang, D. (2011). Dynamic auction for efficient competitive equilibrium under price rigidities. In *Australasian Joint Conference on AI* (pp. 809–818). Springer.

Journal of Mechanism and Institution Design
ISSN: 2399-844X (Print), 2399-8458 (Online)
DOI:10.22574/jmid.2023.12.002

GOLDEN RULE IN COOPERATIVE COMMONS

Sjur Didrik Flåm

University of Bergen, Norway
sjur.flaam@uib.no

ABSTRACT

This paper considers common use of natural, renewable resources. It identifies good prospects for efficiency and welfare. To be precise, a core outcome –hence cooperation – can be secured *over time* by principal planning of total quotas, and *in time* by agents who share these in short-term markets. Information flows in two directions: to the principal as market *prices* and from him as total *quantities*. Of particular interest is eventual convergence to a golden-rule, steady state.

Keywords: Renewable resources, commons, golden rule.

JEL Classification Numbers: D02, D23, D44, P13, Q22.

1. INTRODUCTION

ECONOMIC THEORY, like *political science*, has long regarded use of common resources as largely determined by self-concerned individuals. Then, if such parties see no proper rights or frames, outcomes have often been grossly inefficient.

It's fortunate therefore that the two fields increasingly note, and often criticise, lack of mitigating or welfare-enhancing institutions (March & Olsen, 1989). Both stress that dedicated agencies ought balance competition against coordination and conflict against cooperation. Clearly, many instances are overwhelmingly complex. Yet *two* major features stand out. *First*, since nobody can attend to everything, competencies must be divided and responsibilities delineated. *Second*, for efficiency, short-term concerns must comply with those of the long run.

My thanks are due to B. Harstad, J.M. Nordbotten, N-A. Ekerhovd, S.I. Steinshamn, the editor and the referees for their helpful comments.

Addressing these features, this paper argues for a two-level splitting of roles and tasks. At the upper level, a directorate or *principal* sets aggregate quotas in *macro and over time*. At the lower level, legitimate *agents* recurrently use short-term allowances in *micro and time.* Between the two levels, and between the agents, market-like mechanisms channel information, determine prices, and allocate quantities.

Thus the principal cares, in the aggregate, for the natural resources, their abundance, dynamics and sustainable yields. Lacking knowledge as to minor details, mainly of an economic or operative nature, that agency must leave short-term calculations and small-scale decisions to *resource users*. Each among the latter parties, applying their own quota and technology, appropriates some resource rent or profit, period after period. Set up this way, the organizational design divorces central governance from decentralized discretion. Yet, as argued below, it can coordinate choices and elicit necessary information.

The renewable resources, considered here, are confined to a "region" or local community. By assumption, property and user rights therein are of intermediary nature, between *sole ownership* and *open access*, but regulated or well defined.[1] Multi-agent fishery and forestry are cases in point.[2] Suppose that legitimate parties are so few and foresighted as to honour short and long term commitments. On these premises, the questions addressed here are: *Can the users share resources efficiently and fairly in and across time? Can they operate throughout as would a well informed, highly competent syndicate?*

This paper answers these questions in the positive. Broadly, it decouples long-term control from short-term choices.[3] Further, it aligns the principal's foresight with the agents' myopia. And finally, it marries centralized planning to decentralized sharing. In other terms, a directorate governs an open-ended cooperative game *in the large*. Play unfolds, however, *in the small*, among resource users, like a sequence of competitive equilibria. At each stage, money makes for transferable utility, and it facilitates exchange of quotas or information.

The motivation comes from the prospects for improving efficiency. To that end, the paper goes far by way of decomposition across governance, parties,

[1] Establishing (or enforcing) rights is important, but not an issue discussed here.

[2] See "*Rainforests need laws, not saws*" in The Economist, 4th-10th March, 2023.

[3] Single-agent control dominates parts of the received literature (Gale, 1967; Majumdar & Roy, 2009; Peleg & Ryder, 1974; Rockafellar, 1976).

resources, species and times. The framing fits a resource-based society in which rights are respected, supervised and well defined. There is centralized control of aggregate take-outs. The economy moves iteratively, by efficient sharing of single-period quotas, and it heads towards a long-run steady state. Quota allocations are determined *as if* supported by market-clearing, price-taking behaviour (Flåm, 2020).[4] Consequently, all along the system trajectory, competitive equilibria prevail in quota markets, hence so do Pareto optimality and core solutions as well.[5]

These novelties appear importnt. They indicate constitutional and institutional reforms. However, the optic doesn't directly or easily fit public cost-benefit analysis (Millner & Heal, 2018).

This paper addresses economists concerned with management and use of renewable resources.[6] The main thrusts and novelties are *threefold*: *First*, invoking convex preferences, convoluted criteria and money, it indicates, by way of decomposition, good prospects for overall efficiency.[7] *Second*, it emphasises the chief roles of agencies: decentralized trade of quotas unfolds besides centralized programming of aggregate take-outs. *Third*, the paper adds to the single-agent, single-stock, institution-free setting which dominates much received literature.

The rest of the paper goes as follows. Section 2 presents preliminaries. Section 3 considers single-period sharing of total take-out. It is organized around efficient, *short-term allocation,* implemented and held up by shadow pricing of available resources. Such allocation is realized period after period. Most important, the outcomes form a sequence of contingent *competitive equilibria*.
Section 4 introduces the ecosystem, its dynamics and principal manager. Being science-based, the latter casts planning as a syndicated problem of dynamic programming. Section 5 considers a long-run steady state: a *golden rule in cooperative commons*. Section 6 concludes.

---

[4] Transactions can, but need not, be mediated by markets. Efficiency is described here *as though* facilitated and upheld by clearing prices. In terms of strategic behavior, single-period equilibrium might emerge as a Coase or "folk" theorem (Dutta, 1995).

[5] In principle, there is *open access* to quota markets but *not* directly to the natural resources at hand. Resource rent won't be dissipated but rather restored and shared.

[6] It connects to game theory, optimization and operation research, but no links are spelled out.

[7] For brevity and readability - and for political science - technical prerequisites are minimal.

2. PRELIMINARIES

Each vector space is real, finite-dimensional - with standard partial order $\leq$, "dot $\cdot$" inner product, and associated norm $\|\cdot\|$. Multi-dimensional *quotas* form such a space $\mathbb{Q}$.[8] Any vector *"quantity"* $q \in \mathbb{Q}$ is seen as a bundle of consumption (or harvest), taken out, during a single season, from one and the same local habitat of natural, renewable resources.

$q^* \in \mathbb{Q}^*$ is shorthand for a linear *price* system on $\mathbb{Q}$. Thus, under price-taking behavior, take-out $q \in \mathbb{Q}$ is worth $q^*q := q^*(q) := q^* \cdot q \in \mathbb{R}$.

Each economic agent, considered below, operates with a *pecuniary payoff function* $\pi : \mathbb{Q} \to \mathbb{R} \cup \{-\infty\}$ of his own. The value $\pi(q) = -\infty$ serves as a conceptual but convenient device to account for implicit constraints. It precludes infeasible or nonsensical choice q. Declare a payoff function $\pi(\cdot)$ *proper* iff finite-valued somewhere, meaning $dom(\pi) := \{q \mid \pi(q) \in \mathbb{R}\}$ is non-empty.

As usual, optimality conditions invoke differentiability. That notion is generalized here - and global in nature:

Definition 1. (Generalized derivatives) *A payoff criterion $\pi : \mathbb{Q} \to \mathbb{R} \cup \{-\infty\}$ has a **supgradient** $q^* \in \mathbb{Q}^*$ at $q \in \mathbb{Q}$, written $q^* \in \partial\pi(q)$, iff q maximizes the function $\hat{q} \mapsto \pi(\hat{q}) - q^*\hat{q}$ with finite value. The set $\partial\pi(q)$ of all such supgradients is called the **supdifferential**.*

$\pi(q) = -\infty$ makes $\partial\pi(q)$ empty. Otherwise, provided $\pi(q)$ be finite,

$$q^* \in \partial\pi(q) \Longleftrightarrow \pi(\hat{q}) \leq \pi(q) + q^*(\hat{q} - q) \text{ for all } \hat{q} \in \mathbb{Q}. \quad (1)$$

By (1), the set $\partial\pi(q)$ is closed convex. When a concave $\pi(\cdot)$ is classically differentiable, $\partial\pi(q)$ reduces to the customary gradient $\pi'(q)$. For both practical computation and realism, it is important though, to accommodate non-smooth instances:

Example 1. (Generating payoff by programs) *Suppose that a resource user, while holding bundle $q \in \mathbb{Q}$ as input, faces a linear primal problem:*

$$\pi(q) := \max_y \{y^*y \mid Ay \leq q \;\&\; y \in \mathbb{Y}_+\},$$

presumed solvable with finite optimal value. His output y belongs to a Euclidean space $\mathbb{Y}$, and $y^* \in \mathbb{Y}^*$ prices output linearly. The matrix A maps $\mathbb{Y}$

[8] In geometrical terms, $\mathbb{Q}$ is Euclidean.

into $\mathbb{Q}$. If A^* denotes the transposed matrix, a price $q^* \in \mathbb{Q}^*$, not necessarily unique, belongs to $\partial\pi(q)$ iff it solves the dual problem:

$$\pi(q) = \min_{q^*} \left\{ q^* q \;\middle|\; A^* q^* \geq y^* \;\&\; q^* \in \mathbb{Q}^*_+ \right\}. \tag{2}$$

For any subset $\mathscr{R}$ of the real numbers $\mathbb{R}$, its *supremum*, denoted $\sup\mathscr{R}$ is the smallest $\bar{r} \in \mathbb{R} \cup \{\pm\infty\}$ which is $\geq$ each $r \in \mathscr{R}$. [By convention here, $\sup\varnothing = -\infty$.] *Infimum*, denoted $\inf\mathscr{R}$, equals $-\sup(-\mathscr{R})$. The following derives from elementary convex analysis:

Proposition 1. (On price-taking profit) *Suppose that a resource user, who has a proper and pecuniary payoff function $\pi : \mathbb{Q} \to \mathbb{R} \cup \{-\infty\}$, faces cost regime $q^* \in \mathbb{Q}^*$ when taking out any $q \in \mathbb{Q}$.*
• Then, his (price-taking) ***competitive profit function*** *is given by*

$$q^* \in \mathbb{Q}^* \mapsto \pi^*(q^*) := \sup\left\{ \pi(q) - q^* q \mid q \in \mathbb{Q} \right\}, \tag{3}$$

is closed[9] *and convex;*
• The supremum $\pi^*(q^*)$ in (3) is attained as maximum at q iff $q^* \in \partial\pi(q)$. Then, $q^* \in \partial\pi(q) \Longleftrightarrow \pi(q) = \pi^*(q^*) + q^* q$, meaning that payoff $\pi(q)$ splits into profit $\pi^*(q^*)$ plus competitive factor cost $q^* q$;
• If the payoff function $\pi(\cdot)$ is concave and bounded below near $q \in \mathbb{Q}$, the supdifferential $\partial\pi(q)$ is non-empty;
• Reasonably, $\pi(0) \geq 0$, and in that case, $\pi^*(q^*) \geq 0$ for any $q^* \in \mathbb{Q}^*$;
• If the agent faces user cost q^*, and holds private "property" $\underline{q} \in \mathbb{Q}$, he may aim at taking home $\pi^*(q^*) + q^* \underline{q} = \pi(q) + q^*(\underline{q} - q)$, composed of pure profit $\pi^*(q^*)$ plus "resource rent" $q^* \underline{q}$.

3. EFFICIENCY WITHIN SEASON

The paper attempts to decompose a *long-term* problem of resource management into *short-term* parts. This section begins by considering allocation of the seasonal total quota(s) across legitimate parties. Such quotas are presumed perfectly divisible, marketable and transferable - with no externalities, fees or frictions.

[9] It's *lower semicontinuous*, meaning that $\{q^* \in \mathbb{Q}^* \mid \pi^*(q^*) \leq r\}$ comes closed for each $r \in \mathbb{R}$.

Emphasis is here on efficient sharing during any fixed season (Flåm, 2020). An impatient reader, chiefly concerned with intertemporal allocation and long run, may first skip this section and later return to it.

Accommodated henceforth is a fixed finite ensemble $\mathscr{I}$ of economic agents, $\#\mathscr{I} \geq 2$, construed as resource users. In each period, member $i \in \mathscr{I}$ applies the same proper, pecuniary *payoff function* $\pi_i : \mathbb{Q} \to \mathbb{R} \cup \{-\infty\}$, and he targets maximal profit.

Absent externalities *in season*, arguments as to short-term *allocative efficiency* will revolve around *convolution* (4) of individual payoffs:

Definition 2. (Supremal convolution) *If each $i \in \mathscr{I}$ has a proper payoff function $q_i \in \mathbb{Q} \mapsto \pi_i(q_i) \in \mathbb{R} \cup \{-\infty\}$, their sup-convoluted payoff function is defined by*

$$q_{\mathscr{I}} \in \mathbb{Q} \mapsto \pi_{\mathscr{I}}(q_{\mathscr{I}}) := \sup_{(q_i)} \left\{ \sum_{i \in \mathscr{I}} \pi_i(q_i) \;\middle|\; \sum_{i \in \mathscr{I}} q_i = q_{\mathscr{I}} \right\}. \tag{4}$$

If some $\pi_i(\cdot)$ increases on $dom\pi_i$, then

$$\pi_{\mathscr{I}}(q_{\mathscr{I}}) = \sup_{(q_i)} \left\{ \sum_{i \in \mathscr{I}} \pi_i(q_i) \;\middle|\; \sum_{i \in \mathscr{I}} q_i \leq q_{\mathscr{I}} \right\}.$$

With each $\pi_i(\cdot)$ (quasi-)concave, $\pi_{\mathscr{I}}(\cdot)$ also becomes (quasi-)concave.

Proposition 2. (On efficient allocation by equal margins (Flåm, 2020, 2021b)) *For any best choice (q_i) in (4) it holds*

$$\partial \pi_{\mathscr{I}}(q_{\mathscr{I}}) \subseteq \cap_{i \in \mathscr{I}} \partial \pi_i(q_i).$$

Conversely, provided $q_{\mathscr{I}} = \sum_{i \in \mathscr{I}} q_i$, it also holds the turned-around inclusion:

$$\partial \pi_{\mathscr{I}}(q_{\mathscr{I}}) \supseteq \cap_{i \in \mathscr{I}} \partial \pi_i(q_i).$$

If moreover, $\cap_{i \in \mathscr{I}} \partial \pi_i(q_i)$ is non-empty, then (q_i) solves (4).

In this section, but this only, function $\pi_{\mathscr{I}}(\cdot)$ (4) is mainly a conceptual construct, serving analysis. No member $i \in \mathscr{I}$ states problem (4). Indeed, maybe not knowing more than his own criterion $\pi_i(\cdot)$ - or, seeing just a local approximation to it - he can neither identify $\pi_{\mathscr{I}}(\cdot)$ nor solve any particular instance $\pi_{\mathscr{I}}(q_{\mathscr{I}})$ (4). It suffices though, that the agents together solve problem (4) themselves; see Remark 2 below.

It is worth mentioning that Proposition 2 presumes no concavity of objectives. It just needed that $\partial\pi_{\mathscr{I}}(q_{\mathscr{I}})$ be non-empty.[10]

Added here are some results on *Pareto optimality*, *core outcome* and *competitive equilibrium*. As is well known, these concepts connect to theory on games and economic welfare (Luenberger, 1995; Osborne & Rubinstein, 1994). Their relevance for resource economics appears less noticed:

Proposition 3. (On single-season Pareto optimality, core solution and competitive equilibrium (Flåm, 2021b, 2023))

• *Any solution* (q_i) *to (4) is a* ***Pareto optimal*** *allocation of* $q_{\mathscr{I}}$.

• *If agent* $i \in \mathscr{I}$ *already "owns"* $\underline{q}_i$, *"coalition"* $I \subseteq \mathscr{I}$ *can - by going alone, in autarky - aim at no less joint payoff than*

$$\underline{\pi}_I(\underline{q}_I) := \sup_{(q_i)} \left\{ \sum_{i\in I} \pi_i(q_i) \;\middle|\; \sum_{i\in I} q_i = \sum_{i\in I} \underline{q}_i =: \underline{q}_I \right\}.$$

So, for any ***shadow price*** $q^* \in \partial\pi_{\mathscr{I}}(q_{\mathscr{I}})$ *on actual resource use, the cash payment profile* $i \in \mathscr{I} \mapsto \kappa_i(q^*) := \pi_i^*(q^*) + q^*\underline{q}_i$ *constitutes a* ***core solution*** *in that*

$$\sum_{i\in I} \kappa_i(q^*) \geq \underline{\pi}_I(\underline{q}_I)$$

for each $I \subset \mathscr{I}$ *with equality for* $I = \mathscr{I}$ *and* $\underline{q}_{\mathscr{I}} = q_{\mathscr{I}}$.

• *Still suppose agent* $i \in \mathscr{I}$ *"owns"* $\underline{q}_i$ *with* $\sum_{i\in\mathscr{I}} \underline{q}_i = \underline{q}_{\mathscr{I}}$. *Then, any shadow price* $q^* \in \partial\pi_{\mathscr{I}}(\underline{q}_{\mathscr{I}})$, *alongside any optimal allocation* (q_i) *of* $\underline{q}_{\mathscr{I}} = q_{\mathscr{I}}$ *to (4) constitutes a* ***competitive equilibrium*** *in that the quota market clears, and agent i takes home maximal total profit* $\pi_i^*(q^*) + q^*\underline{q}_i$.[11]

Clearly, activities, quotas q_i or rents $q^*\underline{q}_i$ can be distributed rather unevenly.

Concluding this section is a summary on how resource users enter a season and proceed therein:

[10] For that, presence of many and minor agents may help. While (4) preserves concavity, it may also contribute to "create" it.

[11] These items could - or ought - be taxed.

Assumption 1. (On sharing of seasonal quotas)
(I) At the very beginning of a season, ***prior to any trade of quotas****, the principal sets a total quota $q_{\mathscr{I}}$. Immediately thereafter $q_{\mathscr{I}}$ is split among legitimate users by some time-invariant rule*

$$i \in \mathscr{I},\ q_{\mathscr{I}} \in \mathbb{Q} \mapsto \underline{q}_i = Q_i(q_{\mathscr{I}}), \quad \sum_{i \in \mathscr{I}} Q_i(q_{\mathscr{I}}) = q_{\mathscr{I}}. \tag{5}$$

(II) Subsequently, but ***prior to any use of quotas****, the said users settle on a profile $i \in \mathscr{I} \mapsto q_i \in \mathbb{Q}$ which solves (4).*

Remark 1. (On allotted shares) $\underline{q}_i$ in (5) can reflect grand-fathered or traditional rights; see Flåm (2020). If $\underline{q}_i \neq 0$ and $q_i = 0$, agent i just owns rights, but uses none. He only collects rent. Conversely, if $\underline{q}_i = 0$ and $q_i \neq 0$, being "propertyless," he fully rents his user rights. When $\underline{q}_i \neq 0$ and $q_i \neq 0$, agent i acts in twin capacities: as owner and user. Part (I) of Assumption 1 is *constitutional*. It relates to established law, presumed here, but not discussed.

Remark 2. (On allocation of allotted shares) Part (II) of Assumption 1 is *institutional*. It points to auctions, bargaining, barter *quid pro quo*, direct deals or markets. These institutions or platforms may help agents to solve (4) by themselves, but no mechanism is singled out or modelled here; see Flåm (2021b, 2023).[12]

4. EFFICIENCY ACROSS SEASONS

Henceforth, suppose quotas will be traded, *in each season*, for money, up to price-supported Pareto efficiency (Proposition 2) - in fact, up to single-period, competitive equilibrium (Proposition 3). On that premise, this section considers how a harvest might be allocated *across seasons*?

The ecosystem comprises a finite set S of renewable resources or species $s \in S$. Let the vector space $\mathbb{X} = \mathbb{R}^S$ comprise all bundles $x = (x_s)$. A *system state* $x = (x_s) \in \mathbb{X}_+$ informs about the actual biomass $x_s \geq 0$ of each species $s \in S$. That state remains in some rectangle set $X := [\underline{x}, \bar{x}] \subset \mathbb{X}$, bounded by sustainable stock levels $0 \leq \underline{x}_s < \bar{x}_s$, $s \in S$.

[12] Circumvented here are issues on control of prices versus quantities (Weitzman, 1974).

Natural growth $x \in \mathbb{X}_+ \mapsto G_s(x) \in \mathbb{R}$ of species $s \in S$ - *after harvest, during the season* - is presumed concave. Let $G(x) := [G_s(x)]_{s \in S} \in \mathbb{X}$ and $g(x) := G(x) - x$.[13]

Most often, *the state affects payoffs*.[14] So, from here onwards, $\pi_i(x, q_i)$ takes the place of the above simplified version $\pi_i(q_i)$. By assumption, $(x, q_i) \mapsto \pi_i(x, q_i)$ is jointly concave. It's separately differentiable and increasing in $x \in X$, but maybe neither in q_i; see Example 1.

Naturally, for the quota space, introduced earlier, let $\mathbb{Q} := \mathbb{X}$, and posit $Q := X$ for the set of short-term total quotas.

Time $t \in T := \{0, 1, ..\}$ is discrete, with open horizon. An initial point $x_{-1} \in X$ is specified.

Assumption 2. (On the principal's long-term program) *Given initial point* $x_{-1} \in X$, *discount factor* $\delta \in (0,1)$, *and convoluted, single-period payoff function*

$$t \in T \mapsto (x_t, q_t) \in X \times Q \mapsto \pi(x_t, q_t) := \sup \left\{ \sum_{i \in \mathscr{I}} \pi_i(x_t, q_{it}) \;\middle|\; \sum_{i \in \mathscr{I}} q_{it} = q_t \right\}, \quad (6)$$

the principal will

$$\textit{maximize present value } \sum_{t \in T} \delta^t \pi(x_t, q_t) \;\; \textit{s.t.} \;\; x_{t+1} \leq G(x_t) - q_t \;\; \forall t \in T. \quad (7)$$

Proposition 4. (Existence of optimal profiles) *Suppose single-period payoff* π *in (6) and growth* g *both be upper semicontinuous. If feasible, problem (7) has a best solution.*

Program (7) is an instance of deterministic, discrete-time optimal control. For discussion, let multiplier vector $\lambda_t \in \mathbb{R}^S_+$ value the time-t *excess* $G(x_t) - x_{t+1} - q_t \geq 0$. Thus emerges a *Lagrangian*

$$L(\mathbf{x}, \mathbf{q}, \lambda) := \sum_{t \in T} \delta^t \left\{ \pi(x_t, q_t) + \lambda_t [G(x_t) - x_{t+1} - q_t] \right\}. \quad (8)$$

In (8), the *primal planning profiles* $\mathbf{x} := (x_t)$ and $\mathbf{q} = (q_t)$ both belong to the space l^∞ of bounded sequences in $\mathbb{X}$.[15] So, any *dual price profile* λ should

[13] Ecological interdependence is possible and quite likely. Reasonably, $x_s < \underline{x}_s \Longrightarrow g_s(x) \leq 0$, whence the system collapses. (Moreover, $g_s(\cdot)$ needs not be concave when $x_s \in [0, \underline{x}_s]$; see Majumdar & Roy (2009) and references therein.) By contrast, $x_s > \bar{x}_s \Longrightarrow g_s(x) < 0$ so that viable biomasses and stocks stay bounded. If left alone, a pristine system would evolve by the law $\Delta x := x_{+1} - x = g(x)$.

[14] For instance, use a density-dependent matrix $A(x)$ in Example 1.

[15] In fact, each $(x_t, q_t) \in X \times Q$.

be a continuous linear mapping from l^∞ into $\mathbb{R}$. As a matter of natural and reasonable modelling, take $\lambda = (\lambda_t)$ to be a member of the linear space

$$l^1 := \left\{ \lambda = (\lambda_t) \;\middle|\; \sum_{t \in T} \delta^t \|\lambda_t\| < +\infty \right\}. \tag{9}$$

It comprises precisely those price regimes $\lambda = (\lambda_t)$ that associate finite present value $\lambda \mathbf{x} := \sum_{t \in T} \delta^t \lambda_t x_t$ to any bounded sequence $\mathbf{x} = (x_t)$ in $\mathbb{X}$.[16]

For the subsequent argument, introduce the *Hamiltonian*

$$H(x,q,\lambda) := \pi(x,q) + \lambda[g(x) - q] \tag{10}$$

to account for aggregate payoff plus shadow pricing λ of net savings $g(x) - q$.

Theorem 1. (On the principal's long-term control) *Suppose $t \in T \mapsto (x_t, q_t) \in X \times Q$ solves (7). Then there is an adjoint, dual trajectory of shadow prices $t \in T \mapsto \lambda_t \in \mathbb{R}^S_+$ on resources such that for every $t \in T$ and (x_t, q_t) it holds*

$$\begin{cases} \textit{the system dynamics:} & x_{t+1} & \leq & G(x_t) - q_t, \\ \textit{the adjoint equation:} & \lambda_t - \delta^{-1}\lambda_{t-1} & \in & -\frac{\partial}{\partial x} H(x_t, q_t, \lambda_t), \textit{ and} \\ \textit{the maximum condition:} & q_t & \in & \arg\max H(x_t, \cdot, \lambda_t), \end{cases} \tag{11}$$

with complementarity $\lambda_t[G(x_t) - x_{t+1} - q_t] = 0$, and specified initial point x_{-1}. Moreover, at each time $t \in T$, the valuation vector $\lambda_t \in \mathbb{R}^S_+$ - on resources saved in situ - equals a market equilibrium price vector $q^*_t \in \mathbb{Q}^*$ on resources actually consumed. Thus, Assumption 1, part II, is satisfied with

$$\lambda_t = q^*_t \in \frac{\partial}{\partial q_t} \pi(x_t, q_t) = \cap_{i \in \mathscr{I}} \frac{\partial}{\partial q_{it}} \pi_i(x_t, q_{it}). \tag{12}$$

Proof. Recall Hamiltonian H (10) to rewrite the Lagrangian (8) as

$$L(\mathbf{x}, \mathbf{q}, \lambda) = \sum_{t \in T} \delta^t \left\{ H(x_t, q_t, \lambda_t) - \lambda_t (x_{t+1} - x_t) \right\}.$$

[16] By the *Riesz representation theorem* l^1 has l^∞ as dual space, but not conversely (Dudley, 1989).

By $\frac{\partial}{\partial \lambda_t} L(\mathbf{x}, \mathbf{q}, \lambda) \geq 0$, the system dynamics hold for all $t \in T$. Also, the complementarity conditions come up as usual. Note that L features state x_t just *two* times. The corresponding terms are singled out next:

$$L(\mathbf{x}, \mathbf{q}, \lambda) = \cdots +$$

$$\delta^{t-1} \{H(x_{t-1}, q_{t-1}, \lambda_{t-1}) - \lambda_{t-1}(x_t - x_{t-1})\} +$$

$$\delta^t \{H(x_t, q_t, \lambda_t) - \lambda_t(x_{t+1} - x_t)\} + \cdots .$$

From this and $\frac{\partial}{\partial x_t} L(\mathbf{x}, \mathbf{q}, \lambda) = 0$, the adjoint equation follows. Also, because $L(\mathbf{x}, \mathbf{q}, \lambda)$ is additively separable with respect to $q_t, t \in T$, and should be maximized in each q_t, the maximum condition is immediate. Taken together, $\frac{\partial}{\partial q_t} H(x_t, q_t, \lambda_t) = 0$ and Proposition 2 give (12). □

Remark 3. (On qualified constraints) Normally, proper use of the Lagrangian requires some *constraint qualification*. Granted concave functions π and G, as here, the most convenient one - called the *Slater condition* - amounts to strict feasibility. Specifically, suppose the resource system be *productive*, meaning that some state $x \in X$ can be reached at which $G_s(x) > x_s$ for each $s \in S$.

Remark 4. (On value functions, differentiability and duality) Letting

$$\pi(x, x_{+1}) := \sup_q \{\pi(x, q) \mid x_{+1} \leq G(x) - q\},$$

the overall, joint program has optimal value function

$$x \mapsto V(x) := \sup_{\mathbf{x}} \sum_{t \in T} \delta^t \pi(x_t, x_{t+1}), \;\; x_0 = x.$$

Under standard conditions, it holds the Bellman equation

$$V(x) = \max_{x_{+1}} \{\pi(x, x_{+1}) + \delta V(x_{+1}) \mid x_{+1} \in X\},$$

and $\partial V(x) = \partial_x \pi(x, x_{+1})$ with supposedly unique and optimal continuation x_{+1} (Benveniste & Scheinkman, 1979).

5. GOLDEN RULE

Does problem (7) have a globally stable, steady state? For this, it's expedient that both functions π, G be *strongly concave*. For G, this means that $[g'(x)-g'(\tilde{x})]\cdot[x-\tilde{x}] \leq -\mu_g \|x-\tilde{x}\|^2$ with some modulus $\mu_g > 0$. Further, as stressed by Weitzman (1998), δ should be "large enough" - or tuned to - the associated moduli μ_π, μ_g of strong concavity (Rockafellar, 1976).[17] Here, for brevity, stability is simply presumed:

Assumption 3. (On asymptotic stability) *Optimal control leads to a* ***golden steady state*** $x = \lim_{t\to+\infty} x_t$ *- a long-term, fixed point of the system - presumed unique and well defined.*

To characterize the golden state, let

$$\mathscr{H}(x,\lambda) := \max\{H(x,q,\lambda) : q \in \mathbb{Q}\} \tag{13}$$

be the *reduced* Hamiltonian. Define its composite differential by

$$(x^*,\lambda^*) \in \partial\mathscr{H}(x,\lambda) := \partial_x\mathscr{H}(x,\lambda) \times \partial_\lambda\mathscr{H}(x,\lambda) \Longleftrightarrow$$

$$\begin{cases} \mathscr{H}(\hat{x},\lambda) \leq \mathscr{H}(x,\lambda) + x^*(\hat{x}-x) & \text{for all } \hat{x} \in \mathbb{X}, \\ \mathscr{H}(x,\hat{\lambda}) \geq \mathscr{H}(x,\lambda) + \lambda^*(\hat{\lambda}-\lambda) & \text{for all } \hat{\lambda} \in \Lambda. \end{cases}$$

The first inequality captures that $\mathscr{H}(x,\lambda)$ is concave in x. Hence ∂_x denotes a partial supdifferential. The second inequality tells that $\mathscr{H}$ is convex in λ.[18] In these terms, by Theorem 1 and (13), the controlled system moves by

$$x_{t+1} - x_t \in \partial_\lambda \mathscr{H}(x_t,\lambda_t) \quad \& \quad \lambda_t - \delta^{-1}\lambda_{t-1} \in -\partial_x\mathscr{H}(x_t,\lambda_t). \tag{14}$$

Hamiltonian dynamics (14) resemble classical studies of primal-dual programming. As said, stability and limit behavior link to strong concavity (Rockafellar, 1976), notably in state variable x (Hauswith et al., 2020), derived from growth G and payoffs π.

[17] Invoking uncertainty as to the far distant future, (Weitzman, 1998) argues that the largest possible discount factor should be used.

[18] Accordingly, $\partial_\lambda[\cdot] := -\partial_\lambda[-\cdot]$ is the partial *subdifferential*.

Proposition 5. (On existence of a steady state) *Under natural assumptions - on π, g being closed concave on compact, convex domains - system (14) has at least one fixed point (x, λ) corresponding to the stationary version*

$$[-(1-\delta^{-1})\lambda, 0] \in \partial \mathscr{H}(x, \lambda)$$

of (14) - with steady harvest $q = g(x)$, and discount factor $\delta = 1/(1+\rho)$ defined by interest rate $\rho > 0$. Further, it holds Hotelling's modified rule:

$$\rho\lambda \in \frac{\partial}{\partial x}\mathscr{H}(x, \lambda). \tag{15}$$

Proof. From (14), the convex-valued point-to-set correspondence $(x, \lambda) \rightrightarrows (x_+, \lambda_+)$ defined by

$$x_+ \in x + \partial_\lambda \mathscr{H}(x, \lambda) \quad \& \quad \lambda_+ \in \delta^{-1}\lambda - \partial_x \mathscr{H}(x, \lambda) \tag{16}$$

is well defined, with closed graph and non-empty values, on some compact convex suitably chosen state space $\Xi \subset \mathbb{X}_+ \times \mathbb{X}_+$. Each "value" - that is, each right hand side of (16) - intersects Ξ. Hence by Kakutani's theorem the said correspondence has a fixed point.[19] $\square$

Proposition 6. (Decomposed golden rule) *The steady state x gives constant user cost $q^* \in \frac{\partial}{\partial q}H(x, q, \lambda)$, shadow price $\lambda = q^* \in \frac{\partial}{\partial x}H_x(x, q, \lambda)/\rho$ on resources saved, and fixed total take-out $g(x) = q \in \arg\max H(x, \cdot, \lambda)$. Any optimal allocation (q_i) of the stable aggregate take-out q solves*

$$\pi(x, q) := \sup_{(q_i)} \left\{ \sum_{i \in \mathscr{I}} \pi_i(x, q_i) \;\middle|\; \sum_{i \in \mathscr{I}} q_i = q \right\}. \tag{17}$$

Whatever price $q^* \in \partial_q \pi(x, q)\big|_{q=g(x)}$ on resource use, agent i gets price-taking maximal profit

$$\pi_i^*(x, q^*) := \sup_{\hat{q}_i} \{\pi_i(x, \hat{q}_i) - q^*\hat{q}_i \mid \hat{q}_i \in \mathbb{Q}\} = \pi_i(x, q_i) - q^* q_i.$$

If he can claim property or user right to $\underline{q}_i$, $\sum_{i \in \mathscr{I}} \underline{q}_i = g(x)$, he takes home overall profit $\pi_i^*(x, q^*) + q^*\underline{q}_i$. In the steady state, an owner of right $\underline{q}_i$ holds capital value $q^*\underline{q}_i/(1-\delta)$.

[19] For existence of a steady state, see also Peleg & Ryder (1974).

Returning to the non-reduced Hamiltonian $H(x,q,\lambda)$ (10), a steady solution (x,q,λ), with $\lambda \in \mathbb{R}^S_{++}$, features

$$\begin{cases} \textit{stationary stock levels:} & g(x) = q = \sum_{i\in\mathscr{I}} q_i, \\ \textit{Hotelling's rule:} & \rho\lambda \in \partial_x H(x,q,\lambda), \ \textit{and} \\ \textit{balanced shadow pricing:} & \lambda \in \partial_q \pi(x,q) \subseteq \cap_{i\in\mathscr{I}} \partial_{q_i}\pi_i(x,q_i). \end{cases} \tag{18}$$

If a function is differentiable, the inclusion sign in (18) should be replaced by equality.[20]

Clearly, with fixed agent ensemble $\mathscr{I}$, payoff functions π_i, state x, and seasonal quota q, convoluted outcome $\pi(x,q)$ (17) *is unaffected by the property rights* in Assumption 1 (I). However, for any specified state x, admission of extra agents, each with $\pi_i(x,0) \geq 0$ - that is, an enlargement of $\mathscr{I}$ - cannot but increase (or at least maintain) π. Thus, the incumbents might see potential entrants.

Ending this section are some second thoughts on the public cost-benefit criterion (7) - and on the principal's access to necessary information.

Remark 5. (On principal discounting) Suppose that agent $i \in \mathscr{I}$ prefers discount factor $\delta_i \in (0,1)$. If $\mathscr{D} := \{\delta_i \mid i \in \mathscr{I}\}$ reduces to a singleton, no issues emerge. Otherwise, many papers consider collective choices among time profiles (Chambers & Echenique, 2018; Harstad, 2020; Jackson & Yariv, 2015; Millner & Heal, 2018). Main objects of study there are social preference orders $\succsim$ on bounded flows $\mathbf{r} = (r_t)$ of some *single and common commodity*. This one-dimensional perspective may fit if the time-t common item $r_t \in \mathbb{R}$ denotes *joint* revenue - or, it represents the amount of some single resource. Here, however, monetary revenue $r_{it} = \pi_i(x_t, q_{it})$ is *private* - or, there are several resources. So, the setting doesn't directly fit *public* cost-benefit analysis.

It's desirable, though, that the principal's order be representative. For that, suppose agent $i \in \mathscr{I}$ enjoys increasing, twice continuously differentiable utility $u_i(r)$ of single-period revenue $r \in \mathbb{R}$. Then, in case all u_i are equal,

[20] System (18) captures directly the steady state. It would be vastly more challenging to find first, the entire value function $x \in X \mapsto V(x) = \max\{\pi(x,q) + \delta V(G(x) - q)\}$ by dynamic programming, invariably using convoluted payoff (17), and finally, to require that $g(x) = q$. Usually, computation requires discretization of spaces, whence clustering techniques.

Jackson & Yariv (2015) prove that the principal may use a discount factor $\delta = \sum_{i \in \mathscr{I}} w_i \delta_i$, with weights $(w_i) > 0$. summing to one. Thereby *time-consistency* obtains, and *unanimity* prevails, in that $\mathbf{r} \succsim \hat{\mathbf{r}} \Rightarrow \sum_t \delta_i^t u_i(r_t) \geq \sum_t \delta_i^t u_i(\hat{r}_t)$ for each $i \in \mathscr{I}$.

By contrast, when the single-period utility functions u_i differs, Jackson & Yariv (2015) show that time-consistency requires a dictatorial (or paternalistic) choice $\delta \in \mathscr{D}$.

For these reasons, following Weitzman (1998), patience speaks for itself. That is, a prudent principal could well choose $\delta = \max \mathscr{D}$.[21] This choice also fits concerns with stability, mentioned above.

Remark 6. (On principal information) For his planning the principal needs a firm grip on the function $(x,q) \mapsto \pi(x,q)$. Clearly, if each agent $i \in \mathscr{I}$ honestly hands in his function π_i - say, by a double auction - once and for all,[22] no problems emerge. Otherwise, the principal must somehow estimate, know or synthesize π in (4).

Ignored here, or taken as given, is strategic communication. Even when such mode of play is absent or unimportant, challenges remain. To illustrate, returning to Example 1, suppose agent i faces, time and again, the program form

$$\pi_i(x, q_i) := \sup \{ y_i^* y_i \mid A_i(x) y_i \leq q_i \ \& \ y_i \in \mathbb{Y}_{i+} \}.$$

He knows $y_i^* \in \mathbb{Y}_i^*$ and the state-dependent "technology" $A_i(x) : \mathbb{Y}_i \to \mathbb{R}^S$. By contrast, if the principal knows or sees neither, he can hardly synthesize the corresponding criterion:

$$(x,q) \mapsto \pi(x,q) = \sup \left\{ \sum_{i \in \mathscr{I}} y_i^* y_i \;\middle|\; \sum_{i \in \mathscr{I}} A_i(x) y_i \leq q \text{ and } y_i \in \mathbb{Y}_{i+} \right\}.$$

Clearly, his task simplifies when the stock x impacts no technology hence no payoff. In that case, the principal might use time series $t \mapsto [q_t,\ \pi(q_t),\ q_t^* \in \partial \pi(q_t)]$, observed so far - say, up to time τ - to overestimate $\pi(\cdot)$, step by step and "on line," as a one-sided, time-t envelope:

$$\text{For } \tau \in T \text{ and } q \in Q \text{ posit } \pi_\tau(q) := \min_{t \leq \tau} \{ \pi(q_t) + q_t^*(q - q_t) \} \approx \pi(q),$$

---

[21] It's implicit here that rotation of (finitely lived) principals cause no time inconsistency (Harstad, 2020)

[22] Given any relevant $x \in X$, the function $q_i \mapsto \pi_i(x, q_i)$ could stem from parametric optimization on the part of agent i.

thus $\pi_\tau \geq \pi$.

6. CONCLUDING REMARKS

Considering management of common-property, renewable resources, this paper argues that three agencies can play chief and complementary roles.[23] At the upper level, a competent principal decides total take-outs over time. At the lower level, legitimate users share short-term aggregate quotas in time. Between the two levels, various mechanisms channel information and value resources.

This paper indicates that efficiency and stability may obtain. Indeed, it appears that private and public interests can be aligned, competitive equilibria being constituent components. To this end, the paper has assumed that

• information on payoffs is communicated to the principal by diverse market mechanisms;[24]
• user rights must be clearly defined, perfectly divisible, marketable and transferable (Flåm, 2020);
• qualified users operate, directly or indirectly, with monetary criteria (Flåm, 2021a; Luenberger, 1995);
• there are no single-period externalities and transaction costs;
• aggregate and optimal quotas obtain, via dynamic programming;
• and finally, discounting must be moderate (Rockafellar, 1976; Weitzman, 1998).

Then, modulo oversight and policing, a golden steady state may emerge as limit of short-term, competitive equilibria in quota markets. Clearly, questions remain as to convergence and stability. Others queries include: Who are qualified to which rights - and then, on what grounds (Flåm, 2020)? Who controls compliance or metes out penalties? When viewed as an integrated enterprise, might not the *principal* - or the greater public - tax resource rent?

Among other concerns, important but not considered here, *three issues* merit further study. *First*, do long-lived investments play crucial roles (Clark et al., 1979)?[25] *Second*, how and where should uncertainty be accounted for (Mitra & Roy, 2023; Stokey & Lucas, 1989)? *Third*, can eventual lack

[23] Likewise, *three* fields also enter: *Economic theory* on expediency of markets; *mathematical optimization* on relaxation methods, and *political science* on separation on powers.
[24] Double auctions appear most attractive (Flåm, 2021b).
[25] Ownership to highly specialized capital can be divorced from rental use.

of concavity hence occasional presence of increasing margins in payoffs or growth affect management considerably (Majumdar & Roy, 2009)?

References

Benveniste, L., & Scheinkman, J. A. (1979). On the differentiability of the value function in dynamic models of economics. *Econometrica*, *47*(3), 727–32.

Chambers, C. P., & Echenique, F. (2018). On multiple discount rates. *Econometrica*, *86*(4), 1325–46.

Clark, C. W., Clarke, F. H., & Munro, G. R. (1979). The optimal exploitation of renewable resource stocks: Problems of irreversible investment. *Econometrica*, *47*(1), 25–47.

Dudley, R. M. (1989). *Real Analysis and Probability*. Belmont, Calif: Wadsworth & Brooks/Cole.

Dutta, P. (1995). A folk theorem for stochastic games. *Journal of Economic Theory*, *66*, 1–32.

Flåm, S. D. (2020). Rights and rents in local commons. *Journal of Mechanism and Institution Design*, *5*(1), 119–40. doi: 10.22574/jmid.2020.12.005

Flåm, S. D. (2021a). *Market equilibria and money*. Fixed Point Theory Algorithms for Science and Engineering.

Flåm, S. D. (2021b). Towards competitive equilibrium by double auctions. *Pure and Applied Functional Analysis*, *6*, 1211–25.

Flåm, S. D. (2023). *By bid-ask spreads towards competitive equilibrium* (Vol. forthcoming). Journal of Convex Analysis.

Gale, D. (1967). On optimal development in a multisector economy. *Review of Economic Studies*, *34*, 1–19.

Harstad, B. (2020). Technology and time inconsistency. *Journal of Political Economy*, *128*, 2653–89.

Hauswith, A., Ortmann, L., Bolgnani, S., & D”orfler, F. (2020). Limit behavior and the role of augmentation in projected saddle flows for convex optimization. *IFAC-PapersOnLine,*, *53*, 5511–17.

Jackson, M. O., & Yariv, L. (2015). The necessity of time consistency. *American Economic Journal: Microeconomics*, *7*, 150–78.

Luenberger, D. G. (1995). *Microeconomic Theory*. New York: McGrawhill.

Majumdar, M., & Roy, S. (2009). Non-convexity, discounting and infinite horizon optimization. *Journal of Nonlinear and Convex Analysis*, *10*, 1–18.

March, J. G., & Olsen, J. P. (1989). *Rediscovering Institutions*. Macmillan, New York: The Free Press.

Millner, A., & Heal, G. (2018). Time consistency and time invariance in collective intertemporal choice. *Journal of Economic Theory*, *176*, 158–69.

Mitra, T., & Roy, S. (2023). Stochastic growth, conservation of capital and convergence to a positive steady state. *Economic Theory*, *76*, 311–51.

Osborne, M. J., & Rubinstein, A. (1994). *A Course in Game Theory*. London: MIT Press.

Peleg, B., & Ryder, H. E. (1974). The modified golden rule of a multi-sector economy. *Journal of Mathematical Economincs*, *1*, 193–98.

Rockafellar, R. T. (1976). Saddle points of hamiltonian systems in convex lagrange problem with a non-zero discount rate. *Journal of Economic Theory*, *12*, 71–113.

Stokey, N. L., & Lucas, R. (1989). *Recursive Methods in Economic Dynamics*. Harvard University Press.

Weitzman, M. L. (1974). Prices vs. quantities. *Review of Economic Studies*, *41*(4), 477–91.

Weitzman, M. L. (1998). Why the far-distant future should be discounted at its lowest possible rate. *Journal of Environmental Economics and Management*, *36*, 201–8.

Journal of Mechanism and Institution Design
ISSN: 2399-844X (Print), 2399-8458 (Online)
DOI:10.22574/jmid.2023.12.003

A MARKET DESIGN SOLUTION FOR MULTI-CATEGORY HOUSING ALLOCATION PROBLEMS

Gian Caspari

ZEW — Leibniz Centre for European Economic Research, Germany
gian.caspari@zew.de

ABSTRACT

We study multi-category housing allocation problems: A finite set of objects, which is sorted into categories of equal size, has to be allocated to a finite set of individuals, such that everyone obtains exactly one object from each category. We show that, in the large class of category-wise neutral and non-bossy mechanisms, any strategy-proof mechanism can be constructed by simply letting individuals choose an object from each category one after another following some priority order. We refer to these mechanisms as multi-category serial dictatorships and advocate for selecting priority orders across categories as fairly as possible.

Keywords: Matching, envy-free, multi-category housing allocation

JEL Classification Numbers: D44, D50.

1. INTRODUCTION

CONsider the problem of allocating $m \times n$ objects to n individuals based on the individuals' reported preference information over objects, such that every individual obtains a bundle containing m objects. Different solutions to this problem have been proposed by both researchers and practitioners.

I declare that there are no conflicts of interest. I thank the editor and the referees for their time and effort. I am especially thankful to the anonymous referee, who helped to improve the clarity of the proofs for Lemma 1 and Theorem 1. Any errors are my own.

On one hand, the theoretical literature has focused almost exclusively on strategy-proof mechanisms.[1] In this line of research (serial and sequential) dictatorship mechanisms stand out, as they are the only mechanisms that are simultaneously strategy-proof and (Pareto) efficient (Pápai, 2001; Klaus & Miyagawa, 2002; Ehlers & Klaus, 2003; Hatfield, 2009; Monte & Tumennasan, 2015):

Under a dictatorship mechanism individuals are assigned their most preferred bundle of m objects — among the remaining objects — one after another following some choosing order.

On the other hand, practitioners have proposed mechanisms that focus less on strategy-proofness and more on balancing both fairness and efficiency of the resulting allocation. In this context, an important class of fair and approximately (Pareto) efficient mechanisms are so called the draft mechanisms:

Under a draft mechanism individuals are assigned their most preferred object — among the remaining objects — one after another following some choosing order, which is reversed in each subsequent round until everyone has obtained m objects.[2]

However, how do these two classes of mechanisms compare? Using data on individuals reported as well as true preferences for the Harvard Business School course allocation, Budish & Cantillon (2012) show that draft mechanisms are indeed manipulated in practice and that these manipulations cause meaningful welfare losses. At the same time, they also find that, despite their shortcomings, draft mechanisms outperform dictatorship mechanisms in terms of welfare.[3]

---

[1] Under a strategy-proof mechanism truthful reporting of preferences over objects is a (weakly) dominant strategy for the individuals — allowing them to avoid costly and risky strategic behavior (Roth, 2008).

[2] Note that both draft and dictatorship mechanisms work in the same way if there are only as many objects as individuals ($m = 1$). For such single-object allocation problems (Hylland & Zeckhauser, 1979), also known as housing allocation problems, dictatorship mechanisms are the natural candidates arising, while if there are predefined property rights (Shapley & Scarf, 1974), also known as housing markets, top trading cycles mechanisms (also known as core mechanisms) are used to find an allocation. For more details, see for example Sönmez & Ünver (2011).

[3] Budish & Cantillon (2012) give the following intuition for this result: Under a dictatorship mechanism, individuals who get to pick early make their last choices independently of whether these objects would be some later-picking individuals' first choices; Individuals "callously disregard" the preferences of those who choose after them. This matters for wel-

These findings suggest that we should look at mechanisms that, akin to dictatorship mechanisms, (i) are strategy-proof while ensuring that, analogous to the allocations produced by draft mechanisms, (ii) the objects are distributed more equally (fairly) and are approximately efficient. Unfortunately, there exists a trilemma for multi-object allocation problems: Any mechanism satisfies at most two out of the three desired properties of strategy-proofness, fairness, and efficiency — even for some sensible weakenings/approximations of these properties (Caspari, 2020).[4]

In this paper, our main contribution is to show a positive result for an important special case of multi-object allocation problems — multi-category housing allocation problems: A total of $n \times m$ indivisible objects, which are sorted into m categories containing n objects each, must be allocated to a set of n individuals, based on the individuals' reported preference information over objects, such that everyone obtains exactly one object from each of the m categories. That is, we show that the presence of categories is sufficient for the existence of strategy-proof mechanisms producing fair and (approximately) efficient allocations:

Starting with one category, individuals are assigned their most preferred object — among the remaining objects — one after another following some choosing order which is reversed in each subsequent round and category until everyone has m objects — one from each category.

Intuitively, letting individuals choose objects one after another — analogous to draft mechanisms — can be implemented in a strategy-proof manner, as restricting individuals to choose from a given category each round removes any potential gains from strategic behavior. At the same time, efficiency is not lost as individuals actually want to have an object from every category. As a practical application, we consider the problem of allocating teaching-assistant positions to graduate students, where everyone has to assist exactly one spring-semester and one fall-semester course. Monte & Tumennasan (2015) analyse

fare as the benefit to the early-picking individuals from these last choices will generally be small relative to the harm these choices cause to the later-picking individuals.

4 One notable exception concerns large markets: Budish (2011) describes the approximate competitive equilibrium from equal incomes (ACEEI) mechanisms which are approximately efficient, fair, and strategy-proof if the size of the market makes participants price takers. ACEEI has been applied to course allocation problems (Sönmez & Ünver, 2010) and was successfully implemented at Wharton Business School (Othman et al., 2010; Budish et al., 2016).

a multi-category housing allocation problem with two categories, and discuss several practical applications, including benefit and assistance programs and the allocation of new physicians in the United Kingdom. Overall, depending on the application in mind, categories can either be interpreted as containing different types of objects, containing the same set of objects for different time periods, or some combination of both.

We formally introduce multi-category housing allocation problems in Section 2. Moreover, Section 2.1 introduces the necessary framework for individuals to report rankings over categories as opposed to their full preference relation over all possible bundles. Any particularities stemming from this model choice are discussed there.[5]

Section 3 contains our main theoretical result: Any strategy-proof, category-wise neutral, and non-bossy mechanism can be obtained by specifying a choosing order — referred to as a priority order — for each category. We refer to this class of mechanisms as multi-category serial dictatorships.[6]

Section 3.1 takes a look at two ways to select priority orders for multi-category serial dictatorships: On one hand, analogous to (serial) dictatorships, one can choose an identical order for each category — referred to as the subclass of identical priority multi-category serial dictatorships. Alternatively, analogous to the draft mechanism, one can select a priority order that is reversed in every other category, referred to as the subclass of fair priority multi-category serial dictatorships. As both are strategy-proof, we can solely compare these mechanisms in terms of fairness and efficiency. We show that identical priority multi-category serial dictatorships are (Pareto) efficient but extremely unfair (Proposition 1), while identical priority multi-category serial dictatorships achieve maximal fairness while being approximately efficient (Proposition 2). Moreover, we provide a discussion, that places Proposition 1 and 2 into the broader context of the literature on dictatorship mechanisms.

Section 4 discusses the implementation of a fair priority multi-category serial dictatorship to allocate spring-semester and fall-semester teaching positions to graduate students, while Section 5 concludes.

[5] The overall approach relates to Brams & Fishburn (2000), Brams et al. (2003), and Edelman & Fishburn (2001).

[6] The result can be seen as a generalization of a well-known characterization result by Svensson (1999) from housing allocation problems to multi-category housing allocation problems. Related, Monte & Tumennasan (2015) have shown that in multi-category housing allocation problems any strategy-proof, non-bossy, and Pareto efficient mechanism is a sequential dictatorship.

2. THE MULTI-CATEGORY HOUSING ALLOCATION PROBLEM

Let I be a finite set of $|I| = n$ **individuals**, O be a finite set of $|O| = m \times n$ **objects**, and K be a finite set of $|K| = m$ types or time periods. The set of objects can be partitioned into m different **categories** $(O^k)_{k \in K}$, each containing $|O^k| = n$ distinct objects of type k, or $|O^k| = n$ distinct objects from time period k, respectively.

Each individual has a **preference** relation $\succsim_i$ comparing all sets of objects that contain exactly one object from the same categories. Formally, let $\succsim_i$ be a partial order, such that we either have $O' \succsim O''$, $O'' \succsim O'$, or both if and only if $|O' \cap O^k| = |O'' \cap O^k| \leq 1$ for all $k \in K$.[7] Note that, we have implicitly assumed that preferences over objects (singleton sets) within each category are strict, while objects (singleton sets) from different categories cannot be directly compared with each other.

We want to distribute the available objects among the individuals such that every individual is assigned exactly one object from each category, and no two distinct individuals are assigned the same object. That is, a feasible **allocation** $A = (A_i)_{i \in I}$ assigns every individual $i \in I$ a set of objects A_i with $|A_i \cap O^k| = 1$ for all $k \in K$, and $A_i \cap A_j = \emptyset$ for all $i \in I$, $j \in I \setminus \{j\}$. Let $\mathscr{A}$ denote the set of all feasible allocations. Moreover, let $a_i^k \in O^k \cap A_i$ denote the object from allocation A in category O^k that is assigned to individual i.

We assume that preferences over allocations are **separable** in terms of categories: For all $A, A' \in \mathscr{A}$ and $i \in I$, if $\{a_i^k\} \succsim_i \{a_i'^k\}$ for all $k \in K$ then $A_i \succsim_i A_i'$, and if $\{a_i^k\} \succ_i \{a_i'^k\}$ for at least one $k \in K$ then $A_i \succ_i A_i'$. We denote the set of all separable preferences $\succsim_i$ for individual i by $\mathscr{Q}_i$.

As it is generally done, we assume there are no externalities, i.e., for all $i \in I$ we have that $A_i \succsim_i A_i'$ implies $A \succsim_i A'$. That is, any individual's preference over allocations solely depends on the object the individual is assigned.

[7] A partial order is reflexive, transitive, and antisymmetric binary relation. If either $O' \succsim O''$, $O'' \succsim O'$, or both we say that O' and O'' are comparable. Moreover, for any $O', O'' \subset O$ such that $O' \succsim_i O''$ but $O'' \not\succsim_i O'$, we say that O' is strictly preferred to O'' and write $O' \succ_i O''$. Finally, note that, anti-symmetry implies that for any $O', O'' \subset O$ such that $O' \succsim_i O''$ and $O'' \succsim_i O'$ we have $O' = O''$. In other words, preferences are strict whenever two comparable sets are not identical.

2.1. Rankings instead of Preferences

We assume that, instead of having to report their entire preferences, individuals simply need to report a separate **ranking** for each category. That is, for each $k \in K$, each individual $i \in I$ reports a transitive, asymmetric, and complete order P_i^k over O^k. We denote the associated transitive, antisymmetric, and strongly complete order by R_i^k.[8] The list containing all rankings of a single individual i is denoted by $P_i = (P_i^k)_{k \in K}$ and analogously the list containing all rankings of all individuals is denoted by $P = (P_i)_{i \in I}$. We let $P_{-i} = (P_j)_{j \in I \setminus \{i\}}$ denote the list containing all individuals' rankings, except individual i's rankings. Finally, the set of all possible lists of rankings for a single individual, all individuals, and all but one of the individuals are denoted by $\mathscr{P}_i$, $\mathscr{P}$, and $\mathscr{P}_{-i}$ respectively.

Abstracting away from truthful revelation of rankings for the moment, if P_i is reported we can narrow down the possible preference profiles i might have. That is, a preference $\succsim_i$ is **consistent** with the reported rankings P_i if it ranks objects in each category in the same way as the reported ranking. Formally, $\succsim_i \in \mathscr{Q}_i$ is consistent with P_i if for all $k \in K$, $o \in O^k$, $o' \in O^k$ we have $o \; R_i^k \; o'$ if and only if $\{o\} \succsim_i \{o'\}$. Let $\mathscr{Q}_{P_i} \subset \mathscr{Q}_i$ denote the subset of all separable preferences that are consistent with the reported rankings P_i.

Following the introduction of rankings, we now limit our attention to strategy-proof mechanisms that take a profile of rankings as input. Formally, a **ranking mechanism** $\psi : \mathscr{P} \to \mathscr{A}$ selects an allocation $A \in \mathscr{A}$ for any reported list of rankings $P \in \mathscr{P}$. Given a list of reported rankings P, we let $\psi(P)_i$ denote the set of objects obtained by individual i under mechanism ψ, and slightly abusing notation, we let $\psi(P)_i^k$ denote both the object as well as the singleton set containing the object obtained by individual i in category k under mechanism ψ. Moreover, a ranking mechanism is strategy-proof if an individual having preferences $\succsim_i \in \mathscr{Q}_{P_i}$ consistent with rankings P_i cannot benefit from reporting a different list of rankings $\hat{P}_i$ instead of P_i. That is, a ranking mechanism ψ is **strategy-proof** if for any $i \in I$, $P_i \in \mathscr{P}_i$, $\hat{P}_i \in \mathscr{P}_i$, $P_{-i} \in \mathscr{P}_{-i}$ we have

$$\psi(P)_i \succsim_i \psi(\hat{P}_i, P_{-i})_i \text{ for all } \succsim_i \in \mathscr{Q}_{P_i}.$$

[8] That is, unlike P_i^k which is asymmetric, R_i^k also compares any object in O^k with itself. Otherwise, both relations rank any two distinct objects in O^k in the same way. See, for example, Roberts (1985) for an overview on binary relations and their properties.

2.1.1. Dominance Relation

In this part, we discuss what we can infer about the preference of an individual based on her reported list of rankings. To fix ideas, consider any list of rankings P_i and two allocations A and A' such that for every object $a_i^k \in A_i$ of category O^k we can find a weakly lower ranked object from the same category in the other allocation $a_i'^k \in A_i'$. Note that, in this case A must be preferred to A' under any (separable) preference $\succsim_i \in \mathscr{Q}_{P_i}$ consistent with P_i.

We can generalize this idea by defining a partial order $\geq_{P_i}$ — referred to as a **dominance relation** — over the set of allocations $\mathscr{A}$ for any list of rankings P_i: Formally, fix any P_i, then for all $A, A' \in \mathscr{A}$ we have $A_i \geq_{P_i} A_i'$ if and only if $a_i^k \, R_i^k \, a_i'^k$ for all $k \in K$.[9] In a second step, we show the following result:

Lemma 1. *Fix any P_i and $A, A' \in \mathscr{A}$. We have that $A_i \succsim_i A_i'$ for all $\succsim_i \in \mathscr{Q}_{P_i}$ if and only if $A_i \geq_{P_i} A_i'$.*

Lemma 1 states that if an individual reports a list of rankings P_i then any separable preference $\succsim_i \in \mathscr{Q}_{P_i}$ that is consistent with P_i will order any allocations in the same way as the dominance relation $\geq_{P_i}$ constructed from the same list of rankings P_i. Simultaneously, if two allocations are not comparable by the dominance relation then not all separable preferences that are consistent with P_i will rank the two allocations in the same way.

The dominance relation together with Lemma 1 is necessary for characterizing the set of strategy-proof ranking mechanisms. Moreover, we use the dominance relation later to define a weaker efficiency notion.

3. CHARACTERIZATION OF STRATEGY-PROOF MECHANISMS

We now show that, under two additional mild requirements, all strategy-proof mechanisms can be constructed by simply picking a priority order for each category, i.e., choosing an order specifying the sequence in which individuals are assigned an object following their reported rankings. Formally, a priority order is a bijection $f : I \mapsto \{1, \ldots, n\}$, with a lower number $f(i)$ indicating a higher priority. For a given list of priority orders for every category $f = (f^k)_{k \in K}$, the **multi-category serial dictatorship** mechanism is defined for $\ell \in \{1, \ldots, n\}$ as follows:

[9] We use $A_i >_{P_i} A_i'$ to denote that $A_i \geq_{P_i} A_i'$ but $A_i' \not\geq_{P_i} A_i$. Similarly we use $A_i =_{P_i} A_i'$ whenever $A_i \geq_{P_i} A_i'$ and $A_i' \geq_{P_i} A_i$— in which case $A_i = A_i'$. Note that $\geq_{P_i}$ is a partial order, i.e., a *reflexive*, *antisymmetric*, and *transitive* binary relation.

Step ℓ. *For each $k \in \{1,\dots,m\}$, consider the ℓth highest priority individual $(f^k)^{-1}(\ell)$ according to f^k.*[10] *Then, following $P_{(f^k)^{-1}(\ell)}$, assign individual $(f^k)^{-1}(\ell)$ her most preferred object in category O^k among the remaining objects.*

For the characterization result to go through, we are left with defining two requirements: Nonbossiness and category-wise neutrality. Nonbossiness requires that no individual can influence the allocation of another individual without affecting her own allocation, while category-wise neutrality requires that the mechanism is immune to a relabeling of the object within each category. First, a ranking mechanism ψ is **nonbossy** if for all $P_i, \hat{P}_i \in \mathscr{P}_i$, and $P_{-i} \in \mathscr{P}_{-i}$ we have

$$\psi(P)_i = \psi(\hat{P}_i, P_{-i})_i \implies \psi(P) = \psi(\hat{P}_i, P_{-i}).$$

Second, let $\pi : O \to O$ be a permutations s.t. if $o \in O^k$ then $\pi[o] \in O^k$ — with Π denoting the set of all such permutations. We permute a list of simple orders P, denoted by $\pi[P]$, as follows: For all $i \in I$, $k \in K$ and $o, o \in O^k$ we have $\pi[o]\ \pi[P_i^k]\ \pi[o']$ if and only if $o\ P_i^k\ o'$. We say a ranking mechanism ψ is **category-wise neutral** if for all $k \in K$, $i \in I$, and $\pi \in \Pi$ we have

$$\pi[\psi(P)_i^k] = \psi(\pi[P])_i^k.$$

Now we go through the characterization result step-by-step. First, let us state an alternative strategy-proofness definition, which we will use to proof the next lemma. Formally, a ranking mechanism ψ is **strongly strategy-proof** if for any $i \in I$, $P_i \in \mathscr{P}_i$, $\hat{P}_i \in \mathscr{P}_i$, $P_{-i} \in \mathscr{P}_{-i}$ we have

$$\psi(P)_i \geq_{P_i} \psi(\hat{P}_i, P_{-i})_i.$$

The following statement is a corollary of Lemma 1. Additionally, note that Corollary 1 holds for ranking mechanisms but not necessarily in general.

Corollary 1. *A ranking mechanism is strategy-proof if and only if it is strongly strategy-proof.*

[10] Since f^k is a bijection, the function f^k is invertible. That is, $(f^k)^{-1} : \{1,\dots,n\} \mapsto I$ with $f^{-1}(\ell)$ giving the ℓth highest priority individual under f^k.

Second, Lemma 2 guarantees that if an individual changes her reported list of rankings from P_i to $\hat{P}_i$, the outcome of a strategy-proof and non-bossy ranking mechanism ψ cannot change unless some objects, ranked lower than those in the same category assigned under $\psi(P)_i$, are now ranked higher under $\hat{P}_i$. In other words, one can reorder objects in category O^k without affecting the allocation of strategy-proof and nonbossy ranking mechanisms, as long as for any object $o \in O^k$ such that $\psi(P)_i^k \; P_i^k \; o$ we have $\psi(P)_i^k \; \hat{P}_i^k \; o$. Lemma 2 is adapted from Svensson (1999). Using Lemma 1, in the proof we substitute strategy-proofness for strong strategy-proofness and use the fact that $A_i =_{P_i} A_i'$ implies $A_i = A_i'$.

Lemma 2. *Let ψ be a nonbossy and strategy-proof ranking mechanism. Consider any $P_i \in \mathscr{P}_i$, $P_{-i} \in \mathscr{P}_{-i}$, and some $\hat{P}_i \in \mathscr{P}_i$ such that for all $A \in \mathscr{A}$ where $\psi(P)_i \geq_{P_i} A_i$ we have $\psi(P)_i \geq_{\hat{P}_i} A_i$. Then $\psi(P) = \psi(\hat{P}_i, P_{-i})$.*

Third, Lemma 3 establishes that for identical rankings — all individuals submit an identical ranking in each category — any category-wise neutral mechanism can be obtained through a multi-category serial dictatorship.

Formally, the set of all **identical rankings** is defined as $\mathscr{I} = \{P \in \mathscr{P} : P_j^k = P_i^k \text{ for all } i, j \in I \text{ and } k \in K\}$.

Lemma 3. *Let ψ be a ranking mechanism that is category-wise neutral. For every identical ranking $P \in \mathscr{I}$, $k \in K$, and $\ell \in \{1, \dots, n\}$ the same individual $i_\ell^k \in I$ is assigned the ℓth-highest ranked object in O^k according to ranking P_i^k.*

Finally, it remains to be shown what happens for arbitrary rankings $P \in \mathscr{P}$. We will invoke Lemma 2 to show that for any arbitrary preference profile $P \in \mathscr{P} \setminus \mathscr{I}$ there exists an identical preference profile $P \in \mathscr{I}$ leading to the same outcome.

Theorem 1. *For any multi-category housing allocation problem a ranking mechanism ψ is strategy proof, nonbossy, and category-wise neutral if an only if ψ is a multi-category serial dictatorship.*

3.1. Two Subclasses

In the class of multi-category serial dictatorship mechanisms, two subclasses stand out. As the name suggests, the subclass of **identical priority multi-category serial dictatorships** ψ^{IPD} specify identical priority orders across all categories, i.e., $f = f^k$ for all $k \in K$. On the contrary, if priority orders are selected as fairly as possible — for all $i, j \in I$ we have $|\{k \in K : f^k(i) < f^k(j)\}| \geq \lfloor \frac{m}{2} \rfloor$ — we refer to this subclass as **fair priority multi-category serial dictatorships** ψ^{FPD}.

We show that, identical priority multi-category serial dictatorships are Pareto efficient. That is, a mechanism ψ is **Pareto efficient** if for all $P \in \mathscr{P}$,

$$\nexists A \in \mathscr{A} \text{ s.t. } A_i \succsim_i \psi(P)_i \text{ for all } i \in I \text{ and } A_i \succ_i \psi(P)_i \text{ for at least some } i \in I.$$

In comparison, fair priority multi-category serial dictatorship mechanisms are not Pareto efficient but satisfy a weaker form of efficiency: It rules out efficiency improvements that can directly be inferred from the reported rankings and is referred to as Pareto possibility (Budish, 2011). Formally, a mechanism ψ is **Pareto possible** if for all $P \in \mathscr{P}$,

$$\nexists A \in \mathscr{A} \text{ s.t. } A_i \geq_{P_i} \psi(P)_i \text{ for all } i \in I, \text{ and } A_i >_{P_i} \psi(P)_i \text{ for at least some } i \in I.$$

Next, we formulate a straightforward fairness notion to capture the trade off between fairness and efficiency when comparing identical priority with fair priority multi-category serial dictatorships. That is, for any two individuals $i \in I$ and $j \in I \setminus \{i\}$, we simply count the number of categories where j obtains a better object than i — following i's reported ranking P_i — to calculate i's envy toward j. We then say that a mechanism is ℓ envy-free if any individual i's envy toward any other individual j is at most ℓ for any possible resulting allocation. Formally, a mechanism ψ is **ℓ envy-free** if for all $i \in I$, $j \in I \setminus \{i\}$, and $P \in \mathscr{P}$ we have

$$|\{k \in K : \psi(P)_j^k \, P_i^k \, \psi(P)_i^k\}| \leq \ell.$$

Two observations follow immediately: First, in the multi-category housing allocation problem, any mechanism is at best $\lceil \frac{m}{2} \rceil$ envy-free and at the very least m envy-free. Second, fair priority multi-category serial dictatorships are $\lceil \frac{m}{2} \rceil$ envy-free, while identical priority multi-category serial dictatorships fail envy-freeness for any $\ell < m$. Therefore, the minor improvement in efficiency when using identical priority multi-category serial dictatorships instead of fair priority multi-category serial dictatorships comes at the highest possible fairness cost.

Proposition 1. *Identical priority multi-category serial dictatorships are Pareto efficient (Corollary of Monte & Tumennasan (2015) Theorem 2 for $m = 2$) and not ℓ envy-free for any $\ell < m$.*

Proposition 2. *Fair priority multi-category serial dictatorships are Pareto possible and* $\lceil \frac{m}{2} \rceil$ *envy-free.*

Next, we discuss Propositions 1 and 2 in the context of the literature on dictatorship mechanisms, designed for allocation problems with more objects than individuals (Pápai, 2001; Klaus & Miyagawa, 2002; Ehlers & Klaus, 2003; Hatfield, 2009; Monte & Tumennasan, 2015). With two exceptions, the class of multi-category serial dictatorships cannot be directly compared with other definitions in this literature, as it is specifically designed with multi-category housing allocation problems in mind — although, as do the other definitions, the class of multi-category serial dictatorships generalizes the class of serial dictatorships for single-object assignment problems, e.g., as defined in Svensson (1999).

One exception is Monte & Tumennasan (2015) who discus a class of sequential dictatorships that generalize the subclass of identical priority multi-category serial dictatorships for the case of two categories. In that sense, Pareto efficiency of identical priority multi-category serial dictatorships for two categories can be seen as a corollary of Theorem 2 in Monte & Tumennasan (2015) — with the caveat that this paper analyzes ranking mechanisms while they look at direct mechanisms, and therefore their analysis does not specify how sequential dictatorships would work with rankings instead of preferences as inputs.

The other exception is Caspari (2020) who discusses the class of booster draft (ranking) mechanisms. This class of mechanisms generalizes the idea of fair priority multi-category serial dictatorships to multi-object allocation problems, by creating an arbitrary partition of the objects into categories — referred to as boosters. Moreover, if we specify the same priority order for every booster, we can also generalize the idea of identical priority multi-category dictatorships to multi-object allocation problems. Note that for this class of generalized multi-category dictatorship mechanisms to be strategy-proof, we would have to specify the partition into boosters/categories prior to the elicitation of preferences. Furthermore, for any given specification of priority orders, even if we could choose the partition into boosters/categories after observing the reported rankings, this class of mechanisms will not satisfy Pareto possibility when preferences are separable — as most of the time, there will not exist a partition, such that every individual will prefer any subset, containing exactly one object from each subset of the created partition, to any other every other subset. When it comes to the fairness notion, while envy-freeness for multi-object allocation problems (Budish, 2011; Budish & Cantillon, 2012; Caspari, 2020) is strongly related to our fairness notion discussed here, they are not identical. This relates to the fact that objects are not directly comparable across different categories, while in the more general problem individuals can directly compare all the available objects with each other. As a consequence, in multi-object allocation problems we can find mechanisms that are 1-envy free, while for multi-category housing allocation problems, the best achievable fairness for any class of mechanisms is $\lceil \frac{m}{2} \rceil$ envy-freeness. Therefore, even though booster draft mechanisms are shown to be $\lceil \frac{m}{2} \rceil$ envy-free (Theorem 4 in Caspari (2020)), this does not directly imply that fair priority multi-category serial dictatorships are $\lceil \frac{m}{2} \rceil$ envy-free. Finally, readers interested in an example that contrasts the two classes of mechanisms which illustrate our theoretical framework can find Example 1 in the appendix.

4. AN APPLICATION: TEACHING ASSIGNMENTS FOR GRADUATE STUDENTS

In this section, we examine the allocation of teaching positions to graduate students at the economics department of Boston College: From 2019 until the present, following the proposal of this paper, a fair priority multi-category serial dictatorship has been in place and thus replaced the previous allocation system — kick-started by multiple complaints from graduate students over their final assignments in 2018.

We have used the rankings submitted by graduate students for the 2018 academic year, to compare fair priority multi-category serial dictatorships with identical priority multi-category serial dictatorships as well as the actual allocation made that year. That is, analogous to our theoretical part, there are as many students as teaching positions in each semester, and everyone submits a separate ranking for both the fall and spring semester. Moreover, as there are multiple versions of the same position, students end up having to rank only seven different options for each semester.[11] Then, based on 10,000 randomly

[11] The same positions were available in each semester (category), with the number in brackets giving the total number available for each semester: ta (teaching assistant) principles (12,12), ta statistics (3,3), ta econometrics (3,3), lab (laboratory) stats (5,5), lab econometrics (4,4), tf (teaching fellow) principles (4,4), tf statistics (1,1), and "special arrangements" (5,5). In our data 5 out of 37 students made "special arrangements" outside the available positions,e.g., having received a fellowship that freed them of work for one semester. These 5 students

generated priority orders, we have simulated the resulting allocations under both fair and identical priority multi-category serial dictatorships.

First, under the actual allocation, the number of graduate students envying both assignments of at least one other graduate student was roughly 29% — providing a potential reason for the complaints following the 2018 allocation. Surprisingly, even under identical priority multi-category serial dictatorships, on average, only 15% of graduate students would have envied both assignments of at least one other graduate student — while obviously amounting to 0% under any fair priority multi-category serial dictatorship.

Second, to obtain a grasp of how well graduate students like their assignment, we simply took the average rank of their assignment as a proxy — with the best value being 2 (first choice in both semesters) and the worst value being 14 (last choice in both semesters). We found that both classes of mechanisms lead to an expected average rank of 3.21, which is a stark improvement over the 4.72 of the actual 2018 allocation. We note that this measure does not capture the existence of potential inefficiencies under a fair priority multi-category serial dictatorship, where two graduate students would like to trade their bundles with each other. However, even though knowledge of the assignments is publicly available and a cohort of economic graduate students is generally aware of the concept of Pareto improving trades, no one has come forth suggesting a trade of assignments. This suggests that these trades are not particularly relevant in this application.

Third, the standard deviation in the average rank is roughly 2.2 for identical priority multi-category serial dictatorships compared to 1.6 for fair priority multi-category serial dictatorships. That is, while under an identical priority multi-category serial dictatorship, graduate students have a better chance to get their first two choices compared to a fair priority multi-category serial dictatorship, they also have a higher probability of ending up with a much worse average rank — with the worst possible assignment (assignment with positive probability to realize) under the former being 12 and under the latter being 9. Assuming that graduate students are at least mildly risk averse, the last two

automatically rank their special arrangement first, in the respective semester, while all other students rank them last — ensuring these students end up with their "special arrangements." Apart from this exception, students then had to simply rank the seven positions for each semester. As there was a new director of graduate studies in charge of the allocation for 2018 it was therefore unknown how the reported rankings would translate into the final allocation, one would reasonably expect that graduate students reported their rankings truthfully.

points imply that fair priorities lead to more preferable lotteries than identical priorities, i.e., lotteries with the same expected average rank but a lower variance.

5. CONCLUSION

We consider the problem of allocating a set of objects, which is sorted into categories of equal size, to a set of individuals, such that everyone obtains exactly one object from each category. Our main theoretical result shows that, in the large class of category-wise neutral and non-bossy mechanisms, any strategy-proof mechanism can be constructed by simply letting individuals choose an object from each category one after another, following some priority order. In this class of mechanisms two ways of selecting priority orders stick out: Either choose an identical priority order for each category or select a priority order that is reversed in every other category. Both intuition and the sparse empirical literature (Budish & Cantillon, 2012) seem to suggest that the second variant should lead to better results. This research also aligns with the discussion for future research of Monte & Tumennasan (2015), suggesting a need to look into solution concepts other than Pareto efficiency, due to its restrictiveness when applied to multi-category housing allocation problems.

A. MATHEMATICAL APPENDIX

Proof of Lemma 1. **If.** Fix any P_i, A, A' and suppose that $A_i \geq_{P_i} A'_i$.
Given the definition of the dominance relation, $A_i \geq_{P_i} A'_i$ implies $a_i^k \; R_i^k \; a_i^{k\prime}$ for all $k \in K$.
Pick any $\succsim_i \in \mathscr{Q}_{P_i}$, we have that $a_i^k \succsim_i a_i^{k\prime}$ for all $k \in K$.
Finally, by separability it follows that $A_i \succsim_i A'_i$ for all $\succsim_i \in \mathscr{Q}_{P_i}$, concluding the proof.
Only if. Fix any P_i, A, A' and suppose that $A_i \ngeq_{P_i} A'_i$.
Note that, $A_i \ngeq_{P_i} A'_i$ implies that A_i and A'_i are distinct allocations. Therefore, for any $\succsim_i \in \mathscr{Q}_{P_i}$ such that $A_i \succsim_i A'_i$, we also have $A'_i \not\succsim_i A_i$ as $\succsim_i$ is antisymmetric — that is, $A_i \succsim_i A'_i$ implies $A_i \succ_i A'_i$ and $A'_i \succsim_i A_i$ implies $A'_i \succ_i A_i$.
We want to show that there exists at least one $\succsim_i \in \mathscr{Q}_{P_i}$ such that $A'_i \succ_i A_i$. Start by randomly selecting a preference $\succsim_i \in \mathscr{Q}_{P_i}$. If $A'_i \succsim_i A_i$ — which implies $A'_i \succ_i A_i$, as A_i and A'_i are distinct and $\succsim_i$ is antisymmetric — we are done. Otherwise, consider a preference $\succsim'_i$ constructed as follows. First, recall that

any two sets of objects O' and O'' are comparable under any preference $\succsim_i \in \mathscr{Q}$ if and only if $|O' \cap O^k| = |O'' \cap O^k| \leq 1$ for all $k \in K$. Then, for any comparable O' and O'' such that $O' = O''$ let $O' \succsim_i' O''$ and $O'' \succsim_i' O'$. More importantly, for any two distinct and comparable sets of objects O' and O'', let $O' \succ_i' O''$ if $O' \geq_{P_i} O''$, let $O'' \succ_i' O'$ if $O'' \geq_{P_i} O'$, and otherwise let $O' \succ_i' O''$ if $O'' \succ_i O'$. By definition we have $A_i' \not\geq_{P_i} A_i$ and $A_i \succ_i A_i'$, and hence $A_i' \succ_i A_i$. It remains to be shown that $\succsim_i' \in Q_{P_i}$. First, for any two comparable, singleton sets $O' = \{o\}$, $O'' = \{o'\}$ — where by definition $\{o\}$ and $\{o'\}$ are in the same category — we have $\{o\} \succsim_i' \{o'\}$ if and only if $\{o\} \succsim_i \{o'\}$. That is, since $\succsim_i$ is consistent with P_i, $\succsim_i'$ is also consistent with P_i.

It remains to be shown that $\succsim_i'$ is separable. For any O' and O'' such that $O' = O''$ this is trivially satisfied. Now, suppose by contradiction that $\succsim_i'$ violates separability for two distinct comparable sets of objects O' and O''. That is, $O'' \succsim_i' O'$ but $o^{k\prime} \succsim_i' o^{k\prime\prime}$ for all $k \in \{k \in K : |O' \cap O^k| = |O'' \cap O^k|\}$. As $\succsim_i'$ is consistent with P_i, we have $o^{k\prime} P_i o^{k\prime\prime}$ for all $k \in \{k \in K : |O' \cap O^k| = |O'' \cap O^k|\}$ and therefore $O' \geq_{P_i} O''$. By construction $O' \geq_{P_i} O''$ implies $O' \succ_i' O''$ — a contradiction with $O'' \succsim_i' O'$.

We have shown that, if $A_i \not\geq_{P_i} A_i'$, then for any $\succsim_i \in Q_{P_i}$ with $A_i \succsim_i A_i'$ — which implies $A_i \succ_i A_i'$ — we can construct another preference $\succsim_i' \in Q_{P_i}$ such that $A_i' \succ_i' A_i$, concluding the proof.

□

Proof of Lemma 2. By Lemma 1 we can substitute strategy-proofness for strong strategy-proofness.

By strong strategy-proofness we have $\psi(P)_i \geq_{P_i} \psi(\hat{P}_i, P_{-i})_i$.

By the assumption of the lemma we have $\psi(P)_i \geq_{\hat{P}_i} \psi(\hat{P}_i, P_{-i})_i$.

Using strong strategy-proofness again we get $\psi(\hat{P}_i, P_{-i})_i \geq_{\hat{P}_i} \psi(P)_i$.

Combining the second and third line we get $\psi(\hat{P}_i, P_{-i})_i \geq_{\hat{P}_i} \psi(P)_i$ which implies that $\psi(\hat{P}_i, P_{-i})_i = \psi(P)_i$.

By nonbossiness it directly follows that $\psi(P) = \psi(\hat{P}_i, P_{-i})$ — if i's outcome did not change no-ones outcome changes. □

Proof of Lemma 3. Consider the outcome of any category-wise neutral ranking mechanism ψ for any two identical preference profiles $P \in \mathscr{I}$ and $\hat{P} \in \mathscr{I}$. Let us define the ℓth best choice in O^k under the identical preference profile P as well as $\hat{P}$: For all $\ell \in \{1, \ldots, n\}$ and $k \in K$, let o_ℓ^k denote $o \in O^k$ s.t. $|\{o' \in O^k : o' \; R_i^k \; o\}| = \ell$ respectively $\hat{o}_\ell^k$ denote $o \in O^k$ s.t. $|\{o' \in O^k : o' \; \hat{R}_i^k \; o\}| = \ell$.

Consider the individual i^k_l that is assigned o^k_l under P, i.e. $\psi(P)^k_{i^k_\ell} = o^k_\ell$. We want to show that the same individual gets the ℓth best choice in O^k under any other identical preference profile $\psi(\hat{P})^k_{i^k_\ell} = \hat{o}^k_\ell$. Consider the following permutation $\hat{\pi}$ defined for all $k \in \{1,\dots,m\}$ and $\ell \in \{1,\dots,n\}$ as $\hat{\pi}[o^k_\ell] = \hat{o}^k_\ell$. By construction, for this particular permutation we have that $\hat{\pi}[P^k] = \hat{P}^k$ for all $k \in \{1,\dots,m\}$. In other words, we have $P^k : o^k_1 - o^k_2 - \cdots - o^k_n$ and $\hat{\pi}[P^k]$: $\hat{\pi}[o^k_1] - \hat{\pi}[o^k_2] - \cdots - \hat{\pi}[o^k_n]$ which is nothing else than $\hat{\pi}[P^k] : \hat{o}^k_1 - \hat{o}^k_2 - \cdots - \hat{o}^k_n$, so $\hat{P}^k = \hat{\pi}[P^k]$ for all $k \in \{1,\dots,m\}$.
By neutrality and the construction above, we get $\hat{\pi}[\psi(P)^k_{i^k_\ell}] = \psi((\hat{\pi}[P]))^k_{i^k_\ell} = \psi(\hat{P})^k_{i^k_\ell}$. Moreover, by the definition of the permutation $\hat{\pi}$ we have $\hat{\pi}[\psi(P)^k_{i^k_\ell}] = \hat{\pi}[o^k_\ell] = \hat{o}^k_\ell$. Combining both leads to the desired conclusion that the same individual gets the ℓth best object in set O^k for any two identical preference profiles $\psi(\hat{P})^k_{i^k_\ell} = \hat{o}^k_l$ — both $\hat{o}^k_l$ and o^k_l are assigned to the same individual i^k_ℓ.
$\square$

Proof of Theorem 1. **If.** It is obvious that any multi-category serial dictatorship is category-wise neutral and nonbossy. For (strong) strategy-proofness, suppose by contradiction that there exists $\psi(P)_i \not\succsim_{P_i} \psi(P'_i, P_{-i})_i$. Then there exists at least one category O^k such that $\psi(P'_i, P_{-1})^k_i \; P_i \; \psi(P)^k_i$. However, as P_{-i} is fixed, all individuals with higher priority will pick identical items in category k independent of i reporting P_i or P'_i, so i gets to choose from the same set of remaining objects. Hence, we have that the obtained item under P_i is weakly preferred to any item obtained by reporting another ranking, i.e. $\psi(P)^k_i \; R_i \; \psi(P'_i, P_{-i})^k_i$ for all $k \in K$ contradicting the initial statement.
Only if. We now show that any (strongly) strategy-proof, nonbossy, and category-wise neutral mechanism ψ is a multi-category serial dictatorship.
Start by randomly selecting any identical preference profile $P \in \mathscr{I}$, and consider any strategy-proof, nonbossy, and category-wise neutral ranking mechanism ψ. Then, construct a priority order f^k over individuals I for each category $k \in K$ as follows:

$$f^k(i) = |o \in O^k : o \; R^k_i \; \psi(P)^k_i|$$

That is, the individual with the best object in category k under ψ has priority 1 in this category, the individual with the second best object has priority 2 in this category, and so on. Let ψ^{FP}_f denote the multi-category serial dictatorship

mechanism with priority orders $f = (f^k)_{k \in K}$ as constructed above. By Lemma 3 the mechanism ψ assigns the same individual $i^k_\ell \in I$ the ℓth best object in O^k according to ranking P^k across every identical ranking $P \in \mathscr{I}$. It is therefore easy to check that, $\psi^{FP}_f(P) = \psi(P)$ for any $P \in \mathscr{I}$ — as i^k_ℓ, the uniquely identifiable individual with the ℓth highest priority in category k under mechanism ψ, is also the individual with the ℓth highest priority in category k under ψ^{FP}_f, i.e., $i^k_\ell = (f^k)^{-1}(\ell)$. That is, for each strategy-proof, nonbossy, and category-wise neutral ranking mechanism ψ, we can construct a unique multi-category serial dictatorship mechanism ψ^{FP}_f, such that $\psi^{FP}_f(P) = \psi(P)$ for all $P \in \mathscr{I}$.
Given ψ, it remains to be shown that ψ^{FP}_f gives the same allocation as ψ for any arbitrary preference profile. Start by randomly selecting any preference profile $P \in \mathscr{P}$ and construct an identical preference profile $\hat{P} \in \mathscr{I}$ based on P as follows. For each category $k \in K$, let $\hat{P}^k$ rank object $\psi^{FP}_f(P)^k_{i^k_1} = \psi(P)^k_{i^k_1}$ first, object $\psi^{FP}_f(P)^k_{i^k_2} = \psi(P)^k_{i^k_2}$ second, and so on, with object $\psi^{FP}_f(P)^k_{i^k_n} = \psi(P)^k_{i^k_n}$ ranked last. Note that, for any $A \in \mathscr{A}$ and $i \in I$, such that $\psi(P)_i \geq_{\hat{P}_i} A$ we also have $\psi(P)_i \geq_{P_i} A$. Therefore, by Lemma 2, we can change i's ranking from $\hat{P}_i$ to P_i without changing the outcome of ψ. Recursively applying Lemma 2 for each $i \in I$ we get that $\psi(P) = \psi(\hat{P})$. In a similar fashion, it is easy to check that, under the two preference profiles P and $\hat{P}$ we have $\psi^{FP}_f(P) = \psi^{FP}_f(\hat{P})$. Combining these two observations, we have $\psi^{FP}_f(P) = \psi^{FP}_f(\hat{P}) = \psi(\hat{P}) = \psi(P)$, and therefore $\psi^{FP}_f(P) = \psi(P)$ for all preference profiles $P \in \mathscr{P}$, concluding the proof. □

Proof of Proposition 1. **Identical priority multi-category serial dictatorships are Pareto efficient.**

Consider any Identical priority multi-category serial dictatorship ψ^{IPD}, and let $\mathscr{A}^1 = \mathscr{A} \setminus \{\psi^{IPD}\}$ be the set of allocations potentially Pareto dominating allocation ψ^{IPD}. Note that, the highest priority individual $i_1 = f^{-1}(1)$ gets her m best objects, i.e., for all $P \in \mathscr{P}$ and $k \in K$ we have $\psi^{IPD}(P)^k_{i_1}\ P^k_{i_1}\ o$ for all $o \in O^k \setminus \{\psi^{IPD}(P)^k_{i_1}\}$.
By the definition of the dominance relation we get $\psi^{IPD}(P)_{i_1} \geq_{P_{i_1}} A_i$ for all $A \in \mathscr{A}^1$.
By lemma 1 it follows that $\psi^{IPD}(P)_{i_1} \succsim_{i_1} A_{i_1}$ for all $A \in \mathscr{A}^1$ and for all $\succsim_{i_1} \in \mathscr{Q}_{P_{i_1}}$.
It follows that any allocation Pareto dominating ψ^{IPD} must assign $\psi^{IPD}(P)_{i_1}$

to i_1. That is, the set of allocations potentially Pareto dominating allocation ψ^{IPD} becomes $\mathscr{A}^2 = \{A \in \mathscr{A} : A_{i_1} = \psi^{IPD}(P)_{i_1}\} \setminus \{\psi^{IPD}\}$.
Invoking an analogous argument for $i_2 = f^{-1}(2)$, we get that $\psi^{IPD}(P)_{i_2} \succsim_{i_2} A_{i_2}$ for all $A \in \mathscr{A}^2$ and for all $\succsim_{i_2} \in \mathscr{Q}_{P_{i_2}}$, and therefore the set of allocations potentially Pareto dominating allocation ψ^{IPD} becomes $\mathscr{A}^3 = \{A \in \mathscr{A} : A_{i_1} = \psi^{IPD}(P)_{i_1}$ and $A_{i_2} = \psi^{IPD}(P)_{i_2}\}$.
Iterative applying an analogous argument for individuals $i_3 = f^{-1}(3)$ to $i_{n-1} = f^{-1}(n-1)$ we get that the the set of allocations potentially Pareto dominating allocation ψ^{IPD} becomes $\mathscr{A}^{n-1} = \{A \in \mathscr{A} : A_{i_1} = \psi^{IPD}(P)_{i_1}$ and $\ldots$ and $A_{i_{n-1}} = \psi^{IPD}(P)_{i_{n-1}}\} \setminus \{\psi^{IPD}\} = \emptyset$, concluding the proof. $\square$

Proof of Proposition 1. **Identical priority multi-category serial dictatorships are not $\boldsymbol{\ell}$ envy-free for any $\boldsymbol{\ell < m}$.**

Consider any identical priority multi-category serial dictatorship ψ^{IPD} and some $\ell < m$. Pick any reported list of rankings in the set of identical rankings $P \in \mathscr{I}$. Consider $i_1 = f^{-1}(1)$ and any $j \in I \setminus \{i_1\}$. By the definition of the identical multicategory serial dictatorship, it immediately follows that $|\{k \in K : \psi^{IPD}(P)^k_{i_1}\ P^k_j\ \psi^{IPD}(P)^k_j\}| = m > \ell$, concluding the proof. $\square$

Proof of Proposition 2. **Fair priority multi-category serial dictatorships are Pareto possible.**

Consider any fair priority multi-category serial dictatorship ψ^{FPD} and suppose by contradiction there exists $A \in \mathscr{A}$ such that $A_i \geq_{P_i} \psi^{FPD}(P)_i$ for all $i \in I$ holding strictly for at least one individual. By the definition of the dominance relation $A_i \geq_{P_i} \psi^{FPD}(P)_i$ implies $a^k_i\ R^k_i\ \psi^{FPD}(P)^k_i$ for all $k \in K$ and for all $i \in I$ holding strictly for at least some k and i. Now, pick the highest priority individual i in the first category O^k such that $a^k_i P^k_i \psi^{FPD}(P)^k_i$. As individuals report strict rankings over categories, all higher priority individuals in that category get the same object as before, i.e., $a^k_j = \psi^{FPD}(P)^k_j$ for all $j \in \{j \in I : f^k(j) > f^k(i)\}$. It follows that a^k_i is still available when its i's turn to choose an object form category O^k, and thus $\psi^{FPD}(P)^k_i\ R^k_i\ a^k_i$ contradicting $a^k_i\ P^k_i\ \psi^{FPD}(P)^k_i$. $\square$

Proof of Proposition 2. **Fair priority multi-category serial dictatorships are $\boldsymbol{\lceil \frac{m}{2} \rceil}$ envy-free.**

Consider any fair priority multi-category serial dictatorship ψ^{FPD}. By the definition of any fair priority multi-category serial dictatorship, for any $i \in I$ and $j \in J \setminus \{i\}$ we have that $|\{k \in K : f^k(i) < f^k(j)\}| \geq \lfloor \frac{m}{2} \rfloor$ and therefore

$|\{k \in K : \psi^{FPD}(P)_i^k \, P_i^k \, \psi^{FPD}(P)_j^k\}| \geq \lfloor \frac{m}{2} \rfloor$ for any $P \in \mathscr{P}$. It follows that the maximum envy any $i \in I$ can have is $m - |\{k \in K : \psi^{FPD}(P)_i^k \, P_i^k \, \psi^{FPD}(P)_j^k\}| \leq |\{k \in K : \psi^{FPD}(P)_i^k \, P_j^k \, \psi^{FPD}(P)_i^k\}| \leq \lceil \frac{m}{2} \rceil$ for all $P \in \mathscr{P}$, concluding the proof. $\square$

Example 1. *Consider two individuals (graduate students) $I = \{i_1, i_2\}$. Suppose that they have to work as teaching assistants for a spring semester course $O^1 = \{micro^1, macro^1\}$ and a fall semester course $O^2 = \{micro^2, stats^2\}$. Note that, if both individuals want different teaching assignments within each category, the chosen priority order does not matter. Therefore, more interesting cases are those where both compete for the same objects. In particular, assume both individuals are interested in microeconomics and thus report identical rankings, i.e., for $i \in I$ we have*

$$P_i^1 : micro^1 - macro^1, and$$
$$P_i^2 : micro^2 - stats^2.$$

The dominance relation $\geq_{P_i}$ tells us that both individuals $i \in I$ (strictly) prefer — under any preference consistent with the reported ranking $\succsim_i \in \mathscr{Q}_{P_i}$ — $\{micro^1, micro^2\}$ to both $\{micro^1, stats^2\}$ and $\{macro^1, micro^2\}$ which they in turn prefer to $\{macro^1, stats^2\}$. Observe that the rankings give no insight into how either one compares $\{micro^1, stats^2\}$ to $\{macro^1, micro^2\}$, i.e., whether $\{micro^1, stats^2\} \succ_i \{macro^1, micro^2\}$, $\{macro^1, micro^2\} \succ_i \{micro^1, stats^2\}$, or $\{micro^1, stats^2\} \sim_i \{macro^1, micro^2\}$.

Now, consider any fair priority multi-category serial dictatorship ψ^{FPD} where $i \in I$ gets to choose first from the spring assignments O^1 and $j \in I \setminus \{i\}$ gets to choose first from the fall assignments O^2. This leads to an assignment of $\psi_i^{FPD} = \{micro^1, stats^2\}$ to i and $\psi_i^{FPD} = \{macro^1, micro^2\}$ to j — which in some sense is a natural way to allocate these objects.

In contrast, it has been suggested that one ought to use identical priority multi-category serial dictatorships — or similar mechanisms like sequential dictatorships — due to them being Pareto efficient, while the weaker notion of Pareto possibility does not rule out all possible inefficiencies. In particular, in the example there is one possible inefficiency where $\{macro^1, micro^2\} \succsim_i \{micro^1, stats^2\}$ and $\{micro^1, stats^2\} \succsim_j \{macro^1, micro^2\}$ with at least one preference holding strictly. However, the way this inefficiency is resolved under an identical priority multi-category serial dictatorship strikes us as unsatisfactory: That is, the only strategy-proof way to avoid this inefficiency is to assign one of the two individuals their absolute best bundle leaving the other to pick up the remains — which gets even more problematic the more categories there are. In this case, the identical priority multi-category serial dictatorship assigns $\psi_i^{IPD} = \{micro^1, micro^2\}$ to i $\psi_i^{FPD} = \{macro^1, stats^2\}$ to j or vice versa.

From an economic perspective, why one might find identical priority multi-category serial dictatorship mechanisms problematic in the example above, stems from the observation that envy-freeness combined with Pareto possibility might very well be a better proxy for the welfare of the resulting allocation than Pareto efficiency. To gain an intuition, consider the following utilities, where the discussed inefficiency occurs and nonetheless fair priority multi-category serial dictatorship leads to higher welfare:[12]

	$\{micro^1, micro^2\}$	$\{micro^1, stats^2\}$	$\{macro^1, micro^2\}$	$\{macro^1, stats^2\}$
u_i	60	40	45	10
u_j	60	45	40	10

Here, the fair priority multi-category serial dictatorship (in its worst case) leads to a welfare of 80 and best case to a welfare of 90, while the identical priority multi-category serial dictatorship leads to a welfare of 70. Moreover, if the market designer wants to maximize Rawlsian welfare (Rawls, 1971), the clear winner is the fair priority multi-category serial dictatorship leading to either 40 or 45 while the identical priority multi-category serial dictatorship leads to a Rawlsian welfare of 10. The utilities in the example reflect the intuition provided by Budish et al. (2016), i.e., that moving from a "bad bundle" to a "medium bundle" leads to higher utility gains compared to moving from a "medium bundle" to a "good bundle". This provides a reasonable explanation as to why mechanisms ensuring that individuals' realized resources are roughly equal might strike us as more appealing and seem to outperform their counterparts in practice.

References

Brams, S. J., Edelman, P. H., & Fishburn, P. C. (2003). Fair division of indivisible

[12] Utilities $u_i : \mathscr{A} \mapsto \mathbb{R}$ represent the underlying worth of an assignment for a given individual — representing cardinal values as opposed to the ordinal interpretation of preferences. The welfare of an assignment is then simply the total sum of the resulting utilities.

items. *Theory and Decision*, *55*(2), 147–180.
Brams, S. J., & Fishburn, P. C. (2000). Fair division of indivisible items between two people with identical preferences: Envy-freeness, pareto-optimality, and equity. *Social Choice and Welfare*, *17*(2), 247–267.
Budish, E. (2011). The combinatorial assignment problem: Approximate competitive equilibrium from equal incomes. *Journal of Political Economy*, *119*(6), 1061–1103.
Budish, E., Cachon, G. P., Kessler, J. B., & Othman, A. (2016). Course match: A large-scale implementation of approximate competitive equilibrium from equal incomes for combinatorial allocation. *Operations Research*, *65*(2), 314–336.
Budish, E., & Cantillon, E. (2012). The multi-unit assignment problem: Theory and evidence from course allocation at harvard. *American Economic Review*, *102*(5), 2237–71.
Caspari, G. (2020). Booster draft mechanism for multi-object assignment. *ZEW-Centre for European Economic Research Discussion Paper*(20-074).
Edelman, P., & Fishburn, P. (2001). Fair division of indivisible items among people with similar preferences. *Mathematical Social Sciences*, *41*(3), 327–347.
Ehlers, L., & Klaus, B. (2003). Coalitional strategy-proof and resource-monotonic solutions for multiple assignment problems. *Social Choice and Welfare*, *21*(2), 265–280.
Hatfield, J. W. (2009). Strategy-proof, efficient, and nonbossy quota allocations. *Social Choice and Welfare*, *33*(3), 505–515.
Hylland, A., & Zeckhauser, R. (1979). The efficient allocation of individuals to positions. *Journal of Political Economy*, *87*(2), 293–314.
Klaus, B., & Miyagawa, E. (2002). Strategy-proofness, solidarity, and consistency for multiple assignment problems. *International Journal of Game Theory*, *30*(3), 421–435.
Monte, D., & Tumennasan, N. (2015). Centralized allocation in multiple markets. *Journal of Mathematical Economics*, *61*, 74–85.
Othman, A., Sandholm, T., & Budish, E. (2010). Finding approximate competitive equilibria: Efficient and fair course allocation. In *Proceedings of the 9th International Conference on Autonomous Agents and Multiagent Systems: Volume 1* (pp. 873–880).
Pápai, S. (2001). Strategyproof and nonbossy multiple assignments. *Journal of Public Economic Theory*, *3*(3), 257–271.
Rawls, J. (1971). *A Theory of Justice*. Harvard university press.
Roberts, F. S. (1985). *Measurement Theory*. Cambridge University Press.
Roth, A. E. (2008). What have we learned from market design? *Economic Journal*, *118*(527), 285–310.

Shapley, L., & Scarf, H. (1974). On cores and indivisibility. *Journal of Mathematical Economics*, *1*(1), 23–37.

Sönmez, T., & Ünver, M. U. (2010). Course bidding at business schools. *International Economic Review*, *51*(1), 99–123.

Sönmez, T., & Ünver, M. U. (2011). Matching, allocation, and exchange of discrete resources. In *Handbook of Social Economics* (Vol. 1, pp. 781–852). Elsevier.

Svensson, L.-G. (1999). Strategy-proof allocation of indivisible goods. *Social Choice and Welfare*, *16*(4), 557–567.

Journal of Mechanism and Institution Design
ISSN: 2399-844X (Print), 2399-8458 (Online)
DOI:10.22574/jmid.2023.12.004

OPTIMAL MECHANISM DESIGN WITH APPROXIMATE INCENTIVE COMPATIBILITY AND MANY PLAYERS

Pathikrit Basu

Freelance Researcher
pathikritbasu@gmail.com

ABSTRACT

We consider a setting in which a mechanism designer must choose the appropriate social alternative depending on the state of nature. We study the problem of optimal design and demonstrate that a mechanism which allocates resources so as to achieve the social optimum and assigns payments equal to the posterior expected utility of the agent at the social optimum, is an ε-optimal mechanism for environments with many players.

Keywords: Mechanism design , incentive compatibility, statistical decision theory

JEL Classification Numbers: D60, D61, D62.

1. INTRODUCTION

ONE of the main aspects of the study of mechanism design is aggregating private information in order to reach a socially optimal objective. Since the agents may benefit from particular social alternatives being chosen, they have a willingness to pay, so far as social choice aligns with their preferences. This allows the mechanism designer to extract revenue in the form of payments from the players in the mechanism while ensuring players have the incentive to participate truthfully (Hurwicz, 1960; Gibbard, 1973; Maskin, 1999; Vickrey, 1961; Clarke, 1971; Groves, 1973; Myerson, 1981; Myerson & Satterthwaite, 1983). This paper provides a construction of an optimal mechanism in a setting with many players. The expected payments of the players from the mechanism have a *what you give is what you get* interpretation. The problem studied in the paper is of the nature of statistical decision problems (Wald, 1950; Blackwell & Girshick, 1979; Berger, 2013; Ferguson, 2014; Pratt et al., 1995; DeGroot, 2005). The result in the paper hinges on the intuition that for settings with many players, an individual player's opinion (type reported) does not affect the aggregate, achieving approximate incentive compatibility. Hence, the socially optimal mechanism allows the mechanism designer to extract all the surplus. This feature in incentive compatibility relates the paper to social learning and herd formation models (Banerjee, 1992; Bikhchandani et al., 1992; Chamley, 2004; Smith & Sørensen, 2000; Fudenberg et al., 2021). Further, the paper is also related to the models in De Condorcet (1785) and Roberts & Postlewaite (1976). Approximate incentive compatibility in mechanism and market design contexts has also been considered in Azevedo & Budish (2019), Balcan et al. (2019), Epasto et al. (2018) and Lee (2016). The result in the paper may also be viewed as a general result in a standard setting in which optimal payments may be characterized as agents paying their expected value which perhaps interestingly, is in contrast with the externality based payment scheme of Pigou (1920) and VCG payments.

2. MODEL

An *environment* is a tuple $\mathscr{E} =< N, \Omega, (S_i)_{i\in N}, A, (u_i)_{i\in N}, \pi_0, \mu >$. The set N is a finite set and it is the set of all players in the environment. The set Ω is the set of possible states of nature, assumed to be finite. For each $i \in N$, the set S_i (finite) is the set of possible signals that player i may receive regarding the true state $\omega \in \Omega$. Finally, the set A (finite) is the set of alternatives. Player $i \in N$ has a state-dependent utility function $u_i : A \times \Omega \to \mathbb{R}$. In terms of the information structure present in the environment, $\pi_0 \in \Delta(\Omega)$ is a common prior and $\mu = \mu(.|\omega)_{\omega\in\Omega} \subseteq \Delta(S)$ is the state-dependent signal distribution for the players, with the set of all possible signal profiles being $S = \prod_{i\in N} S_i$. We assume that $\pi_0(\omega) > 0$ and $\mu(s|\omega) > 0$ for each $s \in S$ and $\omega \in \Omega$. We denote as $\pi_0 \otimes \mu$, the joint distribution on the set $\Omega \times S$, generated by the prior π_0 and signal distribution μ. For $s_i \in S_i$ and $s \in S$, we write $\pi(s_i)$ to be the posterior belief in $\Delta(\Omega)$ conditional on player i's signal s_i and $\pi(s)$ to be the posterior belief conditional on the signal profile s i.e. the private signals of all the players in the environment.

A *mechanism* is a tuple (σ, q), in which $\sigma : S \to A$ is a social choice function and $q = (q_i)_{i \in N}$ is a collection of payment functions, the payment function for $i \in N$ is a function $q_i : S \to \mathbb{R}$.

We now state and provide some more definitions.

Definition 1. *Let $\varepsilon > 0$. A mechanism (σ, q) is said to be ε-Bayesian incentive compatible if for each $i \in N$ and $s_i, t_i \in S_i$,*

$$\mathbb{E}_{\pi_0 \otimes \mu}[u_i(\sigma(s_i, s_{-i}), \omega) - q_i(s_i, s_{-i})|s_i] \geq \mathbb{E}_{\pi_0 \otimes \mu}[u_i(\sigma(t_i, s_{-i}), \omega) - q_i(t_i, s_{-i})|s_i] - \varepsilon.$$

Definition 2. *A mechanism (σ, q) is said to be Bayesian individually rational if for each $i \in N$ and $s_i \in S_i$,*

$$\mathbb{E}_{\pi_0 \otimes \mu}[u_i(\sigma(s_i, s_{-i}), \omega) - q_i(s_i, s_{-i})|s_i] \geq 0.$$

For any given mechanism (σ, q), we define the revenue $Q(\sigma, q)$ from the mechanism as the expected sum of payments derived from the mechanism i.e.

$$Q(\sigma, q) = \mathbb{E}_{\pi_0 \otimes \mu}[\sum_{i \in N} q_i(s_i, s_{-i})].$$

We now state the definition of an ε-optimal mechanism.

Definition 3. *Let $\varepsilon > 0$. A mechanism (σ', q') is said to be ε-optimal if*

1. *(σ', q') is ε-Bayesian incentive compatible and Bayesian individually rational.*

2. *For any other mechanism (σ, q) that is ε-Bayesian incentive compatible and Bayesian individually rational,*

$$Q(\sigma', q') \geq Q(\sigma, q).$$

We define the following mechanism (σ^*, q^*), which is the main mechanism proposed by the paper. It implements the social optimum and prescribes payments that are equal to the posterior expected utility of the agent.

1. For each $s \in S$,

$$\sigma^*(s) \in \arg\max_{a \in A} \sum_{\omega \in \Omega} \pi(s)(\omega) \sum_{i \in N} u_i(a, \omega).$$

2. For each $s \in S$, for each $i \in N$,

$$q_i^*(s) = \sum_{\omega \in \Omega} \pi(s)(\omega) u_i(\sigma^*(s), \omega).$$

We prove the main theorem.

Theorem 1. *Suppose we set A and Ω to be the set of alternatives and the set of states of nature. Suppose X (finite) is a signal space. Let $\pi_0 \in \Delta(\Omega)$ be a common prior. Let $\{\nu(.|\omega)\}_{\omega \in \Omega} \subseteq \Delta(X)$ be a signal distribution such that $\nu(x|\omega) > 0$ for each $x \in X$, $\omega \in \Omega$ and $\nu(.|\omega) \neq \nu(.|\omega')$ for $\omega \neq \omega'$.*

Let $\mathscr{U}$ be a finite set of utility functions $u : A \times \Omega \to \mathbb{R}$ such that for each $\lambda \in \Delta(\mathscr{U})$ and for each $\omega \in \Omega$, there exists $a \in A$ (unique maximiser) such that

$$\sum_{u \in \mathscr{U}} \lambda(u) u(a, \omega) > \sum_{u \in \mathscr{U}} \lambda(u) u(b, \omega), \tag{1}$$

for each $b \in A \backslash \{a\}$.

Let $\varepsilon > 0$. Then, there exists $n_0 \in \mathbb{N}$ such that for any environment $\mathscr{E} =< N, \Omega', (S_i)_{i \in N}, A', (u_i)_{i \in N}, \pi_0', \mu >$ satisfying

1. *$|N| > n_0$;*
2. *$\Omega' = \Omega$; $\pi_0' = \pi_0$; $S_i = X$, for each $i \in N$; $\mu(.|\omega) = \nu^N(.|\omega)$, for each $\omega \in \Omega$ (the probability measure $\nu^N(.|\omega)$ is the product probability measure in $\Delta(S)$ with index set N, for each $\omega \in \Omega$);*
3. *$u_i \in \mathscr{U}$, for each $i \in N$, the mechanism (σ^*, q^*) is ε-optimal.*

Proof. For each $\lambda \in \Delta(\mathscr{U})$ and any $\omega \in \Omega$, define the set

$$E(\lambda; \omega) = \Big\{ e \in [0,1] : \text{ for each } \pi \in \Delta(\Omega), \text{ if } \pi(\omega) > e, \text{ then}$$
$$\arg\max_{a \in A} \sum_{\omega' \in \Omega} \pi(\omega') \sum_{u \in \mathscr{U}} \lambda(u) u(a, \omega') = \arg\max_{a \in A} \sum_{u \in \mathscr{U}} \lambda(u) u(a, \omega) \Big\}.$$

The above defines a correspondence taking input values $\lambda \in \Delta(\mathscr{U})$ and outputs the set $E(\lambda; \omega) \subseteq [0,1]$ i.e. $E(.; \omega) : \Delta(\mathscr{U}) \rightrightarrows [0,1]$, a correspondence

$E(.;\omega)$ for each $\omega \in \Omega$. $E(\lambda;\omega)$ is the set of all values e in $[0,1]$ such that if a belief π assigns probability greater than e on the state ω, then the optimal action for the mixture utility function $\sum_{u\in\mathscr{U}} \lambda(u)u(a,\omega)$ in state ω is the same as in under the belief π. Given that condition 1 in the statement of the theorem is satisfied, it follows that $\arg\max_{a\in A} \sum_{u\in\mathscr{U}} \lambda(u)u(a,\omega)$ is a singleton as there is a unique maximiser. Further, since the expression $\sum_{\omega'\in\Omega} \pi(\omega') \sum_{u\in\mathscr{U}} \lambda(u)u(a,\omega')$ is continuous (linear) in both λ and π, it follows that the correspondence $E(\lambda;\omega)$ is both upper and lower hemicontinuous in λ. Further, by definition, $E(\lambda;\omega)$ is convex and closed since it is always a closed interval of the form $[e',1]$. Hence, by applying the theorem of the maximum (Charalambos & Aliprantis, 2013), we may prove that $e(\lambda;\omega) := \min_{e\in E(\lambda;\omega)} e$ is continuous for each $\omega \in \Omega$. Now, let $e = \max_{\omega\in\Omega} \max_{\lambda\in\Delta(\mathscr{U})} e(\lambda;\omega)$. Since it may be argued that $e(\lambda;\omega) < 1$ for each $\lambda \in \Delta(\mathscr{U})$ and $\omega \in \Omega$, it follows that $e < 1$.

Let $\varepsilon > 0$. Then, let $\delta \in (0,1)$ such that

$$5\delta \max_{u\in\mathscr{U}} \max_{a\in A} ||u(a,.)|| < \varepsilon. \tag{2}$$

Let X^∞ be the space of all sequences in X. For each $\omega \in \Omega$, let $\nu^\infty(.|\omega)$ be the product probability measure in $\Delta(X^\infty)$. For $x^n \in X^n$, let $\pi(x^n)$ be the posterior belief on the state of nature conditional on signals in x^n. Define the following events in X^∞, one for each $\omega \in \Omega$,

$$X^\infty_\omega := \{x^\infty \in X^\infty : \lim_{n\to\infty} \pi(x^n)(\omega) = 1\}.$$

Then, it follows that

$$\nu^\infty(X^\infty_\omega|\omega) = 1. \tag{3}$$

Given $n \in \mathbb{N}$, define the following sets in X^n.

$$P^n_\omega = \{x^n \in X^n : \forall y \in X, \pi(x^n,y)(\omega) > e\}.$$
$$T^n = \{x^n \in X^n : \forall y,z \in X, ||\pi(x^n,y) - \pi(x^n,z)|| < \delta\}.$$

Then, it follows from (3) that there exists $n_0 \in \mathbb{N}$ such that for each $n \geq n_0$ and for each $\omega \in \Omega$,

$$\nu^n(P^n_\omega \cap T^n|\omega) \geq 1-\delta, \tag{4}$$

given the n-fold product measure $\nu^n(.|\omega)$ in $\Delta(X^n)$.

The chosen n_0 is the one we pick.

Suppose $\mathscr{E} =< N, \Omega, (S_i)_{i\in N}, A, (u_i)_{i\in N}, \pi_0, \mu >$ is an environment that satisfies the properties 1), 2) and 3). Then, we show that the mechanism (σ^*, q^*) is ε-optimal.

We first show that (σ^*, q^*) is ε-Bayesian incentive compatible. Let $i \in N$ and $s_i, t_i \in S_i$. Then, we have that

$$\begin{aligned}\mathbb{E}_{\pi_0\otimes\mu}[(u_i(\sigma^*(t_i,s_{-i}),\omega) - q_i^*(t_i,s_{-i})) - (u_i(\sigma^*(s_i,s_{-i}),\omega) - q_i^*(s_i,s_{-i}))|s_i] \\ \leq \delta \max_{u\in\mathscr{U}} \max_{a\in A} ||u(a,.)|| + 4\delta \max_{u\in\mathscr{U}} \max_{a\in A} ||u(a,.)||.\end{aligned}$$

The above inequality follows since for the given conditional expectation, from (4), with probability at least $1-\delta$, two things happen simultaneously : i) the social optimum does not change with the unilateral deviation from s_i to t_i i.e. $\sigma^*(s_i,s_{-i}) = \sigma^*(t_i,s_{-i})$ hence $u_i(\sigma^*(s_i,s_{-i}),\omega) = u_i(\sigma^*(t_i,s_{-i}),\omega)$ and ii) the change in posterior belief is at most of distance δ i.e. $||\pi(s_i,s_{-i}) - \pi(t_i,s_{-i})|| < \delta$, hence this means that $q_i^*(s_i,s_{-i}) - q_i^*(t_i,s_{-i}) \leq \delta \max_{u\in\mathscr{U}} \max_{a\in A} ||u(a,.)||$. Further, with probability at most δ, we get a difference of four terms

$$(u_i(\sigma^*(t_i,s_{-i}),\omega) - q_i^*(t_i,s_{-i})) - (u_i(\sigma^*(s_i,s_{-i}),\omega) - q_i^*(s_i,s_{-i}))$$

that takes a value of at most $4\max_{u\in\mathscr{U}} \max_{a\in A} ||u(a,.)||$, by the definition of q^*.

Hence, it follows from (2) that

$$\mathbb{E}_{\pi_0\otimes\mu}[(u_i(\sigma^*(t_i,s_{-i}),\omega) - q_i^*(t_i,s_{-i})) - (u_i(\sigma^*(s_i,s_{-i}),\omega) - q_i^*(s_i,s_{-i}))|s_i] < \varepsilon,$$

which implies that (σ^*, q^*) is ε-Bayesian incentive compatible.

Next, we show that (σ^*, q^*) is Bayesian individually rational. Let $i \in N$ and $s_i \in S_i$. Then,

$$\begin{aligned}\mathbb{E}_{\pi_0\otimes\mu}[u_i(\sigma^*(s_i,s_{-i}),\omega)|s_i] &= \mathbb{E}_{s_{-i}}[\sum_{\omega\in\Omega} \pi(s_i,s_{-i})(\omega) u_i(\sigma^*(s_i,s_{-i}),\omega)|s_i] \\ &= \mathbb{E}_{s_{-i}}[q_i^*(s_i,s_{-i})|s_i] \\ &= \mathbb{E}_{\pi_0\otimes\mu}[q_i^*(s_i,s_{-i})|s_i], \qquad (5)\end{aligned}$$

implying that (σ^*, q^*) is Bayesian individually rational.

Finally, we prove that (σ^*, q^*) is ε-optimal. Let (σ, q) be any other mechanism that is ε-Bayesian incentive compatible and Bayesian individually rational. We will show that $Q(\sigma^*, q^*) \geq Q(\sigma, q)$ i.e. $\mathbb{E}_{\pi_0 \otimes \mu}[\sum_{i \in N} q_i^*(s_i, s_{-i})] \geq \mathbb{E}_{\pi_0 \otimes \mu}[\sum_{i \in N} q_i(s_i, s_{-i})]$.

Since (σ, q) is Bayesian individually rational, we have that for each $i \in N$ and $s_i \in S_i$,

$$\mathbb{E}_{\pi_0 \otimes \mu}[u_i(\sigma(s_i, s_{-i}), \omega) | s_i] \geq \mathbb{E}_{\pi_0 \otimes \mu}[q_i(s_i, s_{-i}) | s_i].$$

Hence, taking the unconditional expectation and summing over all players, we get that

$$\mathbb{E}_{\pi_0 \otimes \mu}[\sum_{i \in N} u_i(\sigma(s_i, s_{-i}), \omega)] \geq \mathbb{E}_{\pi_0 \otimes \mu}[\sum_{i \in N} q_i(s_i, s_{-i})].$$

Since, the social choice function σ^* implements the social optimum it follows that

$$\mathbb{E}_{\pi_0 \otimes \mu}[\sum_{i \in N} u_i(\sigma^*(s_i, s_{-i}), \omega)] \geq \mathbb{E}_{\pi_0 \otimes \mu}[\sum_{i \in N} u_i(\sigma(s_i, s_{-i}), \omega)].$$

By applying (5), taking unconditional expectations and summing over all players, we obtain

$$\mathbb{E}_{\pi_0 \otimes \mu}[\sum_{i \in N} q_i^*(s_i, s_{-i})] = \mathbb{E}_{\pi_0 \otimes \mu}[\sum_{i \in N} u_i(\sigma^*(s_i, s_{-i}), \omega)].$$

Hence, it follows by the previous conclusions, that $Q(\sigma^*, q^*) \geq Q(\sigma, q)$. Thus, we have proved that the mechanism (σ^*, q^*) is ε-optimal. □

With the regard to the above theorem, some remarks are in order. Firstly, by applying standard results on convergence rates for Bayesian posteriors (Ibragimov et al., 1981; Le Cam, 1986; Ghosal et al., 2000), we may further derive a threshold on the number of agents $n_0 = N(\varepsilon, \nu)$ of the order of $O(\frac{1}{\varepsilon^2})$. Secondly, the condition on $\mathscr{U}$ implies that each $u \in \mathscr{U}$ admits a unique maximiser, for each state. The condition would also be satisfied if all agents exhibit the same ordinal preference over alternatives. However, agents could disagree in their state-dependent ordinal ranking over states and yet, condition (1) may be satisfied. The finiteness of $\mathscr{U}$ is applied in the proof of the theorem to ensure that $e < 1$ indeed exists as the set $\Delta(\mathscr{U})$ is compact. Of course, in this setting, $\lambda(u)$ is the proportion of agents having utility function u. Hence, it is essentially the normalised welfare weight on u in the term for aggregate social welfare.

We next discuss some examples. Consider a situation involving single object assignment in which there exist different sets of agent $\{N_u\}_{u\in\mathscr{U}}$ (pairwise disjoint), where $N = \cup_{u\in\mathscr{U}} N_u$ and each agent in N_u has utility function u. Further, the set of alternatives is defined as $A = \mathscr{U}$, meaning that the object is assigned to exactly one of $\{N_u\}_{u\in\mathscr{U}}$. Whichever N_u is assigned the object, an agent in N_u derives state-dependent utility according to u and would have an expected payment equal to the expected value of obtaining the object for N_u. Hence, if not obtaining the object has no value, this means the agent does not pay anything in the mechanism. Perhaps interestingly, the theorem application above would only need $\sum_{u\in\mathscr{U}} |N_u|$ to go to infinity and hence we may have that one set of agents is large relative to other sets of agents. For another example, one may set aside optimality and instead consider ε-Budget Balanced mechanisms (Myerson & Satterthwaite, 1983), where the expected sum of payments would be close to zero. This would be the case, when the distribution of utilities over agents (i.e. profile of utility functions $(u_i)_{i\in N}$) is such that the expected welfare is close to zero, hence the expected sum of payments (for the given payment scheme) would be close to zero. Hence, we would get a mechanism that would satisfy the properties of being ε-BIC, BIR and ε-Budget Balanced. This would demonstrate a situation that would prove to be in contrast to Myerson & Satterthwaite (1983).

References

Azevedo, E. M., & Budish, E. (2019). Strategy-proofness in the large. *Review of Economic Studies*, *86*(1), 81–116.

Balcan, M.-F., Sandholm, T., & Vitercik, E. (2019). Estimating approximate incentive compatibility. *arXiv preprint arXiv:1902.09413*.

Banerjee, A. V. (1992). A simple model of herd behavior. *Quarterly Journal of Economics*, *107*(3), 797–817.

Berger, J. O. (2013). *Statistical Decision Theory and Bayesian Analysis*. Springer.

Bikhchandani, S., Hirshleifer, D., & Welch, I. (1992). A theory of fads, fashion, custom, and cultural change as informational cascades. *Journal of Political Economy*, *100*(5), 992–1026.

Blackwell, D. A., & Girshick, M. A. (1979). *Theory of Games and Statistical Decisions*. Courier Corporation.

Chamley, C. P. (2004). *Rational Herds: Economic Models of Social Learning*. Cambridge University Press.

Charalambos, D., & Aliprantis, B. (2013). *Infinite Dimensional Analysis: A Hitchhiker's Guide*. Springer.

Clarke, E. H. (1971). Multipart pricing of public goods. *Public Choice*, 17–33.

De Condorcet, N. (1785). *Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix*.

DeGroot, M. H. (2005). *Optimal Statistical Decisions* (Vol. 82). John Wiley & Sons.

Epasto, A., Mahdian, M., Mirrokni, V., & Zuo, S. (2018). Incentive-aware learning for large markets. In *Proceedings of the 2018 World Wide Web Conference* (pp. 1369–1378).

Ferguson, T. S. (2014). *Mathematical Statistics: A Decision Theoretic Approach* (Vol. 1). Academic press.

Fudenberg, D., Lanzani, G., & Strack, P. (2021). Pathwise concentration bounds for misspecified bayesian beliefs. *Available at SSRN 3805083*.

Ghosal, S., Ghosh, J. K., & Van Der Vaart, A. W. (2000). Convergence rates of posterior distributions. *Annals of Statistics*, 500–531.

Gibbard, A. (1973). Manipulation of voting schemes: a general result. *Econometrica*, 587–601.

Groves, T. (1973). Incentives in teams. *Econometrica*, 617–631.

Hurwicz, L. (1960). Optimality and informational efficiency in resource allocation processes. *Mathematical Methods in the Social Sciences*.

Ibragimov, I., Minskii, H., & Zalmanovich, R. (1981). *Statistical Estimation: Asymptotic Theory* (Vol. 16). Springer.

Le Cam, L. (1986). *Asymptotic Methods in Statistical Decision Theory*. Springer.

Lee, S. (2016). Incentive compatibility of large centralized matching markets. *Review of Economic Studies*, *84*(1), 444–463.

Maskin, E. (1999). Nash equilibrium and welfare optimality. *Review of Economic Studies*, *66*(1), 23–38.

Myerson, R. B. (1981). Optimal auction design. *Mathematics of Operations Research*, *6*(1), 58–73.

Myerson, R. B., & Satterthwaite, M. A. (1983). Efficient mechanisms for bilateral trading. *Journal of Economic Theory*, *29*(2), 265–281.

Pigou, A. C. (1920). Economics of Welfare. *MacMillan*.

Pratt, J. W., Raiffa, H., Schlaifer, R., et al. (1995). *Introduction to Statistical Decision Theory*. MIT press.

Roberts, D. J., & Postlewaite, A. (1976). The incentives for price-taking behavior in large exchange economies. *Econometrica*, 115–127.

Smith, L., & Sørensen, P. (2000). Pathological outcomes of observational learning. *Econometrica*, *68*(2), 371–398.

Vickrey, W. (1961). Counterspeculation, auctions, and competitive sealed tenders. *Journal of Finance*, *16*(1), 8–37.

Wald, A. (1950). *Statistical Decision Functions*. Wiley.

Journal of Mechanism and Institution Design
ISSN: 2399-844X (Print), 2399-8458 (Online)
DOI:10.22574/jmid.2023.12.005

REDUCING INCENTIVE CONSTRAINTS IN BIDIMENSIONAL SCREENING

Braulio Calagua

Federal University of Rio Grande do Norte, Brazil
braulio.calagua@gmail.com

ABSTRACT

This paper studies screening problems with quasilinear preferences, where agents' private information is two-dimensional and the allocation instrument is one-dimensional. We define a preorder to compare types based on their marginal valuation to the instrument, which facilitates the reduction of incentive compatibility constraints that must be checked. With this approach, the discretized problem becomes computationally tractable. As an application, we numerically solve a problem introduced by Lewis & Sappington (1988b).

Keywords: Two-dimensional screening, incentive compatibility, regulation of a monopoly.

JEL Classification Numbers: D82, L51, C69.

1. INTRODUCTION

SCREENING problems arise in several economic situations, including optimal taxation, nonlinear pricing, regulation of monopolists, and auctions. While much of the existing literature has focused on modeling private information with a one-dimensional parameter, a more nuanced approach with at least two dimensions is required to represent agents' characteristics accurately. For example, in the context of nonlinear pricing, customers may exhibit differences in demand intensity (intercept of demand) and price sensitivity (slope of

I am grateful to Aloisio Araujo, Alexandre Belloni, Mauricio Villalba, Carolina Parra, Victor Perez, and three anonymous referees for their helpful comments. This research was developed during my Ph.D. at IMPA and supported by Conselho Nacional de Desenvolvimento Cientifico e Tecnologico (CNPq).

demand), as in Laffont et al. (1987). Regarding the regulation of monopolists, the regulator may be uncertain about both the marginal and fixed cost of the company, as in Rochet (2009), or may be uncertain about the cost and demand functions of the company, as in Lewis & Sappington (1988b). In auctions, bidders may have private information about both the valuation of the good and their financial constraints, as in Che & Gale (1998).

Compared to one-dimensional models, the analysis of models with two-dimensional agent characteristics and a one-dimensional principal's instrument has been relatively limited. This is because of the inherent difficulty of finding explicit solutions and the challenging task of obtaining numerical approximations. For example, Laffont et al. (1987), Basov (2001), Deneckere & Severinov (2017), and Araujo et al. (2022), have developed some techniques to obtain the solution. However, with these techniques, we are limited to using simple forms of agents' valuation functions and commonly assuming a uniform distribution of types.

This article aims to provide a methodology that enables the numerical solution of a wide variety of bidimensional adverse selection problems. The central assumption is that the agent's marginal valuation can be ranked under each private information parameter. Our main contribution lies in proving that it is sufficient to consider the incentive compatibility constraint with types over a unidimensional set for each agent type rather than the entire bidimensional set as required by definition. Consequently, a significant number of incentive constraints are ruled out in the discretized problem, rendering it computationally tractable even with a relatively fine discretization. In this manner, we can have well-educated predictions about some features of the solution, such as optimal quality, agents' surplus, and optimal tariff (if we refer to the monopolist's problem), as well as the participation region and how types are bunched.

In contrast to most numerical approaches that approximate the solution, this study does not consider the incomplete or relaxed problem in which the constraints are obtained from the necessary conditions of the original problem (Wilson, 1996; Tarkiainen & Tuomala, 1999). Moreover, it does not focus on the Lagrange multipliers associated with the incentive compatibility constraints (Berg & Ehtamo, 2009), nor does it appeal to optimization methods for the complete problem (Judd et al., 2017). Instead, this study explores the idea -from the unidimensional case- that a priori eliminates some constraints when the set of types is finite, and the Spence-Mirrlees condition holds.

It is well known that in one dimension, under the Spence-Mirrlees or single-crossing condition, the types can be ordered by their marginal valuation for the principal's instrument. However, in multiple dimensions, the lack of an exogenous order among types presents challenges in solving the problem. McAfee & McMillan (1988) proposed a generalization of the Spence-Mirrlees condition, according to which the bunching of types must be linear. Most other generalizations have extended the injectivity of marginal valuation as a function of types. Generalizing the Spence-Mirrlees condition has, therefore, proven difficult. One of the key innovations of this article is the introduction of a preorder in the set of types, established by comparing the marginal valuation for the principal's instrument. With this definition and by considering the possible shape of the contour lines of an implementable assignment, we show that it is sufficient to ensure that each type does not wish to imitate another type on a subset of the boundary of the type space.

As an application, we numerically solve the regulation model introduced by Lewis & Sappington (1988b). Armstrong (1999) reviewed this model and showed that Lewis and Sappington's solution was incorrect. Since this model lacks a known analytical solution, obtaining a numerical approximation becomes essential. In addition, Armstrong (1999) conjectured that it is optimal to exclude a positive mass of agents, as is the case for the nonlinear pricing setting. However, the numerical solution suggests that the exclusion should not be optimal in this case. The analysis of this model exemplifies the importance of having an approximation of the solution to make robust predictions.

The plan of the paper is as follows: Section 2 describes the model (in the style of Mussa & Rosen (1978)). Section 3 explains the reduction of incentive constraints, establishes the discretized problem, and studies asymptotic properties when discretization becomes finer. The regulation model of Lewis & Sappington (1988b) is numerically solved −for particular parameters− in Section 4, where I also provide some considerations about the optimality of exclusion. Appendix A tests the method by comparing the numerical solutions with the analytical solutions for some models from the literature. All proofs are included in Appendix B.

2. MODEL

Consider a monopolistic firm that produces a single product with quantity (or quality) $q \in Q \subset \mathbb{R}_+$ at cost $C(q)$. The set $Q \subset \mathbb{R}_+$ represents the production set of the firm. Customers' characteristics, reflecting their preferences over the product, are captured by a bidimensional vector $(a,b) \in [0,1] \times [0,1]$, which is labeled as its *type*. This type refers to the private information for each customer, but the firm knows the probability distribution over $[0,1]^2$ defined by a differentiable density function $\rho(a,b) > 0$. The utility of the customer of type (a,b) is quasi-linear $v(q,a,b) - t$, where $v(q,a,b)$ is the value for consumption $q \in \mathbb{R}_+$, and $t \in \mathbb{R}_+$ is the monetary transfer.

The firm designs a menu of options to offer the customer specifying the quantity and the corresponding payment according to the customer's type revealed. Due to the revelation principle (Myerson, 1979), it is sufficient to restrict attention to contracts where truth-telling is the best response for customers. Thus, to maximize expected revenue, the monopolist's problem is

$$\max_{q(\cdot,\cdot),t(\cdot,\cdot)} \int_0^1 \int_0^1 \big(t(a,b) - C(q(a,b))\big)\rho(a,b)\,da\,db \tag{MP}$$

subject to

$$v(q(a,b),a,b) - t(a,b) \geq v(q(a',b'),a,b) - t(a',b') \tag{IC}$$

and

$$v(q(a,b),a,b) - t(a,b) \geq 0 \tag{IR}$$

Labels (IC) and (IR) refer to *incentive compatibility* and *individual rationality* constraints. We assume that the reservation utility is type independent and normalized at zero. For an incentive-compatible contract $(q(\cdot,\cdot),t(\cdot,\cdot))$, the informational rent is defined as

$$V(a,b) = v(q(a,b),a,b) - t(a,b)$$

This variable V is used to eliminate monetary transfer. The monopolist's problem can now be set as

$$\max_{q(\cdot,\cdot),V(\cdot,\cdot)} \int_0^1 \int_0^1 \big(v(q(a,b),a,b) - C(q(a,b)) - V(a,b)\big)\rho(a,b)\,da\,db \tag{1}$$

subject to

$$\text{(IC)} \quad V(a,b) - V(\widehat{a},\widehat{b}) \geq v(q(\widehat{a},\widehat{b}),a,b) - v(q(\widehat{a},\widehat{b}),\widehat{a},\widehat{b})$$
$$\text{(IR)} \quad V(a,b) \geq 0$$

Assumptions. *Agent's valuation function v is three times differentiable,*[1] *and cost function C is differentiable. Additionally, by denoting q^{out} the exit option, the following is assumed:*

1. $v(q^{\text{out}}, \cdot, \cdot)$ *is constant.*

2. $v_{qa} > 0$ *and* $v_{qb} < 0$.

3. $\frac{d}{dq}\left(\frac{v_{qa}}{v_{qb}}\right) \geq 0$ *and* $\frac{d}{da}\left(\frac{v_{qa}}{v_{qb}}\right) \geq 0$.

Assumption 1 is usually presented as $v(0, \cdot, \cdot) = 0$ because, in nonlinear pricing, the exit option is $q^{\text{out}} = 0$, and all customers assign it zero value. However, in other adverse selection problems, q^{out} could take other values. Assumption 2 is the single-crossing condition in each direction, i.e., the agent's marginal valuation can be ranked with respect to each private information parameter. We have assumed the particular signs mentioned above, but what truly matters is the uniform signs of v_{qa} and v_{qb}. As a consequence, it requires that an implementable $q(\cdot, \cdot)$ be nondecreasing with respect to a and nonincreasing with respect to b. Technical assumptions 3 are given to avoid pathological cases in our main result.

If, in addition, the following condition is satisfied:

$$v_a \geq 0 \quad \text{and} \quad v_b \leq 0 \tag{*}$$

then the informational rent V is nondecreasing with respect to a, and is nonincreasing with respect to b (since $V(a,b)$ is the optimal value of the agent's maximization problem, and according to the envelope theorem, $V_a(a,b) = v_a(q(a,b),a,b)$ and $V_b(a,b) = v_b(q(a,b),a,b)$; hence, $V_a \geq 0$ and $V_b \leq 0$). Thus, it will be sufficient to impose $V(0,1) = 0$ to fulfill all the IR constraints.

Unless otherwise stated, condition (*) will not be treated as an assumption for our model because the main issue concerns IC constraints.

From this point onwards, we focus solely on piecewise almost everywhere twice-differentiable and continuous contracts $(q(\cdot,\cdot), t(\cdot,\cdot))$.

The following result is from Araujo et al. (2022), and it will be useful for further analysis because it provides information about the shape of isoquants, that is, the set of types who are assigned the same quantity.

[1] Due to $v = v(q,a,b)$, we denote by v_q, v_a and v_b the first, second and third partial derivatives of v.

Proposition 1. *If $q(\cdot,\cdot)$ is implementable, then at any point $(a,b) \in [0,1]^2$ in the participation region the vector $(-\frac{v_{qb}}{v_{qa}}(q(a,b),a,b),1)$ is a tangent to the contour line of $q(\cdot,\cdot)$ at level $q(a,b)$.*

As a consequence, taking into account Assumption 2, we have that any isoquant curve in the participation region has a positive slope at any point. In Appendix B: Mathematical Proofs, we present the methodology for deriving this conclusion, which involves the analysis of a partial differential equation obtained from the incentive compatibility restrictions.

3. REDUCTION OF INCENTIVE CONSTRAINTS

The main challenge in numerically solving the problem (1) is associated with the significant number of incentive compatibility constraints. This issue arises when discretizing the set of types $[0,1] \times [0,1]$ into a grid of n points over each axis, leading to $n^4 - n^2$ IC constraints that can lead to memory storage problems, especially with fine discretization. To overcome this challenge, I present a methodology that allows us to reduce the number of IC constraints. This approach draws inspiration from the idea to address IC constraints in the unidimensional case with a finite set of types and when the single-crossing condition holds (see Laffont & Martimort (2001)).

Recall that if a one-dimensional parameter $\theta \in [\underline{\theta}, \overline{\theta}]$ describes the agents' characteristics, the Spence-Mirrlees or single-crossing condition $v_{q\theta} > 0$ allows us to rank the agents by their marginal valuation for consumption in view of v_q is an increasing function on parameter θ. Then, from the point of view of the monopolist, θ_1 is worse than θ_2 whenever $\theta_1 < \theta_2$. In addition, when the set of types is finite, we can a priori eliminate all IC constraints with better types (and verify them ex-post) because better types aim to mimic the worst types rather than vice versa. In fact, provided that $q(\cdot)$ is nondecreasing, it is sufficient for each type to verify the IC constraint with the first worse type (the one adjacent to the left). We extend and apply these ideas to the bidimensional case.

However, in two dimensions, is there any way to know a priori what types are better or worse from the monopolist's viewpoint than an (a,b) type? In the discrete setting, what could it mean that a fixed type (a,b) verifies the IC constraint with the first worst type?

To respond to these questions, we work in the continuous setting. First, we introduce some kind of generalization of the single-crossing condition because, for our purpose, there is no satisfactory generalization of such a condition. Most of the extensions have focused on the injectivity of the function[2] $v_q(q,\cdot)$, but they have not considered the implicit order of types given by the marginal valuation on the instrument. Thus, to compare two types at least partially, we define the following binary relation:

Definition 1. *Given* $(a,b),(\widehat{a},\widehat{b}) \in [0,1]^2$, (a,b) *is* worse than $(\widehat{a},\widehat{b})$, *denoted by* $(a,b) \preceq (\widehat{a},\widehat{b})$, *if and only if*

$$v_q(q,a,b) \leq v_q(q,\widehat{a},\widehat{b}) \quad \forall\, q \in Q$$

Note that $\preceq$ is a preorder (reflexive and transitive) on $[0,1]^2$. This definition attempts to capture the idea that when $(a,b) \preceq (\widehat{a},\widehat{b})$, the (a,b)-agent has no incentive to announce the type $(\widehat{a},\widehat{b})$ because for any $q \in Q$, the $(\widehat{a},\widehat{b})-$agent has greater marginal valuation for consumption and is willing to pay more for each additional unit of the product. Thus, from the point of view of the monopolist, type (a,b) is worse than type $(\widehat{a},\widehat{b})$.

As a direct consequence of Assumption 2, we note that type (a,b) is worse than any other type in the southeast. Figure 1 illustrates these points.

Proposition 2. *Given* (a,b), *if* $\widehat{a} > a$ *and* $\widehat{b} < b$, *then* $(a,b) \preceq (\widehat{a},\widehat{b})$

Then, for a fixed type, we a priori exclude the IC constraints with any type in the southeast, and ex-post verify that these constraints are fulfilled. Specifically, given $(a,b) \in [0,1]^2$, the following IC constraints are omitted:

$$V(a,b) - V(\widehat{a},\widehat{b}) \geq v(q(\widehat{a},\widehat{b}),a,b) - v(q(\widehat{a},\widehat{b}),\widehat{a},\widehat{b}) \quad \forall\, \widehat{a} > a \;,\; \widehat{b} < b \tag{2}$$

To address the second question, we denote the northeast boundary of the set of types for a given $(a,b) \in [0,1]^2$ by

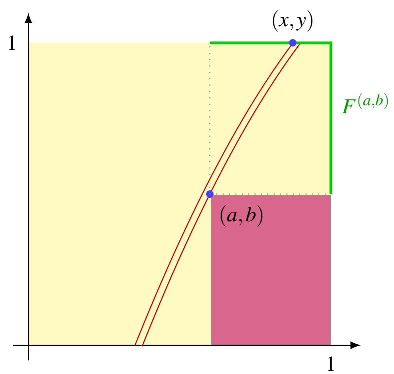
$$F^{(a,b)} := \{(s,1) \mid a \leq s \leq 1\} \cup \{(1,s) \mid b \leq s \leq 1\} \tag{3}$$

Now, consider the isoquant passing through (a,b) which, as understood from Proposition 1, has a positive slope at any point and hence intersects $F^{(a,b)}$. The first worse type is any type located over the closest isoquant on the left, as these agents obtain an inferior level of quantity. Instead of dealing with

[2] See Levin (1999), Chiappori et al. (2010), Figalli et al. (2011)

IC constraints with all the types of the isoquant, we can focus on the type (x,y) located on $F^{(a,b)}$ because it is sufficient to verify the IC constraint with a representative type over each isoquant (this is shown within the proof of Theorem 1 below). Since we do not know the exact shape of the isoquant, we must consider an IC constraint with all the types of $F^{(a,b)}$ in order to guarantee that the (a,b) type verifies the IC constraint with the first worst type.[3] Figure 1 illustrates this situation.

Figure 1: The red curves represent two isoquants. It would be sufficient to impose that the (a,b) type does not mimic the (x,y) type, but since those curves are endogenously determined, the IC constraint with all types represented by the green lines is needed. On the other hand, the purple region represents the types with which the IC constraints are a priori omitted.



Theorem 1. *Let $(q(\cdot,\cdot),V(\cdot,\cdot))$ be such that $q_a \geq 0$, $q_b \leq 0$ and for any $(a,b) \in [0,1]^2$:*

$$V(a,b) - V(x,y) \geq v(q(x,y),a,b) - v(q(x,y),x,y) \quad , \quad \forall\ (x,y) \in F^{(a,b)}$$

Then, $(q(\cdot,\cdot),V(\cdot,\cdot))$ satisfies all the incentive compatibility constraints.

[3] The IC constraint with the horizontal point $(1,b)$ could be ignored, and we do this in the discretized problem. However, in the continuous problem, it is considered because it will be used in the proof of Theorem 1 below.

This result could be understood as analogous to the affirmation *local IC constraints imply global IC constraints*, which is correct in the unidimensional case when single-crossing holds.[4]

The proof is based on three lemmas. The first lemma justifies that it is sufficient to verify the IC constraints with a representative type over an isoquant, and then the constraint will be satisfied with all the types over that isoquant. As a consequence, any type has no incentive to locally misrepresent its true type. Therefore, the contract $(q(\cdot,\cdot),V(\cdot,\cdot))$ satisfies the necessary conditions related to the envelope theorem:

$$V_a(a,b) = v_a(q(a,b),a,b) \quad \wedge \quad V_b(a,b) = v_b(q(a,b),a,b)$$

The second lemma demonstrates that a contract satisfying the previous conditions and monotonicity $q_a \geq 0$, $q_b \leq 0$ will fulfill the a priori omitted IC constraints in (2) with types in the southeast of any given type. The third lemma is more technical and roughly states that if a given type verifies the IC constraint with a particular type over an isoquant, then this constraint is also verified with all the types on the left of that isoquant. This yields the result.

Note that any information about the shape of isoquants could help us eliminate more IC constraints. In this sense, the following proposition provides a sufficient condition of $(a,b) \preceq (\widehat{a},\widehat{b})$ for a particular valuation function:

Proposition 3. *Assume that v_q is concave in the second argument and convex in the third argument. Let (a,b) and $(\widehat{a},\widehat{b})$ be in $[0,1]^2$ such that $a < \widehat{a}$, $b < \widehat{b}$.*
If $\forall\, q \in Q \,:\, \dfrac{\widehat{b}-b}{\widehat{a}-a} \leq \dfrac{-v_{qa}}{v_{qb}}(q,\widehat{a},b)$ *then* $(a,b) \preceq (\widehat{a},\widehat{b})$

To interpret Proposition 3, recall that $\frac{-v_{qa}}{v_{qb}}(q(x,y),x,y)$ is the slope of the isoquant at level $q(x,y)$ passing through point (x,y). Thus, if the slope of any possible isoquant passing through $(\widehat{a},b)$ is greater than the slope between (a,b) and $(\widehat{a},\widehat{b})$, we can be sure that $(a,b) \preceq (\widehat{a},\widehat{b})$. In example (5.1) of Appendix A: Testing the method, we use this proposition to further reduce the number of IC constraints.

3.1. Discretized problem

By Theorem 1, it is sufficient for each type to satisfy the IC constraint with all the points over a unidimensional set instead of the whole square. This

[4] See Section 3.2.1 in Salanié (1997).

result allows us to approximate the solution of the continuous problem by discretizing the set of types. This section is devoted to establishing the discrete problem and discussing its limitations.

Let $X_n = \{0, \frac{1}{n-1}, \frac{2}{n-1}, \dots, 1\} \times \{0, \frac{1}{n-1}, \frac{2}{n-1}, \dots, 1\}$ be the grid of n^2 points on $[0,1]^2$. For a fixed (a,b) with $a<1$ and $b<1$, let $\widetilde{F}^{(a,b)} := F^{(a,b)} \cap X_n$, where $F^{(a,b)}$ is defined in (3). Because for points over the line $x=1$ or $y=1$, the constraints with the points on the northeast cannot be written, we consider equivalently

$$\widetilde{F}^{(a,1)} = \big(\{(0,s) : 0 \leq s \leq 1\} \cup \{(s,0) : 0 \leq s < a\}\big) \cap X_n$$

$$\widetilde{F}^{(1,b)} = \big(\{(0,s) : 0 \leq s \leq b\} \cup \{(s,0) : 0 \leq s < 1\}\big) \cap X_n$$

Thus, $\widetilde{F}^{(a,b)}$ is the set of types with which (a,b) must satisfy an IC constraint.

We approximate the integral in the monopolist's objective by the trapezoidal rule.[5] Let $w(i,j)$ be the associated weight for each point $(a_i,b_j) \in X_n$. Additionally, denote $q_{i,j} := q(a_i,b_j)$ and $V_{i,j} := V(a_i,b_j)$. We are interested in solving the following discretized problem:

$$\begin{aligned}
&\max_{\{q_{i,j},V_{i,j}\}} \sum_{i=1}^{n}\sum_{j=1}^{n} w(i,j)\big(v(q_{i,j},a_i,b_j) - V_{i,j} - C(q_{i,j})\big)\rho(a_i,b_j) \\
&\text{s.t.} \\
&\quad (\widetilde{\text{IC}}) \quad V_{i,j} - V_{k,l} \geq v(q_{k,l},a_i,b_j) - v(q_{k,l},x_k,y_l) \ \ \forall\, (x_k,y_l) \in \widetilde{F}^{(a_i,b_j)} \\
&\quad (\text{IR}) \quad V_{i,j} \geq 0 \\
&\quad (\text{M}) \quad q_{i,j} \leq q_{i+1,j} \ , \ \ q_{i,j} \leq q_{i,j-1}
\end{aligned} \tag{4}$$

Remarks:

1. In the original discretized problem, there are $n^4 - n^2$ (maybe nonlinear) IC constraints. After our methodology, the number of IC constraints is in the order n^3.

2. If case condition (*) is verified, all IR constraints can be replaced by $V_{1,n} = 0$.

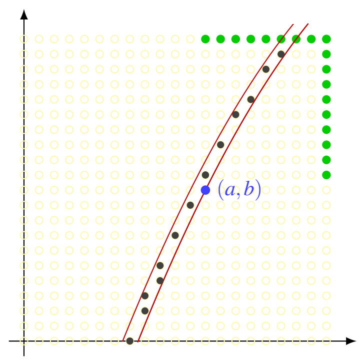
[5] Note that we consider the discretization of the objective function of the continuous formulation of the problem. We do not consider a discretized density function on the set of types.

3. When the valuation function has the special *multiplicative separable* form $v(q,a,b) = \psi(q) + \alpha(a,b)q + \beta(a,b)$, the IC constraints become linear in $q_{i,j}$. Therefore, since the IC constraints are linear in $V_{i,j}$ (regardless of v), if the objective function is strictly concave,[6] there exists a unique solution allowing us to rely on numerical approximations.

Asymptotic analysis

Due to discretization, it is impossible to ensure that all IC constraints are fulfilled for each type $(a,b) \in X_n$. This is because there could be some points between the isoquant passing through (a,b) and the closer to the left isoquant curve intersecting $\widetilde{F}^{(a,b)}$. Figure 2 illustrates this issue.

Figure 2: By discretization, IC constraints are not ensured with black points.



Nevertheless, we proceed by following Belloni et al. (2010),[7] and prove that the violations of the IC constraints (i.e., the terms for which these constraints are not satisfied) uniformly converge to zero with finer discretizations, and the sequence of optimal values converges to the optimal value of the continuous problem.

For this purpose, denote (Q^n, V^n) to the solution of the discretized problem (4). Since these functions are defined on X_n, we define the extensions $\widetilde{Q}^n, \widetilde{V}^n$:

[6] This will be the case if $\psi'' - C'' < 0$

[7] Belloni et al. (2010) considered a linear model including multiple agents and border constraints. These constraints are related to the allocation treated as a probability since, in their model, there are N buyers and J degrees of product quality.

$[0,1]^2 \to \mathbb{R}$ by

$$\widetilde{Q}^n(x,y) := Q^n(a,b) \quad , \quad \widetilde{V}^n(x,y) := V^n(a,b)$$

where $(a,b) \in X_n$ is such that $a \le x < a + \frac{1}{n-1}$ and $b - \frac{1}{n-1} < y \le b$.

We then define $\delta^*(\widetilde{Q}^n, \widetilde{V}^n)$ as the supremum (in absolute value) of the violations of the IC constraints by the pair $(\widetilde{Q}^n, \widetilde{V}^n)$.[8] That is, although some constraints are not fulfilled, we can be sure that for any $(a,b),(a',b') \in [0,1]^2$:

$$\widetilde{V}^n(a,b) - \widetilde{V}^n(a',b') \ge v(\widetilde{Q}^n(a',b'),a,b) - v(\widetilde{Q}^n(a',b'),a',b') - \delta^*(\widetilde{Q}^n, \widetilde{V}^n)$$

The next proposition shows the asymptotic feasibility of the extensions $(\widetilde{Q}^n, \widetilde{V}^n)$. That is, all IC constraint violations uniformly converge to zero.

Proposition 4. *We have* $\delta^*(\widetilde{Q}^n, \widetilde{V}^n) \le O(\frac{1}{n-1})$.

The following proposition shows the asymptotic optimality of the numerical solutions. That is, the objective function value of the discretized problem converges to the objective function value of the continuous problem as the discretization becomes finer.

Proposition 5. *Let OPT_n be the optimal value of the discretized problem, and let OPT^* be the optimal value of the continuous problem. Then, $\liminf_{n\to\infty} OPT_n \ge OPT^*$. Additionally, if $\lim_{n\to\infty} \widetilde{Q}^n(a,b)$ and $\lim_{n\to\infty} \widetilde{V}^n(a,b)$ exists for any $(a,b) \in [0,1]^2$, then $\lim_{n\to\infty} OPT_n = OPT^*$.*

4. NUMERICAL SOLUTION: REGULATING A MONOPOLIST FIRM

Lewis & Sappington (1988b) studied the design of regulatory policy when the regulator imperfectly knows both the costs and the demand functions of the monopolist firm under regulation. This model was then reviewed by Armstrong (1999) who observed that Lewis and Sappington's solution for a particular

[8] If type (a,b) does not satisfy the IC restriction with type (a',b') then $\widetilde{V}^n(a,b) - \widetilde{V}^n(a',b') - (v(\widetilde{Q}^n(a',b'),a,b) - v(\widetilde{Q}^n(a',b'),a',b')) < 0$. Thus, $\delta^*(\widetilde{Q}^n, \widetilde{V}^n)$ is formally defined as

$$-\delta^*(\widetilde{Q}^n, \widetilde{V}^n) := \inf\{\widetilde{V}^n(a,b) - \widetilde{V}^n(a',b') - (v(\widetilde{Q}^n(a',b'),a,b) - v(\widetilde{Q}^n(a',b'),a',b'))\}$$

where the infimum is taken over all pairs $(a,b),(a',b') \in [0,1]^2$.

example was incorrect; however, he did not find the correct solution. Here, I present the numerical solution of the problem and discuss a conjecture made by Armstrong (1999) about the optimality of excluding a positive mass of agents.

Lewis and Sappington's model

In the framework of the regulation of a monopolistic company, Lewis & Sappington (1988b) considered that the demand for the firm's product $q = Q(p,a)$ and the costs of producing output q, $C(q,b)$, involve the firm's private information parameters (a,b) distributed over $\Theta = [\underline{a},\overline{a}] \times [\underline{b},\overline{b}]$ according to a strictly positive density function $f(a,b)$.

The regulator offers the firm a menu of unit prices p and a corresponding subsidy t conforming to the firm's type is revealed. The profit of the firm of type (a,b) is $pQ(p,a) - C(Q(p,a),b) + t$. The profit reservation level is type independent and normalized at zero. It is assumed that the regulator can ensure that the firm serves all demand at the established prices. The regulator's objective function is the expected consumer surplus net of the transfer to the firm

$$\int_{\underline{a}}^{\overline{a}} \int_{\underline{b}}^{\overline{b}} \{\Pi(Q(p(a,b),a),a) - p(a,b)Q(p(a,b),a) - t(a,b)\} f(a,b)\,db\,da \quad (5)$$

where $\Pi(Q,a) = \int_0^Q P(\xi,a)d\xi$, and $P(\cdot)$ denotes the inverse demand curve.

The regulator's problem is to design the menu of contracts $(p(a,b),t(a,b))$ to maximize (5) subject to individual rationality

$$p(a,b)Q(p(a,b),a) - C(Q(p(a,b),a),b) + t(a,b) \geq 0$$

and incentive compatibility constraints

$$\begin{aligned} p(a,b)Q(p(a,b),a) - C(Q(p(a,b),a),b) + t(a,b) \geq \\ p(\widehat{a},\widehat{b})Q(p(\widehat{a},\widehat{b}),a) - C(Q(p(\widehat{a},\widehat{b}),a),b) + t(\widehat{a},\widehat{b}) \end{aligned}$$

Lewis & Sappington derived a solution for the particular example

$$Q(p,a) = \alpha - p + a \ , \ C(q,b) = K + (c_0 + b)q \quad (6)$$

with α, K and c_0 positive constants and a uniform distribution over $\Theta = [0,1]^2$. However, in Armstrong (1999), the author noted that Lewis and Sappington's solution for this example was incorrect, but without solving it correctly himself.

Armstrong's conjecture

Furthermore, Armstrong (1999) argued that excluding a positive mass of types should be optimal, like in the nonlinear pricing setting where it is optimal for a monopolist to exclude a positive mass of customers. However, due to the change in the variables he used, the type set is not convex, and his exclusion argument cannot strictly be applied. Armstrong also expressed the following:

"*Nevertheless, I believe that the condition that the support be convex is strongly sufficient and that it will be the usual case that exclusion is optimal...*"

Numerical solution

Before solving the problem numerically, we introduce a change of variables to fit the standard problem (1) from Section 2. By setting

$$\begin{aligned} v(p,a,b) &= pQ(p,a) - C(Q(p,a),b) \\ H(p,a) &= pQ(p,a) - \Pi(Q(p,a),a) \\ V(a,b) &= v(p(a,b),a,b) + t(a,b) \end{aligned}$$

the regulator's problem can be written as (notice that the new variable V is the firm's profit)

$$\max_{p(\cdot),V(\cdot)} \int_{\underline{a}}^{\overline{a}} \int_{\underline{b}}^{\overline{b}} \{v(p(a,b),a,b) - H(p(a,b),a) - V(a,b)\} f(a,b)\, db\, da \quad \text{(RP)}$$

subject to

(IR) $V(a,b) \geq 0 \qquad \forall\, (a,b) \in \Theta$

(IC) $V(a,b) - V(\widehat{a},\widehat{b}) \geq v(p(\widehat{a},\widehat{b}),a,b) - v(p(\widehat{a},\widehat{b}),\widehat{a},\widehat{b}) \qquad \forall\, (a,b),(\widehat{a},\widehat{b}) \in \Theta$

We focus on solving the problem with $Q(p,a)$ and $C(q,b)$ as in (6) and the uniform distribution of types over $\Theta = [0,1]^2$.

Since $v(p,a,b) = (\alpha + a - p)(p - c_0 - b) - K$, then $v_{pa} = v_{pb} = 1$. In this case, $p(\cdot,\cdot)$ will be nondecreasing in a and b. Additionally, since $\frac{-v_{pb}}{v_{pa}} < 0$, from Proposition 1 we find that any isoprice curve has a negative slope at any point and hence intersects the northwest boundary of the set of types.

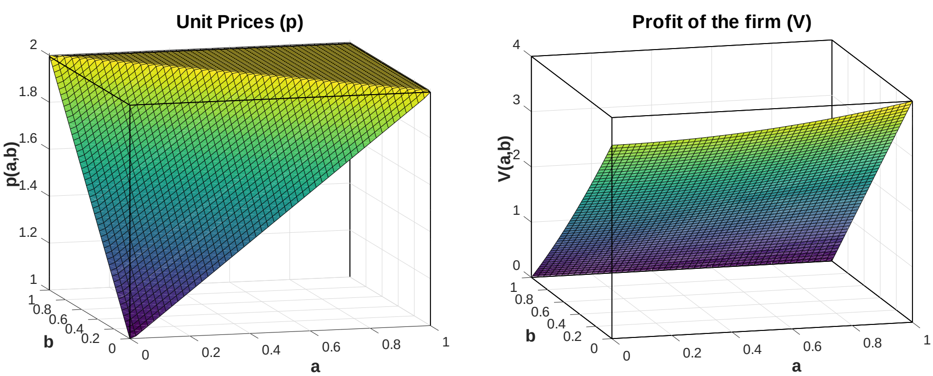
Following the same considerations as in Section 3, it will be sufficient that each $(a,b)-$agent verifies the IC constraints with all the points over the set

$$F^{(a,b)} := \{(0,s) \mid b \leq s \leq 1\} \cup \{(s,1) \mid 0 \leq s \leq a\}$$

Note that the discretized problem has a unique solution: the multiplicative separable form of v implies linearity of the IC constraints (see remark 3 on page 117) and the objective function is strictly concave. Note also that the signs of v_a and v_b are endogenously determined; hence, condition (*) cannot be verified and all the IR constraints must be considered.

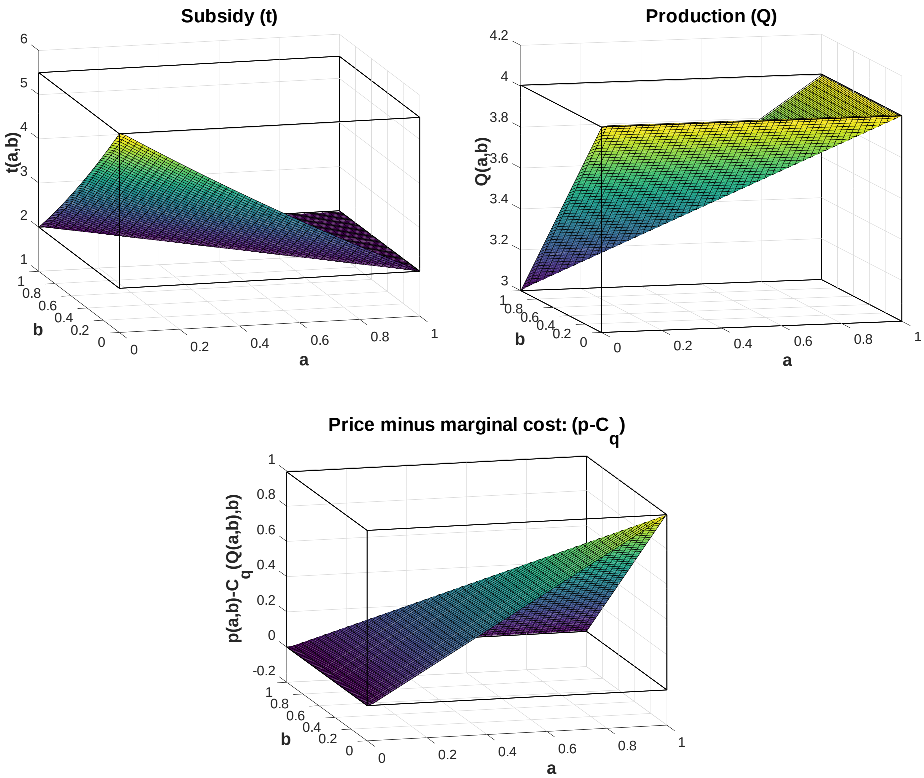
We numerically solved the problem for the case of $c_0 = 1$, $\alpha = 5$, and $K = 2$ with $n = 51$ points over each direction. The numerical solution $(p^n(\cdot,\cdot), V^n(\cdot,\cdot))$ was obtained via Knitro/AMPL by using the active set algorithm. The optimal value was $OPT_n = 4.21$, and the maximum violation of the IC constraints (in absolute value) was $\delta^* = 8.403 \times 10^{-11}$. Figure 3 shows the graphs of $p^n(\cdot,\cdot)$ and $V^n(\cdot,\cdot)$.

Figure 3: Graphs of optimal prices and firm's profit for the case $c_0 = 1$, $\alpha = 5$, and $K = 2$ with $n = 51$.



From the solution, we calculated the optimal subsidies $t(\cdot,\cdot)$, the quantity produced by the firm $Q(p(\cdot,\cdot),\cdot)$, and the difference between unit prices and marginal costs $p(\cdot,\cdot) - C_q(q(\cdot,\cdot),\cdot)$. The graphs of these functions are shown in Figure 4.

Figure 4: Graphs of optimal subsidies, production of the firm, and the difference between unit prices and marginal costs for the case $c_0 = 1$, $\alpha = 5$, and $K = 2$ with $n = 51$.



Some insights from the numerical solution

This example gives rise to an optimization problem with linear constraints, and it possesses a unique solution, making numerical methods for solving highly efficient. Additionally, the solution for different values of c, α, and K presents the same features. As a result, the following statements are considered reliable:

1. It seems that, at the optimum, all types (a,b) with $a+b \geq 1$ are bunching at unit price $c_0 + 1$, and their subsidy is the fixed cost K. Additionally, the unit price assigned to type $(0,0)$ seems to be c_0. In fact, I conjecture the optimum price p to be $p(a,b) = c_0 + a + b$ when $a+b \leq 1$ and

$p(a,b) = c_0 + 1$ when $a+b > 1$.

2. At the optimum, the firm produces a positive quantity regardless of its type.

3. In view of the numerical difference $p - C_q$, the regulator induces the firm to price above marginal costs for almost all (a,b) types rather than $a = 0$ or $b = 1$ (i.e., such types with the a priori lowest demand or such types with the highest costs).[9]

4. The type of firm with the highest cost function receives zero profit.

5. The firm's numerical profit V suggests that there is no exclusion.

Some additional conclusions could be made if our conjecture of optimal unit prices were true. In this case, in the region of types (a,b) with $a+b \leq 1$, the isoprices are lines of negative slope. Given $s \in [0,1]$, the *adjusted marginal cost* along the isoprice of vertical intercept s is defined as

$$AMC(s) = \frac{\int_0^s C_q(Q(p(\tilde{a}, s-\tilde{a}), \tilde{a}), s-\tilde{a}) f(\tilde{a}, s-\tilde{a}) d\tilde{a}}{\int_0^s f(\tilde{a}, s-\tilde{a}) d\tilde{a}} + \frac{\int_0^s F(\tilde{a}, s-\tilde{a}) d\tilde{a}}{\int_0^s f(\tilde{a}, s-\tilde{a}) d\tilde{a}}$$

where $F(a,b) = \int_0^b f(a,\tilde{b})\, d\tilde{b}$. Then, for all types (a,b) with $a+b = s \leq 1$, the optimal price equals the adjusted marginal cost. That is,

$$p(a,b) = AMC(s)$$

Lewis & Sappington (1988b) introduced the previous definition of $AMC(s)$ as the sum of the expected marginal costs given s and the bidimensional version of the inverse hazard rate. They also provided the following interpretation of the second term in $AMC(s)$: "*Intuitively, this term captures the optimal mark-up of price above expected marginal cost. The mark-up balances the expected losses from inefficiently low output with the expected gains from reduced information rents that accrue to the firm because of its private knowledge.*"

Baron & Myerson (1982) obtained a conclusion related to '*price equals adjusted marginal cost*' in their analysis of a model where the regulator faces

[9] In the model of Lewis & Sappington (1988a), the regulator is uncertain only about the position of the demand curves. In that model, if $C''(q) \geq 0$ (similar to here), setting prices at the level of marginal costs for the reported demand is optimal ($p = C_q$).

uncertainty solely about the firm's cost function. According to their findings, at the optimum, prices exceed marginal costs for all cost realizations except the lowest one.

Moreover, as previously mentioned, when considering types (a,b) with $a+b>1$, bunching is observed at the same price, and the same subsidy, which is the firm's fixed cost, is obtained. This subsidy feature in half of the realizations aligns with the traditional policy regulation approach used in situations without information asymmetry.

In any case, the optimal price is set above the realized marginal cost for almost all types of firms, as clearly indicated by the numerical results.

About the optimality of exclusion

In screening problems with multi-dimensional types, the optimality of exclusion refers to the Principal's optimal contract design where a positive mass of agents chooses not to participate. Armstrong (1996) first established this result in the nonlinear pricing setting.

Although the optimality of exclusion is a generic property (that is, valid for almost all parameter's values), perhaps the most intriguing insight from the numerical solution is the possibility that, in this regulation model, non-exclusion of a positive mass of types should be optimal; this is contrary to Armstrong's conjecture stated previously.

Furthermore, in Barelli et al. (2014), the authors relaxed the strong conditions from Armstrong (1996) (strict convexity and homogeneity of degree one) and proved a more general result regarding the generic desirability of exclusion. For the example we have analyzed, they considered prices to belong to $[c_0+1,\alpha]$ to conclude that their result can be applied, and hence confirm Armstrong's conjecture. However, the numerical results indicate that prices do not belong to $[c_0+1,\alpha]$. Therefore, their theorem should not be applied.

Next, I will provide one technical argument explaining why Armstrong's Theorem regarding the desirability of exclusion, formulated in the context of nonlinear pricing, could not be extended to this regulation model.

In nonlinear pricing, the customer's exit option is $q^{\text{out}}=0$ and $t^{\text{out}}=0$. Hence, the natural assumptions $v(0,a,b)=0$ and $C(0)=0$ imply that the monopolist's revenue $v(0,a,b)-C(0)-V(a,b)$ is zero when $V(a,b)=0$ (that is, when type (a,b) is excluded). Then, the monopolist's penalty for causing some customers to exit the market is to not receive income from them.

On the other hand, in the regulation model, the exit option of firm (a,b) is the unit price p^{out} and subsidy t^{out} at which profit $V(a,b)$ is zero. For such a firm, in the previous example given by (6), the regulator's benefit is $(\alpha + a - c_0 - b)Q(p^{\text{out}}, a) - (Q(p^{\text{out}}, a))^2/2 - K$ (in fact, if there is no production, this amount is $-K$).

Therefore, in contrast with the monopolist, the regulator could face the necessity of assuming a negative penalty when excluding a firm. Thus, Armstrong's argument of comparing benefits (more income from customers still in the market) versus penalties (zero income from customers excluded) might not be applicable to this model.

In addition, based on the model formulation, it becomes apparent that the regulator has no interest in excluding the firm. This situation may arise if, for example, the monopolist firm is already operational and having zero production is deemed undesirable for the economy. The concept of individual rationality reflects that regardless of the characteristics of costs and demand facing the firm, whether it is producing or not, the subsidy provided by the regulator must be adjusted such that the firm's profit is at least zero.

In a regulation model where the firm refrains from production when it cannot make a positive profit and the regulator pays zero, an additional variable is necessary, indicating the probability that the firm will be allowed to produce (such as the $r(\cdot)$ variable in Baron & Myerson (1982) or Rochet (2009)). The optimal value of this variable will indicate whether a type of firm is excluded.

Appendix A: Testing the method

In this appendix, we compare the numerical solution of problem (4) with the analytical solution of the following three models from the literature. Laffont et al. (1987) considered that a monopolist faces customers with linear demand curves and is uncertain about both the slope and intercept of such linear demand, leading to a linear-quadratic valuation from the customers $v(q,a,b) = aq - (1+b)q^2/2$. Basov (2001) proposed the Hamiltonian approach and solved the generalization $v(q,a,b) = aq - (1+b)q^{\gamma}/\gamma$ with $\gamma \geq 2$, for which demand curves are concave. Araujo et al. (2022) analyzed a case of convex demand curves, for which the customers' valuation is $v(q,a,b) = (c-b)\log(aq+1)$.

Two criteria are presented for comparing our approximations. The first one involves computing the *average quadratic error (a.q.e.)* between analytical quantity q^{exact} and numerical quantity q^{num}, and between analytical and numerical informational rent:

$$a.q.e.(q^{\text{num}}, q^{\text{exact}}) := \frac{1}{n^2} \sum_{i,j=1}^{n} (q_{i,j}^{\text{num}} - q_{i,j}^{\text{exact}})^2$$

$$a.q.e.(V^{\text{num}}, V^{\text{exact}}) := \frac{1}{n^2} \sum_{i,j=1}^{n} (V_{i,j}^{\text{num}} - V_{i,j}^{\text{exact}})^2$$

We also provide the distance between the profit of the analytical solution and the profit of the numerical solution: $\left| \text{profit}^{num} - \text{profit}^{exact} \right|$.

The second criterion involves a visual comparison. Although it may not be as formal as the first criterion, in practice, the numerical approximation helps in formulating predictions about the functional form of the solution, such as the participation set or the contour lines (i.e., how types are bunching). We provide graphs of the quantity and the informational rent for both the numerical and analytical solutions, along with contour lines and cross-sections.

Some graphs are included to determine whether a fixed type (a_i, b_j) (shown in blue) satisfies the IC constraint with the other types in X_n. A green point is drawn if such a restriction is fulfilled and a red point is drawn if it is not. Due to numerical optimization, as well as the limitations of the discretization noted in the remarks of subsection 3.1, it is not surprising that red points exist in some graphs; however, the violations may be considered small. The value of $\delta = \delta_{i,j}$ in each of the graphs indicates the maximum value (in absolute value) of the violation of IC constraints between (a_i, b_j) and the red points. That is, for any $tol > \delta_{i,j}$ we have that $\forall\ (x,y) \in X_n$:

$$V^{\text{num}}(a_i, b_j) - V^{\text{num}}(x,y) - (v(q^{\text{num}}(x,y), a_i, b_j) - v(q^{\text{num}}(x,y), x, y)) > -tol$$

Furthermore, by defining $\delta^* = \max_{i,j}\{\delta_{i,j}\}$, if $tol > -\delta^*$, a similar version of the previous inequality is valid for any two points in the grid X_n. The value of δ^* is also provided in each example.

Numerical solutions were performed via Knitro/AMPL using the Active Set Algorithm. The optimization process stopped if one of the following tolerances was achieved: `maxit`$= 10^4$, `feastol`$= 10^{-15}$, `xtol`$= 10^{-15}$, and `opttol`$= 10^{-15}$, where `maxit` is the maximum number of iterations, `feastol` refers to feasibility tolerance, `xtol` is the relative change of decision variables and `opttol` is the optimality KKT stopping tolerance. In all examples, `xtol` were achieved first.

5.1. Example 1. Linear Demand

In Laffont et al. (1987), the authors have solved the monopolist's problem for the data

$$v(q,a,b) = aq - \frac{(1+b)}{2}q^2 \ , \ C(q) = 0 \ , \ f(a,b) = 1$$

The solutions q and T they found are:

$$q(a,b) = \begin{cases} 0 & , \ a \leq \frac{1}{2} \\ \dfrac{4a-2}{4b+1} & , \ \dfrac{1}{2} \leq \dfrac{a+2b}{4b+1} \leq \dfrac{3}{5} \\ \dfrac{3a-1}{2+3b} & , \ \dfrac{3}{5} \leq \dfrac{2a+b}{2+3b} \leq 1 \end{cases}$$

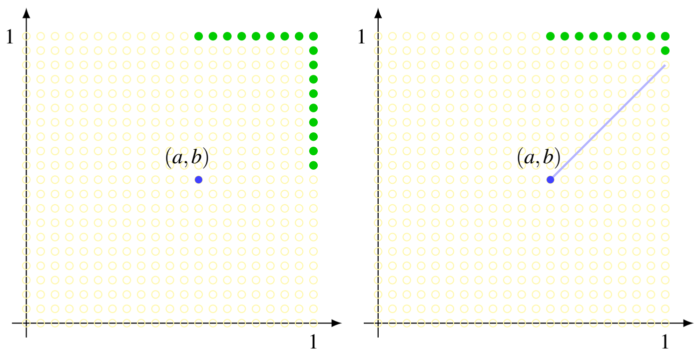
$$T(q) = \begin{cases} \dfrac{q}{2} - \dfrac{3q^2}{8} & , \ q \leq \dfrac{2}{5} \\ \dfrac{q}{3} - \dfrac{q^2}{6} + \dfrac{1}{30} & , \ \dfrac{2}{5} \leq q \leq 1 \end{cases}$$

Note that v_q is linear in each a and b variable. Then, by Proposition 3, the number of constraints can be reduced even further. Since the better type $(1,0)$ has no distortion with respect to the contract over complete information, we must have $v_q(q(1,0),1,0) = 0$ (marginal utility equals marginal cost), which implies $q(1,0) = 1$. Then, $Q = [0,1]$ because the output option is $q^{\text{out}} = 0$. Thus, $\frac{-v_{qa}}{v_{qb}} = \frac{1}{q} \geq 1$. Therefore, for any (a,b), $(\widehat{a},\widehat{b})$ with $\widehat{a} > a$, $\widehat{b} > b$, it is sufficient that $\frac{\widehat{b}-b}{\widehat{a}-a} \leq 1$ to ensure that $(a,b) \preceq (\widehat{a},\widehat{b})$. Figure 5 illustrates the difference in IC constraints considered.

The exact number of IC constraints is $\frac{1}{2}(3n^3 - 3n^2 - 4n + 4)$ instead of $n^4 - n^2$ as in the original problem. The problem was solved with $n = 36$. For this value, 67 970 IC constraints were considered whereas 1 610 350 were ruled out.

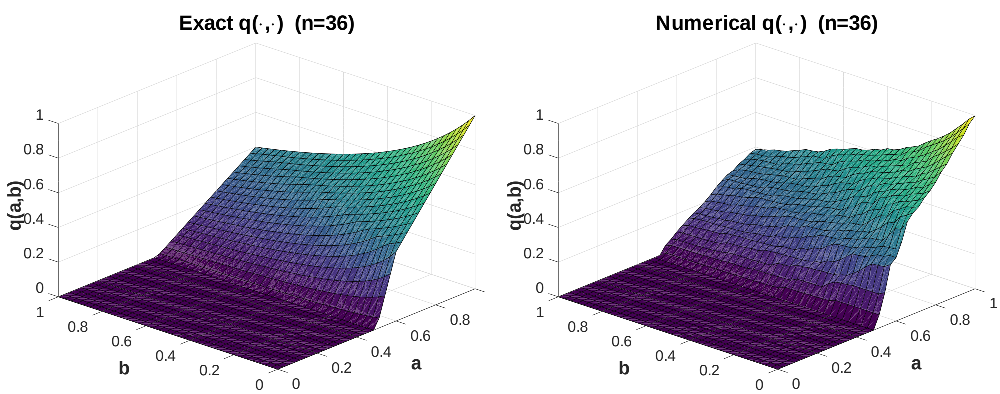
Next, the numerical solution is compared with the analytical (exact) solution.

Figure 5: The green points in both graphics represent the types with which the IC constraint is required. The graphic on the left corresponds to the original formulation, whereas the graphic on the right is after applying Proposition 3. The slope of the blue line is 1.

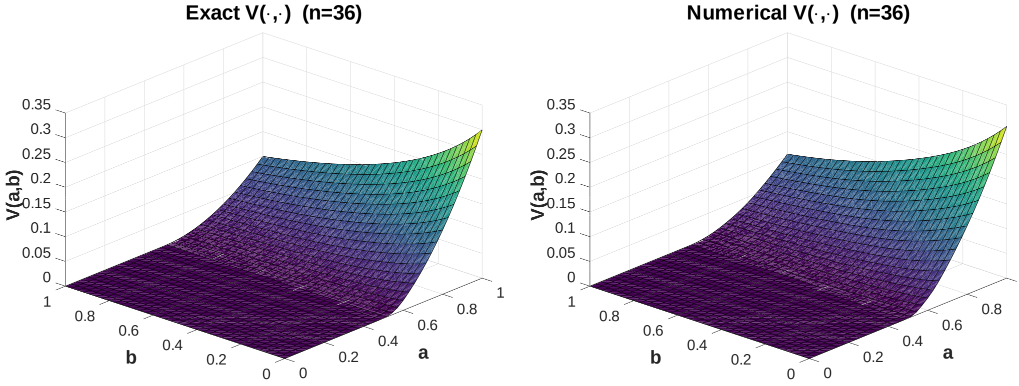


$$\begin{aligned} a.q.e.(q^{num}, q^{exact}) &= 3.6442 \times 10^{-4} \\ a.q.e.(V^{num}, V^{exact}) &= 0.1149 \times 10^{-4} \\ \left| \text{profit}^{num} - \text{profit}^{exact} \right| &= 9.6668 \times 10^{-4} \\ \delta^* &= 1.70804 \times 10^{-4} \end{aligned}$$

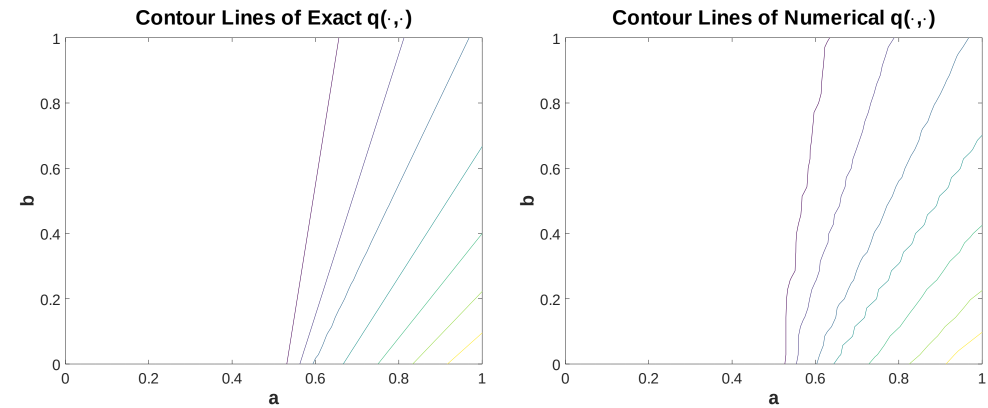
Comparing Quantity



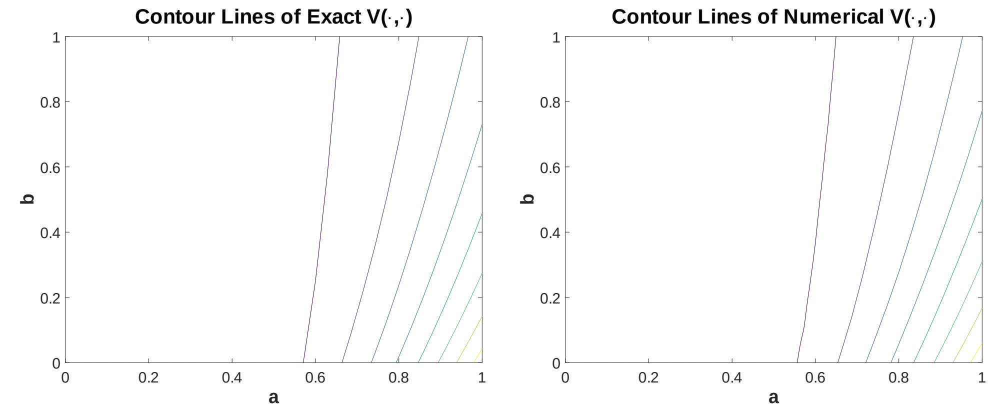
Comparing Informational Rent



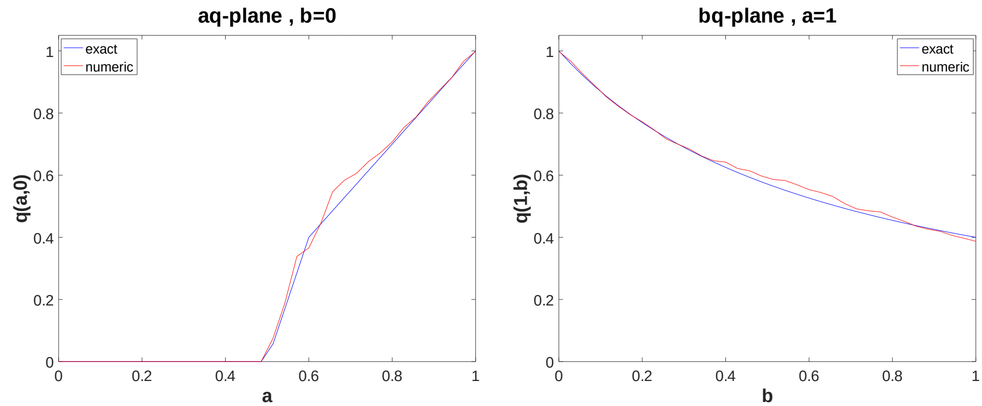
Contour Lines of Quantity



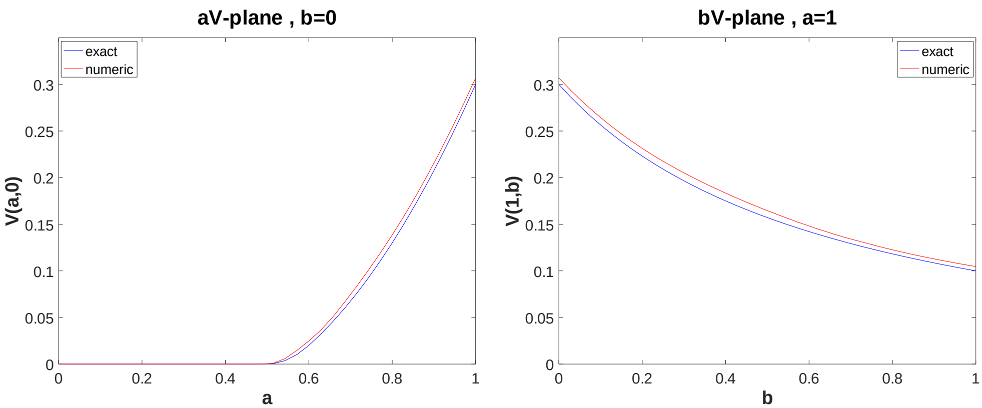
Contour Lines of Informational Rent



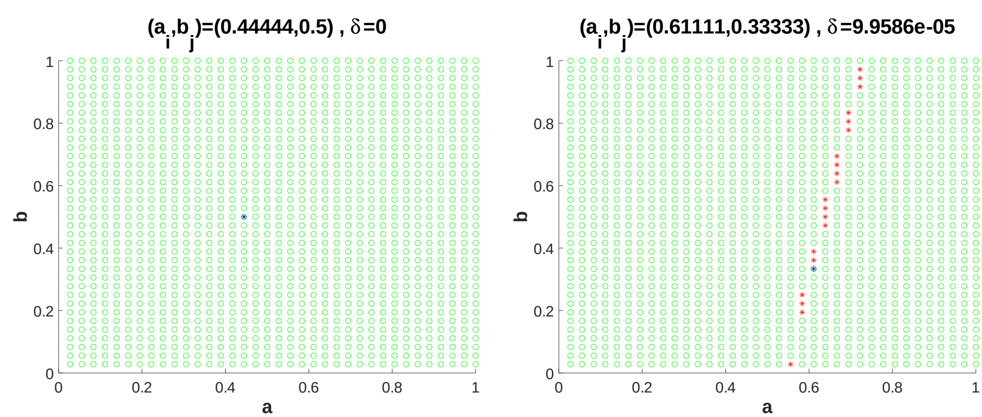
Cross-sections of Quantity

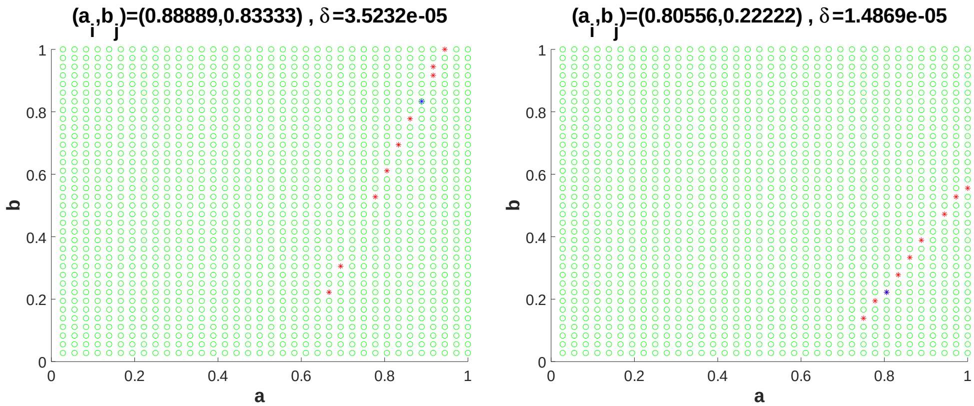


Cross-sections of Informational Rent



Verifying IC constraints





5.2. Example 2. Concave Demand

In Basov (2001), the author solved the original problem for the data

$$v(q,a,b) = aq - \frac{(c+b)}{\gamma} q^{\gamma} \ , \ C(q) = 0 \ , \ f(a,b) = 1$$

where $c > \frac{1}{2}$ and $\gamma \geq 2$ are constants. The solutions q and T he have found are:

$$q(a,b) = \begin{cases} 0 & , \ a \leq \frac{1}{2} \\ \left(\frac{4a-2}{4b+2c-1}\right)^{\frac{1}{\gamma-1}} & , \ (3+2c)a - 2b \leq 2c+1 \\ \left(\frac{3a-1}{3b+2c}\right)^{\frac{1}{\gamma-1}} & , \ (3+2c)a - 2b > 2c+1 \end{cases}$$

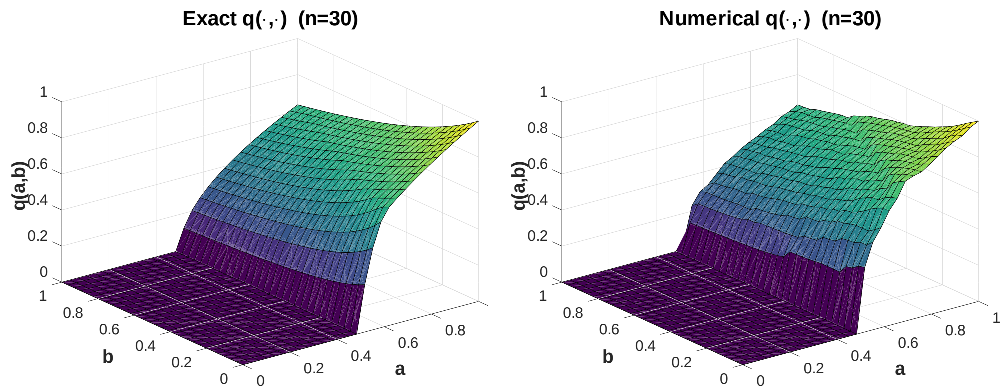
$$T(q) = \begin{cases} \frac{q}{2} - \frac{\left(\frac{c}{2}+\frac{1}{4}\right)}{\gamma} q^{\gamma} & , \ q \leq \left(\frac{2}{3+2c}\right)^{\frac{1}{\gamma-1}} \\ \frac{1}{6}\left(\frac{2}{3+2c}\right)^{\frac{1}{\gamma-1}} - \frac{\left(\frac{c}{6}+\frac{1}{4}\right)}{\gamma}\left(\frac{2}{3+2c}\right)^{\frac{\gamma}{\gamma-1}} + \frac{q}{3} - \frac{c}{3\gamma} q^{\gamma} & , \ q > \left(\frac{2}{3+2c}\right)^{\frac{1}{\gamma-1}} \end{cases}$$

The discretized problem was solved for the case $c = 1$ and $\gamma = 3$ with $n = 30$ points over each axis. For this value, 39 092 incentive compatibility constraints were considered, and 770 008 were eliminated.

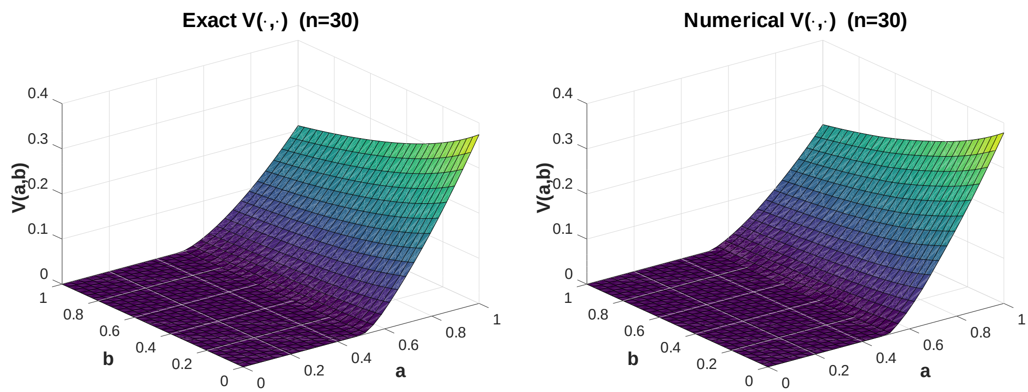
Next, the numerical solution is compared with the analytical (exact) solution.

$$
\begin{aligned}
a.q.e.(q^{num}, q^{exact}) &= 4.5853 \times 10^{-4} \\
a.q.e.(V^{num}, V^{exact}) &= 0.0384 \times 10^{-4} \\
\left| \text{profit}^{num} - \text{profit}^{exact} \right| &= 2.5717 \times 10^{-3} \\
\delta^* &= 7.82371 \times 10^{-4}
\end{aligned}
$$

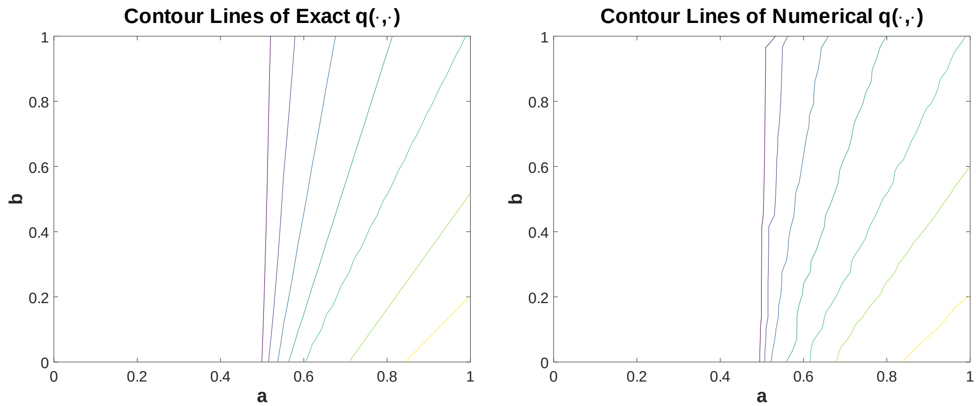
Comparing Quantity



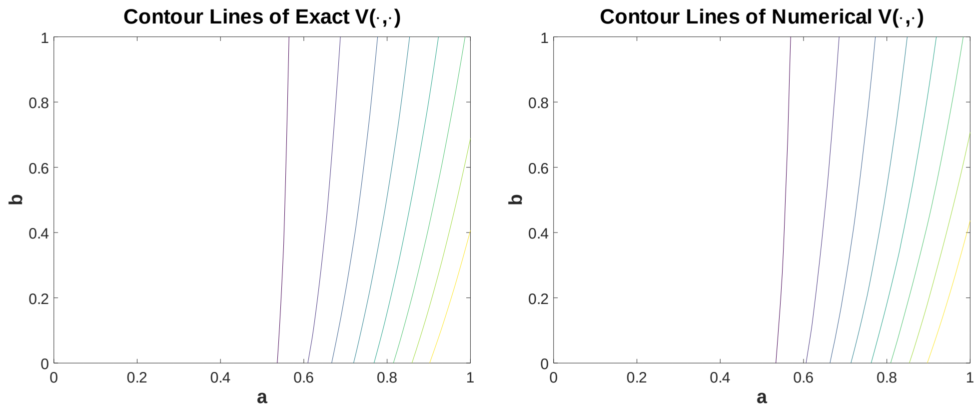
Comparing Informational Rent



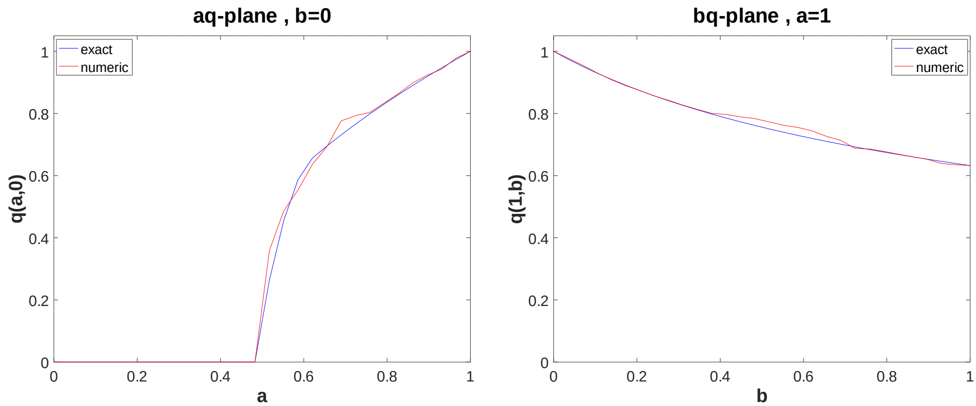
Contour Lines of Quantity



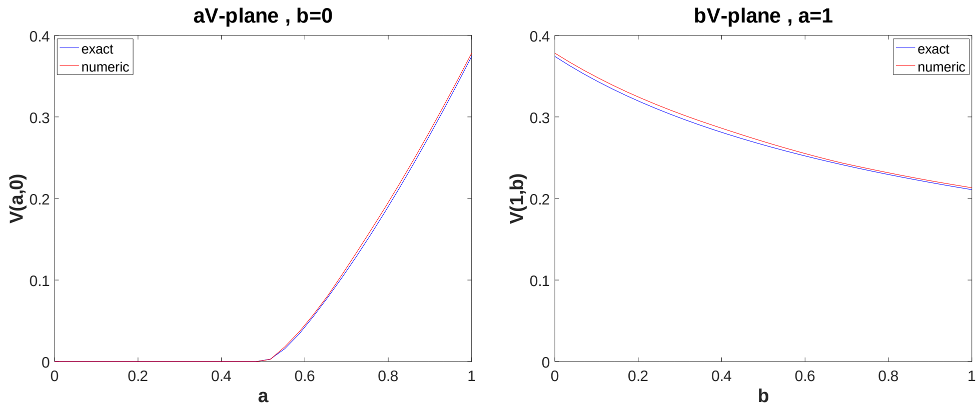
Contour Lines of Informational Rent



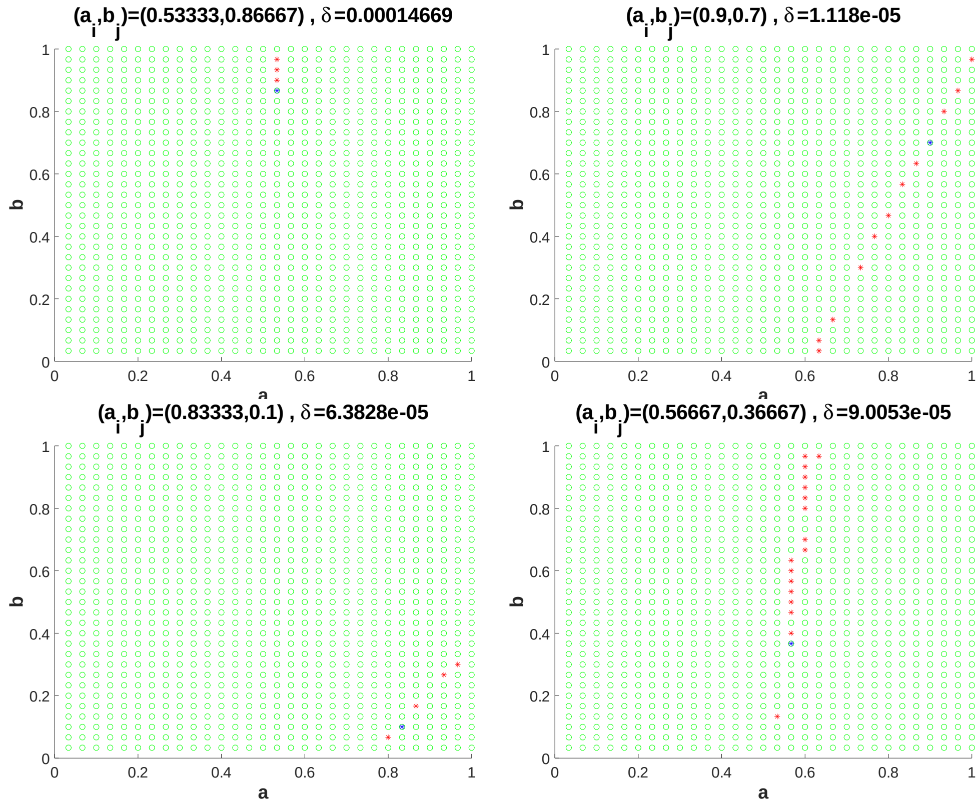
Cross-sections of Quantity



Cross-sections of Informational Rent



Verifying IC constraints



5.3. Example 3. Convex Demand

In Araujo et al. (2022), the authors analyzed the monopolist's problem for the case

$$v(q,a,b) = (c-b)\log(aq+1) \;\;,\;\; C(q) = \lambda q \;\;,\;\; f(a,b) = 1$$

where $c \geq 1$ and $\lambda \in (0,1)$ are given. In this case, the solution proposed is not given in a closed form. To express the analytical solution, define

$$\begin{aligned} D(r) &= \lambda r(1-r) \\ E(r) &= \lambda(1-r) - \lambda r \log(r) - cr(1-r) \\ F(r) &= 2cr\log(r) + c(1-r) - \lambda \log(r) \end{aligned}$$

and let $\underline{r} \in]0,1[$ be the solution of

$$(2cr-\lambda)\log(r) + c(1-r) = 0$$

Additionally, define

$$\phi(r) = \frac{-E(r) + \sqrt{E(r)^2 - 4D(r)F(r)}}{2D(r)} \quad , \quad \forall\, r \in]\underline{r},1[$$

Finally, given $(a,b) \in [0,1]^2$, $q(a,b)$ is defined as follows:

- If $b \geq c - (c\underline{r})/a$, let $q(a,b) := 0$.
- If $b < c - (c\underline{r})/a$, let $r(a,b) \in]\underline{r},1[$ be the solution of

$$\frac{c-b}{br} - \frac{c}{ab} = \frac{-E(r) + \sqrt{E(r)^2 - 4D(r)F(r)}}{2D(r)}$$

  such that $\phi(r(a,b)) > 0$ and $\phi'(r(a,b)) > 0$, and let $q(a,b) := \phi(r(a,b))$.

Furthermore, the tariff as a function of r over $]\underline{r},1[$ can be expressed as

$$T(r) = \int_{\underline{r}}^{r} v_q(\phi(\tilde{r}),\tilde{r},0)\phi'(\tilde{r})d\tilde{r}$$

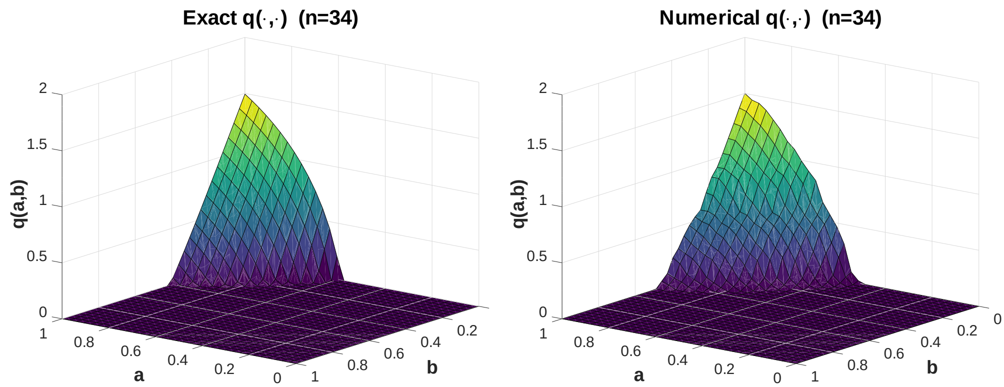
Agent type (a,b) has to transfer $t(a,b) = T(r(a,b))$ to the monopolist, which determines $V(a,b)$. In this way, the variables q and V are defined over $[0,1]^2$.

On the other hand, the discretized problem was solved for the case $c = 1$ and $\lambda = 0.4$ with $n = 34$ points over each axis.

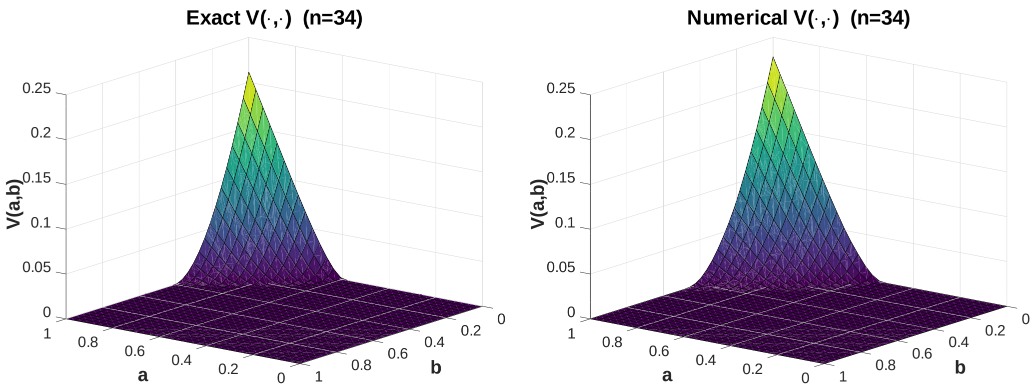
Next, the numerical solution is compared with the analytical (exact) solution.

$$
\begin{aligned}
a.q.e.(q^{num}, q^{exact}) &= 2.6300\times10^{-3}\\
a.q.e.(V^{num}, V^{exact}) &= 2.6064\times10^{-5}\\
\left|\ \text{profit}^{num} - \text{profit}^{exact}\right| &= 2.3191\times10^{-2}\\
\delta^{*} &= 5.77989\times10^{-4}
\end{aligned}
$$

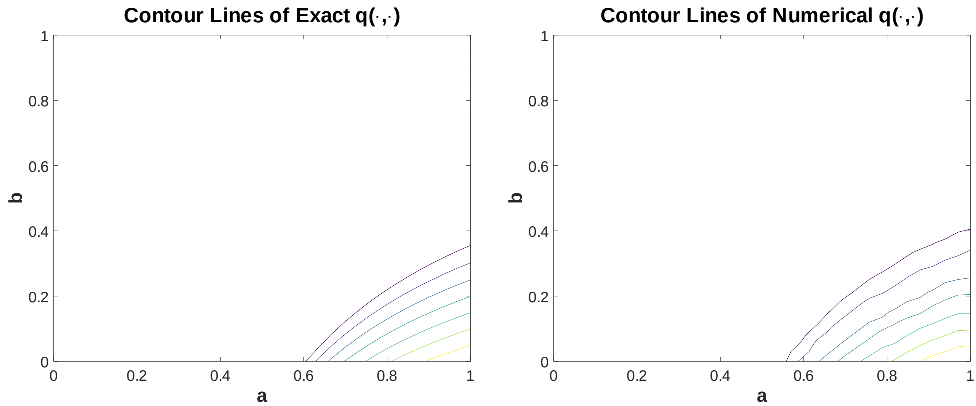
Comparing Quantity



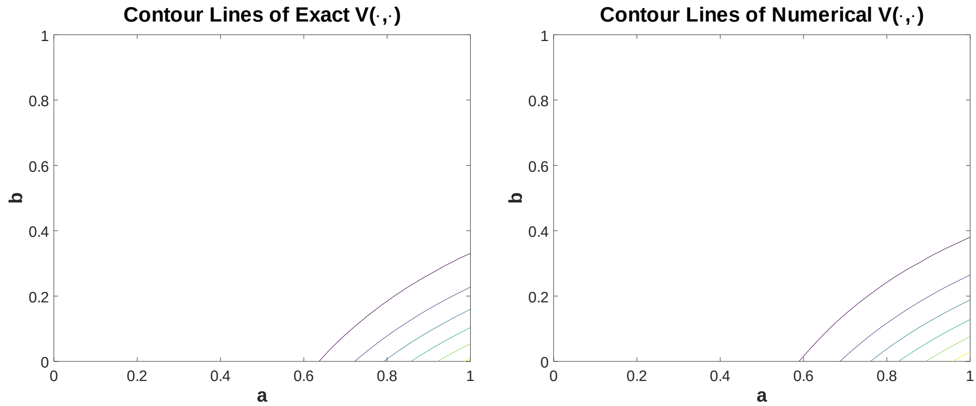
Comparing Informational Rent



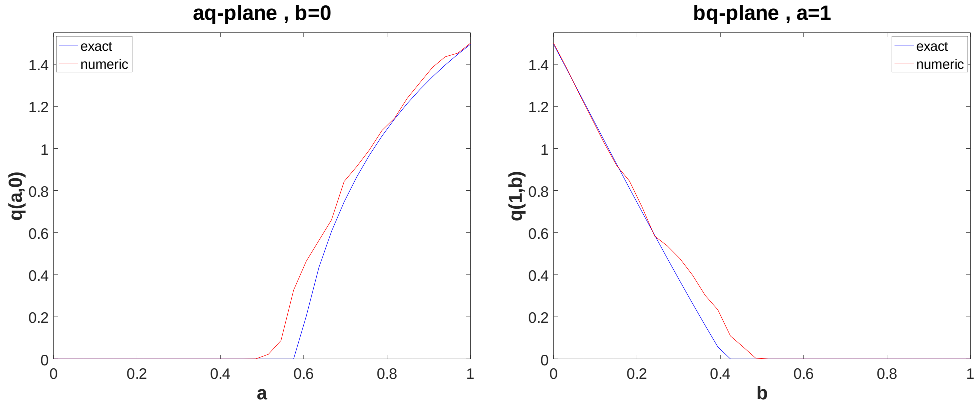
Contour Lines of Quantity



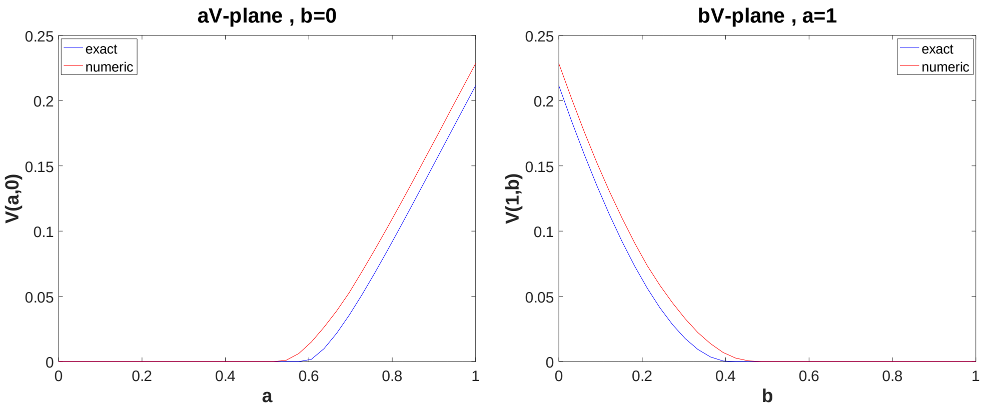
Contour Lines of Informational Rent



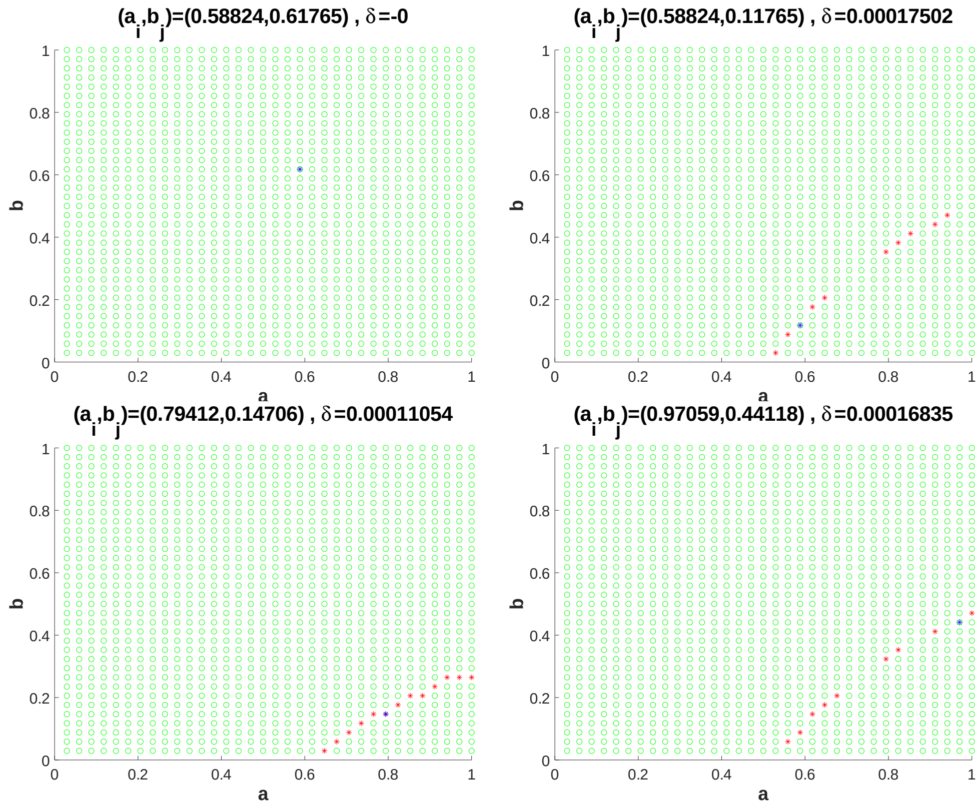
Cross-sections of Quantity



Cross-sections of Informational Rent



Verifying IC constraints



Appendix B: Mathematical Proofs

Notes:

- The expression “ (a,b) *is IC with* $(\widehat{a},\widehat{b})$ ” means that

$$V(a,b) - V(\widehat{a},\widehat{b}) \geq v(q(\widehat{a},\widehat{b}),a,b) - v(q(\widehat{a},\widehat{b}),\widehat{a},\widehat{b})$$

- The *characteristic curve* passing through (a,b), denoted by $CC(a,b)$, is the isoquant of $q(\cdot,\cdot)$ passing through (a,b). The reason for this name will be clear after the proof of Proposition 1 below.

- With some nomenclature flexibility, we say that a characteristic curve or isoquant is *increasing*, meaning that the curve has, at any point, a positive slope.

Proof of Proposition 1. We explain the methodology used by Araujo et al. (2022) to obtain the result.

Let $(q(\cdot,\cdot),t(\cdot,\cdot))$ be an incentive compatible contract, and let $(x,y) \in [0,1]^2$ be a point of twice-differentiability. Then, (x,y) must solve the problem

$$\max_{[0,1]^2}\{v(q(\cdot,\cdot),x,y) - t(\cdot,\cdot)\} \tag{7}$$

by the first-order necessary conditions, we have

$$\begin{aligned} v_q(q(x,y),x,y)q_a(x,y) &= t_a(x,y) \\ v_q(q(x,y),x,y)q_b(x,y) &= t_b(x,y) \end{aligned} \tag{8}$$

From the equations in (8), the cross derivatives t_{ab} and t_{ba} can be calculated. Then, by using Schwarz’s integrability condition $t_{ab} = t_{ba}$, the following quasi-linear PDE is derived

$$-\frac{v_{qb}}{v_{qa}}q_a + q_b = 0 \qquad \text{a.e. in } [0,1]^2 \tag{9}$$

Define $\Gamma_0 = \{(r,0) : r \in [0,1]\}$, let ϕ be a function defined over $[0,1]$, and consider the initial value problem

$$\begin{aligned} -\frac{v_{qb}}{v_{qa}}q_a + q_b &= 0 \\ q|_{\Gamma_0} &= \phi(r) \end{aligned} \tag{10}$$

Following the method of characteristic curves to solve[10] (10), there is a family of plane characteristic curves $(a(r,s),b(r,s))$ defined as the solution of:

[10] See Evans (1998) for a description of the method.

$$
\begin{aligned}
a_s(r,s) &= -\frac{v_{qb}}{v_{qa}}(\phi(r),a(r,s),b(r,s)) \quad , \quad a(r,0)=r \\
b_s(r,s) &= 1 \qquad\qquad\qquad\qquad\qquad\qquad\quad , \quad b(r,0)=0
\end{aligned} \tag{11}
$$

The meaning of characteristic curves is that, for a fix $r \in [0,1]$, the curve $(a(r,s),b(r,s))$ parameterized by $s \in [0,\bar{s}(r)]$ is a contour curve of q at level[11] $\phi(r)$. These characteristic curves determine a reparameterization of the set of types in terms of (r,s).

The applicability of the method of characteristics is assured, at least locally, in view of

$$
\begin{vmatrix} a_r(r,0) & b_r(r,0) \\ a_s(r,0) & b_s(r,0) \end{vmatrix} = \begin{vmatrix} 1 & 0 \\ -\frac{v_{qb}}{v_{qa}} & 1 \end{vmatrix} \neq 0
$$

That is, at any point the tangent vector of the initial curve Γ_0 is not parallel with the tangent vector of the characteristic curve. It is assumed that, in general,

$$
\begin{vmatrix} a_r & b_r \\ a_s & b_s \end{vmatrix} \neq 0
$$

This guarantees that we can return to the original variables

$$
a = a(r,s) \quad , \quad b = b(r,s)
$$

such that $q(a,b) = q(a(r,s),b(r,s)) = \phi(r)$.

Therefore, the tangent vector of a contour curve at the point $(a,b) = (a(r,s),b(r,s))$ is given by

$$
\begin{aligned}
(a_s(r,s),b_s(r,s)) &= (-\frac{v_{qb}}{v_{qa}}(\phi(r),a(r,s),b(r,s)),1) \\
&= (-\frac{v_{qb}}{v_{qa}}(q(a,b),a,b),1)
\end{aligned}
$$

□

Proof of Proposition 2. Fix $q \in Q$; by Assumption 2, $v_q(q,\cdot,b)$ is strictly increasing and $v_q(q,\widehat{a},\cdot)$ is strictly decreasing, so $\widehat{a} > a$ and $\widehat{b} < b$ imply $v_q(q,a,b) < v_q(q,\widehat{a},b)$ and $v_q(q,\widehat{a},b) < v_q(q,\widehat{a},\widehat{b})$, respectively. Thus, $v_q(q,a,b) < v_q(q,\widehat{a},\widehat{b})$. □

[11] Such a function ϕ must be optimally determined, which endogenously defines isoquants of $q(\cdot)$. This idea was developed in Araujo et al. (2022).

Proof of Theorem 1. The proof is based on the following three lemmas:

Lemma 1. *Let* $(a,b),(\widehat{a},\widehat{b}) \in [0,1]^2$ *be such that* (a,b) *is IC with* $(\widehat{a},\widehat{b})$*. Then* (a,b) *is IC with* (x,y) *,* $\forall\ (x,y) \in CC(\widehat{a},\widehat{b})$

Lemma 2. *Assume* $(q(\cdot,\cdot),V(\cdot,\cdot))$ *is such that* $q_a \geq 0$ *,* $q_b \leq 0$ *and for any* $(a,b) \in [0,1]^2$*:*

$$V_a(a,b) = v_a(q(a,b),a,b)\ \wedge\ V_b(a,b) = v_b(q(a,b),a,b)$$

Then, fixed $(a,b) \in [0,1]^2$*, the constraints given in (2) are satisfied.*

Lemma 3. *Let* $(x,y),(\widehat{a},\widehat{b}),(a,b) \in [0,1]^2$ *be such that* (a,b) *is IC with* $(\widehat{a},\widehat{b})$ *and* $(\widehat{a},\widehat{b})$ *is IC with* (x,y)*. If* $(\widehat{a},\widehat{b}) \preceq (a,b)$ *and* $q(x,y) \leq q(\widehat{a},\widehat{b})$*, then* (a,b) *is IC with* (x,y)*.*

We return to the proof of Theorem 1. Fix any $(a,b),(\widehat{a},\widehat{b}) \in [0,1]^2$. Let us prove that (a,b) is IC with $(\widehat{a},\widehat{b})$.

If $q(\widehat{a},\widehat{b}) = q^{\text{out}}$ (that is, if type $(\widehat{a},\widehat{b})$ is excluded), we have $V(\widehat{a},\widehat{b}) = 0$, so from the IR constraint $V(a,b) \geq 0$, we can write

$$V(a,b) - V(\widehat{a},\widehat{b}) \geq v(q^{\text{out}},a,b) - v(q^{\text{out}},\widehat{a},\widehat{b})$$

in view of $v(q^{\text{out}},a,b) = v(q^{\text{out}},\widehat{a},\widehat{b})$ by Assumption 1. Therefore, (a,b) is IC with $(\widehat{a},\widehat{b})$.

If $q(\widehat{a},\widehat{b}) \neq q^{\text{out}}$, since $CC(\widehat{a},\widehat{b})$ is increasing, there are three possible cases:

Case 1 $CC(\widehat{a},\widehat{b})$ intersects $F^{(a,b)}$:

Let (x,y) be the point of intersection. Because (a,b) is IC with (x,y) and $(x,y) \in CC(\widehat{a},\widehat{b})$, by Lemma 1, (a,b) is IC with $(\widehat{a},\widehat{b})$.

Note that the previous case implies that the IC constraint is fulfilled with all types in some small enough neighborhood of (a,b) because all the characteristic curves passing through this neighborhood intersect $F^{(a,b)}$. Consequently, there exist two intervals $I_a =]a-\varepsilon, a+\varepsilon[$ and $I_b =]b-\varepsilon, b+\varepsilon[$ for some small enough $\varepsilon > 0$, such that the (a,b) type verifies the IC constraints with all types over $I_a \times \{b\}$ and $\{a\} \times I_b$. That is, the agent type (a,b) has no incentive to locally misrepresent its true type over each direction. Therefore, we have

$$V(a,b) = \max_{\widehat{a} \in I_a}\{v(q(\widehat{a},b),a,b) - t(\widehat{a},b)\}$$

$$V(a,b) = \max_{\widehat{b} \in I_b} \{v(q(a,\widehat{b}),a,b) - t(a,\widehat{b})\}$$

which implies

$$V_a(a,b) = v_a(q(a,b),a,b) \;\wedge\; V_b(a,b) = v_b(q(a,b),a,b) \tag{12}$$

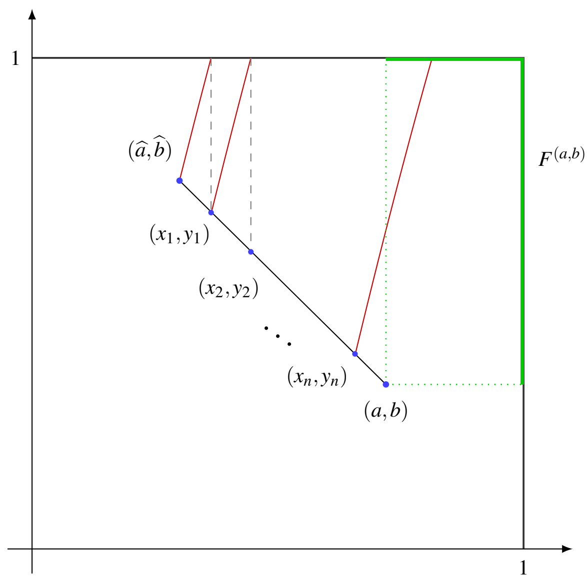
Since these relations hold for any point over $[0,1]^2$ and we have assumed $q_a \geq 0$, $q_b \leq 0$, by the Lemma 2, we have that the IC constraints with all types southeast of the given point (a,b) are fulfilled.

Case 2 $CC(\widehat{a},\widehat{b})$ intersects $\{(1,s) : 0 \leq s < b\}$:
Let $(1,\widehat{s})$ with $\widehat{s} \in [0,b[$ be the point of intersection. Since this point is southeast of (a,b), then (a,b) is IC with $(1,\widehat{s}) \in CC(\widehat{a},\widehat{b})$. Then, by Lemma 1, (a,b) is IC with $(\widehat{a},\widehat{b})$.

Case 3 $CC(\widehat{a},\widehat{b})$ intersects $\{(s,1) : 0 \leq s \leq a\}$ (Figure 6 illustrates for this case):
Since $CC(\widehat{a},\widehat{b})$ is increasing, we must have $\widehat{a} < a$. Without the loss of generality, we consider that[12] $\widehat{b} > b$. That is, $(\widehat{a},\widehat{b})$ is northwest of (a,b). Let $(x_1,1) \in CC(\widehat{a},\widehat{b}) \cap \{(s,1) : 0 \leq s \leq a\}$, and let[13] $(x_1,y_1) \in \{(x_1,y) : y \in \mathbb{R}\} \cap \text{conv}\{(\widehat{a},\widehat{b}),(a,b)\}$. Thus, the point (x_1,y_1) is southeast of $(\widehat{a},\widehat{b})$ and northwest of (a,b). By monotonicity, we have $q(\widehat{a},\widehat{b}) \leq q(x_1,y_1)$ and by Proposition 2 $(x_1,y_1) \preceq (a,b)$. Since (x_1,y_1) is IC with $(x_1,1)$ (due to $(x_1,1) \in F^{(x_1,y_1)}$), by Lemma 1, (x_1,y_1) is IC with $(\widehat{a},\widehat{b})$. Then, by Lemma 3, it will be sufficient that $CC(x_1,y_1) \cap F^{(a,b)} \neq \emptyset$ to conclude that (a,b) is IC with $(\widehat{a},\widehat{b})$. If that is not the case, i.e., $CC(x_1,y_1) \cap F^{(a,b)} = \emptyset$, repeat the procedure taking $(x_2,1) \in CC(x_1,y_1) \cap \{(s,1) : 0 \leq s \leq a\}$ and $(x_2,y_2) \in \{(x_2,y) : y \in \mathbb{R}\} \cap \text{conv}\{(x_1,y_1),(a,b)\}$. Similar to the above, we have $q(x_1,y_1) \leq q(x_2,y_2)$, $(x_2,y_2) \preceq (a,b)$ and (x_2,y_2) is IC with (x_1,y_1). Then, by Lemma 3, it will be sufficient that $CC(x_2,y_2) \cap F^{(a,b)} \neq \emptyset$ to conclude that (a,b) is IC with (x_1,y_1), and therefore, by Lemma 3, (a,b) is IC with $(\widehat{a},\widehat{b})$. If $CC(x_2,y_2) \cap F^{(a,b)} = \emptyset$, we set up the point (x_3,y_3), and so on. Note that $\frac{d}{dq}\left(\frac{v_{qa}}{v_{qb}}\right) \geq 0$ and $\frac{d}{da}\left(\frac{v_{qa}}{v_{qb}}\right) \geq 0$

[12] Otherwise, replace $(\widehat{a},\widehat{b})$ for any point in $CC(\widehat{a},\widehat{b})$ to the northwest of (a,b).
[13] The set $\{(x_1,y) : y \in \mathbb{R}\} \cap \text{conv}\{(\widehat{a},\widehat{b}),(a,b)\}$ is nonempty in view of $\widehat{a} < x_1 \leq a$.

Figure 6: Illustration of Theorem 1 proof.



(Assumption 3) imply $\frac{d}{dr}a_s(r,1) \geq 0$ in view of

$$\begin{aligned}\frac{d}{dr}a_s(r,1) &= \frac{d}{dr}\Big(-\frac{v_{qb}}{v_{qa}}(q(r,1),r,1)\Big)\\ &= \Big(\frac{v_{qb}}{v_{qa}}\Big)^2\Big[\frac{d}{dq}\Big(\frac{v_{qa}}{v_{qb}}\Big)\times q_a(r,1)+\frac{d}{da}\Big(\frac{v_{qa}}{v_{qb}}\Big)\Big]\end{aligned}$$

That is, the slope of the characteristic curves at the border $(r,1)$ is nondecreasing which guarantees that, for large enough n, $CC(x_n,y_n)\cap F^{(a,b)} \neq \emptyset$ because (x_n,y_n) will be close to (a,b) and $CC(x_n,y_n)$ is increasing. Thus, applying Lemma 3 n times, (a,b) is IC with $(\widehat{a},\widehat{b})$. □

Proof of Lemma 1. If $(x,y) \in CC(\widehat{a},\widehat{b})$, then $q(\widehat{a},\widehat{b}) = q(x,y)$. Therefore, by the taxation principle, $t(\widehat{a},\widehat{b}) = T(q(\widehat{a},\widehat{b})) = T(q(x,y)) = t(x,y)$. Because (a,b) is IC with $(\widehat{a},\widehat{b})$, we have

$$\begin{aligned}v(q(a,b),a,b)-t(a,b) &\geq v(q(\widehat{a},\widehat{b}),a,b)-t(\widehat{a},\widehat{b})\\ &= v(q(x,y),a,b)-t(x,y)\end{aligned}$$

that is, (a,b) is IC with (x,y). $\square$

Proof of Lemma 2. Fix $(\widehat{a},\widehat{b}) \in [0,1]^2$ such that $a < \widehat{a}$ and $b > \widehat{b}$. Define the auxiliary function

$$f(x,y) := V(x,y) - v(q(\widehat{a},\widehat{b}),x,y) \;\; \forall\, (x,y) \in [0,\widehat{a}] \times [\widehat{b},1]$$

The partial derivatives of f with respect to the first and second argument are:[14]

$$\begin{aligned} f_1(x,y) &= V_a(x,y) - v_a(q(\widehat{a},\widehat{b}),x,y) \\ &= v_a(q(x,y),x,y) - v_a(q(\widehat{a},\widehat{b}),x,y) \\ f_2(x,y) &= V_b(x,y) - v_b(q(\widehat{a},\widehat{b}),x,y) \\ &= v_b(q(x,y),x,y) - v_b(q(\widehat{a},\widehat{b}),x,y) \end{aligned}$$

where we use the envelope conditions.
Given $(x,y) \in [0,\widehat{a}] \times [\widehat{b},1]$, by the monotonicity conditions $q_a \geq 0$ and $q_b \leq 0$, we have that $q(x,y) \leq q(\widehat{a},\widehat{b})$. From Assumption 2, v_a increases in the first argument and v_b decreases in the first argument. Then, $v_a(q(x,y),x,y) \leq v_a(q(\widehat{a},\widehat{b}),x,y)$ and $v_b(q(x,y),x,y) \geq v_b(q(\widehat{a},\widehat{b}),x,y)$. Hence, we have $f_1 \leq 0$ and $f_2 \geq 0$. Finally, $a < \widehat{a}$ implies $f(a,b) \geq f(\widehat{a},b)$, and $b > \widehat{b}$ implies $f(\widehat{a},b) \geq f(\widehat{a},\widehat{b})$. Therefore, $f(a,b) \geq f(\widehat{a},\widehat{b})$. That is, (a,b) is IC with $(\widehat{a},\widehat{b})$. $\square$

Proof of Lemma 3. Since (a,b) is IC with $(\widehat{a},\widehat{b})$ and $(\widehat{a},\widehat{b})$ is IC with (x,y), we have

$$\begin{aligned} &V(a,b) - V(x,y) + v(q(x,y),x,y) \geq \\ &\qquad v(q(\widehat{a},\widehat{b}),a,b) - v(q(\widehat{a},\widehat{b}),\widehat{a},\widehat{b}) + v(q(x,y),\widehat{a},\widehat{b}) \end{aligned} \tag{13}$$

On the other hand, due to $q(x,y) \leq q(\widehat{a},\widehat{b})$ and $v_q(q,\widehat{a},\widehat{b}) \leq v_q(q,a,b)\; \forall\, q \in Q$, by integrating we have

$$\int_{q(x,y)}^{q(\widehat{a},\widehat{b})} v_q(q,\widehat{a},\widehat{b})dq \;\leq\; \int_{q(x,y)}^{q(\widehat{a},\widehat{b})} v_q(q,a,b)dq$$

[14] Recall that we denote V_a and V_b to the partial derivatives of V with respect to the first and second arguments, respectively, since $V = V(a,b)$. Similarly, v_a and v_b denote the partial derivatives of $v = v(q,a,b)$ with respect to the second and third argument.

Then,

$$v(q(\widehat{a},\widehat{b}),a,b) - v(q(\widehat{a},\widehat{b}),\widehat{a},\widehat{b}) + v(q(x,y),\widehat{a},\widehat{b}) \geq v(q(x,y),a,b) \tag{14}$$

Therefore, from (13) and (14), (a,b) is IC with (x,y). $\square$

Proof of Proposition 3. Fix any $q \in Q$. Since $v_q(q,\cdot,b)$ is concave and $v_q(q,\widehat{a},\cdot)$ is convex:

$$v_q(q,\widehat{a},b) - v_q(q,a,b) \geq v_{qa}(q,\widehat{a},b)(\widehat{a}-a)$$

$$v_q(q,\widehat{a},\widehat{b}) - v_q(q,\widehat{a},b) \geq v_{qb}(q,\widehat{a},b)(\widehat{b}-b)$$

then

$$v_q(q,\widehat{a},\widehat{b}) - v_q(q,a,b) \geq v_{qa}(q,\widehat{a},b)(\widehat{a}-a) + v_{qb}(q,\widehat{a},b)(\widehat{b}-b) \tag{15}$$

In addition, if $\frac{\widehat{b}-b}{\widehat{a}-a} \leq \frac{-v_{qa}(q,\widehat{a},b)}{v_{qb}(q,\widehat{a},b)}$ with $\widehat{a} > a$ and $v_{qb} < 0$ then

$$v_{qa}(q,\widehat{a},b)(\widehat{a}-a) + v_{qb}(q,\widehat{a},b)(\widehat{b}-b) \geq 0 \tag{16}$$

hence, from (15) and (16), $v_q(q,\widehat{a},\widehat{b}) - v_q(q,a,b) \geq 0$ for any $q \in Q$, that is $(a,b) \preceq (\widehat{a},\widehat{b})$. $\square$

Proof of Proposition 4. The proof is based on the two following lemmas.

Lemma 4. *Given $(a,b) \in X_n$, $\forall\ (x,y) \in F^{(a,b)}$, we have*

$$\widetilde{V}^n(a,b) - \widetilde{V}^n(x,y) \geq v(\widetilde{Q}^n(x,y),a,b) - v(\widetilde{Q}^n(x,y),x,y) - O(\tfrac{1}{n-1})$$

That is, since $(a,b) \in X_n$ verifies IC with all points in $\widetilde{F}^{(a,b)} = F^{(a,b)} \cap X_n$, it satisfies a relaxed IC version with all points in the continuous set $F^{(a,b)}$ with some tolerance that is asymptotically zero. The next lemma shows that between any two points on grid X_n, the same relaxed IC version holds.

Lemma 5. *Given $(a,b),(\widehat{a},\widehat{b}) \in X_n$, we have*

$$V^n(a,b) - V^n(\widehat{a},\widehat{b}) \geq v(Q^n(\widehat{a},\widehat{b}),a,b) - v(Q^n(\widehat{a},\widehat{b}),\widehat{a},\widehat{b}) - O(\tfrac{1}{n-1})$$

We return to the proof of Proposition 4. Given $(a,b),(a',b')\in[0,1]^2$, it will be sufficient to prove that

$$\widetilde{V}^n(a,b)-\widetilde{V}^n(a',b')-(v(\widetilde{Q}^n(a',b'),a,b)-v(\widetilde{Q}^n(a',b'),a',b'))\geq -O(\tfrac{1}{n-1})$$

Let $(\widehat{a},\widehat{b}),(\widehat{a}',\widehat{b}')\in X_n$ be such that $\widehat{a}\leq a<\widehat{a}+\frac{1}{n-1}$, $\widehat{b}-\frac{1}{n-1}<b\leq\widehat{b}$ and $\widehat{a}'\leq a'<\widehat{a}'+\frac{1}{n-1}$, $\widehat{b}'-\frac{1}{n-1}<b'\leq\widehat{b}'$. Let $q=\widetilde{Q}^n(a',b')=Q^n(\widehat{a}',\widehat{b}')$. Since $\widetilde{V}^n(a,b)=V^n(\widehat{a},\widehat{b})$, $\widetilde{V}^n(a',b')=V^n(\widehat{a}',\widehat{b}')$ we have

$$\begin{aligned}\widetilde{V}^n(a,b)-\widetilde{V}^n(a',b')-(v(q,a,b)-v(q,a',b'))=&\\ &V^n(\widehat{a},\widehat{b})-V^n(\widehat{a}',\widehat{b}')-(v(q,\widehat{a},\widehat{b})-v(q,\widehat{a}',\widehat{b}'))\\ &+v(q,\widehat{a},\widehat{b})-v(q,a,b)+v(q,a',b')-v(q,\widehat{a}',\widehat{b}')\end{aligned}\tag{17}$$

Since $(\widehat{a},\widehat{b}),(\widehat{a}',\widehat{b}')\in X_n$, by Lemma 5,

$$V^n(\widehat{a},\widehat{b})-V^n(\widehat{a}',\widehat{b}')-(v(q,\widehat{a},\widehat{b})-v(q,\widehat{a}',\widehat{b}'))\geq -O(\tfrac{1}{n-1})\tag{18}$$

In addition, since v is differentiable, $v(q,\cdot,\cdot)$ is Lipschitz over $[0,1]^2$ (with constant L), then

$$\left|v(q,\widehat{a},\widehat{b})-v(q,a,b)\right|\leq L||(\widehat{a},\widehat{b})-(a,b)||\leq O(\tfrac{1}{n-1})$$

which implies

$$v(q,\widehat{a},\widehat{b})-v(q,a,b)\geq -O(\tfrac{1}{n-1})\tag{19}$$

Similarly,

$$\left|v(q,a',b')-v(q,\widehat{a}',\widehat{b}')\right|\leq L||(a',b')-(\widehat{a}',\widehat{b}')||\leq O(\tfrac{1}{n-1})$$

which implies

$$v(q,a',b')-v(q,\widehat{a}',\widehat{b}')\geq -O(\tfrac{1}{n-1})\tag{20}$$

Therefore, from (17) using (18), (19) and (20) we obtain

$$\widetilde{V}^n(a,b)-\widetilde{V}^n(a',b')-(v(q,a,b)-v(q,a',b'))\geq -O(\tfrac{1}{n-1})$$

□

Proof of Lemma 4. Let $(x,y) \in F^{(a,b)}$ be such that $x = 1$ (case $y = 1$ is analogous), and let $\widehat{b}$ be such that $\widehat{b} - \frac{1}{n-1} < y \leq \widehat{b}$. Since (Q^n, V^n) are the solutions of problem (4), (a,b) satisfies IC with $(1,\widehat{b})$

$$V^n(a,b) - V^n(1,\widehat{b}) \geq v(Q^n(1,\widehat{b}),a,b) - v(Q^n(1,\widehat{b}),1,\widehat{b})$$

By definition, $\widetilde{Q}^n(x,y) = Q^n(1,\widehat{b})$ and $\widetilde{V}^n(x,y) = V^n(1,\widehat{b})$. Additionally, in view of $(a,b) \in X_n$, we have $\widetilde{V}^n(a,b) = V^n(a,b)$. Then,

$$\widetilde{V}^n(a,b) - \widetilde{V}^n(x,y) \geq v(\widetilde{Q}^n(x,y),a,b) - v(\widetilde{Q}^n(x,y),1,\widehat{b}) \tag{21}$$

On the other hand, since v is Lipschitz,

$$\left| v(\widetilde{Q}^n(x,y),1,\widehat{b}) - v(\widetilde{Q}^n(x,y),x,y) \right| \leq L||(1,\widehat{b}) - (x,y)|| = O(\tfrac{1}{n-1})$$

Then,

$$-v(\widetilde{Q}^n(x,y),1,\widehat{b}) \geq -v(\widetilde{Q}^n(x,y),x,y) - O(\tfrac{1}{n-1}) \tag{22}$$

Therefore, from (21) and (22),

$$\widetilde{V}^n(a,b) - \widetilde{V}^n(x,y) \geq v(\widetilde{Q}^n(x,y),a,b) - v(\widetilde{Q}^n(x,y),x,y) - O(\tfrac{1}{n-1})$$

□

Proof of Lemma 5. If $CC(\widehat{a},\widehat{b}) \cap F^{(a,b)} = (x,y)$, we apply Lemma 4 for (a,b) with (x,y), and considering that $Q^n(\widehat{a},\widehat{b}) = Q^n(x,y)$ and $t(x,y) = t(\widehat{a},\widehat{b})$, we conclude.

Other cases are treated analogously as in the proof of Theorem 1. □

Proof of Proposition 5. Let $(\overline{Q},\overline{V})$ denote the solution for the continuous problem, and let $(\overline{Q}^n,\overline{V}^n)$ be their restriction on the grid X_n. If (Q^n,V^n) are the solutions of the discretized problem and OPT_n is the optimal value, we have

$$\begin{aligned}
OPT_n &\geq \sum_{i=1}^{n}\sum_{j=1}^{n} w(i,j)(v(\overline{Q}^n_{i,j},a_i,b_j) - \overline{V}^n_{i,j} - C(\overline{Q}^n_{i,j}))\rho(a_i,b_j) \\
&= \int_0^1\int_0^1 (v(\overline{Q}(a,b),a,b) - \overline{V}(a,b) - C(\overline{Q}(a,b)))\rho(a,b)\,da\,db - O(\frac{1}{n}) \\
&= OPT^* - O(\frac{1}{n})
\end{aligned}$$

Then, $\liminf_{n\to\infty} OPT_n \geq OPT^*$.

On the other hand, if $\exists \lim_{n\to\infty} \widetilde{Q}^n(a,b)$ and $\lim_{n\to\infty} \widetilde{V}^n(a,b)$ for any $(a,b) \in [0,1]^2$, define

$$\widehat{Q}(a,b) := \lim_{n\to\infty} \widetilde{Q}^n(a,b) \ , \ \widehat{V}(a,b) := \lim_{n\to\infty} \widetilde{V}^n(a,b)$$

By Proposition 4, $(\widehat{Q},\widehat{V})$ is feasible. Hence

$$\begin{aligned}
OPT^* &\geq \int_0^1 \int_0^1 (v(\widehat{Q}(a,b),a,b) - \widehat{V}(a,b) - C(\widehat{Q}(a,b)))\rho(a,b)\,da\,db \\
&= \lim_{n\to\infty} \int_0^1 \int_0^1 (v(\widetilde{Q}^n(a,b),a,b) - \widetilde{V}^n(a,b) - C(\widetilde{Q}^n(a,b)))\rho(a,b)\,da\,db \\
&= \lim_{n\to\infty} \sum_{i=1}^n \sum_{j=1}^n w(i,j)(v(\widetilde{Q}^n_{i,j},a_i,b_j) - \widetilde{V}^n_{i,j} - C(\widetilde{Q}^n_{i,j}))\rho(a_i,b_j) + O(\tfrac{1}{n-1}) \\
&= \lim_{n\to\infty} \sum_{i=1}^n \sum_{j=1}^n w(i,j)(v(Q^n_{i,j},a_i,b_j) - V^n_{i,j} - C(Q^n_{i,j}))\rho(a_i,b_j) + O(\tfrac{1}{n-1}) \\
&= \lim_{n\to\infty} \left(OPT_n + O(\tfrac{1}{n-1})\right)
\end{aligned}$$

where equalities are true owing to the dominated convergence theorem (each $\widetilde{Q}^n$ and $\widetilde{V}^n$ are bounded), by the finite approximation of the integral, by the definition of $\widetilde{Q}^n$ and $\widetilde{V}^n$, and because (Q^n,V^n) is the solution of the discretized problem. $\square$

References

Araujo, A., Vieira, S., & Calagua, B. (2022). A necessary optimality condition in two-dimensional screening. *Economic Theory*, 1–26.

Armstrong, M. (1996). Multiproduct nonlinear pricing. *Econometrica*, 51–75.

Armstrong, M. (1999). Optimal regulation with unknown demand and cost functions. *Journal of Economic Theory*, *84*(2), 196–215.

Barelli, P., Basov, S., Bugarin, M., & King, I. (2014). On the optimality of exclusion in multi-dimensional screening. *Journal of Mathematical Economics*, *54*, 74–83.

Baron, D. P., & Myerson, R. B. (1982). Regulating a monopolist with unknown costs. *Econometrica*, 911–930.

Basov, S. (2001). Hamiltonian approach to multi-dimensional screening. *Journal of Mathematical Economics*, *36*(1), 77–94.

Belloni, A., Lopomo, G., & Wang, S. (2010). Multidimensional mechanism design: Finite-dimensional approximations and efficient computation. *Operations Research*, *58*(4-part-2), 1079–1089.

Berg, K., & Ehtamo, H. (2009). Learning in nonlinear pricing with unknown utility functions. *Annals of Operations Research*, *172*(1), 375.

Che, Y.-K., & Gale, I. (1998). Standard auctions with financially constrained bidders. *Review of Economic Studies*, *65*(1), 1–21.

Chiappori, P.-A., McCann, R. J., & Nesheim, L. P. (2010). Hedonic price equilibria, stable matching, and optimal transport: Equivalence, topology, and uniqueness. *Economic Theory*, *42*, 317–354.

Deneckere, R., & Severinov, S. (2017). A solution to a class of multi-dimensional screening problems: Isoquants and clustering. *Technical report*.

Evans, L. (1998). *Partial Differential Equations* (Vol. 19). American Mathematical Society.

Figalli, A., Kim, Y.-H., & McCann, R. J. (2011). When is multidimensional screening a convex program? *Journal of Economic Theory*, *146*(2), 454–478.

Judd, K., Ma, D., Saunders, M. A., & Su, C.-L. (2017). Optimal income taxation with multidimensional taxpayer types. *Working Paper*.

Laffont, J.-J., & Martimort, D. (2001). *The Theory of Incentives: The Principal-Agent Model*. Princeton University Press.

Laffont, J.-J., Maskin, E., & Rochet, J.-C. (1987). Optimal nonlinear pricing with two-dimensional characteristics. *Information, Incentives and Economic Mechanisms*, 256–266.

Levin, V. (1999). Abstract cyclical monotonicity and Monge solutions for the general Monge–Kantorovich problem. *Set-Valued Analysis*, *7*(1), 7–32.

Lewis, T. R., & Sappington, D. E. (1988a). Regulating a monopolist with unknown demand. *American Economic Review*, 986–998.

Lewis, T. R., & Sappington, D. E. (1988b). Regulating a monopolist with unknown demand and cost functions. *RAND Journal of Economics*, 438–457.

McAfee, R. P., & McMillan, J. (1988). Multidimensional incentive compatibility and mechanism design. *Journal of Economic Theory*, *46*(2), 335–354.

Mussa, M., & Rosen, S. (1978). Monopoly and product quality. *Journal of Economic theory*, *18*(2), 301–317.

Myerson, R. B. (1979). Incentive compatibility and the bargaining problem. *Econometrica*, 61–73.

Rochet, J.-C. (2009). Monopoly regulation without the Spence–Mirrlees assumption. *Journal of Mathematical Economics*, *45*(9-10), 693–700.

Salanié, B. (1997). *The Economics of Contracts*. Cambridge, MA: MIT press.

Tarkiainen, R., & Tuomala, M. (1999). Optimal nonlinear income taxation with a two-dimensional population: A computational approach. *Computational Economics*, *13*(1), 1–16.

Wilson, R. (1996). Nonlinear pricing and mechanism design. *Handbook of Computational Economics*, *1*, 253–293.

How to prepare a paper for submission

Before submitting a paper to **this Journal**, the authors are advised to follow closely the following instructions to prepare the paper.

1. Papers submitted to this Journal must be unpublished and original work that is neither under review elsewhere nor will be submitted elsewhere for publication without withdrawing from this Journal.

2. This Journal requires that all results (experimental, empirical and computational) be replicable. All underlying data necessary to replicate results must be made available to the Journal.

3. Papers must be submitted in electronic form (preferably as pdf files) and the size of the text font must be at least 12 point. Each figure and table must be included on the relevant page of the paper and should not be collected at the end of the paper. The list of references should appear after any appendices, as the last part of the paper.

4. There is no restriction on the number of pages of any submitted manuscript. We seek to process any first submission within 3 months. However, very long manuscripts may take considerably more time to review.

5. Each submitted paper should have an abstract of no more than 150 words containing no mathematical formulas, complete with no more than 3 suggested keywords and JEL codes. We also encourage authors to make the introduction of their submitted articles understandable to the widest audience possible.

6. The first page of each submitted paper should contain every author's e-mail address, phone number and affiliation.

7. The editors or the publishing Society will not hold any responsibility for views expressed by authors in this Journal.

How to submit a paper

Papers should be submitted electronically in PDF to the Journal of Mechanism and Institution Design through the website `http://www.mechanism-design.org/`.

Aims & Scope of the Journal

The Journal of Mechanism and Institution Design aims to publish original articles that deal with the issues of designing, improving, analysing and testing economic, financial, political, or social mechanisms and institutions. It welcomes theoretical, empirical, experimental, historical and practical studies. It seeks creative, interesting, rigorous, and logical research and strives for clarity of thought and expression. We particularly encourage less experienced researchers such as recent PhD graduates to submit their work to the Journal and are sympathetic towards those papers that are novel and innovative but which have been unsuccessful elsewhere. We hope that the published articles will be interesting and valuable to a broad audience from the areas of economics, finance, politics, law, computer science, management, history, mathematics, government, and related disciplines. The journal is an open-access, independent, peer-reviewed, non-profit, English-language journal with the purpose of disseminating and sharing the latest knowledge and understanding of the subject widely.

In order for any work that is published by the Journal of Mechanism and Institution Design to be freely accessible to the widest audience possible, when a paper is accepted by this Journal for publication, its author(s) will be asked to release the paper under the Creative Commons Attribution-Non-Commercial license. While the authors retain the copyright of their published work, this license permits anyone to copy and distribute the paper for non-commercial purposes provided that both the author(s) of the article and the Journal are properly acknowledged. For details of the copyrights, please see the “human-readable summary” of the license.