Happy 50th birthday TTC! & Characterizing the top trading cycles rule for housing markets with lexicographic preferences when externalities are limited

Bettina Klaus[*]

Plenary Talk: Online Stochastic Matching
September 24th 2024

[*] Department of Economics, HEC, University of Lausanne, CH

Example: vacation apartments

A set of families (1, 2, 3, and 4) from different parts of Switzerland with similar apartments and low travel budgets contemplates what to do during their summer vacations.

The initial allocation is that each family stays home during the vacations:

- (H_1, H_2, H_3, H_4).

The families have preferences over where to spend their vacations and the following allocation would be better:

- (H_2, H_3, H_1, H_4).

This is a classical Shapley-Scarf housing market situation and we know how to deal with it.

Example: vacation apartments (continued)

Now suppose that the families know a bit more about each other

- they are all friends from high school that stayed in touch,

e.g., family 1 likes to listen to heavy metal music loudly, family 3 is known to live in an apartment block with old neighbors that like to have a quiet life and go to bed early, etc.

Now the following allocation may be better:

- (H_2, H_1, H_4, H_3).

This is a housing market situation *with limited externalities*.

Survey: The point of departure (50 years ago!)

The classical housing market model by Shapley and Scarf (1974): "Cores and Indivisibilities," *Journal of Mathematical Economics* 1.

- n agents: $N = \{1, \ldots, n\}$;
- n indivisible and heterogeneous items (e.g., houses): $H = \{H_1, \ldots, H_n\}$;
- endowments: agent $i \in N$ owns house $H_i \in H$;
- preferences: agent $i \in N$ has weak preferences R_i over houses H (notation P_i, I_i, and R_i); and
- agents can trade their houses with no transfers of money.

In their seminal article, Shapley and Scarf model "trading in commodities that are inherently indivisible" as NTU games.

Notation for such a game is V.

Shapley-Scarf housing markets

- A preference profile $R = (R_1, \ldots, R_n)$;
- $\mathcal{W}$ denotes the set of all weak preference profiles and
- $\mathcal{S}$ denotes the set of all strict preference profiles.
- A Shapley-Scarf housing market (N, H, R) will for short be denoted by R.

  Thus, the sets $\mathcal{W}$ / $\mathcal{S}$ also denote the sets of Shapley-Scarf housing markets with weak / strict preferences.
- An allocation $a = (a_1, \ldots, a_n)$ is a feasible (re)assignment of houses to agents;
- a_i denotes the allotment of agent i
- and for a coalition $S \subseteq N$, $a(S) = \bigcup_{i \in S} a_i$ and $h(S) = \bigcup_{i \in S} H_i$.

(Weak) core allocations (Shapley and Scarf, 1974)

A coalition $S \subseteq N$ strongly blocks allocation a if there exists an allocation b such that:

a. $b(S) = h(S)$ and

b. for all $i \in S$, $b_i \, P_i \, a_i$.

Allocation a is a (weak) core allocation ($a \in C(R)$) if it is not strongly blocked by any coalition.

Strong core allocations (Roth and Postlewaite, 1977)

A coalition $S \subseteq N$ weakly blocks allocation a if there exists an allocation b such that a. and

b'. for all $i \in S$, $b_i \, R_i \, a_i$ and for some $j \in S$, $b_j \, P_j \, a_j$.

Allocation a is a **strong core allocation** ($a \in SC(R)$ if it is not weakly blocked by any coalition.

Weak domination (and the strong / strict core) was considered by

Roth and Postlewaite (1977):“Weak Versus Strong Domination in a Market with Indivisible Goods,” *Journal of Mathematical Economics* 4.

Shapley and Scarf (1974), Section 4

V is a balanced game; hence the market in question has a non-empty core.

Scarf (2009): “My introduction to top-trading cycles,” *Games and Economic Behavior* 66.

Herbert Scarf recalls: “I was very eager to find a cooperative game without transferable utility, whose core could be shown to be non empty using my theorem and not by any other method”

“I still find it quite remarkable that this game is balanced, without any additional assumptions on the preferences, and therefore has a non-empty core.”

“I gave a talk at Berkeley on this subject several months later, and spoke with great pleasure on the use of my theorem on balanced games. David [Gale] was in the audience and came up to chat with me after my talk was over. He said to me, rather diffidently, "I think that I have another argument that the core is not empty."”

“And he told me about the notion of ‘top trading cycles,’ and their use in providing a simple algorithm which always found a permutation in the core. . . . When I returned, I sent David a letter with the phrase "I’m terribly sorry to have tell you that there is nothing wrong with your argument." ”

The top trading cycles algorithm

Input. A Shapley-Scarf housing market $R \in \mathcal{S}$

($R \in \mathcal{W} \Rightarrow$ break ties).

Steps 1,...

- let each house point to its owner;
- let each agent point to his top house;
- execute the top trading cycles that form, and remove all the involved agents / houses;
- repeat this process until no agent / house is left.

Output. The TTC algorithm terminates when each agent in N is assigned a house in H (it takes at most $|N|$ steps). We denote the obtained allocation by $\mathrm{TTC}(R)$.

Happy 50th birthday TTC!

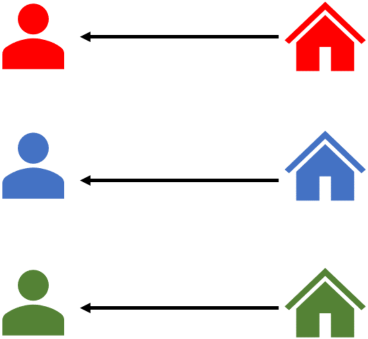
Two recent surveys to celebrate the 50th anniversary of the Shapley and Scarf paper (and the TTC algorithm) are

- Afacan, Hu, and Li (2024): “Housing Markets since Shapley and Scarf,” *Journal of Mathematical Economics* 111

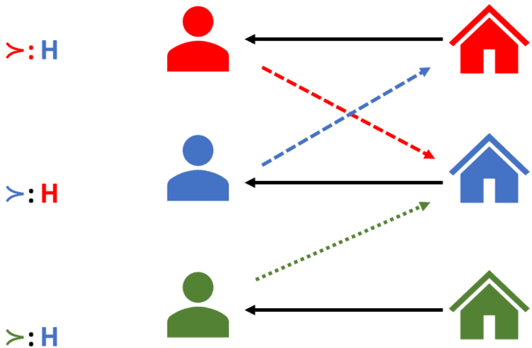
  and

- Morrill and Roth (2024): “Top trading cycles,” *Journal of Mathematical Economics* 111.

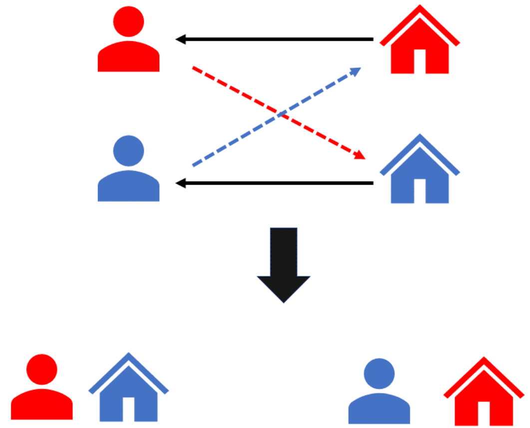
Example TTC, 1/4



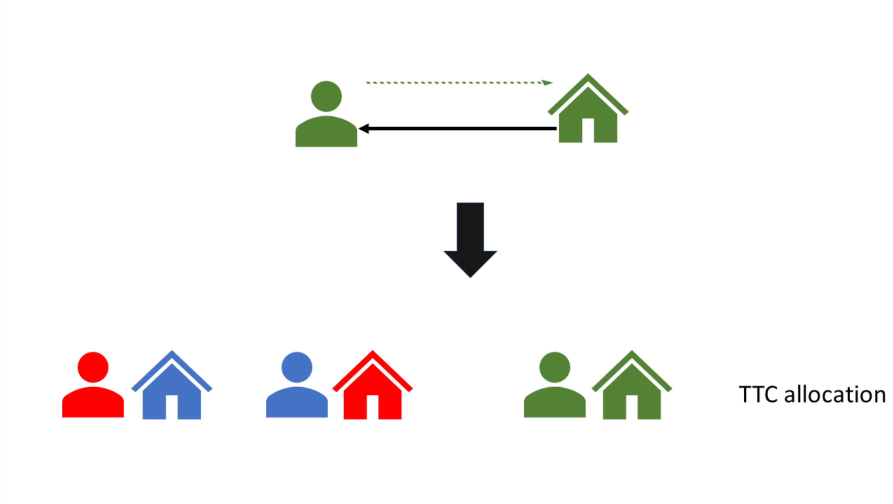
Example TTC, 2/4



Example TTC, 3/4



Example TTC, 4/4



Positive results for classical Shapley-Scarf housing markets

Shapley and Scarf (1974), main result

For each $R \in \mathcal{W}$, the set of allocations resulting from the TTC algorithm coincides with the set of competitive allocations $CA(R)$ and

$$\emptyset \neq CA(R) \subseteq C(R).$$

Roth and Postlewaite (1977): In particular, when preferences are strict, $|\mathrm{TTC}(R)| = 1$ and

$$\{\mathrm{TTC}(R)\} = SC(R).$$

Furthermore, $\mathrm{TTC}(R)$ is

- Pareto efficient and
- individually rational.

Positive results for classical Shapley-Scarf housing markets

A rule f assigns to each Shapley-Scarf housing market $R \in \mathcal{S} \,/\, \mathcal{W}$ an allocation $f(R)$.

The top trading cycles (TTC) rule assigns to each Shapley-Scarf housing market $R \in \mathcal{S} \,/\, \mathcal{W}$ the allocation $\mathrm{TTC}(R)$ (fixed tie-breaking on $\mathcal{W}$).

Roth (1982): "Incentive Compatibility in a Market with Indivisible Goods," *Economics Letters* 9.

Theorem (Roth, 1982)

Under the TTC rule, it is a weakly dominant strategy for each agent to reveal his true preferences. TTC is strategy-proof, i.e., for each (R_i, R_{-i}) and R_i',

$$\mathrm{TTC}_i(R_i, R_{-i}) \; R_i \; \mathrm{TTC}_i(R_i', R_{-i}).$$

Positive results for classical Shapley-Scarf housing markets

Bird (1984): "Group Incentive Compatibility in a Market with Indivisible Goods," *Economics Letters* 14.

Theorem (Bird, 1984)

The TTC rule is group incentive compatible, i.e., no subset of agents can report preferences other than their true preferences and make all members of that subset better off.

TTC is weakly group strategy-proof on $\mathcal{W}$, i.e., there does not exist (R_S, R_{-S}) and R'_S such that for each $i \in S$, $\mathrm{TTC}_i(R'_S, R_{-S})\, P_i\, \mathrm{TTC}_i(R_S, R_{-S})$.

TTC is group strategy-proof on $\mathcal{S}$, i.e., there does not exist (R_S, R_{-S}) and R'_S such that for each $i \in S$, $\mathrm{TTC}_i(R'_S, R_{-S})\, R_i\, \mathrm{TTC}_i(R_S, R_{-S})$ and for some $j \in S$, $\mathrm{TTC}_j(R'_S, R_{-S})\, P_j\, \mathrm{TTC}_j(R_S, R_{-S})$.

Don’t trust old and well-cited papers

Sandholtz and Tai (2024): “Group incentive compatibility in a market with indivisible goods: A comment,” *Economics Letters* 243.

“We note that the proofs of Bird (1984), the first to show group strategy-proofness of top trading cycles (TTC), require correction.

We provide a counterexample to a critical claim and present corrected proofs in the spirit of the originals.”

Positive results for classical Shapley-Scarf housing markets

Ma (1994): "Strategy-Proofness and the Strict Core in a Market with Indivisibilities," *International Journal of Game Theory* 23.

Theorem (Ma, 1994)

For strict preferences $\mathcal{S}$, rule f satisfies individual rationality, Pareto efficiency, and (group) strategy-proofness if and only if $f = \mathrm{TTC}$.

For the subset of weak preferences $\mathcal{W}$ with a non-empty strong core, a correspondence F satisfies individual rationality, Pareto efficiency, and strategy-proofness if and only if F selects from the strong core.

Happy 30th birthday TTC characterization!

Alternative proofs of Ma's characterization result are provided in

- Svensson (1999): "Strategy-Proof Allocation of Indivisible Goods," *Social Choice and Welfare* 16.
- Anno (2015): "A Short Proof for the **Characterization of the Core** in Housing Markets," *Economics Letters* 128.
- Sethuraman (2016): "An Alternative Proof of a **Characterization of the TTC mechanism**," *Operations Research Letters* 44.

Theorem (Sethuraman, 2016)

On $\mathcal{S}$ there is at most one rule that satisfies individual rationality, Pareto efficiency, and strategy-proofness.

Recall that Ma (1994) also showed that on the subset of weak preferences $\mathcal{W}$ with a non-empty strong core,

- the core is essentially single-valued and
- that a correspondence F satisfies individual rationality, Pareto efficiency, and strategy-proofness if and only if F selects from the strong core.

Sönmez (1999): "Strategy-Proofness and Essentially Single-Valued Cores," *Econometrica* 67.

Theorem (Sönmez, 1999)

For generalized indivisible goods allocation problems (no indifferences with endowments, plus domain richness): individual rationality, Pareto efficiency, and strategy-proofness imply essential single-valuedness and selection from the strong core.

Strong core versus TTC rule

Recall also that for strict preferences $\mathcal{S}$, the TTC allocation is the unique strong core allocation.

Research Question:

Do the properties in Ma's characterization characterize the TTC rule or the strong core rule?

I will answer this question for housing markets with limited externalities.

The TTC rule is great, but ...

there are several key assumptions!

- **Strict** versus weak preferences

  e.g., Jaramillo and Manjunath (2012) and Alcalde-Unzu and Molis (2013) independently constructed different classes of rules that satisfy individual rationality, Pareto efficiency, and strategy-proofness.

  These results are unified and extended by various computer science teams: Aziz and de Keijzer (2012); Plaxton (2013); Saban and Sethuraman (2013); Xiong, Wang, and He (2022).

  However, to the best of my knowledge, a characterization à la Ma does not exist.

The TTC rule is great, but ...

- **Unit-demand** versus multi-demand

  Moulin (1995) introduced multiple-type housing markets and Konishi et al. (2001) showed for (separable) multiple-type housing markets that individual rationality, Pareto efficiency, and strategy-proofness are incompatible, i.e., a characterization à la Ma does not hold anymore.

  Feng, Klaus, Klijn (2024a,b) characterize different extension of the TTC rule to multiple-type housing markets (some of these results, Flip will present in his talk just after this one). See Altuntas et al. (2023) and Coreno and Feng (2024) for further results when agents own sets of objects.

- **Selfishness** versus externalities

  This presentation.

Some previous literature on externalities

- Mumcu and Sağlam (*Economics Bulletin*, 2007): general preferences over allocations: the core may be empty.
- Graziano, Meo, and Yannelis (*Journal of Public Economic Theory*, 2020): existence and uniqueness of stable sets for two particular preference domains.
- Hong and Park (*Journal of Mathematical Economics*, 2022): existence of core solutions and the TTC rule for “egocentric” and “hedonic” preferences.
- Aziz and Lee (*Proceedings of the 19th International Conference on Autonomous Agents and Multi-Agent System*, AAMAS ’20, 2020): preferences with limited externalities, computational complexity, partial axiomatic results for TTC.

Housing markets with limited externalities

Klaus and Meo (2023): “The core for housing markets with limited externalities,” *Economic Theory* 76.

- Each agent does not only care about the house he receives and
- he also cares about the recipient of his house (but not about the houses other agents receive);
- hence, the externality is limited.

Example: Vacation Home Exchange

Agents’ care who will be in their vacation home while they are away.

Housing markets with limited externalities

- A finite set $\boldsymbol{N} = \{1, \ldots, n\}$ of $\boldsymbol{n}$ **agents**.
- A finite set $\boldsymbol{H} = \{H_1, \ldots, H_n\}$ of $\boldsymbol{n}$ **houses**.
- An **allocation** is a feasible assignment of houses to agents.

  $\boldsymbol{h = (H_1, \ldots, H_n)}$ is the **endowment allocation**.
- Given an allocation $\boldsymbol{a}$, the **allotment** of agent i is the pair

$$\boldsymbol{(a(i), a^{-1}(H_i)) \in H \times N,}$$

  formed by the object $a(i)$ assigned to agent i and the agent who receives agent i's house, i.e., $a^{-1}(H_i)$.

Endowment allotment is linked - coordinates are not independent

$$a(i) = H_i \text{ if and only if } a^{-1}(H_i) = i.$$

Separable preferences

Agent $i \in N$ has separable preferences if he has

1. a demand preference relation $\succ_i^d$ over the set H of houses;
2. a supply preference relation $\succ_i^s$ over the set N of agents;
3. a preference relation $\succ_i$ over his allotments that is separable, i.e., coordinatewise improvements yield better allotments.

Demand lexicographic preferences

An agent $i \in N$ has demand lexicographic preferences if his preferences are separable with strict demand preferences and he primarily cares about the house he receives and only secondarily about who receives his house, i.e.,

for any two allotments $(h, j), (h', k)$,

$$(h, j) \succ_i (h', k)$$

if and only if

$$h \succ_i^d h' \text{ or } [h = h' \text{ and } j \succ_i^s k].$$

Similarly, we can define supply lexicographic preferences.

Demand lexicographic preferences

Example

Agent 1:	$\succ_1^d$	h_3	h_2	h_1
	$\succ_1^s$	3	2	1

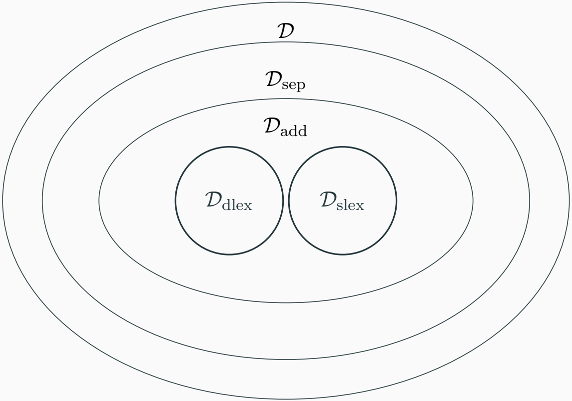
Separable preferences that are demand lexicographic:

$$(h_3, 3) \succ_1 (h_3, 2) \succ_1 (h_2, 3) \succ_1 (h_2, 2) \succ_1 (h_1, 1).$$

Separable preferences that are not demand lexicographic:

$$(h_3, 2) \succ_1 (h_1, 1) \succ_1 (h_3, 3) \succ_1 (h_2, 2) \succ_1 (h_2, 3).$$

Separable preference domains (including separable additive preferences)



Properties of rules

The definitions of rules and Ma's characterization properties remain the same:

- individual rationality,
- Pareto efficiency, and
- (group) strategy-proofness.

However, we also consider the following Shapley-Scarf housing market rule properties.

Pair efficiency

For Shapley-Scarf housing markets, Ekici (2024) weakened Pareto efficiency by requiring that no pair of agents can gain from swapping their assigned houses.

Without externalities, if two agents swapped houses to be better off, the obtained allocation would be a Pareto improvement.

However, in our model with limited externalities, the obtained allocation might not only affect the demand preferences of the two agents that swap, it might at the same time impact other agents' supply preferences, possibly making them worse off.

Therefore, in order to maintain the spirit of an efficiency property, we'll require that after the swap, all agents are weakly better off.

Pair efficiency

An allocation a is pair efficient if there is no pair of agents $i, j \in N$, $i \neq j$, such that allocation b that is obtained from a by agents i and j swapping houses a_i and a_j, Pareto dominates allocation a.

Theorem (Ekici, 2024)

For Shapley-Scarf housing markets with strict preferences $\mathcal{S}$, rule f satisfies individual rationality, pair efficiency, and strategy-proofness if and only if $f = \mathrm{TTC}$.

Ekici (2023): "Pair-efficient reallocation of indivisible objects" *Theoretical Economics* 19.

Adapting Sethuraman's proof strategy for Ma's result, a short proof is provided in Ekici and Sethuraman (2024): "Characterizing the TTC rule via pair-efficiency: A short proof," *Economics Letters* 234.

Stability

Another solution concept based on the idea of "stable exchange" was introduced by Roth and Postlewaite (1977): an allocation is stable when no group of agents can reallocate the houses they have obtained such that each agent in the group is strictly better off.

Roth and Postlewaite (1977) show that for Shapley-Scarf housing markets without externalities, stability is equivalent to Pareto efficiency.

However, when there are externalities, stability is a stronger property than Pareto efficiency.

We weaken stability to pairwise stability by requiring that no two agents i and j can be strictly better off by swapping the houses they have obtained at the corresponding allocation.

Pairwise stability

An allocation and a is pairwise stable if there exists no pair of agents $i, j \in N$, $i \neq j$, such that allocation b that is obtained from a by agents i and j swapping houses a_i and a_j, is strictly better for both agents.

Note that for Shapley-Scarf housing markets, pairwise stability and pair efficiency are equivalent.

This is not the case in our model. By definition, pairwise stability implies pair efficiency, but the converse does not hold.

TTC rule characterizations for demand lex. preferences

The top trading cycles (TTC) rule assigns to each market $(N, h, \succcurlyeq)$ with associated Shapley-Scarf market $(N, h, \succcurlyeq^d)$, the allocation $\mathrm{TTC}(\succcurlyeq^d)$.

Our main result is the following characterization:

Theorem 1 (with pair efficiency)

For demand lexicographic preferences, the TTC rule is the only rule satisfying individual rationality, pair efficiency, and strategy-proofness.

Corollary 1 (with Pareto efficiency)

For demand lexicographic preferences, the TTC rule is the only rule satisfying individual rationality, Pareto efficiency, and strategy-proofness.

TTC rule characterizations for demand lex. preferences

For Shapley-Scarf housing markets with egocentric preferences (a larger preference domain), Hong and Park (2022, Proposition 4) characterize the TTC rule by individual rationality, stability, and strategy-proofness.

Since pairwise stability implies pair efficiency, for our model with demand lexicographic preferences, we obtain a corresponding result with pairwise stability instead of stability.

Corollary 2 (with pairwise stability)

For demand lexicographic preferences, the TTC rule is the only rule satisfying individual rationality, pairwise stability, and strategy-proofness.

Impossibility, e.g., for separable preferences

We obtain an impossibility result when extending the preference domain to include mixed lexicographic preferences (e.g., separable preferences).

Impossibility result

Let $|N| \geq 3$ and $\widetilde{\mathcal{D}}$ be a preference domain that contains the domains of demand and supply lexicographic preferences (e.g., additive separable preferences). Then, no rule defined on $\widetilde{\mathcal{D}}^N$ satisfies individual rationality, pair efficiency, and strategy-proofness.

Clearly, the above impossibility result persists when we replace pair efficiency by Pareto efficiency or (pairwise) stability.

Discussion

Klaus and Meo (2023) showed that for demand lexicographic preferences, the strong core can be multi-valued. Hence,

f is the TTC rule

$\Updownarrow$ Theorem 1 / Corollary 1

f is individually rational, pair/Pareto efficient, strategy-proof

$\not\Uparrow \; \Downarrow$ Klaus and Meo (2023)

f assigns a strong core allocation

Thank You!!!

Please feel free to send me an e-mail (`bettina.klaus@unil.ch`) if you have comments / feedback or would like to receive the current version of the paper.