

# Limited Farsightedness in Priority-Based Matching

Ata Atay\*    Ana Mauleon<sup>†</sup>    Vincent Vannetelbosch<sup>‡</sup>

December 14, 2022

## Abstract

We introduce the horizon- $k$  vNM stable set to study priority-based matching problems with limited farsightedness. We show that, once agents are sufficiently farsighted, the matching obtained from the Top Trading Cycles (TTC) algorithm becomes stable: a singleton set consisting of the TTC matching is a horizon- $k$  vNM stable set if the degree of farsightedness is greater than three times the number of agents in the largest cycle of the TTC. Hence, the TTC mechanism satisfies Pareto efficiency, strategy-proofness and limited farsighted stability. On the contrary, the matching obtained from the Deferred Acceptance (DA) algorithm may not belong to any horizon- $k$  vNM stable set for  $k$  large enough.

Keywords: priority-based matching; top trading cycle; stable sets; limited farsightedness.

---

\*Department of Mathematical Economics, Finance and Actuarial Sciences, University of Barcelona, Spain. E-mail: aatay@ub.edu

<sup>†</sup>CORE/LIDAM and CEREC, UCLouvain, Belgium. E-mail: ana.c.mauleon@gmail.com

<sup>‡</sup>CORE/LIDAM, UCLouvain, Belgium. E-mail: vincent.vannetelbosch@uclouvain.be

# 1 Introduction

Many objects such as houses, school seats, jobs or organs are allocated based on the preferences of the agents and their priorities.<sup>1</sup> Two approaches for analysing the stability of matchings have been proposed in the literature depending on whether and how far agents anticipate that their actions may also induce others to change their matches. On the one hand, standard stability concepts involves fully myopic agents in the sense that they do not anticipate that others might react to their actions. On the other hand, a number of solution concepts involve perfectly farsighted agents who fully anticipate the complete sequence of reactions that results from their own actions. However, experimental evidence suggests that subjects are consistent with an intermediate degree of farsightedness: agents only anticipate a limited number of reactions by the other agents to the actions they take themselves.<sup>2</sup>

Two prominent mechanisms used for priority-based matching are the Gale and Shapley's (1962) Deferred Acceptance (DA) mechanism and the Shapley and Scarf's (1974) Top Trading Cycles (TTC) mechanism, applied across the globe in different forms, such as public school choice, subsidized housing assignment, and cadaver organ allocation. Abdulkadiroğlu and Sönmez (2003) show that both mechanisms are strategy-proof: truthful preference revelation is a weakly dominant strategy for the agents.<sup>3</sup> It is well-established that these two prominent mechanisms may fail to combine two important desiderata of stability and efficiency. On the one hand, the TTC mechanism is Pareto efficient while the DA mechanism may select an inefficient matching. On the other hand, the DA mechanism is stable while the TTC mechanism may select an unstable matching.<sup>4</sup> Shi (2022) points out this tension to design an optimal priority system for priority-based matching problems. Bichler, Hammerl, Morrill and Waldherr (2021) discuss this central problem of trade-off between these two properties by employing a design science approach.

Existing literature that examines the trade-off between stability and efficiency

---

<sup>1</sup>Roth and Sotomayor (1990) and Haeringer (2017) provide a general introduction to matching problems.

<sup>2</sup>See e.g. Kirchsteiger, Mantovani, Mauleon and Vannetelbosch (2016).

<sup>3</sup>Dubins and Freedman (1981) and Roth (1982) were first to show that the DA mechanism satisfies strategy-proofness in one-to-one matching problems.

<sup>4</sup>Reny (2022) introduces the Priority-Efficient (PE) mechanism that always selects a Pareto efficient matching that dominates the DA stable matching, but PE is not strategy-proof. See also Kesten (2006, 2010).

tends to adopt a myopic perspective. However, experimental evidence suggests that agents exhibit an intermediate degree of farsightedness. This leads to two natural questions: First, in one-to-one priority-based matching problems, is it possible to stabilize the matching obtained from the DA or TTC algorithm when agents are limited farsighted? Second, if so, to what extent do agents need to be farsighted? In contrast to previous studies, we explore these questions by examining stability from a farsighted standpoint. In light of the experimental evidence, we investigate priority-based matching problems considering limited farsightedness.

To answer these questions we introduce the notion of horizon- $k$  vNM stable set to study the matchings that are stable when agents are limited in their degree of farsightedness.<sup>5</sup> A horizon- $k$  improving path for priority-based matching problems is a sequence of matchings that can emerge when limited farsighted agents form or destroy matches based on the improvement the  $k$ -steps ahead matching offers them relative to the current matching. A set of matchings is a horizon- $k$  vNM stable set if (Internal Stability) for any two matchings belonging to the set, there is no horizon- $k$  improving path from one matching to the other one, and (External Stability) there always exists a horizon- $k$  improving path from every matching outside the set to some matching within the set.

We show that, once agents are sufficiently farsighted, the matching obtained from the TTC algorithm becomes stable in one-to-one priority-based matching problems. Precisely, a singleton set consisting of the TTC matching is a horizon- $k$  vNM stable set if the degree of farsightedness is greater than three times the number of agents in the largest cycle of the TTC. We provide a constructive proof where we build step by step a horizon- $k$  improving path from any matching leading to the TTC matching. Along the horizon- $k$  improving path, agents move one at a time and agents belonging to cycles sequentially act in the order of the formation of cycles in the TTC algorithm. Looking forward  $k$  steps ahead, agents belonging to a cycle first match one by one to the object that ranks them first on their priority list. Second, they give up one by one their object, and by doing so, vacating the object. Third, they match one by one to the object they are assigned to in the TTC matching. The number of steps in this improving path is at most equal to three times the number

---

<sup>5</sup>See Chwe (1994), Mauleon, Vannetelbosch and Vergote (2011), Ray and Vohra (2015, 2019), Herings, Mauleon and Vannetelbosch (2019, 2020), Luo, Mauleon and Vannetelbosch (2021) for definitions of the farsighted stable set.

of agents in the largest cycle of the TTC. Hence, looking forward such a number of steps ahead allows the agents to anticipate ending up with their TTC matches; and by doing so, they have incentives for engaging a move towards the matches they have in the TTC matching. Finally, we show that, in the special case where each agent owns an object, a singleton set consisting of the TTC matching is the unique horizon- $k$  vNM stable set.

Thus, the matching obtained from the TTC algorithm is not only Pareto efficient and strategy-proof, it is also horizon- $k$  stable. On the contrary, the matching obtained from the DA algorithm may not belong to any horizon- $k$  vNM stable set for  $k$  large enough. Next, we provide a tighter bound on the length of the largest cycle that is sufficient for stabilizing the TTC matching. However, this tighter bound requires more coordination on behalf of the agents. We also show that our main result is robust to an alternative concept for limited farsightedness, which is obtained by adapting Herings, Mauleon and Vannetelbosch (2019) definition of a horizon- $L$  farsighted set of networks to priority-based matching problems. This concept mainly replaces the internal stability condition of the horizon- $k$  vNM stable set by two alternative conditions: deterrence of external deviations and minimality.

Our results complement Atay, Mauleon and Vannetelbosch (2022) who show that for school choice problems with farsighted students, a singleton set consisting of the TTC matching is a farsighted stable set.<sup>6</sup> In addition, they also find that the DA matching may not belong to any farsighted stable set, and so the TTC mechanism provides an assignment that is not only Pareto efficient but also farsightedly stable for many-to-one priority-based matching problems. Notice that Atay, Mauleon and Vannetelbosch (2022) allow group of students (and schools) to move all together. Here, we show that, in one-to-one priority-based matching problems, their main result is robust to one agent moving at a time and limited farsightedness.

There is a recent literature that provides theoretical and/or empirical results supporting the TTC mechanism in one-to-one priority-based matching problems. For instance, Abdulkadiroğlu, Che, Pathak, Roth and Tercieux (2020) show that the TTC mechanism is justified envy minimal in the class of Pareto efficient and strategy-proof mechanisms. In addition, Doğan and Ehlers (2022) find that, for

---

<sup>6</sup>Morrill (2015) and Hakimov and Kesten (2018) propose variations of the TTC for school choice problems. Atay, Mauleon and Vannetelbosch (2022) show that the matchings obtained from the variations are all farsightedly stable. For one-to-one priority-based problems, all variations coincide.

any stability comparison satisfying three basic properties, the TTC mechanism is minimally unstable among Pareto efficient and strategy-proof mechanisms.<sup>7</sup>

Furthermore, priority-based matching problems serve as important examples of broader issues in network formation and coalition formation. Economic desiderata play a crucial role in decision support and decision science literature (Banker and Kauffman, 2004) within this context. Although literature on farsightedness in operations management is still limited, there is a growing body of research in this area. Granot and Sošić (2005) introduce farsightedness in the context of strategic alliance formation in online exchanges, while Sošić (2006) examines farsightedness in the inventory sharing problem among retailers. Additionally, Nagarajan and Sošić (2007) investigate farsighted coalition formation in competitive markets.

The paper is organized as follows. In Section 2, we introduce priority-based matching problems. In Section 3, we provide a formal description of the TTC mechanism and its algorithm. In Section 4, we introduce the notions of horizon- $k$  improving path and horizon- $k$  stable set for priority-based matching problems. In Section 5 we provide our main result. In Section 6, we discuss some extensions: (i) a tighter bound on the degree of farsightedness, (ii) all agents are farsighted, (iii) an alternative concept of limited farsightedness, and (iv) the matching problem where each agent owns an object.

## 2 Priority-based matching problems

A priority-based matching problem is a list  $\langle I, S, P, F \rangle$  where

- (i)  $I = \{i_1, \dots, i_n\}$  is the set of agents,
- (ii)  $S = \{s_1, \dots, s_m\}$  is the set of objects,
- (iv)  $P = (P_{i_1}, \dots, P_{i_n})$  is the preference profile where  $P_i$  is the strict preference of agent  $i$  over the objects and her outside option,
- (v)  $F = (F_{s_1}, \dots, F_{s_m})$  is the strict priority structure of the objects over the agents.

---

<sup>7</sup>Doğan and Ehlers (2021) study efficient and minimally unstable Pareto improvements over the DA mechanism. Ehlers (2008) investigates the impact of truncation strategies that leave the true ranking of the positions unchanged, but might drop some acceptable positions.

Let  $i$  be a generic agent and  $s$  be a generic object. The preference  $P_i$  of agent  $i$  is a linear order over  $S \cup \{i\}$ . Agent  $i$  prefers object  $s$  to object  $s'$  if  $sP_i s'$ . Object  $s$  is acceptable to agent  $i$  if  $sP_i i$  (i.e.  $sP_i i$  means that  $i$  strictly prefers  $s$  to being unassigned). We often write  $P_i = s, s', s''$  meaning that agent  $i$ 's most preferred object is  $s$ , her second best is  $s'$ , her third best is  $s''$  and any other object is unacceptable for her. Let  $R_i$  be the weak preference relation associated with the strict preference relation  $P_i$ .

The priority  $F_s$  of object  $s$  is a linear order over  $I$ . That is,  $F_s$  assigns ranks to agents according to their priority for object  $s$ . The rank of agent  $i$  for object  $s$  is denoted by  $F_s(i)$  and  $F_s(i) < F_s(j)$  means that agent  $i$  has higher priority for object  $s$  than agent  $j$ . For  $s \in S, i \in I$ , let  $\Phi(s, i) = \{j \in I \mid F_s(j) < F_s(i)\}$  be the set of agents who have higher priority than agent  $i$  for object  $s$ .

An outcome of a priority-based matching problem is a matching  $\mu : I \rightarrow S \cup I$  such that for any  $i \in I$  and any  $s \in S$ ,

(i)  $\mu(i) \notin S \Rightarrow \mu(i) = i$ , and

(ii)  $\#(\mu^{-1}(s)) \leq 1$ .

Condition (i) means that agent  $i$  is assigned to object  $s$  under  $\mu$  if  $\mu(i) = s$  and is unassigned under  $\mu$  if  $\mu(i) = i$ . Condition (ii) requires that no object is assigned to more than one agent. The set of all matchings is denoted  $\mathcal{M}$ .<sup>8</sup> For instance,

$$\mu = \begin{pmatrix} i_1 & i_2 & i_3 & i_4 \\ s_2 & s_3 & s_1 & i_4 \end{pmatrix}$$

is the matching where agent  $i_1$  is assigned to object  $s_2$ , agent  $i_2$  is assigned to object  $s_3$ , agent  $i_3$  is assigned to object  $s_1$ , and agent  $i_4$  is unassigned. For convenience, we often write such matching as  $\mu = \{(i_1, s_2), (i_2, s_3), (i_3, s_1), (i_4, i_4)\}$ .

A matching  $\mu'$  Pareto dominates a matching  $\mu$  if  $\mu'(i)R_i\mu(i)$  for all  $i \in I$  and  $\mu'(j)P_j\mu(j)$  for some  $j \in I$ . A matching is Pareto efficient if it is not Pareto dominated by another matching. Let  $\mathcal{E}(I, S, P, F)$  be the set of Pareto efficient matchings. A matching  $\mu$  is stable if

(i) for all  $i \in I$  we have  $\mu(i)R_i i$  (individual rationality),

---

<sup>8</sup>Throughout the paper we use the notation  $\subseteq$  for weak inclusion and  $\subset$  for strict inclusion. Finally,  $\#$  will refer to the notion of cardinality.

- (ii) for all  $i \in I$  and all  $s \in S$ , if  $sP_i\mu(i)$  then  $\#(\mu^{-1}(s)) = 1$  (non-wastefulness),
- (iii) for all  $i, j \in I$  with  $\mu(j) = s$ , if  $\mu(j)P_i\mu(i)$  then  $j \in \Phi(s, i)$  (no justified envy).

Let  $\mathcal{S}(I, S, P, F)$  be the set of stable matchings.

A mechanism systematically selects a matching for any given priority-based matching problem  $\langle I, S, P, F \rangle$ . A mechanism is individually rational (non-wasteful / stable / Pareto efficient) if it always selects an individually rational (non-wasteful / stable / Pareto efficient) matching. A mechanism is strategy-proof if no agent can ever benefit by unilaterally misrepresenting her preferences.

### 3 The Top Trading Cycles algorithm

Abdulkadiroğlu and Sönmez (2003) introduce the Top Trading Cycles (TTC) mechanism for selecting a matching for general priority-based matching problems. In the case of a priority-based matching problem  $\langle I, S, P, F \rangle$ , the TTC mechanism finds a matching by means of the following TTC algorithm.

Step 1. Each agent  $i \in I$  points to the object that is ranked first in  $P_i$ . If there is no such object, then agent  $i$  points to herself and she forms a self-cycle. Each object  $s \in S$  points to the agent that has the highest priority in  $F_s$ . Since the number of agents and objects are finite, there is at least one cycle. A cycle is an ordered list of distinct objects and distinct agents  $(s^1, i^1, s^2, \dots, s^l, i^l)$  where  $s^1$  points to  $i^1$  (denoted  $s^1 \mapsto i^1$ ),  $i^1$  points to  $s^2$  ( $i^1 \mapsto s^2$ ),  $s^l$  points to  $i^l$  ( $s^l \mapsto i^l$ ) and  $i^l$  points to  $s^1$  ( $i^l \mapsto s^1$ ). Each object (agent) can be part of at most one cycle. Every agent in a cycle is assigned to the object she points to and she is removed. Similarly, every agent in a self-cycle is not assigned to any object and is removed. If an object  $s$  is part of a cycle, then  $s$  is removed. Let  $C_1 = \{c_1^1, c_1^2, \dots, c_1^{L_1}\}$  be the set of cycles in Step 1 (where  $L_1 \geq 1$  is the number of cycles in Step 1). Let  $I_1$  be the set of agents who are assigned to some object at Step 1. Let  $S_1$  be the set of objects that are assigned to some agent at Step 1. Let  $m_1^l$  be all the matches from cycle  $c_1^l$  that are formed in Step 1 of the algorithm:

$$m_1^l = \begin{cases} \{(i, s) \mid i, s \in c_1^l \text{ and } i \mapsto s\} & \text{if } c_1^l \neq (j) \\ \{(j, j)\} & \text{if } c_1^l = (j) \end{cases}$$

where  $(j, j)$  simply means that agent  $j$  who is in a self-cycle ends up being definitely unassigned to any object. Let  $M_1 = \cup_{l=1}^{L_1} m_1^l$  be all the matches between agents and objects formed in Step 1 of the algorithm. Let  $\bar{c}_1^l = \#\{i \in I \mid i \in c_1^l\}$  be the number of agents involved in cycle  $c_1^l$ , for  $l = 1, \dots, L_1$ . Let  $c_1^{\max} = \max\{\bar{c}_1^1, \dots, \bar{c}_1^{L_1}\}$ .

Step  $k \geq 2$ . Each remaining agent  $i \in I \setminus \cup_{l=1}^{k-1} I_l$  points to the object  $s \in S \setminus \cup_{l=1}^{k-1} S_l$  that is ranked first in  $P_i$ . If there is no such object, then agent  $i$  points to herself and she forms a self-cycle. Each object  $s \in S \setminus \cup_{l=1}^{k-1} S_l$  points to the agent  $j \in I \setminus \cup_{l=1}^{k-1} I_l$  that has the highest priority in  $F_s$ . There is at least one cycle. Every agent in a cycle is assigned to the object she points to and she is removed. Similarly, every agent in a self-cycle is not assigned to any object and is removed. If an object  $s$  is part of a cycle, then  $s$  is removed. Let  $C_k = \{c_k^1, c_k^2, \dots, c_k^{L_k}\}$  be the set of cycles in Step  $k$  (where  $L_k \geq 1$  is the number of cycles in Step  $k$ ). Let  $I_k$  be the set of agents who are assigned to some object at Step  $k$ . Let  $S_k$  be the set of objects that are assigned to some agent at Step  $k$ . Let  $m_k^l$  be all the matches from cycle  $c_k^l$  that are formed in Step  $k$  of the algorithm:

$$m_k^l = \begin{cases} \{(i, s) \mid i, s \in c_k^l \text{ and } i \mapsto s\} & \text{if } c_k^l \neq (j) \\ \{(j, j)\} & \text{if } c_k^l = (j) \end{cases}$$

Let  $M_k = \cup_{l=1}^{L_k} m_k^l$  be all the matches between agents and objects formed in Step  $k$  of the algorithm. Let  $\bar{c}_k^l = \#\{i \in I \mid i \in c_k^l\}$  be the number of agents involved in cycle  $c_k^l$ , for  $l = 1, \dots, L_k$ . Let  $c_k^{\max} = \max\{\bar{c}_k^1, \dots, \bar{c}_k^{L_k}\}$ .

End. The algorithm stops when all agents have been removed. Let  $\bar{k}$  be the step at which the algorithm stops. Let  $\mu^T$  denote the matching obtained from the Top Trading Cycles mechanism and it is given by  $\mu^T = \cup_{k=1}^{\bar{k}} M_k$ . Let  $\gamma = \max\{c_1^{\max}, \dots, c_{\bar{k}}^{\max}\}$  be the maximum number of agents involved in any cycle of the TTC.

Abdulkadiroğlu and Sönmez (2003) show that the TTC mechanism is Pareto efficient and strategy-proof. TTC is also individually rational and non-wasteful, but it is not stable.

Another mechanism that is commonly adopted all over the world is Gale and Shapley's Deferred Acceptance (DA) algorithm. Let  $\mu^D$  denote the matching ob-



tained from the DA mechanism. The DA mechanism is strategy-proof and stable but not Pareto efficient.<sup>9</sup>

## 4 Horizon- $k$ vNM Stable Set

Is it possible to stabilize the matching obtained from the TTC algorithm once agents become limited farsighted? If yes, how much farsightedness from the agents do we need? To answer this question we propose the notion of horizon- $k$  vNM stable set for priority-based matching problems to study the matchings that are stable when agents are limited in their degree of farsightedness.

A horizon- $k$  improving path for priority-based matching problems is a sequence of matchings that can emerge when limited farsighted agents form or destroy matches based on the improvement the  $k$ -steps ahead matching offers them relative to the current matching. A set of matchings is a horizon- $k$  vNM stable set if (IS) for any two matchings belonging to the set, there is no horizon- $k$  improving path connecting from one matching to the other one, and (ES) there always exists a horizon- $k$  improving path from every matching outside the set to some matching within the set.

Given a matching  $\mu \in \mathcal{M}$  with agent  $i \in I$  assigned to object  $s \in S$ , so  $\mu(i) = s$ , the matching  $\mu'$  that is identical to  $\mu$ , except that the match between  $i$  and  $s$  has been destroyed by either  $i$  or  $s$ , is denoted by  $\mu' = \mu - (i, s)$ . Given a matching  $\mu \in \mathcal{M}$  such that  $i \in I$  and  $s \in S$  are not matched to one another, the matching  $\mu'$  that is identical to  $\mu$ , except that the pair  $(i, s)$  has formed at  $\mu'$  (and some  $j = \mu^{-1}(s)$  becomes unassigned if  $\#\mu^{-1}(s) = 1$ ), is denoted by  $\mu' = \mu + (i, s) - \{(i, \mu(i)) \mid \mu(i) \neq i\} - \{(\mu^{-1}(s), s) \mid \#(\mu^{-1}(s)) = 1\}$ .

**Definition 1.** Let  $\langle I, S, P, F \rangle$  be a priority-based matching problem. A horizon- $k$  improving path from a matching  $\mu \in \mathcal{M}$  to a matching  $\mu' \in \mathcal{M} \setminus \{\mu\}$  is a finite sequence of distinct matchings  $\mu_0, \dots, \mu_L$  with  $\mu_0 = \mu$  and  $\mu_L = \mu'$  such that for every  $l \in \{0, \dots, L-1\}$  either

- (i)  $\mu_{l+1} = \mu_l - (i, s)$  for some  $(i, s) \in I \times S$  such that  $\mu_{\min\{l+k, L\}}(i) P_i \mu_l(i)$ , or
- (ii)  $\mu_{l+1} = \mu_l + (i, s) - \{(i, \mu_l(i)) \mid \mu_l(i) \neq i\} - \{(\mu_l^{-1}(s), s) \mid \#(\mu_l^{-1}(s)) = 1\}$  for some  $(i, s) \in I \times S$  such that  $\mu_{\min\{l+k, L\}}(i) P_i \mu_l(i)$  and  $F_s(i) < F_s(\mu_l^{-1}(s))$  if

---

<sup>9</sup>Che and Tercieux (2019) show that both Pareto efficiency and stability can be achieved asymptotically using DA and TTC mechanisms when agents have uncorrelated preferences.

$$\#(\mu_l^{-1}(s)) = 1.$$

Definition 1 tells us that a horizon- $k$  improving path for priority-based matching problems consists of a sequence of matchings where along the sequence agents form or destroy matches based on the improvement the  $k$ -steps ahead matching offers them relative to the current one. Precisely, along a horizon- $k$  improving path, each time some agent  $i$  is on the move she is comparing her current match (i.e.  $\mu_l(i)$ ) with the match she will get  $k$ -steps ahead on the sequence (i.e.  $\mu_{l+k}(i)$ ) except if the end matching of the sequence lies within her horizon (i.e.  $L < l + k$ ). In such a case, she simply compares her current match (i.e.  $\mu_l(i)$ ) with the end match (i.e.  $\mu_L$ ).

Objects can be assigned to any agent on their priority lists unless they have already been assigned to some agent. When an object  $s \in S$  is already assigned to some agent  $\mu_l^{-1}(s)$  at  $\mu_l$ , this object  $s$  can be reassigned to another agent  $\mu_{l+1}^{-1}(s) \neq \mu_l^{-1}(s)$  at  $\mu_{l+1}$  only if agent  $\mu_{l+1}^{-1}(s)$  has a higher priority than agent  $\mu_l^{-1}(s)$ .

Let some  $\mu \in \mathcal{M}$  be given. If there exists a horizon- $k$  improving path from a matching  $\mu$  to a matching  $\mu'$ , then we write  $\mu \rightarrow_k \mu'$ . The set of matchings  $\mu' \in \mathcal{M}$  such that there is a horizon- $k$  improving path from  $\mu$  to  $\mu'$  is denoted by  $\phi_k(\mu)$ , so  $\phi_k(\mu) = \{\mu' \in \mathcal{M} \mid \mu \rightarrow_k \mu'\}$ .

**Definition 2.** Let  $\langle I, S, P, F \rangle$  be a priority-based matching problem. A set of matchings  $V \subseteq \mathcal{M}$  is a horizon- $k$  vNM stable set if it satisfies:

- (i) Internal stability (IS): For every  $\mu, \mu' \in V$ , it holds that  $\mu' \notin \phi_k(\mu)$ .
- (ii) External stability (ES): For every  $\mu \in \mathcal{M} \setminus V$ , it holds that  $\phi_k(\mu) \cap V \neq \emptyset$ .

Condition (i) of Definition 2 corresponds to internal stability. For any two matchings  $\mu$  and  $\mu'$  in the horizon- $k$  vNM stable set  $V$  there is no horizon- $k$  improving path connecting  $\mu$  to  $\mu'$ . Condition (ii) of Definition 2 expresses external stability. There exists a horizon- $k$  improving path from every matching  $\mu$  outside the horizon- $k$  vNM stable set  $V$  to some matching in  $V$ .<sup>10</sup>

---

<sup>10</sup>Ehlers (2007) and Herings, Mauleon and Vannetelbosch (2017) study vNM stable sets when all agents are myopic in two-sided matching problems.

## 5 Main Result

Remember that  $\gamma = \max\{c_1^{\max}, \dots, c_k^{\max}\}$  is the maximum number of agents involved in any cycle of the TTC.

**Theorem 1.** *Let  $\langle I, S, P, F \rangle$  be a priority-based matching problem and  $\mu^T$  be the matching obtained from the Top Trading Cycles mechanism. The singleton set  $\{\mu^T\}$  is a horizon- $k$  stable set for  $k \geq (3\gamma - 1)$ .*

*Proof.* Since  $\{\mu^T\}$  is a singleton set, internal stability (IS) is satisfied. (ES) Take any matching  $\mu \neq \mu^T$ , we need to show that  $\phi_k(\mu) \ni \mu^T$  for  $k \geq (3\gamma - 1)$ . We build in steps a horizon- $k$  improving path from  $\mu$  to  $\mu^T$  for  $k \geq (3\gamma - 1)$ .

Step 1.1. If  $m_1^1 \subseteq \mu$  and  $1 \neq L_1$  then go to Step 1.2 with  $\mu''_{1,1} = \mu$ . If  $m_1^1 \subseteq \mu$  and  $1 = L_1$  then go to Step 1.End with  $\mu''_{1,L_1} = \mu$ . If  $m_1^1 \not\subseteq \mu$  then go to Step 1.1.A.

Step 1.1.A. If  $\{(i, s) \mid i, s \in c_1^1 \text{ and } s \mapsto i\} \subseteq \mu$  then go to Step 1.1.B with  $\mu'_{1,1} = \mu$ . If  $\{(i, s) \mid i, s \in c_1^1 \text{ and } s \mapsto i\} \not\subseteq \mu$  then there is some agent  $i$  such that  $s \neq \mu(i) \neq \mu^T(i)$  and  $s \mapsto i$  with  $i, s \in c_1^1$ . This agent  $i$  matches with object  $s$  that ranks her first on its priority list. We reach the matching  $\mu'_{1,1,1} = \mu + (i, s) - \{(i, \mu(i)) \mid \mu(i) \neq i\} - \{(\mu^{-1}(s), s) \mid \#(\mu^{-1}(s)) = 1\}$  where  $s \mapsto i$  and  $i, s \in c_1^1$ . If  $\{(i, s) \mid i, s \in c_1^1 \text{ and } s \mapsto i\} \subseteq \mu'_{1,1,1}$  then go to Step 1.1.B with  $\mu'_{1,1} = \mu'_{1,1,1}$ . If  $\{(i, s) \mid i, s \in c_1^1 \text{ and } s \mapsto i\} \not\subseteq \mu'_{1,1,1}$  then there is some agent  $i'$  such that  $s' \neq \mu'_{1,1,1}(i') \neq \mu^T(i')$  and  $s' \mapsto i'$  with  $i', s' \in c_1^1$ . This agent  $i'$  matches with object  $s'$  that ranks her first on its priority list. We reach the matching  $\mu'_{1,1,2} = \mu'_{1,1,1} + (i', s') - \{(i', \mu'_{1,1,1}(i')) \mid \mu'_{1,1,1}(i') \neq i'\} - \{(\mu'^{-1}_{1,1,1}(s'), s') \mid \#(\mu'^{-1}_{1,1,1}(s')) = 1\}$  where  $s' \mapsto i'$  and  $i', s' \in c_1^1$ . We proceed as above until we reach the matching  $\mu'_{1,1} = \mu + \{(i, s) \mid i, s \in c_1^1 \text{ and } s \mapsto i\} - \{(i, \mu(i)) \mid i, s \in c_1^1, s \mapsto i \text{ and } \mu(i) \neq s\} - \{(\mu^{-1}(s), s) \mid i, s \in c_1^1, s \mapsto i \text{ and } \mu^{-1}(s) \neq i\}$  where each agent involved in  $c_1^1$  is matched to the object that ranks her first on its priority list. Step 1.1.A counts at most  $\bar{c}_1^1$  steps.

Step 1.1.B. Let  $\mathcal{I}_1^1 = \{(i^l)\}_{l=1}^{\bar{c}_1^1}$  be such that  $i^l \in c_1^1$  and  $i^l = i_{o_l} \neq i^{l+1} = i_{o_{l+1}}$  with  $o_l < o_{l+1}$  for  $l = 1, \dots, \bar{c}_1^1 - 1$ . That is,  $\mathcal{I}_1^1$  is an ordered set of the agents involved in cycle  $c_1^1$  where  $\bar{c}_1^1 = \#\{i \in I \mid i \in c_1^1\}$  is the number of agents involved in cycle  $c_1^1$ . From the matching  $\mu'_{1,1}$ , agents  $i^1$  to  $i^{\bar{c}_1^1-1}$  successively leave their objects to reach the matching  $\mu''_{1,1} = \mu'_{1,1} - \{(i, s) \mid i, s \in c_1^1, s \mapsto i \text{ and } i \neq i^{\bar{c}_1^1}\}$  where

only agent  $i^{\bar{c}_1^1}$  is still matched to the object that ranks her first on its priority list. Step 1.1.B counts at most  $\bar{c}_1^1 - 1$  steps.

Step 1.1.C. From the matching  $\mu''_{1,1}$ , agent  $i^{\bar{c}_1^1}$  first matches with her top choice object  $s$  to reach the matching  $\mu''_{1,1} + (i^{\bar{c}_1^1}, s) - (i^{\bar{c}_1^1}, \mu''_{1,1}(i^{\bar{c}_1^1}))$  where  $s$  is such that  $i \mapsto s$  and  $i, s \in c_1^1$ . Notice that  $s$  was unassigned at  $\mu''_{1,1}$  while the object that ranks  $i^{\bar{c}_1^1}$  first on its priority list, i.e.  $\mu''_{1,1}(i^{\bar{c}_1^1})$ , is now unassigned. Next, agents  $i^1$  to  $i^{\bar{c}_1^1-1}$  successively match to their top choice object to reach the matching  $\mu'''_{1,1} = \mu''_{1,1} - \{(i, s) \mid i, s \in c_1^1 \text{ and } s \mapsto i\} + \{(i, s) \mid i, s \in c_1^1 \text{ and } i \mapsto s\}$ . Step 1.1.C counts at most  $\bar{c}_1^1$  steps. We have reached  $\mu'''_{1,1}$  with  $m_1^1 \subseteq \mu'''_{1,1}$  and so agents belonging to  $c_1^1$  are assigned to the same object as in  $\mu^T$ . Step 1.1 counts at most  $3\bar{c}_1^1 - 1$  steps. Hence, it is sufficient that the agents who move in Step 1.1 look forward  $3\bar{c}_1^1 - 1$  steps ahead to have incentives for engaging the move towards the matching  $\mu'''_{1,1}$  where they already get the object assigned by the TTC. Once they reach those matches they do not move afterwards. If  $1 \neq L_1$ , then go to Step 1.2. Otherwise, go to Step 1.End with  $\mu'''_{1,L_1} = \mu'''_{1,1}$ .

Step 1.k. ( $k > 1$ ) If  $m_1^k \subseteq \mu'''_{1,k-1}$  and  $k \neq L_1$  then go to Step 1.k+1 with  $\mu'''_{1,k} = \mu'''_{1,k-1}$ . If  $m_1^k \subseteq \mu'''_{1,k-1}$  and  $k = L_1$  then go to Step 1.End with  $\mu'''_{1,L_1} = \mu'''_{1,k-1}$ . If  $m_1^k \not\subseteq \mu'''_{1,k-1}$  then go to Step 1.k.A.

Step 1.k.A. If  $\{(i, s) \mid i, s \in c_1^k \text{ and } s \mapsto i\} \subseteq \mu'''_{1,k-1}$  then go to Step 1.k.B with  $\mu'_{1,k} = \mu'''_{1,k-1}$ . If  $\{(i, s) \mid i, s \in c_1^k \text{ and } s \mapsto i\} \not\subseteq \mu'''_{1,k-1}$  then there is some agent  $i$  such that  $s \neq \mu'''_{1,k-1}(i) \neq \mu^T(i)$  and  $s \mapsto i$  with  $i, s \in c_1^k$ . This agent  $i$  matches with object  $s$  that ranks her first on its priority list. We reach the matching  $\mu'_{1,k,1} = \mu'''_{1,k-1} + (i, s) - \{(i, \mu'''_{1,k-1}(i)) \mid \mu'''_{1,k-1}(i) \neq i\} - \{(\mu'''_{1,k-1}^{-1}(s), s) \mid \#(\mu'''_{1,k-1}^{-1}(s)) = 1\}$  where  $s \mapsto i$  and  $i, s \in c_1^k$ . If  $\{(i, s) \mid i, s \in c_1^k \text{ and } s \mapsto i\} \subseteq \mu'_{1,k,1}$  then go to Step 1.k.B with  $\mu'_{1,k} = \mu'_{1,k,1}$ . If  $\{(i, s) \mid i, s \in c_1^k \text{ and } s \mapsto i\} \not\subseteq \mu'_{1,k,1}$  then there is some agent  $i'$  such that  $s' \neq \mu'_{1,k,1}(i') \neq \mu^T(i')$  and  $s' \mapsto i'$  with  $i', s' \in c_1^k$ . This agent  $i'$  matches with object  $s'$  that ranks her first on its priority list. We reach the matching  $\mu'_{1,k,2} = \mu'_{1,k,1} + (i', s') - \{(i', \mu'_{1,k,1}(i')) \mid \mu'_{1,k,1}(i') \neq i'\} - \{(\mu'_{1,k,1}^{-1}(s'), s') \mid \#(\mu'_{1,k,1}^{-1}(s')) = 1\}$  where  $s' \mapsto i'$  and  $i', s' \in c_1^k$ . We proceed as above until we reach the matching  $\mu'_{1,k} = \mu'''_{1,k-1} + \{(i, s) \mid i, s \in c_1^k \text{ and } s \mapsto i\} - \{(i, \mu'''_{1,k-1}(i)) \mid i, s \in c_1^k, s \mapsto i \text{ and } \mu'''_{1,k-1}(i) \neq s\} - \{(\mu'''_{1,k-1}^{-1}(s), s) \mid i, s \in c_1^k, s \mapsto i \text{ and } \mu'''_{1,k-1}^{-1}(s) \neq i\}$  where each agent involved in  $c_1^k$  is matched to the object that ranks her first on its priority list. Step 1.k.A counts at most  $\bar{c}_1^k$

steps.

Step 1.k.B. Let  $\mathcal{I}_1^k = \{(i^l)\}_{l=1}^{\bar{c}_1^k}$  be such that  $i^l \in c_1^k$  and  $i^l = i_{o_l} \neq i^{l+1} = i_{o_{l+1}}$  with  $o_l < o_{l+1}$  for  $l = 1, \dots, \bar{c}_1^k - 1$ . That is,  $\mathcal{I}_1^k$  is an ordered set of the agents involved in cycle  $c_1^k$  where  $\bar{c}_1^k = \#\{i \in I \mid i \in c_1^k\}$  is the number of agents involved in cycle  $c_1^k$ . From the matching  $\mu'_{1,k}$ , agents  $i^1$  to  $i^{\bar{c}_1^k-1}$  successively leave their objects to reach the matching  $\mu''_{1,k} = \mu'_{1,k} - \{(i, s) \mid i, s \in c_1^k, s \mapsto i \text{ and } i \neq i^{\bar{c}_1^k}\}$  where only agent  $i^{\bar{c}_1^k}$  is still matched to the object that ranks her first on its priority list. Step 1.k.B counts at most  $\bar{c}_1^k - 1$  steps.

Step 1.k.C. From the matching  $\mu''_{1,k}$ , agent  $i^{\bar{c}_1^k}$  first matches with her top choice object  $s$  to reach the matching  $\mu''_{1,k} + (i^{\bar{c}_1^k}, s) - (i^{\bar{c}_1^k}, \mu''_{1,k}(i^{\bar{c}_1^k}))$  where  $s$  is such that  $i \mapsto s$  and  $i, s \in c_1^k$ . Notice that  $s$  was unassigned at  $\mu''_{1,k}$  while the object that ranks  $i^{\bar{c}_1^k}$  first on its priority list, i.e.  $\mu''_{1,k}(i^{\bar{c}_1^k})$ , is now unassigned. Next, agents  $i^1$  to  $i^{\bar{c}_1^k-1}$  successively match to their top choice object to reach the matching  $\mu'''_{1,k} = \mu''_{1,k} - \{(i, s) \mid i, s \in c_1^k \text{ and } s \mapsto i\} + \{(i, s) \mid i, s \in c_1^k \text{ and } i \mapsto s\}$ . Step 1.k.C counts at most  $\bar{c}_1^k$  steps. We have reached  $\mu'''_{1,k}$  with  $m_1^k \subseteq \mu'''_{1,k}$  and so agents belonging to  $c_1^k$  are assigned to the same object as in  $\mu^T$ . Step 1.k counts at most  $3\bar{c}_1^k - 1$  steps. Hence, it is sufficient that the agents who move in Step 1.k look forward  $3\bar{c}_1^k - 1$  steps ahead to have incentives for engaging the move towards the matching  $\mu'''_{1,k}$  where they already get the object assigned by the TTC. Once they reach those matches they do not move afterwards. If  $1 \neq L_1$ , then go to Step 1.k + 1. Otherwise, go to Step 1.End with  $\mu'''_{1,L_1} = \mu'''_{1,k}$ .

Step 1.End. We have reached  $\mu'''_{1,L_1}$  with  $\cup_{l=1}^{L_1} m_1^l = M_1 \subseteq \mu'''_{1,L_1}$ . If  $\mu'''_{1,L_1} = \mu^T$  then the process ends. Otherwise, go to Step 2.1.

Step 2.1. If  $m_2^1 \subseteq \mu'''_{1,L_1}$  and  $1 \neq L_2$  then go to Step 2.2 with  $\mu'''_{2,1} = \mu'''_{1,L_1}$ . If  $m_2^1 \subseteq \mu'''_{1,L_1}$  and  $1 = L_2$  then go to Step 2.End with  $\mu'''_{2,L_2} = \mu'''_{1,L_1}$ . If  $m_2^1 \not\subseteq \mu'''_{1,L_1}$  then go to Step 2.1.A.

Step 2.1.A. If  $\{(i, s) \mid i, s \in c_2^1 \text{ and } s \mapsto i\} \subseteq \mu'''_{1,L_1}$  then go to Step 2.1.B with  $\mu'_{2,1} = \mu'''_{1,L_1}$ . If  $\{(i, s) \mid i, s \in c_2^1 \text{ and } s \mapsto i\} \not\subseteq \mu'''_{1,L_1}$  then there is some agent  $i$  such that  $s \neq \mu'''_{1,L_1}(i) \neq \mu^T(i)$  and  $s \mapsto i$  with  $i, s \in c_2^1$ . This agent  $i$  matches with object  $s \in S \setminus S_1$  that ranks her first on its priority list among agents belonging to  $I \setminus I_1$ . We reach the matching  $\mu'_{2,1,1} = \mu'''_{1,L_1} + (i, s) - \{(i, \mu'''_{1,L_1}(i)) \mid \mu'''_{1,L_1}(i) \neq i\} - \{(\mu'''_{1,L_1}^{-1}(s), s) \mid \#(\mu'''_{1,L_1}^{-1}(s)) = 1\}$  where  $s \mapsto i$  and  $i, s \in c_2^1$ .

If  $\{(i, s) \mid i, s \in c_2^1 \text{ and } s \mapsto i\} \subseteq \mu'_{2,1,1}$  then go to Step 2.1.B with  $\mu'_{2,1} = \mu'_{2,1,1}$ . If  $\{(i, s) \mid i, s \in c_2^1 \text{ and } s \mapsto i\} \not\subseteq \mu'_{2,1,1}$  then there is some agent  $i'$  such that  $s' \neq \mu'_{2,1,1}(i') \neq \mu^T(i')$  and  $s' \mapsto i'$  with  $i', s' \in c_2^1$ . This agent  $i'$  matches with object  $s' \in S \setminus S_1$  that ranks her first on its priority list among agents belonging to  $I \setminus I_1$ . We reach the matching  $\mu'_{2,1,2} = \mu'_{2,1,1} + (i', s') - \{(i', \mu'_{2,1,1}(i')) \mid \mu'_{2,1,1}(i') \neq i'\} - \{(\mu'^{-1}_{2,1,1}(s'), s') \mid \#(\mu'^{-1}_{2,1,1}(s')) = 1\}$  where  $s' \mapsto i'$  and  $i', s' \in c_2^1$ . We proceed as above until we reach the matching  $\mu'_{2,1} = \mu'''_{1,L_1} + \{(i, s) \mid i, s \in c_2^1 \text{ and } s \mapsto i\} - \{(i, \mu'''_{1,L_1}(i)) \mid i, s \in c_2^1, s \mapsto i \text{ and } \mu'''_{1,L_1}(i) \neq s\} - \{(\mu'''_{1,L_1}^{-1}(s), s) \mid i, s \in c_2^1, s \mapsto i \text{ and } \mu'''_{1,L_1}^{-1}(s) \neq i\}$  where each agent involved in  $c_2^1$  is matched to the object that ranks her first on its priority list among agents belonging to  $I \setminus I_1$ . Step 2.1.A counts at most  $\bar{c}_1^2$  steps.

Step 2.1.B. Let  $\mathcal{I}_2^1 = \{(i^l)\}_{l=1}^{\bar{c}_2^1}$  be such that  $i^l \in c_2^1$  and  $i^l = i_{o_l} \neq i^{l+1} = i_{o_{l+1}}$  with  $o_l < o_{l+1}$  for  $l = 1, \dots, \bar{c}_2^1 - 1$ . That is,  $\mathcal{I}_2^1$  is an ordered set of the agents involved in cycle  $c_2^1$  where  $\bar{c}_2^1 = \#\{i \in I \mid i \in c_2^1\}$  is the number of agents involved in cycle  $c_2^1$ . From the matching  $\mu'_{2,1}$ , agents  $i^1$  to  $i^{\bar{c}_2^1-1}$  successively leave their objects to reach the matching  $\mu''_{2,1} = \mu'_{2,1} - \{(i, s) \mid i, s \in c_2^1, s \mapsto i \text{ and } i \neq i^{\bar{c}_2^1}\}$  where only agent  $i^{\bar{c}_2^1}$  is still matched to the object that ranks her first on its priority list among agents belonging to  $I \setminus I_1$ . Step 2.1.B counts at most  $\bar{c}_1^2 - 1$  steps.

Step 2.1.C. From the matching  $\mu''_{2,1}$ , agent  $i^{\bar{c}_2^1}$  first matches with her top choice object  $s \in S \setminus S_1$  to reach the matching  $\mu''_{2,1} + (i^{\bar{c}_2^1}, s) - (i^{\bar{c}_2^1}, \mu''_{2,1}(i^{\bar{c}_2^1}))$  where  $s$  is such that  $i \mapsto s$  and  $i, s \in c_1^2$ . Notice that  $s$  was unassigned at  $\mu''_{2,1}$  while the object that ranks  $i^{\bar{c}_2^1}$  first on its priority list among agents belonging to  $I \setminus I_1$ , i.e.  $\mu''_{2,1}(i^{\bar{c}_2^1})$ , is now unassigned. Next, agents  $i^1$  to  $i^{\bar{c}_2^1-1}$  successively match to their top choice object in  $S \setminus S_1$  to reach the matching  $\mu'''_{2,1} = \mu''_{2,1} - \{(i, s) \mid i, s \in c_2^1 \text{ and } s \mapsto i\} + \{(i, s) \mid i, s \in c_2^1 \text{ and } i \mapsto s\}$ . Step 2.1.C counts at most  $\bar{c}_2^1$  steps. We have reached  $\mu'''_{2,1}$  with  $m_2^1 \subseteq \mu'''_{2,1}$  and so agents belonging to  $c_2^1$  are assigned to the same object as in  $\mu^T$ . Step 2.1 counts at most  $3\bar{c}_2^1 - 1$  steps. Hence, it is sufficient that the agents who move in Step 2.1 look forward  $3\bar{c}_2^1 - 1$  steps ahead to have incentives for engaging the move towards the matching  $\mu'''_{2,1}$  where they already get the object assigned by the TTC. Once they reach those matches they do not move afterwards. If  $1 \neq L_2$ , then go to Step 2.2. Otherwise, go to Step 2.End with  $\mu'''_{2,L_2} = \mu'''_{2,1}$ .

Step 2.k. ( $k > 1$ ) If  $m_2^k \subseteq \mu_{2,k-1}'''$  and  $k \neq L_2$  then go to Step 2.k+1 with  $\mu_{2,k}''' = \mu_{2,k-1}'''$ .  
 If  $m_2^k \subseteq \mu_{2,k-1}'''$  and  $k = L_2$  then go to Step 2.End with  $\mu_{2,L_2}''' = \mu_{2,k-1}'''$ . If  $m_2^k \not\subseteq \mu_{2,k-1}'''$  then go to Step 2.k.A.

Step 2.k.A If  $\{(i, s) \mid i, s \in c_2^k \text{ and } s \mapsto i\} \subseteq \mu_{2,k-1}'''$  then go to Step 2.k.B with  $\mu'_{2,k} = \mu_{2,k-1}'''$ . If  $\{(i, s) \mid i, s \in c_2^k \text{ and } s \mapsto i\} \not\subseteq \mu_{2,k-1}'''$  then there is some agent  $i$  such that  $s \neq \mu_{2,k-1}'''(i) \neq \mu^T(i)$  and  $s \mapsto i$  with  $i, s \in c_2^k$ . This agent  $i$  matches with object  $s$  that ranks her first on its priority list among agents belonging to  $I \setminus I_1$ . We reach the matching  $\mu'_{2,k,1} = \mu_{2,k-1}''' + (i, s) - \{(i, \mu_{2,k-1}'''(i)) \mid \mu_{2,k-1}'''(i) \neq i\} - \{(\mu_{2,k-1}'''^{-1}(s), s) \mid \#(\mu_{2,k-1}'''^{-1}(s)) = 1\}$  where  $s \mapsto i$  and  $i, s \in c_2^k$ . If  $\{(i, s) \mid i, s \in c_2^k \text{ and } s \mapsto i\} \subseteq \mu'_{2,k,1}$  then go to Step 2.k.B with  $\mu'_{2,k} = \mu'_{2,k,1}$ . If  $\{(i, s) \mid i, s \in c_2^k \text{ and } s \mapsto i\} \not\subseteq \mu'_{2,k,1}$  then there is some agent  $i'$  such that  $s' \neq \mu'_{2,k,1}(i') \neq \mu^T(i')$  and  $s' \mapsto i'$  with  $i', s' \in c_2^k$ . This agent  $i'$  matches with object  $s'$  that ranks her first on its priority list among agents belonging to  $I \setminus I_1$ . We reach the matching  $\mu'_{2,k,2} = \mu'_{2,k,1} + (i', s') - \{(i', \mu'_{2,k,1}(i')) \mid \mu'_{2,k,1}(i') \neq i'\} - \{(\mu_{2,k,1}'^{-1}(s'), s') \mid \#(\mu_{2,k,1}'^{-1}(s')) = 1\}$  where  $s' \mapsto i'$  and  $i', s' \in c_2^k$ . We proceed as above until we reach the matching  $\mu'_{2,k} = \mu_{2,k-1}''' + \{(i, s) \mid i, s \in c_2^k \text{ and } s \mapsto i\} - \{(i, \mu_{2,k-1}'''(i)) \mid i, s \in c_2^k, s \mapsto i \text{ and } \mu_{2,k-1}'''(i) \neq s\} - \{(\mu_{2,k-1}'''^{-1}(s), s) \mid i, s \in c_2^k, s \mapsto i \text{ and } \mu_{2,k-1}'''^{-1}(s) \neq i\}$  where each agent involved in  $c_2^k$  is matched to the object that ranks her first on its priority list among agents belonging to  $I \setminus I_1$ . Step 2.k.A counts at most  $\bar{c}_2^k$  steps.

Step 2.k.B. Let  $\mathcal{I}_2^k = \{(i^l)\}_{l=1}^{\bar{c}_2^k}$  be such that  $i^l \in c_2^k$  and  $i^l = i_{o_l} \neq i^{l+1} = i_{o_{l+1}}$  with  $o_l < o_{l+1}$  for  $l = 1, \dots, \bar{c}_2^k - 1$ . That is,  $\mathcal{I}_2^k$  is an ordered set of the agents involved in cycle  $c_2^k$  where  $\bar{c}_2^k = \#\{i \in I \mid i \in c_2^k\}$  is the number of agents involved in cycle  $c_2^k$ . From the matching  $\mu'_{2,k}$ , agents  $i^1$  to  $i^{\bar{c}_2^k-1}$  successively leave their objects to reach the matching  $\mu''_{2,k} = \mu'_{2,k} - \{(i, s) \mid i, s \in c_2^k, s \mapsto i \text{ and } i \neq i^{\bar{c}_2^k}\}$  where only agent  $i^{\bar{c}_2^k}$  is still matched to the object that ranks her first on its priority list among agents belonging to  $I \setminus I_1$ . Step 2.k.B counts at most  $\bar{c}_2^k - 1$  steps.

Step 2.k.C. From the matching  $\mu''_{2,k}$ , agent  $i^{\bar{c}_2^k}$  first matches with her top choice object  $s \in S \setminus S_1$  to reach the matching  $\mu''_{2,k} + (i^{\bar{c}_2^k}, s) - (i^{\bar{c}_2^k}, \mu''_{2,k}(i^{\bar{c}_2^k}))$  where  $s$  is such that  $i \mapsto s$  and  $i, s \in c_2^k$ . Notice that  $s$  was unassigned at  $\mu''_{2,k}$  while the object that ranks  $i^{\bar{c}_2^k}$  first on its priority list among agents belonging to  $I \setminus I_1$ , i.e.  $\mu''_{2,k}(i^{\bar{c}_2^k})$ , is now unassigned. Next, agents  $i^1$  to  $i^{\bar{c}_2^k-1}$  successively match to their top choice object in  $S \setminus S_1$  to reach the matching  $\mu'''_{2,k} = \mu''_{2,k} - \{(i, s) \mid$

$i, s \in c_2^k$  and  $s \mapsto i\}$  +  $\{(i, s) \mid i, s \in c_2^k \text{ and } i \mapsto s\}$ . Step 2. $k$ .C counts at most  $c_2^k$  steps. We have reached  $\mu_{2,k}'''$  with  $m_2^k \subseteq \mu_{2,k}'''$  and so agents belonging to  $c_2^k$  are assigned to the same object as in  $\mu^T$ . Step 2. $k$  counts at most  $3c_2^k - 1$  steps. Hence, it is sufficient that the agents who move in Step 2. $k$  look forward  $3c_2^k - 1$  steps ahead to have incentives for engaging the move towards the matching  $\mu_{2,k}'''$  where they already get the object assigned by the TTC. Once they reach those matches they do not move afterwards. If  $1 \neq L_2$ , then go to Step 2. $k+1$ . Otherwise, go to Step 2.End with  $\mu_{2,L_2}''' = \mu_{2,k}'''$ .

Step 2.End. We have reached  $\mu_{2,L_2}'''$  with  $M_1 \cup M_2 \subseteq \mu_{2,L_2}'''$ . If  $\mu_{2,L_2}''' = \mu^T$  then the process ends. Otherwise, go to Step 3.1.

End. The process goes on until we reach  $\mu_{\bar{k},L_{\bar{k}}}''' = \cup_{k=1}^{\bar{k}} M_k = \mu^T$ .

Given  $k \geq 3\gamma - 1$ , we have that, along the constructed horizon- $k$  improving path, each time an agent  $i$  is on the move she has incentives to do so since her end match (i.e. her TTC match  $\mu^T(i)$ ) is within her horizon.

□

The matching obtained from the TTC algorithm is always Pareto efficient but may not be stable when agents are myopic. Theorem 1 shows that, once agents are sufficiently farsighted (i.e.  $k \geq 3\gamma - 1$ ), the matching obtained from the TTC algorithm becomes stable. Example 1 highlights Theorem 1. In addition, it shows that, once agents are no more myopic, the matching obtained from the Deferred Acceptance (DA) algorithm may become unstable.

**Example 1.** Consider a priority-based matching problem  $\langle I, S, P, F \rangle$  with  $I = \{i_1, i_2, i_3\}$  and  $S = \{s_1, s_2, s_3\}$ . Agents' preferences and objects' priorities are as follows.

Agents			Objects		
$P_{i_1}$	$P_{i_2}$	$P_{i_3}$	$F_{s_1}$	$F_{s_2}$	$F_{s_3}$
$s_1$	$s_1$	$s_2$	$i_3$	$i_2$	$i_2$
$s_3$	$s_2$	$s_1$	$i_1$	$i_1$	$i_3$
$s_2$	$s_3$	$s_3$	$i_2$	$i_3$	$i_1$

Using Example 1 we provide the basic intuition behind Theorem 1 and its proof. In Example 1,  $\mu^T = \{(i_1, s_3), (i_2, s_1), (i_3, s_2)\}$  is the matching obtained from the



TTC algorithm. In the first round of the TTC algorithm, there is one cycle where agent  $i_2$  points to object  $s_1$ , object  $s_1$  points to agent  $i_3$ , agent  $i_3$  points to object  $s_2$  and object  $s_2$  points to agent  $i_2$ . That is,  $C_1 = \{c_1^1\}$  with  $c_1^1 = (s_1, i_3, s_2, i_2)$ . Agent  $i_2$  is assigned to object  $s_1$  and agent  $i_3$  is assigned to object  $s_2$ :  $m_1^1 = \{(i_2, s_1), (i_3, s_2)\}$ , and so  $i_2$  and  $i_3$  exchange their priority. In the second round of the TTC algorithm, there is only one leftover agent,  $i_1$ , who points to object  $s_3$  and one leftover object,  $s_3$ , that points to agent  $i_1$ . That is,  $C_2 = \{c_2^1\}$  with  $c_2^1 = (s_3, i_1)$ . Agent  $i_1$  is assigned to object  $s_3$ :  $m_2^1 = \{(i_1, s_3)\}$ , and so  $\mu^T = m_1^1 \cup m_2^1$ .

From Theorem 1 we know that  $\{\mu^T\}$  is a horizon- $k$  vNM stable set for  $k \geq 3\gamma - 1$ . Indeed, if  $k \geq 3\gamma - 1$ , then from any  $\mu \neq \mu^T$  there exists a horizon- $k$  improving path leading to  $\mu^T$ . In Example 1,  $\gamma = c_1^{\max} = 2$ . Take for instance the matching  $\mu_0 = \{(i_1, s_1), (i_2, s_2), (i_3, s_3)\}$ . We now construct a horizon- $k$  improving from  $\mu_0$  to  $\mu^T = \{(i_1, s_3), (i_2, s_1), (i_3, s_2)\} = \mu_5$  following the steps as in the proof of Theorem 1. First, we consider agents and objects belonging to the cycles in  $C_1$ . Notice that  $m_1^1 = \{(i_2, s_1), (i_3, s_2)\} \cap \mu_0 = \emptyset$ . Looking forward towards  $\mu_4$  and  $\mu^T$  (where  $\mu_4(i_3) = \mu^T(i_3)$ ), agent  $i_3$  matches to the object  $s_1$  that ranks her first on its priority list to reach the matching  $\mu_1 = \{(i_1, i_1), (i_2, s_2), (i_3, s_1)\}$  where agent  $i_3$  is matched to the object in  $c_1^1$  where she has priority. By doing so, agent  $i_1$  is left without object. In  $\mu_0$  (and  $\mu_1$ ), agent  $i_2$  is already matched to the object in  $c_1^1$  where she has priority.<sup>11</sup> Next, agent  $i_3$  leaves her object  $s_1$  to reach the matching  $\mu_2 = \{(i_1, i_1), (i_2, s_2), (i_3, i_3)\}$  where agent  $i_3$  is not assigned to any object. Agent  $i_3$  is temporarily worse off, but she anticipates to end up being matched with  $\mu^T(i_3)$ . Next, agent  $i_2$  matches to  $s_1$  that was left unassigned to reach the matching  $\mu_3 = \{(i_1, i_1), (i_2, s_1), (i_3, i_3)\}$ . Next, agent  $i_3$  matches to  $s_2$  that was left by  $i_2$  to reach the matching  $\mu_4 = \{(i_1, i_1), (i_2, s_1), (i_3, s_2)\}$  with  $m_1^1 = \{(i_2, s_1), (i_3, s_2)\} \subseteq \mu_4$ . Finally, we consider agents and objects belonging to the cycles in  $C_2$ . Since  $m_2^1 = \{(i_1, s_3)\} \cap \mu_4 = \emptyset$ , agent  $i_1$  is assigned to object  $s_3$  to form the match  $(i_1, s_3)$  and to reach the matching  $\mu_5 = \mu^T$ . Thus,  $\mu^T \in \phi_k(\mu_0)$ .

In Example 1,  $\mu^D = \{(i_1, s_3), (i_2, s_2), (i_3, s_1)\}$  is the matching obtained from the Deferred Acceptance (DA) algorithm,  $\mu^B = \{(i_1, s_1), (i_2, s_3), (i_3, s_2)\}$  is the matching obtained from the Immediate Acceptance (IA) algorithm (i.e. the Boston mechanism). Thus,  $\mu^T \neq \mu^D \neq \mu^B$ .

<sup>11</sup>Hence, it is sufficient that agent  $i_3$  looks forward at least 4 (instead of 5) steps ahead for having incentives to engage her first move towards  $\mu_4$ .

Let

$$\begin{aligned}\mu^1 &= \{(i_1, s_1), (i_2, i_2), (i_3, s_2)\}, \\ \mu^2 &= \{(i_1, s_1), (i_2, s_3), (i_3, s_2)\} = \mu^B, \\ \mu^3 &= \{(i_1, s_1), (i_2, s_2), (i_3, i_3)\}, \\ \mu^4 &= \{(i_1, s_1), (i_2, s_2), (i_3, s_3)\}, \\ \mu^5 &= \{(i_1, i_1), (i_2, s_2), (i_3, s_1)\}.\end{aligned}$$

Since agent  $i_1$  is as well off and agents  $i_2$  and  $i_3$  are strictly better off in  $\mu^T$  than in  $\mu^D$ , we have that there is no horizon- $k$  improving path from  $\mu^T$  to  $\mu^D$  for  $k \geq 5$ . That is,  $\mu^D \notin \phi_k(\mu^T)$  for  $k \geq 5$ . Hence,  $\{\mu^D\}$  is not a horizon- $k$  vNM stable set for  $k \geq 5$  since (ES) is violated.

Computing the horizon- $k$  improving paths emanating from  $\mu^T$  for  $k \geq 5$ , we get  $\phi_k(\mu^T) = \{\mu^1, \mu^2, \mu^3, \mu^4\}$ . Notice that  $\mu^5 \notin \phi_k(\mu^T)$  since agent  $i_1$  is worse off in  $\mu^5$  than in  $\mu^T$ . From  $\mu^1, \mu^2, \mu^3, \mu^4$  and  $\mu^5$ , there is a horizon- $k$  improving path to  $\mu^D$ . That is,  $\mu^D \in \phi_k(\mu)$  for  $\mu \in \{\mu^1, \mu^2, \mu^3, \mu^4, \mu^5\}$ . From  $\mu^D$  there is only a horizon- $k$  improving path to  $\mu^T$ ; i.e.  $\phi_k(\mu^D) = \{\mu^T\}$  for  $k \geq 5$ .

Thus, we have that, for  $k \geq 5$ , (i)  $\phi_k(\mu^D) = \{\mu^T\}$ , (ii)  $\mu^D \notin \phi_k(\mu^T)$ , and (iii)  $\mu^D \in \phi_k(\mu)$  for each  $\mu$  such that  $\mu \in \phi_k(\mu^T)$ . It follows then that for  $k \geq 5$ ,  $V = \{\mu^T\} = \{(i_1, s_3), (i_2, s_1), (i_3, s_2)\}$  is the unique horizon- $k$  vNM stable set. First, any  $V' \neq \{\mu^T\}$  such that  $\mu^T \in V'$  violates IS. Hence,  $\mu^T \notin V'$ . Second, any  $V' \not\supseteq \{\mu^T, \mu^D\}$  violates ES. Hence,  $\mu^D \in V'$ . Third, any  $V' \neq \{\mu^D\}$  such that  $\mu^D \in V', \mu^T \notin V'$  violates either IS or ES.

Thus, the DA matching  $\mu^D$  and the IA matching  $\mu^B$  do not belong to any horizon- $k$  vNM stable set for  $k \geq 5$ . Since the matching obtained from the IA algorithm is Pareto efficient, Example 1 also shows that there are priority-based matching problems where some Pareto efficient matching does not belong to any horizon- $k$  vNM stable set for  $k \geq 3\gamma - 1$ .

**Remark 1.** There are priority-based matching problems such that, for  $k \geq 3\gamma - 1$ ,

- (i) the matching obtained from the Deferred Acceptance (DA) algorithm does not belong to any horizon- $k$  vNM stable set;
- (ii) some Pareto efficient matching does not belong to any horizon- $k$  vNM stable set.

What happens if  $k$  becomes small? Computing the horizon- $k$  improving paths emanating from  $\mu^T$  for  $k \leq 4$  in Example 1, we get  $\phi_k(\mu^T) = \{\mu^1, \mu^2, \mu^3, \mu^4, \mu^5, \mu^D\}$ . So, there is now a horizon- $k$  improving path from  $\mu^T$  to  $\mu^D$ . In addition, from  $\mu^1, \mu^2, \mu^3, \mu^4$  and  $\mu^5$ , there is still a horizon- $k$  improving path to  $\mu^D$ . That is,  $\mu^D \in \phi_k(\mu)$  for  $\mu \in \{\mu^1, \mu^2, \mu^3, \mu^4, \mu^5, \mu^T\}$  for  $k \leq 4$ . From  $\mu^D$  there is no horizon- $k$  improving path for  $k \leq 2$ , but there is one for  $3 \leq k \leq 4$ ; i.e.  $\phi_k(\mu^D) = \emptyset$  for  $k \leq 2$  and  $\phi_k(\mu^D) = \{\mu^T\}$  for  $3 \leq k \leq 4$ . It follows then that for  $3 \leq k \leq 4$ , both  $V = \{\mu^D\}$  and  $V' = \{\mu^T\}$  are horizon- $k$  vNM stable sets. However, for  $k \leq 2$ ,  $V = \{\mu^D\}$  is the unique horizon- $k$  vNM stable set. In general, it holds that for any priority-based matching problem, the DA matching  $\mu^D$  belongs to all horizon-1 vNM stable sets.<sup>12</sup>

**Remark 2.** Let  $\langle I, S, P, F \rangle$  be a priority-based matching problem and  $\mu^D$  be the matching obtained from the Deferred Acceptance (DA) mechanism. The matching  $\mu^D$  belongs to all horizon-1 vNM stable sets.

## 6 Discussion

### 6.1 Tighter bound on $k$

We now look whether one could find a tighter bound on  $k$  such that for all  $k' \geq k$ , the singleton set  $\{\mu^T\}$  is a horizon- $k'$  vNM stable set. Consider the proof of Theorem 1. At the end of Step 1.1.A we reach the matching  $\mu'_{1,1} = \mu + \{(i, s) \mid i, s \in c_1^1 \text{ and } s \mapsto i\} - \{(i, \mu(i)) \mid i, s \in c_1^1, s \mapsto i \text{ and } \mu(i) \neq s\} - \{(\mu^{-1}(s), s) \mid i, s \in c_1^1, s \mapsto i \text{ and } \mu^{-1}(s) \neq i\}$  where each agent involved in  $c_1^1$  is matched to the object that ranks her first on its priority list. Remember that Step 1.1.A counts at most  $\bar{c}_1^1$  steps where  $\bar{c}_1^1 = \#\{i \in I \mid i \in c_1^1\}$  is the number of agents involved in cycle  $c_1^1$ . We slightly modify Step 1.1.B as follows.

Step 1.1.B Let  $\bar{\mathcal{I}}_1^1 = \{(i^l)\}_{l=1}^{\bar{c}_1^1}$  be an ordered set of the agents involved in cycle  $c_1^1$  such that  $i^l \in c_1^1$  and  $i^l$  matches before  $i^{l+1}$  to the object that ranks her first on its priority list in Step 1.1.A, for  $l = 1, \dots, \bar{c}_1^1 - 1$ . Agents who were already matched (at the beginning of Step 1.1.A) to the objects that rank them first on their priority lists occupy the first positions of  $\bar{\mathcal{I}}_1^1$ . From the matching

---

<sup>12</sup>Proposition 3 in Luo, Mauleon and Vannetelbosch (2021) provides a characterization of a horizon-1 vNM stable set of networks. Since matchings are a subclass of networks and the DA matching is stable, it follows that the DA matching belongs to all horizon-1 vNM stable sets.

$\mu'_{1,1}$ , agents  $i^1$  to  $i^{\bar{c}_1-1}$  successively leave their objects to reach the matching  $\mu''_{1,1} = \mu'_{1,1} - \{(i, s) \mid i, s \in c_1^1, s \mapsto i \text{ and } i \neq i^{\bar{c}_1}\}$  where only agent  $i^{\bar{c}_1}$  is still matched to the object that ranks her first on its priority list. Step 1.1.B counts at most  $\bar{c}_1^1 - 1$  steps.

Step 1.1.C. From the matching  $\mu''_{1,1}$ , agent  $i^{\bar{c}_1}$  first matches with her top choice object  $s$  to reach the matching  $\mu''_{1,1} + (i^{\bar{c}_1}, s) - (i^{\bar{c}_1}, \mu''_{1,1}(i^{\bar{c}_1}))$  where  $s$  is such that  $i \mapsto s$  and  $i, s \in c_1^1$ . Notice that  $s$  was unassigned at  $\mu''_{1,1}$  while the object that ranks  $i^{\bar{c}_1}$  first on its priority list, i.e.  $\mu''_{1,1}(i^{\bar{c}_1})$ , is now unassigned. Next, agents  $i^1$  to  $i^{\bar{c}_1-1}$  successively match to their top choice object to reach the matching  $\mu'''_{1,1} = \mu''_{1,1} - \{(i, s) \mid i, s \in c_1^1 \text{ and } s \mapsto i\} + \{(i, s) \mid i, s \in c_1^1 \text{ and } i \mapsto s\}$ . Step 1.1.C counts at most  $\bar{c}_1^1$  steps. We have reached  $\mu'''_{1,1}$  with  $m_1^1 \subseteq \mu'''_{1,1}$  and so agents belonging to  $c_1^1$  are assigned to the same object as in  $\mu^T$ . If  $1 \neq L_1$ , then go to Step 1.2. Otherwise, go to Step 1.End with  $\mu'''_{1,L_1} = \mu'''_{1,1}$ .

Step 1.1 counts at most  $3\bar{c}_1^1 - 1$  steps. Since we use now the order in which the agents are matched to the objects that rank them first on their priority lists in Step 1.1.A, there is at most  $2\bar{c}_1^1 + 1$  steps between the first move of an agent in Step 1.1.A and her final move in Step 1.1.C. Hence, it becomes sufficient that the agents who move in Step 1.1 look forward  $2\bar{c}_1^1 + 1$  steps ahead to have incentives for engaging the move towards the matching  $\mu'''_{1,1}$  where they already get the object assigned by the TTC. Once they reach those matches they do not move afterwards.

Thus, the lower bound  $\underline{k} = 2\gamma + 1$  is a tighter bound on  $k$  such that for all  $k' \geq \underline{k}$ , the singleton set  $\{\mu^T\}$  is a horizon- $k'$  vNM stable set. However, it relies on improving paths that require much more coordination on behalf of the agents than the ones associated to the lower bound of Theorem 1, i.e.  $3\gamma - 1$ .

## 6.2 Farsighted stable set

By simply replacing  $\mu_{\min\{l+k, L\}}(i)P_i\mu_l(i)$  by  $\mu_L(i)P_i\mu_l(i)$  in the definition of a horizon- $k$  improving path, we obtain the definition of a farsighted improving path. Let  $\phi_\infty(\mu)$  be the set of matchings that can be reached by means of a farsighted improving path emanating from  $\mu$ . Given the number of possible matchings is finite, there exists  $k^*$  such that for all  $k \geq k^*$ ,  $\phi_k(\mu) = \phi_{k+1}(\mu)$ , and so  $\phi_{k^*}(\mu) = \phi_\infty(\mu)$ .<sup>13</sup>

<sup>13</sup>Mauleon, Vannetelbosch and Vergote (2011) define and characterize the vNM farsighted stable set for two-sided matching problems. Doğan and Ehlers (2023) show the existence of myopic-

**Definition 3.** Let  $\langle I, S, P, F \rangle$  be a priority-based matching problem. A set of matchings  $V \subseteq \mathcal{M}$  is a vNM farsighted stable set if it satisfies:

- (i) Internal stability (IS): For every  $\mu, \mu' \in V$ , it holds that  $\mu' \notin \phi_\infty(\mu)$ .
- (ii) External stability (ES): For every  $\mu \in \mathcal{M} \setminus V$ , it holds that  $\phi_\infty(\mu) \cap V \neq \emptyset$ .

**Corollary 1.** Let  $\langle I, S, P, F \rangle$  be a priority-based matching problem and  $\mu^T$  be the matching obtained from the Top Trading Cycles mechanism. The singleton set  $\{\mu^T\}$  is a vNM farsighted stable set.<sup>14</sup>

### 6.3 Horizon- $L$ farsighted set

An alternative concept for limited farsightedness is obtained by adapting Herings, Mauleon and Vannetelbosch (2019) definition of a horizon- $L$  farsighted set of networks to priority-based matching problems. A horizon- $L$  farsighted set of matchings has to satisfy three conditions: (i) external deviations should be horizon- $L$  deterred, (ii) from any matching outside the set there is a sequence of farsighted improving paths of length smaller than or equal to  $L$  leading to some matching in the set, (iii) there is no proper subset satisfying the first two conditions.

**Definition 4.** Let  $\langle I, S, P, F \rangle$  be a priority-based matching problem. A farsighted improving path of length  $L$  from a matching  $\mu \in \mathcal{M}$  to a matching  $\mu' \in \mathcal{M} \setminus \{\mu\}$  is a finite sequence of distinct matchings  $\mu_0, \dots, \mu_L$  with  $\mu_0 = \mu$  and  $\mu_L = \mu'$  such that for every  $l \in \{0, \dots, L-1\}$  either

- (i)  $\mu_{l+1} = \mu_l - (i, s)$  for some  $(i, s) \in I \times S$  such that  $\mu_L(i)P_i\mu_l(i)$ , or
- (ii)  $\mu_{l+1} = \mu_l + (i, s) - \{(i, \mu_l(i)) \mid \mu_l(i) \neq i\} - \{(\mu_l^{-1}(s), s) \mid \#(\mu_l^{-1}(s)) = 1\}$   
for some  $(i, s) \in I \times S$  such that  $\mu_L(i)P_i\mu_l(i)$  and  $F_s(i) < F_s(\mu_l^{-1}(s))$  if  $\#(\mu_l^{-1}(s)) = 1$ .

---

farsighted stable sets for matching problems where there are farsighted agents only on one side of the market, while there may be myopic agents on both sides.

<sup>14</sup>This result is robust to the incorporation of various forms of maximality in the definition of farsighted improving path, like the strong rational expectations farsighted stable set in Dutta and Vohra (2017) and absolute maximality as in Ray and Vohra (2019). See also Herings, Mauleon and Vannetelbosch (2020). In addition, Theorem 1 and Corollary 1 still hold for priority-based matching problems with multiple copies and fixed tie-breaking rules as in Ehlers (2014).

If there exists a farsighted improving path of length  $L$  from  $\mu$  to  $\mu'$ , then we write  $\mu \longrightarrow_L \mu'$ . For a given matching  $\mu$  and some  $L' \geq 1$ , let  $\widehat{\phi}_{L'}(\mu)$  be the set of matchings that can be reached from  $\mu$  by a farsighted improving path of length  $L \leq L'$ . That is,  $\widehat{\phi}_{L'}(\mu) = \{\mu' \in \mathcal{M} \mid \exists L \leq L' \text{ such that } \mu \longrightarrow_L \mu'\}$ . Let  $\widehat{\phi}_\infty(\mu) = \{\mu' \in \mathcal{M} \mid \exists L \in \mathbb{N} \text{ such that } \mu \longrightarrow_L \mu'\} = \phi_\infty(\mu)$  denote the set of matchings that can be reached from  $\mu$  by some farsighted improving path. From Lemma 1 in Herings, Mauleon and Vannetelbosch (2019) we have that for every  $L \geq 1$ , for every  $\mu \in \mathcal{M}$ , it holds that  $\widehat{\phi}_L(\mu) \subseteq \widehat{\phi}_{L+1}(\mu)$ , and that for  $L \geq k^*$ , for every  $\mu \in \mathcal{M}$ , it holds that  $\widehat{\phi}_L(\mu) = \widehat{\phi}_{L+1}(\mu) = \widehat{\phi}_\infty(\mu) = \phi_\infty(\mu)$ .

The set  $\widehat{\phi}_L^2(\mu) = \widehat{\phi}_L(\widehat{\phi}_L(\mu)) = \{\mu'' \in \mathcal{M} \mid \exists \mu' \in \widehat{\phi}_L(\mu) \text{ such that } \mu'' \in \widehat{\phi}_L(\mu')\}$  consists of those matchings that can be reached by a composition of two farsighted improving paths of length at most  $L$  from  $\mu$ . For  $m \in \mathbb{N}$ , let  $\widehat{\phi}_L^m(\mu)$  be the matchings that can be reached from  $\mu$  by means of  $m$  compositions of farsighted improving paths of length at most  $L$ . Let  $\widehat{\phi}_L^\infty$  denote the set of matchings that can be reached from  $\mu$  by means of any number of compositions of farsighted improving paths of length at most  $L$ .<sup>15</sup>

The notion of a horizon- $L$  farsighted set is based on two main requirements: horizon- $L$  deterrence of external deviations and horizon- $L$  external stability. A set of matchings  $V$  satisfies horizon- $L$  deterrence of external deviations if all possible deviations from any matching  $\mu \in V$  to a matching outside  $V$  are deterred by a threat of ending worse off or equally well off.<sup>16</sup>

**Definition 5.** For  $L \geq 1$ , a set of matchings  $V \subseteq \mathcal{M}$  satisfies horizon- $L$  deterrence of external deviations if for every  $\mu \in V$ ,

(i)  $\forall (i, s) \notin \mu$  such that  $\tilde{\mu} = \mu + (i, s) - \{(i, \mu(i)) \mid \mu(i) \neq i\} - \{(\mu^{-1}(s), s) \mid \#(\mu^{-1}(s)) = 1\}$  and  $\tilde{\mu} \notin V$ , either there exists  $\mu' \in [\widehat{\phi}_{L-2}(\tilde{\mu}) \cap V] \cup [\widehat{\phi}_{L-1}(\tilde{\mu}) \setminus \widehat{\phi}_{L-2}(\tilde{\mu})]$  such that  $\mu R_i \mu'$ , or  $F_s(i) > F_s(\mu^{-1}(s))$  if  $\#(\mu^{-1}(s)) = 1$ ,

(ii)  $\forall (i, s) \in \mu$  such that  $\tilde{\mu} = \mu - (i, s)$  and  $\tilde{\mu} \notin V$ , there exists  $\mu' \in [\widehat{\phi}_{L-2}(\tilde{\mu}) \cap V] \cup [\widehat{\phi}_{L-1}(\tilde{\mu}) \setminus \widehat{\phi}_{L-2}(\tilde{\mu})]$  such that  $\mu R_i \mu'$ .

<sup>15</sup>From Lemma 2 in Herings, Mauleon and Vannetelbosch (2019) we have that for every  $L \geq 1$ , for every  $\mu \in \mathcal{M}$ , it holds that  $\widehat{\phi}_L^\infty(\mu) \subseteq \widehat{\phi}_{L+1}^\infty(\mu)$ , and that for  $L \geq k^*$ , for every  $\mu \in \mathcal{M}$ , it holds that  $\widehat{\phi}_L^\infty(\mu) = \widehat{\phi}_{L+1}^\infty(\mu) = \widehat{\phi}_\infty(\mu)$ .

<sup>16</sup>We use the notational convention that  $\widehat{\phi}_{-1}(\mu) = \emptyset$  for every  $\mu \in \mathcal{M}$ .

Condition (i) in Definition 5 captures that forming the match  $(i, s)$  at  $\mu \in V$  and reaching a matching  $\tilde{\mu}$  outside of  $V$ , is deterred by the threat of ending in  $\mu'$ . Here  $\mu'$  is such that either there is a farsighted improving path of length smaller than or equal to  $L - 2$  from  $\tilde{\mu}$  to  $\mu'$  and  $\mu'$  belongs to  $V$  or there is a farsighted improving path of length equal to  $L - 1$  from  $\tilde{\mu}$  to  $\mu'$  and there is no farsighted improving path from  $\tilde{\mu}$  to  $\mu'$  of smaller length.<sup>17</sup>

A set of matchings  $V$  satisfies horizon- $L$  external stability if from any matching outside of  $V$  there is a sequence of farsighted improving paths of length smaller than or equal to  $L$  leading to some matching in  $V$ .

**Definition 6.** For  $L \geq 1$ , a set of matchings  $V \subseteq \mathcal{M}$  satisfies horizon- $L$  external stability if for every  $\mu' \in \mathcal{M} \setminus V$ ,  $\widehat{\phi}_L^\infty(\mu') \cap V \neq \emptyset$ .

**Definition 7.** For  $L \geq 1$ , a set of matchings  $V \subseteq \mathcal{M}$  is a horizon- $L$  farsighted set if it is a minimal set satisfying horizon- $L$  deterrence of external deviations and horizon- $L$  external stability.

From Herings, Mauleon, and Vannetelbosch (2019) we have that a horizon- $L$  farsighted set of matchings exists.

**Theorem 2.** Let  $\langle I, S, P, F \rangle$  be a priority-based matching problem and  $\mu^T$  be the matching obtained from the Top Trading Cycles mechanism. The singleton set  $\{\mu^T\}$  is a horizon- $L$  farsighted set for  $L \geq 3\gamma$ .

*Proof.* Take  $L \geq 3\gamma$ . We show that  $\{\mu^T\}$  is a horizon- $L$  farsighted set. First,  $\{\mu^T\}$  is a minimal set. Second,  $\{\mu^T\}$  satisfies horizon- $L$  deterrence of external deviations. Any deviation from  $\mu^T$  to  $\tilde{\mu} = \mu^T - (i, s)$  is deterred since agent  $i$  is worse off at  $\tilde{\mu}$  where  $\tilde{\mu}(i) = i$ . An agent  $i$  may only have incentives to deviate from  $\mu^T$  to  $\tilde{\mu} = \mu^T + (i, s) - \{(i, \mu(i)) \mid \mu(i) \neq i\} - \{(\mu^{-1}(s), s) \mid \#(\mu^{-1}(s)) = 1\}$  if she matches to an object  $s$  that was assigned in  $\mu^T$  to some agent  $j$  who belongs to a cycle formed before agent  $i$ 's cycle in the TTC algorithm. Hence, any deviation from  $\mu^T$

---

<sup>17</sup>We distinguish farsighted improving paths of length less than or equal to  $L - 2$  after a deviation from  $\mu$  to  $\tilde{\mu}$  and farsighted improving paths of length equal to  $L - 1$ . In the former case, the reasoning capacity of the agent is not yet reached, and the threat of ending in  $\mu'$  is only credible if it belongs to the set  $V$ . In the latter case, the only way to reach  $\mu'$  from  $\mu$  requires  $L$  steps of reasoning or even more (i.e. one step in the deviation to  $\tilde{\mu}$  and at least  $L - 1$  additional steps in any farsighted improving path from  $\tilde{\mu}$  to  $\mu'$ ). Since this exhausts the reasoning capacity of the agent, the threat of ending in  $\mu'$  is credible, irrespective of whether it belongs to  $V$  or not.

to  $\tilde{\mu} = \mu^T + (i, s) - \{(i, \mu(i)) \mid \mu(i) \neq i\} - \{(\mu^{-1}(s), s) \mid \#(\mu^{-1}(s)) = 1\}$  can be deterred since  $\mu^T - (i, \mu^T(i)) \in \widehat{\phi}_{L-1}(\tilde{\mu})$ . Indeed, from  $\tilde{\mu}$  the agents belonging to agent  $j$ 's cycle in the TTC algorithm can simply follow the steps of Theorem 1's proof to reach  $\mu^T - (i, \mu^T(i))$ , and this farsighted improving path counts at most  $3\gamma - 1$  moves.<sup>18</sup> Third, the horizon- $k$  improving path from Theorem 1's proof can be decomposed in a succession of farsighted improving paths of length smaller than or equal to  $3\gamma - 1$  where each farsighted improving path consists of the formation of the matches between the agents belonging to the same cycle in the TTC algorithm. Hence, for every  $\mu \in \mathcal{M} \setminus \{\mu^T\}$ ,  $\mu^T \in \widehat{\phi}_L^\infty(\mu)$ , and so  $\{\mu^T\}$  satisfies horizon- $L$  external stability.  $\square$

## 6.4 The agents own the objects

Closely related to priority-based matching problems are matching problems where the agents own the object. Let  $\langle I, S, P, F \rangle$  be a matching problem where each agent  $i$  owns an object  $s$ . The strict priority structure  $F$  of the objects over the agents is such that the priority  $F_s$  of object  $s$  only ranks the owner of object  $s$ . Without loss of generality, let agent  $i_l$  be the owner of object  $s_l$ , for  $l = 1, \dots, n$ . Let  $i_s$  be a generic agent who owns object  $s$ .

**Example 2.** Consider a matching problem  $\langle I, S, P, F \rangle$  with  $I = \{i_1, i_2, i_3\}$  and  $S = \{s_1, s_2, s_3\}$ , and where agent  $i_l$  owns object  $s_l$ , for  $l = 1, 2, 3$ . Agents' preferences and endowments are as follows.

Agents				Objects		
Endowment	$s_1$	$s_2$	$s_3$	$F_{s_1}$	$F_{s_2}$	$F_{s_3}$
$P_{i_1}$	$P_{i_2}$	$P_{i_3}$		$i_1$	$i_2$	$i_3$
	$s_3$	$s_3$	$s_2$			
	$s_1$	$s_2$	$s_3$			
	$s_2$	$s_1$	$s_1$			

In Example 2,  $\mu^T = \{(i_1, s_1), (i_2, s_3), (i_3, s_2)\}$  is the matching obtained from the TTC algorithm. In the first round of the TTC algorithm, there is one cycle where agent  $i_2$  points to object  $s_3$ , object  $s_3$  points its owner  $i_3$ , agent  $i_3$  points to object  $s_2$  and object  $s_2$  points its owner  $i_2$ . That is,  $C_1 = \{c_1^1\}$  with  $c_1^1 =$

<sup>18</sup>For  $L > 3\gamma$ ,  $\mu^T \in \widehat{\phi}_{L-1}(\tilde{\mu})$  since agent  $i$  has incentives to match to  $s = \mu^T(i)$  at  $\mu^T - (i, \mu^T(i))$ .



$(s_3, i_3, s_2, i_2)$ . Agent  $i_2$  is assigned to object  $s_3$  and agent  $i_3$  is assigned to object  $s_2$ :  $m_1^1 = \{(i_2, s_3), (i_3, s_2)\}$ , and so  $i_2$  and  $i_3$  exchange their objects. In the second round of the TTC algorithm, there is only one leftover agent,  $i_1$ , who points to object  $s_1$  that she owns and one leftover object,  $s_1$ , that points to its owner  $i_1$ . That is,  $C_2 = \{c_2^1\}$  with  $c_2^1 = (s_1, i_1)$ . Agent  $i_1$  is assigned to her own object  $s_1$ :  $m_2^1 = \{(i_1, s_1)\}$ , and so  $\mu^T = m_1^1 \cup m_2^1$ .

Roth and Postlewaite (1977) show that, for any matching problem  $\langle I, S, P, F \rangle$  where each agent  $i$  owns an object  $s$ , there is always a unique matching that is in the core. Moreover, this matching can be obtained with the TTC algorithm.<sup>19</sup>

**Definition 8.** Let  $\langle I, S, P, F \rangle$  be a matching problem where each agent  $i$  owns an object  $s$ . A horizon- $k$  improving path from a matching  $\mu \in \mathcal{M}$  to a matching  $\mu' \in \mathcal{M} \setminus \{\mu\}$  is a finite sequence of distinct matchings  $\mu_0, \dots, \mu_L$  with  $\mu_0 = \mu$  and  $\mu_L = \mu'$  such that for every  $l \in \{0, \dots, L-1\}$  either

- (i)  $\mu_{l+1} = \mu_l - (i, s)$  for some  $(i, s) \in I \times S$  such that  $\mu_{\min\{l+k, L\}}(i) P_i \mu_l(i)$ , or
- (ii)  $\mu_{l+1} = \mu_l + (i, s) - \{(i, \mu_l(i)) \mid \mu_l(i) \neq i\} - \{(\mu_l^{-1}(s), s) \mid \#(\mu_l^{-1}(s)) = 1\}$  for some  $(i, s) \in I \times S$  such that  $\mu_{\min\{l+k, L\}}(i) P_i \mu_l(i)$  and  $\mu_{\min\{l+k, L\}}(i_s) P_{i_s} \mu_l(i_s)$ .

In the case of matching problems where each agent owns an object, we still require that, along a horizon- $k$  improving path, each time some agent  $i$  is on the move she is better off at the match she will get  $k$ -steps ahead on the sequence compared to her current match. Moreover, if agent  $i$  matches to  $s$ , we also require that the owner of the object (i.e.  $i_s$ ) prefers the match he will get  $k$ -steps ahead compared to his current match. In other words, the owner of the object has a word to say about the assignment of his endowment to some agent.

The set of matchings  $\mu' \in \mathcal{M}$  such that there is a horizon- $k$  improving path from  $\mu$  to  $\mu'$  is denoted by  $\tilde{\phi}_k(\mu)$ , so  $\tilde{\phi}_k(\mu) = \{\mu' \in \mathcal{M} \mid \mu \rightarrow_k \mu'\}$ . Replacing  $\phi_k(\mu)$  by  $\tilde{\phi}_k(\mu)$  in Definition 2 we obtain the definition of a horizon- $k$  vNM stable set for matching problems where each agent owns an object.

**Theorem 3.** Let  $\langle I, S, P, F \rangle$  be a matching problem where each agent  $i$  owns an object  $s$  and  $\mu^T$  is the matching obtained from the Top Trading Cycles mechanism. The singleton set  $\{\mu^T\}$  is the unique horizon- $k$  vNM stable set for  $k \geq 3\gamma - 1$ .

<sup>19</sup>For matching problems with private endowments, Ma (1994) shows that a mechanism is strategy-proof, Pareto efficient and individually rational if and only if it uses the TTC algorithm.

*Proof.* From the proof of Theorem 1 it follows that  $\{\mu^T\}$  is a horizon- $k$  vNM stable set for  $k \geq 3\gamma - 1$ .<sup>20</sup> We now show that  $\{\mu^T\}$  is the unique horizon- $k$  vNM stable set for  $k \geq 3\gamma - 1$  since  $\tilde{\phi}_k(\mu^T) = \emptyset$ . Consider any cycle that is obtained in the first step of the TTC algorithm. All the agents involved in this cycle obtain their most preferred object in  $\mu^T$  and, in this cycle, any agent obtains the endowment of another agent who is also in the cycle. Hence, from  $\mu^T$  they will never engage a move towards another matching. Consider now any cycle that is obtained in the second step of the TTC algorithm. Taking as fixed the matches done in  $\mu^T$  by all agents involved in any cycle of the first step of the TTC, all the agents involved in this cycle of the second step of the TTC obtain their most preferred object in  $\mu^T$  and, in this cycle, any agent obtains the endowment of another agent who is also in the cycle. Knowing that agents from any cycle of the first step of the TTC will never engage a move, agents from any cycle of the second step of the TTC will never engage either a move from  $\mu^T$  towards another matching. Repeating this argument with the matches found in steps 3, 4,  $\dots$  of the TTC leads to the conclusion that  $\tilde{\phi}_k(\mu^T) = \emptyset$ . Hence, any set  $V \neq \{\mu^T\}$  would violate (ES) or (IS), and so  $\{\mu^T\}$  is the unique horizon- $k$  vNM stable set for  $k \geq 3\gamma - 1$ .  $\square$

**Corollary 2.** *Let  $\langle I, S, P, F \rangle$  be a matching problem where each agent  $i$  owns an object  $s$  and  $\mu^T$  is the matching obtained from the Top Trading Cycles mechanism. The singleton set  $\{\mu^T\}$  is the unique vNM farsighted stable set.<sup>21</sup>*

## 6.5 Conclusion

We have considered priority-based matching problems with limited farsightedness. We have shown that, once agents are sufficiently farsighted, the matching obtained from the TTC algorithm becomes stable: a singleton set consisting of the TTC matching is a horizon- $k$  vNM stable set if the degree of farsightedness is greater

---

<sup>20</sup>Notice that along the horizon- $k$  improving path from the proof of Theorem 1, an agent matches either to an unassigned object whose owner is unmatched or to an object that she is the owner. Hence, the owner of the object does not block her move towards the TTC matching  $\mu^T$ .

<sup>21</sup>In an exchange economy with indivisible goods of Shapley and Scarf (1974), Kawasaki (2010) as well as Klaus, Klijn and Walzl (2010) show that there exists a unique vNM farsighted stable set, which coincides with the set of competitive allocations. Thus, they obtain a similar result to Corollary 2 except that they allow for coalitional moves while agents can only move one at a time in our definition of vNM farsighted stable set.

than three times the number of agents in the largest cycle of the TTC. On the contrary, the matching obtained from the DA algorithm may not belong to any horizon- $k$  vNM stable set for  $k$  large enough.

Hence, the TTC mechanism satisfies Pareto efficiency, strategy-proofness and (limited) farsighted stability. Notice that a mechanism is strategy-proof if no agent has incentives to misrepresent her preferences anticipating perfectly the outcome of the TTC algorithm. So, strategy-proofness implicitly presumes some degree of farsightedness on behalf of the agents. Thus, it seems more consistent to look for a mechanism that satisfies strategy-proofness together with (limited) farsighted stability.

## Acknowledgments

Ana Mauleon and Vincent Vannetelbosch are, respectively, Research Director and Senior Research Associate of the National Fund for Scientific Research (FNRS). Financial support from the Fonds de la Recherche Scientifique - FNRS PDR research grant T.0143.18 is gratefully acknowledged. Ata Atay is a Serra Hünter Fellow (Professor Lector Serra Hünter). Ata Atay gratefully acknowledges the support from the Spanish Ministerio de Ciencia e Innovación research grant PID2020-113110GB-100/AEI/10.13039/501100011033, from the Generalitat de Catalunya research grant 2021SGR00306 and from the University of Barcelona research grant AS017672.

## References

- [1] Abdulkadiroğlu, A., Y.K. Che, P. A. Pathak, A.E. Roth, and O. Tercieux (2020), “Efficiency, justified envy, and incentives in priority-based matching”, *American Economic Review Insights* 2, 425-42.
- [2] Abdulkadiroğlu, A., and T. Sönmez (2003), “School choice: A mechanism design approach”, *American Economic Review* 93, 729-747.
- [3] Atay, A., A. Mauleon, and V. Vannetelbosch (2022), “School choice with farsighted students”, CORE/LIDAM Discussion Paper 2022/25, UCLouvain, Louvain-la-Neuve, Belgium.

- [4] Banker, R.D., and R.J. Kauffman (2004), “50th Anniversary article: The evolution of research on information systems: A fiftieth-year survey of the literature in management science”, *Management Science* 50, 281-298.
- [5] Bichler, M., A. Hammerl, T. Morrill, and S. Waldherr (2021), “How to assign scarce resources without money: Designing information systems that are efficient, truthful, and (pretty) fair”, *Information Systems Research* 32(2), 335-355.
- [6] Che, Y.K., and O. Tercieux (2019), “Efficiency and stability in large matching markets”, *Journal of Political Economy* 127, 2301-2342.
- [7] Chwe, M. S.-Y. (1994), “Farsighted coalitional stability”, *Journal of Economic Theory* 63, 299-325.
- [8] Doğan, B., and L. Ehlers (2021), “Minimally unstable Pareto improvements over deferred acceptance”, *Theoretical Economics* 16, 1249-1279.
- [9] Doğan, B., and L. Ehlers (2022), “Robust minimal instability of the top trading cycles mechanism”, *American Economic Journal: Microeconomics* 14, 556-582.
- [10] Doğan, B., and L. Ehlers (2023), “Existence of myopic-farsighted stable sets in matching markets”, available at SSRN:<http://dx.doi.org/10.2139/ssrn.4354768>
- [11] Dubins, L.E., and D.A. Freedman (1981), “Machiavelli and the Gale-Shapley algorithm”, *American Mathematical Monthly* 88, 485-494.
- [12] Dutta, B., and R. Vohra (2017), “Rational expectations and farsighted stability”, *Theoretical Economics* 12, 1191-1227.
- [13] Ehlers, L. (2007), “Von Neumann-Morgenstern stable sets in matching problems”, *Journal of Economic Theory* 134, 537-547.
- [14] Ehlers, L. (2008), “Truncation strategies in matching markets”, *Mathematics of Operations Research* 33, 327-335.
- [15] Ehlers, L. (2014), “Top trading with fixed tie-breaking in markets with indivisible goods”, *Journal of Economic Theory* 151, 64-87.
- [16] Gale, D., and L.S. Shapley (1962), “College admissions and the stability of marriage”, *American Mathematical Monthly* 69, 9-15.

- [17] Granot, D., and G. Sošić (2005), “Formation of alliances in internet-based supply exchanges”, *Management Science* 51(1), 92-105.
- [18] Haeringer, G. (2017), Market design: auctions and matching, MIT Press, Cambridge, MA.
- [19] Hakimov, R., and O. Kesten (2018), “The equitable top trading cycles mechanism for school choice”, *International Economic Review* 59, 2219-2258.
- [20] Herings, P.J.J., A. Mauleon, and V. Vannetelbosch (2017), “Stable sets in matching problems with coalitional sovereignty and path dominance”, *Journal of Mathematical Economics* 71, 14-19.
- [21] Herings, P.J.J., A. Mauleon, and V. Vannetelbosch (2019), “Stability of networks under horizon- $K$  farsightedness”, *Economic Theory* 68, 177-201.
- [22] Herings, P.J.J., A. Mauleon, and V. Vannetelbosch (2020), “Matching with myopic and farsighted players”, *Journal of Economic Theory* 190, 105125.
- [23] Kawasaki, R. (2010), “Farsighted stability of the competitive allocations in an exchange economy with indivisible goods”, *Mathematical Social Sciences* 9 (2010) 46-52.
- [24] Kesten, O (2006), “On two competing mechanisms for priority-based allocation problems ”, *Journal of Economic Theory* 127, 155-171.
- [25] Kesten, O (2010), “School choice with consent ”, *The Quarterly Journal of Economics* 125, 1297-1348.
- [26] Kirchsteiger, G., M. Mantovani, A. Mauleon, and V. Vannetelbosch (2016), “Limited farsightedness in network formation”, *Journal of Economic Behavior and Organization* 128, 97-120.
- [27] Klaus, B., F. Klijn, and M. Walzl (2010), “Farsighted house allocation”, *Journal of Mathematical Economics* 46, 817-824.
- [28] Luo, C., A. Mauleon, and V. Vannetelbosch (2021), “Network formation with myopic and farsighted players”, *Economic Theory* 71, 1283-1317.
- [29] Ma, J. (1994), “Strategy-proofness and the strict core in a market with indivisibilities ”, *International Journal of Game Theory* 23, 75-83.

- [30] Mauleon, A., V. Vannetelbosch and W. Vergote (2011), “Von Neumann - Morgenstern farsightedly stable sets in two-sided matching”, *Theoretical Economics* 6, 499-521.
- [31] Morrill, T. (2015), “Two simple variations of top trading cycles”, *Economic Theory* 60, 123-140.
- [32] Nagarajan, M., and G. Sošić (2007), “Stable farsighted coalitions in competitive markets”, *Management Science* 53(1), 29-45.
- [33] Ray, D. and R. Vohra (2015), “The farsighted stable set”, *Econometrica* 83, 977-1011.
- [34] Ray, D., and R. Vohra (2019), “Maximality in the farsighted stable set”, *Econometrica* 87, 1763-1779.
- [35] Reny, P.J. (2022), “Efficient matching in the school choice problem”, *American Economic Review* 112, 2025-43.
- [36] Roth, A.E. (1982), “The economics of matching: stability and incentives”, *Mathematics of Operations Research* 7, 617-628.
- [37] Roth, A.E. and A. Postlewaite (1977), “Weak versus strong domination in a market with indivisible goods”, *Journal of Mathematical Economics* 4, 131-137.
- [38] Roth, A.E. and M.A.O. Sotomayor (1990), Two-sided matching, a study in game-theoretic modeling and analysis, Econometric Society Monographs No.18, Cambridge University Press, Cambridge, UK.
- [39] Shapley, L.S., and H. Scarf (1974), “On cores and indivisibility”, *Journal of Mathematical Economics* 1, 23-37.
- [40] Shi, P. (2022), “Optimal priority-based allocation mechanisms”, *Management Science* 68(1), 171-188.
- [41] Sošić, G. (2006), “Transshipment of inventories among retailers: Myopic vs. farsighted stability”, *Management Science* 52(10), 1493-1508.