Top Trading Cycles in Endogenous Information Acquisition

Annika Johnson*

11 April 2017

Abstract
Consider a housing problem in which each agent arrives at the market with an endowment but is unsure of the value of others’ objects and is unwilling to exchange without learning more. An individually rational, Pareto optimal and strategyproof exchange requires Gale’s Top Trading Cycles but the ability to investigate others’ endowment must also be introduced. For the instance in which each agent has only the resources to learn about one other object, I show how the agents’ decisions over what to learn about impact the nature of the cycles that could form. Large cycles are risky so no cycle containing more than two agents can exist in equilibrium. Any set of cycles which is stable will also yield the maximum ex-ante welfare in equilibrium. Furthermore, when objects are ex-ante non-identical, the unique set of cycles which maximise ex-ante welfare in equilibrium is identical to the unique set of stable cycles.

1 Introduction
Consider a unilateral matching problem in which each agent is endowed with an indivisible object that may be exchanged with other such agents. Upon arriving at the market, each agent is unsure about the value of other objects. It is infeasible for each agent to investigate every object but without investigation each agent prefers to keep his own endowment. What results is a housing problem with an additional stage of endogenous information acquisition. Indeed the problem could be demonstrated in terms of council house exchange programmes or college dormitory swapping policy where investigation is a precarious and time-intensive process. However, the need for an endogenous acquisition stage becomes readily apparent if one uses the terminology of one of the most

*Annika.Johnson.2007@live.rhul.ac.uk; Royal Holloway, University of London. This work is supported by the Economic and Social Research Council [grant number ES/J500148/1].
prominent applications: kidney exchange (Roth et al. (2004)).

When a person requires a new kidney three cases may arise. It might be that the patient has no donor at all and remains on the waiting list for a cadaver organ or that they are successful in finding a compatible living donor organ. This model is applicable in the final case in which each patient can find only an incompatible living donor. An obvious candidate mechanism for organising such an exchange is Gale’s Top Trading Cycles (GTT) (Shapley and Scarf (1974)). In this environment it is the unique, strategy proof, Pareto optimal and individually rational mechanism (Ma (1994), Roth (1982)). It is clear in this application however, that before any exchange can take place, the organs must first be tested. A kidney can never be safely transplanted without first confirming its compatibility with the patient. This requirement means that the conventional GTT mechanism must be modified to account for a stage of endogenous information acquisition. Without such a stage, organised matches can easily fail. An analysis of the deterministic mechanism used by the United Network for Organ Sharing (UNOS) (Dickerson et al. (2013)) found that only 7% of matches ever resulted in a completed transplant. Since not all costly testing procedures were performed upfront, 16% of matches failed explicitly because the organ was ultimately found to be incompatible. It then seems advisable to incorporate a learning stage into the mechanism in such a way that it reduces the incidence of failed matches.

In this paper I introduce an opportunity for learning prior to the GTT matching stage and analyse its impact on equilibrium and what proves to be the mutually compatible objectives of stability and ex-ante welfare. In particular, across two different environments I allow each agent to learn the value of one object other than their own endowment. In the first environment, all objects are ex-ante identical and agents have common, but limited information about others’ objects. This is a simpler environment but it suffices to illustrate many of the key ideas. In the second, objects are ex-ante non-identical and though each agent cannot be certain which objects they will find acceptable, they are aware that some objects have a higher ex-ante value than others. For example, in the case of kidney exchange the patient is aware that one kidney comes from a healthy twenty year old while another is classified as coming from extended donor criteria and is an older, less reliable kidney that is likely to fail sooner. After learning, agents form their ex-post preferences and are matched via the GTT mechanism which even under this new domain, retains the properties of individual rationality, Pareto optimality and strategyproofness. In order to analyse the effects of the endogenous information acquisition, I examine the problem in the context of learning cycles. These are the cycles which
could form under GTT as a result of the agents’ learning decisions.

In Theorems 1 and 2, I characterise the large set of equilibria and show that in both the identical and the non-identical object environments, the largest cycle that exists in equilibrium contains only two agents. This arises from the increased risk of failure that is inherent in large cycles before learning. Since an agent will find an object acceptable only with some probability, the more agents involved in any one cycle, the higher the probability it will ultimately fail. This means that in any larger cycle there is always at least one agent who can improve his expected utility by learning the value of an alternative object and as a result only small cycles form. This is in contrast with the kidney exchange literature, where due to the constraints of hospitals, exchange is often exogenously restricted to only two agent-donor pairs (Roth et al. (2004), (2005)). In this paper, since the two-agent cycles arise endogenously, such a restriction would not affect the results.

To discriminate between the potentially large set of equilibria, I use ex-ante welfare. The least desirable equilibria will allow for only one exchange to take place, while the most desirable will maximise the number of possible exchanges. The welfare-maximising equilibria can only be reached if learning is organised to prevent the wasteful learning that occurs when many agents all learn the value of the same object. At first glance it may seem desirable to create some mechanism with two stages that firstly allows agents to express their preferences over which objects they wish to learn about and secondly which objects they wish to match with ex-post. However, if the goal is ex-ante welfare maximisation then in these environments such a mechanism proves ultimately unnecessary. When objects are identical, agents are ex-ante indifferent between all objects and so any equilibrium where the number of two-agent learning cycles is maximised will also maximise ex-ante welfare. When objects are non-identical, a two-stage mechanism would still be unnecessary since there is a unique set of learning cycles which maximise the ex-ante welfare that can be achieved in equilibrium.

It should then be questioned whether such a set of learning cycles can be enforced. Returning for a moment to the kidney exchange application, it is well known that hospitals act strategically on behalf of the patient to maximise their own number of matches (Roth (2002), Ashlagi and Roth (2012)). Indeed, in the environments discussed here many learning allocations are unstable. When objects are identical, only a relatively small subset of the equilibria are stable and when objects are non-identical then the stable set of learning cycles is, once again, unique. That is, there is a unique way to
pair agents together such that no two agents wish to deviate from their allocation and learn each other’s respective objects. For instance, if the two agents with the two objects of highest ex-ante value are not paired together then clearly they will both profit from ignoring the designer’s learning allocation and learning about their respective endowments. A patient in need of a kidney is unlikely to want to ignore a donor with a high probability match who is also interested in mutual exchange.

In Theorems 3 and 4, I show that these goals of welfare maximisation and stability are not in conflict in either environment. Any stable set of learning cycles will also be an ex-ante welfare maximising equilibrium and when objects are non-identical, such a set of learning cycles is unique. Allowing agents with high value objects to match, yields the welfare dominant equilibrium outcome in which the incentive for any group of agents to deviate is also eliminated. Whether this is the ‘best’ outcome may depend upon the precise application. If the goal is to prioritise matches between specific patients, such as those for whom it is difficult to find a match, then the results here are of little assistance. In contrast, if the aim of kidney exchange is to maximise the number of matches and possible life expectancy and to reduce numbers of patients on dialysis then the discussion over which tests to perform or how to construct a two-stage mechanism is somewhat redundant. The stable allocation perfectly coincides with the learning allocation that maximises ex-ante welfare in equilibrium and this is the only allocation that achieves these objectives.

The literature on endogenous information acquisition within the matching field is not extensive but this is by no means the first. Bade (2015) shows that when learning is costly, serial dictatorship is the unique ex-ante Pareto optimal, strategyproof and non-bossy allocation mechanism when information is endogenous. Harless and Manjunath (2017) find that top trading cycles dominate priority rules in progressive measures of social welfare under costless but restricted learning. The key difference in this paper is that each agent arrives at the problem with an endowment already in place and so priority rules such as serial dictatorship cannot be applied without losing individual rationality. In the kidney exchange literature, Dickerson et. al (2013) use random graph models to try and increase the number of successful matches in algorithmic programs and Blum et al. (2013) use such models to show that the problem of maximising the number of expected exchanges with two crossmatch opportunities is NP complete. As in this paper, learning is restricted and an exchange only takes place with some probability. However, the models do not examine individual incentives and stability. Ignoring incentives allows them to move outside the GTT mechanism but neglecting the strategic
components is not without cost. Besides failed crossmatches, another significant factor in failed matches is hospitals withholding easy matches from the centralised system. Since without cumbersome and impractical legislation it is not possible to prevent this problem (Ashlagi and Roth (2012), Ashlagi et al. (2013)) it seems prudent to build this strategic behaviour into the model.

This paper is organised as follows. Section 2 introduces the model and definitions and then section 3 characterises the equilibrium. Section 4 examines the problem of maximising ex-ante welfare amongst the large set of equilibria and creating stable learning allocations. In both section 3 and 4, the result is first stated using the model with identical objects and then non-identical objects. Despite being a more restricted version of the model, the identical object environment suffices to illustrate the key points of all proofs for the non-identical case as well. In section 5, the ex-ante welfare of the equilibrium outcomes are compared to the maximum possible ex-ante welfare achievable when all strategic concerns are put aside. Section 6 consists of remarks on the effects on the results of relaxing some of the model’s constraints.

2 Model

A finite set of agents $N = \{1, \ldots, n\}$ is such that each agent $i$ is endowed with object $i$ in the set $K = \{1, \ldots, n\}$. An agent values his own endowment at zero but does not know his value for any other object. All the agent knows is the value is drawn from some distribution with an expected value less than that of his endowment. All agents are expected utility maximisers, so without any further information no agent wants to trade and prefers instead to keep his own endowment. More formally, there exists a state space $\Omega$ consisting of profiles of values $\omega = (\omega_k^i)_{i \in N, k \in K}$, where $\omega_k^i$ is the value of object $k$ to agent $i$ in state $\omega$. Agents know the value of their own endowment so $\omega_i^i = 0$ for all $\omega \in \Omega$. For all other objects, $\omega_k^i$ is an independent draw from some distribution $f_k^i$ with some support not containing zero\(^1\) and such that $E(\omega_k^i) < 0$ for all $i \in N, k \in K$. This means agent $i$ knows both the probability $\pi_k^i = \pi(\omega_k^i > 0)$ that $k$ has a higher value than his own endowment and its expected value if it is higher; $E_k^i = E(\omega_k^i | \omega_k^i > 0)$.

Each agent $i$ chooses to learn the value $\omega_k^i$ of one object $k \neq i$. If $kR_i k'$ but not $k'R_i k$ then it is denoted $kP_i k'$. The ex-post preference profile is $R = (R_i)_{i \in N}$ and the set of

\(^1\)This is to prevent indifferences in what follows. Alternatively, zero can be allowed in the support and a tie breaking rule introduced.
all possible ex-post preference profiles is $\mathcal{R}$. If $i$ chooses to learn about object $k$ then $kR_i i$ if $\omega^i_k > 0$. Since each agent tests only one object and any untested objects have an expected value below $i$’s endowment, there is at most one $k \neq i$ such that $kP_i i$. There may, however, be many untested objects $k' \neq i$ such that $iR_i k'$ and it is possible that an agent is indifferent between two such objects.

A matching is a bijection $\mu : N \rightarrow K$. The set of all matchings is $\mathcal{M}$. A matching is individually rational if $\mu(i)R_i i$ for all $i \in N$. A matching $\mu'$ Pareto dominates $\mu$ if $\mu'(i)R_i \mu(i)$ for all $i \in N$ and $\mu'(i^*)P_i \mu(i^*)$ for at least one $i^* \in N$. If a matching is not Pareto dominated then it is Pareto optimal. A mechanism, $M : \mathcal{R} \rightarrow \mathcal{M}$, is individually rational and Pareto optimal if it always results in an individually rational and Pareto optimal matching. A mechanism is strategyproof if $M(R)(i)R_i M(R'_i; R_{-i})(i)$ for all $i \in N, R'_i$.

Agents are matched via Gale’s Top Trading Cycles mechanism, $GTT : \mathcal{R} \rightarrow \mathcal{M}$ which works as follows:

**Step r:** Each unmatched agent points at his most preferred object from amongst those remaining. Each object points at its owner. At least one cycle forms. All agents in a cycle receive the object they are pointing at and are removed. If at least one agent remains then proceed to step $r + 1$. If not then end.

Since $N$ is finite, $GTT$ ends when all agents have been matched with an object. The domain of the preferences $\mathcal{R}$ considered here, differs from the conventional $GTT$ domain in which preferences over all objects are strict. However, the indifference between some objects permitted under $\mathcal{R}$ does not affect the mechanism’s function since if agent $i$ is indifferent between object $k$ and $k'$ then it must be that $iP_k k$ and $iP_k k'$. Therefore, $i$ will always point at and be matched to his own object before needing to choose between objects to which he is indifferent. $GTT$ is used because it is the unique strategyproof, individually rational and Pareto optimal mechanism for the domain $\mathcal{R}$.

Since $GTT$ is strategyproof, I assume that each agent truthfully reports their ex-post preferences and so the only decision each agent need take is which object to learn about. For this reason, I model $i$’s strategy space as $A_i = K \setminus \{i\}$ and a strategy profile is $a = (a_i)_{i \in N}$ such that $a \in A = \times_{i \in N} A_i$. The agents’ ex-post preferences will depend

---

2The proof in Ma (1994), which applies to the domain of linear orders can be extended to the domain considered here.
on this strategy as well as the profile of object values and so the ex-post preference is a function \( R : A \times \Omega \rightarrow \mathbb{R} \). Agent \( i \) will choose \( a_i \) to maximise his ex-ante expected utility, which depends not only on the expected value of the object \( E^i_{a_i} \), but also the probability that \( i \) is matched to \( a_i \), \( \pi(GTT(R(a, \omega))(i) = a_i) \):

\[
U_i(a) = \pi(GTT(R(a, \omega))(i) = a_i) \cdot E^i_{a_i}
\]

A strategy profile \( a \) is an equilibrium if \( U_i(a) \geq U_i(a'_{i}, a_{-i}) \) for all \( i \in N \).

A learning cycle is a vector \((k_1, ..., k_m)\) such that \( a_{k_i} = k_{i+1} \) for all \( i < m \) and \( a_m = k_1 \). An \( m \)-cycle is a learning cycle that contains \( m \) agents. Let \( o(a) \) be the set of learning cycles that forms under \( a \). Since 2-cycles play a prominent role in what follows, let \( C(a) \) be the set of agents in 2-cycles and \( B(a) = N \setminus C(a) \) the set of agents not in 2-cycles under \( a \).

### 2.1 Environments

I make several assumptions about the distributions from which agents draw their values of the objects. In the first environment, IDO, I assume that \( f_{i,k} \) is identically and independently distributed across all agents and objects. Though this may seem overly restrictive it is sufficient to highlight the key features of the model. In the second environment, NIDO, I relax this assumption and allow heterogeneous objects, making some more (ex-ante) desirable than others.

**IDO (Identical Objects and Agents):** When agents and objects are ex-ante identical then \( f_{i,k} = f \) for all \( i \in N, k \in K \) and agent \( i \)'s value of object \( k \neq i \), \( \omega_k^i \) is an iid draw from \( f \). Since \( f \) is constant across objects let \( E := \mathbb{E}(\omega_k^i | \omega_k^i > 0) \) and \( \pi := \mathbb{P}(\omega_k^i > 0) \) for all \( i \in N, k \in K \).

**Observation 1.** For all \( k \in N \), \( \max_{a} U_k(a) = \pi^2 E \).

Observation 1 follows from:

\[
U_k(a) = \begin{cases} 
\pi^m E & \text{if } k \text{ is in an } m \text{-cycle} \\
0 & \text{otherwise}
\end{cases}
\]

Since an agent cannot form a learning cycle alone, the smallest \( m \)-cycle that can form is a 2-cycle. This means \( \pi^2 E \) is the maximum utility an agent can achieve. If not in
a cycle, then an agent cannot trade and so keeps his endowment which has a value of zero.

**NIDO (Nonidentical Objects and Identical Agents):** When agents are ex-ante identical but objects are not then $f_k^i = f_k$ for all $i \in N, k \in K$ such that $i \neq k$ and $\omega_k^i$ is an iid draw from $f_k$. Since $f_k$ is non-identical across objects let $E_k := E(\omega_k^i | \omega_k^i > 0)$ and $\pi_k := \pi(\omega_k^i > 0)$ for all $i \in N, k \in K$. Unless otherwise specified I assume that agent $i$ is endowed with the $i$th best object such that $E_1 > E_2 > \ldots > E_n$ and $\pi_1 > \pi_2 > \ldots > \pi_n$. This is not without loss of generality and in section 5 I show how changes in this assumption can affect the results discussed here.

**Observation 2.** For all $k \in N$, $\max_a U_k(a) = \pi_k E_k^\hat{k}$ where $\hat{k} = \min N \setminus \{k\}$

Observation 2 follows from:

$$U_k(a) = \begin{cases} \prod_{i \in T} \pi_i E_{ai} & \text{if } k \text{ is in an } m\text{-cycle with } T \subset N \\ 0 & \text{otherwise} \end{cases}$$

Since the smallest cycle that can form is a 2-cycle and $\hat{k} = \min N \setminus \{k\}$, $\pi_k E_k^\hat{k}$ is the maximum utility an agent can achieve.

The assumption in both IDO and NIDO that $f_k^i$ is identical across all agents is an important one. If this assumption is relaxed then the results that follow do not apply. In particular, in example 1 at the end of section 3, I show how the characterisation of the equilibrium does not hold if different agents face different distributions.

### 3 Equilibrium

The nature of equilibrium is very similar in the IDO and NIDO environments. In both cases, it is not possible for large cycles to exist in equilibrium and as a result only 2-cycles persist. Since each cycle results in an executable trade only with some probability, the more agents involved in any one cycle, the greater the probability it will fail. In IDO it is unimportant which agents are in the 2-cycles in equilibrium and it is possible that some agents are not in any cycle at all. As shown below, NIDO differs only in as much that it does matter which agents are in the 2-cycles.
3.1 Identical Objects (IDO)

**Theorem 1.** If \( a \) is an equilibrium then all learning cycles are 2-cycles. Fix \( C \subset N \) and \( B = N \setminus C \) such that \(|C|\) is even. Then there exists an equilibrium \( a \) such that \( C = C(a) \) and \( B = B(a) \).

**Proof.** The fact that one cycle forms under \( a \) follows directly from the finiteness of \( N \) and \( K \). To see that in equilibrium every learning cycle is a 2-cycle consider a single \( m \)-cycle \((k, k', ..., k''_m)\) such that \( m \geq 3 \). Such a cycle is illustrated in figure 1(a). For agent \( k' \):

\[
U_{k'}(a) = \pi^m E
\]

But for \( a'_{k'} = k \):

\[
U_{k'}(a'_{k'}, a_{-k'}) = \pi^2 E
\]

This is shown in figure 1(b). Since \( \pi^2 > \pi^m \), \( a \) cannot be an equilibrium.

![Figure 1](image1.png)

To see that an equilibrium \( a \) exists such that \( C(a) = C \) and \( B(a) = B \), let all agents in \( C \) form some set of 2-cycles such that \( C(a) = C \). Let \( a_b = c^* \) for all \( b \in B \) and some \( c^* \in C \). Then no \( b \in B \) is in a cycle and \( B = B(a) \). An example is shown in figure 2.

By observation 1, \( U_c(a) \geq U_c(a'_c, a_c) \) for all \( a'_c \in A_c, c \in C \). Since no agent learns about the endowment of any agent in \( B \), all agents in \( B \) have an expected utility of zero.
regardless of their strategy. That is, since $a_k \neq b$ for any $k \in N$, $b \in B$, $U_{k'}(a^*) = 0 = U_{k'}(a'_{k'}, a^*_{-k'})$ for all $k' \in B(a^*)$ and all $a'_{k'} \in A_k$.

3.2 Non-identical Objects (NIDO)

**Theorem 2.** If $a$ is an equilibrium then all learning cycles are 2-cycles. Fix $C \subset N$ and $B = N \setminus C$ such that $|C|$ is even and $1, 2 \in C$. Then there exists an equilibrium $a$ such that $C(a) = C$ and $B(a) = B$.

**Proof.** To see that in equilibrium every learning cycle is a 2-cycle consider a single $m$-cycle $(k, k', k^*, ..., k^{m'})$ that forms between some $S = \{k, k', k^*, ..., k^{m'}\}$ such that $m \geq 3$ and $k^* \geq i$ for all $i \in S$.

This is illustrated in figure 3(a). For agent $k'$:

$$U_{k'}(a) = \prod_{i \in S} \pi_i E_{k^*}$$

But for $a'_{k'} = k$:

$$U_{k'}(a'_{k'}, a_{-k'}) = \pi_{k'} \pi_k E_k$$

This is shown in figure 3(b). Since $k^* > k$, by the NIDO assumptions, $E_k > E_{k^*}$. But then $U_{k'}(a) < U_{k'}(a'_{k'}, a_{-k'})$ and $a$ cannot be an equilibrium.

To see there exists an equilibrium $a$ such that $C(a) = C$ and $B(a) = B$, let all agents in $C$ form some set of 2-cycles under $a$ such that $(1, 2) \in o(a)$ and $C(a) = C$. Let $a_b = 1$.

---

3If $|S| = 3$ then $k^* = k''$
for all $b \in B$. Then no $b \in B$ is in a cycle and so $C = C(a)$ and $B = B(a)$. An example is shown in figure 4.

Every agent in $C$ is in a 2-cycle so no $c \in C \setminus \{1, 2\}$ can form a cycle with any agent other than $a_c$ since $a_k \neq c$ for all $k \in N \setminus \{a_c\}$. By the same argument, neither agent 1 nor 2 can form a cycle other than (1,2) with any agent in $C$. It is possible for agents 1 and 2 to form 2-cycles or 3-cycles with agents in $B$. Following observation 2, however, all such cycles would yield a lower utility than their existing (1,2) cycle. This means that $U_c(a'_c, a_{-c}) \leq \pi_c \pi_{a_c} E_{a_c}$ for all $a'_c \in A_c, c \in C$. Since $a_k \neq b$ for any $k \in N, b \in B$, no agent in $B$ can form a cycle regardless of his strategy and so $U_b(a'_b, a_{-b}) = 0$ for all $a'_b \in A_b, b \in B$. \hfill \square

![Figure 3](image-url)  
Figure 3: (a) The $m$-cycle $(k, k', k^*, \ldots, k''')$ (b) The alternative 2-cycle $(k, k')$

![Figure 4](image-url)  
Figure 4: An equilibrium for a set of 2-cycles in the NIDO environment
3.3 Non-identical Objects and Agents

In both the IDO and NIDO environments, all agents have the same ex-ante value for any given object other than one’s own endowment. However, without this assumption the above results do not hold. In particular, if $f_k^i$ is non-identical across both agents and objects then, as example 1 shows, learning cycles larger than just 2-cycles can exist in equilibrium.

**Example 1:** First suppose three agents, $N = \{1, 2, 3\}$, all disagree over which objects are the most and least (ex-ante) desirable. Agent 1 knows he ex-ante prefers agent 2’s endowment to 3’s since $E_{12} = 10 > E_{13} = 1$ and $\pi_1^2 = 0.9 > \pi_1^3 = 0.1$. Similarly, agent 2 ex-ante prefers 3’s endowment and 3 ex-ante prefers 1’s: $E_{23} = E_{31} = 10 > E_{11} = E_{22} = 1$ and $\pi_2^3 = \pi_3^1 = 0.9 > \pi_2^1 = \pi_3^2 = 0.1$. Then $a_1 = 2, a_2 = 3$ and $a_3 = 1$ is an equilibrium (as shown in Figure 5(a)). Under $a$, $U_i(a) = (0.9)^3 \times 10 = 7.29$ for all $i \in N$ and $U_i(a'_i, a_{-i}) = 0.9 \times 0.1 = 0.09$ for all $a'_i \neq a_i$ and $i \in N$. Since $U_i(a) > U_i(a'_i, a_{-i})$ for all $i \in N$, $a'_i \in A_i$ is an equilibrium.

![Figure 5(a)](image1)

![Figure 5(b)](image2)

**Figure 5:** (a) A 3-cycle between all agents. (b) An $n$-cycle between all agents.

As the number of agents increases, so too does the maximum possible cycle size. In fact, Example 2 demonstrates that for any $n$ it is possible to construct an $n$-cycle that can exist in equilibrium. To do this, I deviate from the standard approach by first fixing an
n-cycle between all agents and then finding parameter values such that the n-cycle can indeed be maintained in equilibrium.

**Example 2:** For some \( N = \{1, \ldots, n\} \), let \( a_i = i + 1 \) for all \( i \in N \setminus \{n\} \) and \( a_n = 1 \). An n-cycle then forms between all agents as shown in Figure 5(b). Let the ex-ante values of the objects be such that \( E_{i+1}^j > E_{i-1}^j > E_j^i \) and \( \pi^i_{i+1} > \pi^i_{i-1} > \pi^i_j \) for all \( j \in N \setminus \{i - 1, i + 1\}, i \in N \setminus \{1, n\} \). For agent 1, \( E_2^1 > E_n^1 > E_j^1 \) and \( \pi^1_2 > \pi^1_n > \pi^1_j \) for all \( j \in N \setminus \{2, n\} \) and for agent \( n \), \( E_1^n > E_{n-1}^n > E_j^n \) and \( \pi^n_1 > \pi^n_{n-1} > \pi^n_j \) for all \( j \in N \setminus \{1, n - 1\} \). Then \( U_i(a) = E_{a_i}^i \prod_{j=1}^{n-2} \pi^i_{a_j} \) and \( U_i(a'_i, a_{-i}) \leq E_{i'-1}^i \pi^i_{i'-1} \pi^i_i \) for all \( a'_i \in A_i, i \in N \setminus \{1\} \) and \( U_1(a'_i, a_{-i}) \leq E_n^n \pi^n_i \pi^n_i \) for all \( a'_i \in A_1 \). So for a sufficiently large \( E_1^n \) and \( E_{i+1}^j \) for all \( i \in N \setminus \{n\} \) and a sufficiently small \( E_n^n \) and \( E_{i-1}^i \) for all \( i \in N \setminus \{1\} \), \( U_i(a) > U_i(a'_i, a_{-i}) \) for all \( a'_i \in A_i \).

## 4 Ex-ante Welfare and Stability

### 4.1 More definitions

**Ex-ante welfare** is the sum of all agents’ expected utilities under a given strategy profile: \( W(a) := \sum_{i \in N} U_i(a) \). In the context of this type of object allocation problem, stability is normally defined with respect to the final allocation of objects. An allocation is stable if there is no agent matched with an object that would rather not be and if there is no subset of agents who would each have higher utility if they redistributed their allocated objects amongst themselves. It is well established that Gale’s Top Trading Cycles produces stable outcomes (Shapley and Scarf, 1974). However, in the problem at hand, agents’ final allocations are also dependent on the outcome of the agents’ learning decisions. Stability then is defined here over learning cycles and agents’ expected utility of the final allocation they expect to be realised. A set of learning cycles is stable if there is no agent in a cycle who would rather not be and if there is no subset of agents who would rather form a different learning cycle amongst themselves. More formally, a set of cycles \( o(a) \) is **stable** if:

(i) There exists no \( k \in N \) and \( a'_k \) such that \( k \) is in a cycle under \( o(a) \), \( k \) is in not in a cycle under \( o(a'_k, a_{-k}) \) and \( U_k(a) < U_k(a'_k, a_{-k}) \).

(ii) There exists no \( S \subset N \) and \( a'_S \) such that \( o(a) \neq o(a'_k, a_{-k}) \) and \( U_k(a) < U_k(a'_S, a_{-S}) \) for all \( k \in S \).
Note that part (i) of this definition is trivially satisfied since an agent’s expected utility is positive if and only if an agent is in a learning cycle. If \( k \) is not in a learning cycle then \( k \)’s expected utility is zero and as so \( k \)’s expected utility will always be higher when \( k \) is in a learning cycle. If (ii) holds for all \( S \subset N \) such that \(|S| = 2\) then \( o(a) \) is pairwise stable. However, as Lemma 1 shows, if \( o(a) \) is pairwise stable, then it is also stable for any \(|S| > 2\). Lemma 1 is applicable in both the IDO and NIDO environments.

**Lemma 1.** If \( o(a) \) is pairwise stable then \( o(a) \) is stable

Proof. Suppose \( o(a) \) is pairwise stable but not stable. Then there exists an \( a' \) under which an \( m \)-cycle, \((b, c, \ldots, d)\) forms between some \( S = \{b, c, \ldots, d\} \) such that \( U_k(a'_{S'}, a_{-S'}) > U_k(a) \) for all \( k \in S \). Since \( o(a) \) is pairwise stable, \(|S| > 2\). W.l.o.g let \( b < d < k \) for all \( k \in S \). Since \( o(a) \) is not stable via \( a' \), \( U_b(a'_{S'}, a_{-S'}) = \prod_{k \in S} \pi_k E_c > U_b(a) \) and \( U_d(a'_{S'}, a_{-S'}) = \prod_{k \in S} \pi_k E_b > U_d(a) \). But then there also exists a strategy profile \( a'' \) under which the learning cycle \((b, d)\) forms such that \( U_b(a''_{\{b, d\}}, a'_{S\setminus\{b, d\}}, a_{-S}) = \pi_b \pi_d E_b \) and \( U_d(a''_{\{b, d\}}, a'_{S\setminus\{b, d\}}, a_{-S}) = \pi_b \pi_d E_d \). Then, \( U_k(a''_{\{b, d\}}, a'_{S\setminus\{b, d\}}, a_{-S}) > U_k(a'_{S'}, a_{-S}) > U_k(a) \) for \( k \in \{b, d\} \). Since \( b \) and \( d \)’s expected utility under \( a'' \) is independent of all other agents’ strategies, it must also be that \( U_b(a''_{\{b, d\}}, a_{-(b,d)}) > U_b(a) \) and \( U_d(a''_{\{b, d\}}, a_{-(b,d)}) > U_d(a) \). But this implies that \( o(a) \) is pairwise unstable which is a contradiction. \( \square \)

As was the case with the equilibria, in what follows I also show that the nature of any set of stable learning cycles is very similar in both environments. For IDO and NIDO, a set of learning cycles is stable only if all cycles are 2-cycles and there is never more than one agent not in any cycle at all. The number of sets of stable learning cycles may be much smaller than the number of sets of equilibrium learning cycles. However, whilst ex-ante welfare may differ across equilibria, all sets of stable learning cycles attain the maximum possible ex-ante welfare for a set of equilibrium learning cycles. Where the two environments differ is in the size of these sets; whilst in IDO there may be many such sets of stable learning cycles, in NIDO that set is unique.

### 4.2 Identical Objects (IDO)

**Theorem 3.** Any stable set of learning cycles \( o(a) \) is also a set of ex-ante welfare maximising equilibrium learning cycles.

**Lemma 2.** \( o(a) \) is stable if and only if the number of 2-cycles is maximised.
Figure 6: A stable set of learning cycles in (a) the IDO environment, (b) the NIDO environment.

Proof. Let \( o(a) \) maximise the number of 2-cycles (see figure 6(a)). For all \( k \in C(a) \), \( U_k(a) = \pi^2 E \) and since the number of 2-cycles is maximised, there is at most one \( k' \in B(a) \). Then by observation 1 there exists no \( S \subset N \) such that \( U_k(a_S', a_{-S}) > U_k(a) \).

If the number of 2-cycles is not maximised then \( |B(a)| \geq 2 \). Let \( b, b' \in B(a) \). Then \( b \) and \( b' \) are either not in a cycle or in an \( m \)-cycle. In either case, \( U_k(a) < \pi^2 E \) for \( k \in \{b, b'\} \). But for \( a' \) such that \((b, b') \in o(a')\), \( U_k(a'_{(b,b')}, a_{-(b,b')}) = \pi^2 E > U_k(a) \) for \( k \in \{b, b'\} \) and so \( a \) is unstable. \( \square \)

Lemma 3. The number of 2-cycles is maximised in the ex-ante welfare maximising equilibrium.

Proof. Since \( \frac{|C(a)|}{2} \) is the number of 2-cycles that form under any equilibrium \( a \) and \( W(a) = |C(a)|\pi^2 E \) then \( a \in \arg\max W(a) \) if and only if \( |C(a)| \) is maximised. \( \square \)

The proof of Theorem 3 follows directly from Lemmas 2 and 3. All sets of stable learning cycles will be similar to that shown in Figure 6(a).
4.3 Non-identical Objects (NIDO)

**Theorem 4.** The unique set of stable learning cycles is the unique set of ex-ante welfare maximising equilibrium learning cycles

The proof of Theorem 4 is via Lemmas 4 and 5. Both rely on a unique set of learning cycles. This set may be generated by more than one \( a \in A \) but if this is so then each such \( a \) would differ only in the learning decision of agent \( n \) who (by the NIDO assumptions) is endowed with the lowest ex-ante valued object. Define \( A^* = \{a|o(a) = \{(1,2),(3,4),...,(k,k')\}\} \) where if \( n \) is even, \( k = n-1 \) and \( k' = n \), if \( n \) is odd, \( k = n-2 \) and \( k' = n-1 \). Then if \( n \) is even, all \( k \in N \) are in 2-cycles and \( A^* \) is a singleton. If \( n \) is odd then all \( k \in N \setminus \{n\} \) are in 2-cycles and agent \( n \) is not in any learning cycle, regardless of \( n \)’s strategy. Then \(|A^*| = n-1\). In either case, the number of 2-cycles is maximised. An example of an \( a \in A^* \) is shown in Figure 6(b). The key difference between this and the IDO environment is that it is now important which agents are in which 2-cycles.

**Lemma 4.** The unique set of stable learning cycles is \( o(a^*) \), where \( a \in A^* \).

**Proof.** I first show that any \( o(a^*) \) with \( a^* \in A^* \) is stable. If \( a^* \) is not stable then by Lemma 1 there is some pair of agents \( c \) and \( c' \) such that \( (c,c') \notin o(a^*) \) and for \( a_{c'}^c = c' \) and \( a_{c'}^c = c, U_i(a^*) = U_i(a_{(c,c')}^i), a^*_{(c,c')} \) for \( i \in \{c,c'\} \). By the construction of \( A^* \), under \( o(a^*) \) \( c \) is in a cycle with either \( c = 1 \) or \( c + 1 \) and so either \( U_c(a^*) = \pi_c \pi_{c-1} E_{c-1} \) or \( U_c(a^*) = \pi_c \pi_{c+1} E_{c+1} \). In either case since \( U_c(a^*) < U_c(a_{(c,c')}^i), a^*_{(c,c')} \) it must be that \( c' < c \). But by the same argument for \( c' \) it must be that \( c < c' \). This is a contradiction and so \( o(a^*) \) is stable.

To see that any other set of 2-cycles is unstable, fix some \( a' \notin A^* \). Suppose \( o(a') \) does not maximise the number of 2-cycles, then \(|B(a')| \geq 2\). Let \( b,b' \in B(a') \) be such that \( b = \max B(a') \) and \( b' = \max B(a') \setminus \{b\} \). Since \( b,b' \in B(a') \), \( b \) and \( b' \) are either not in a cycle or in an \( m \)-cycle such that \( m > 2 \). In either case, \( U_k(a') < \pi_{b} \pi_{b'} E_{b'} \) for \( k \in \{b,b'\} \). But for some \( \tilde{a} \) such that \( (b,b') \in o(\tilde{a}), U_k(\tilde{a}_{(b,b')}), a_{(b,b')}^k \geq \pi_{b} \pi_{b'} E_{b'} \) for \( k \in \{b,b'\} \) and so \( o(a') \) is not stable. So it must be that if \( o(a') \) is stable then the number of 2-cycles are maximised.

Now suppose \( o(a') \) does maximise the number of 2-cycles. Since \( a' \notin A^* \), there must be some agents who are in a cycle under any \( a^* \in A^* \) but in a different cycle (or no cycle)
Lemma 4. To see that $o$ is non-empty and so $q^*$ exists and $x$ and $y$ are such that $x, y > q^* + 1$. But then ex-ante welfare realised under $a'$ can be increased by replacing $(q^*, x)$ and $(q^* + 1, y)$ under $a'$. Since the number of 2-cycles is maximised there can be at most one agent not in a cycle under $a'$. If there exists an agent $c$ such that $c \in C(a^*)$ but $c \notin C(a')$ then by the construction of $A^*$, $c \neq n$. Since the number of 2-cycles is maximised there must also be some $(n, c') \in o(a')$ but this makes $o(a')$ unstable via the cycle $(c, c')$. Then if there exists some $k \in N$ such that $k \notin C(a')$, it must be that $k = n$ and all other agents are in 2-cycles.

Define $Q := \{q \mid (q, q + 1) \in o(a^*)$ and $(q, q + 1) \notin o(a')\}$. Note that $Q$ contains only odd numbered agents but since $a' \notin A^*$ and all cycles in any $o(a^*)$ such that $a^* \in A^*$ contain one odd numbered and one even numbered agent, $Q$ is nonempty. Let $q^* = \min Q$. Since $q^*$ is not in a cycle with $q^* + 1$, let $(q^*, x), (q^* + 1, y) \in o(a')$. Since $q^* = \min Q$, all $i \in N$ such that $i < q^*$ are in the same cycles in $o(a^*)$ and $o(a')$. Since $(q^*, x), (q^* + 1, y) \in o(a')$ and $(q^*, x), (q^* + 1, y) \notin o(a^*)$, $x, y > q^* + 1$. But then,

$$U_{q^*}(a') = \pi_{q^*} \pi_x E_x < \pi_{q^*} \pi_{q^* + 1} E_{q^* + 1} = U_{q^*}(a_{(q^*, q^* + 1)}, a_{- (q^*, q^* + 1)})$$

$$U_{q^* + 1}(a') = \pi_{q^* + 1} \pi_y E_y < \pi_{q^* + 1} \pi_{q^* + 2} E_{q^* + 2} = U_{q^* + 1}(a_{(q^*, q^* + 1)}, a_{- (q^*, q^* + 1)})$$

which implies $o(a')$ is unstable.

\[\square\]

**Lemma 5.** The unique set of ex-ante welfare maximising equilibrium learning cycles is $o(a^*)$, where $a^* \in A^*$.

**Proof.** By Theorem 2, if $a$ is an equilibrium then all cycles in $o(a)$ are 2-cycles under $a$. Fix an equilibrium $a' \notin A^*$ such that for any $k \notin C(a')$, $k$ is not in any cycle in $o(a')$. If $a'$ is ex-ante welfare maximising then $|B(a')| < 2$. Otherwise, if $k, k' \in B(a')$ then $o(a') \cup \{(k, k')\}$ ex-ante welfare dominates $o(a')$ by observation 2. So $|B(a')| < 2$ and $o(a')$ must maximise the number of 2-cycles.

To see that $o(a^*)$ ex-ante welfare dominates $o(a')$, define $Q := \{q \mid (q, q + 1) \in (a^*)$ and $(q, q + 1) \notin o(a')\}$. Let $(q^*, x), (q^* + 1, y) \in o(a')$ and $q^* = \min Q$. By the same arguments as in Lemma 4, $Q$ is non-empty and so $q^*$ exists and $x$ and $y$ are such that $x, y > q^* + 1$. But then ex-ante welfare realised under $a'$ can be increased by replacing $(q^*, x)$ and $(q^* + 1, y)$
with \((q^*, q^* + 1) \in o(a^*)\) and \((x, y)\). If this were not the case then,
\[
\pi_{q^*} \pi_x E_x + \pi_{q^* + 1} \pi_y E_y + \pi_x \pi_{q^*} E_{q^*} + \pi_y \pi_{q^* + 1} E_{q^* + 1} \geq \\
\pi_{q^*} \pi_{q^* + 1} E_{q^* + 1} + \pi_{q^* + 1} \pi_{q^*} E_{q^*} + \pi_x \pi_y E_y + \pi_y \pi_x E_x
\]
\[
\Rightarrow \pi_x (\pi_{q^*} E_{q^*} - \pi_y E_y) + \pi_y (\pi_{q^* + 1} E_{q^* + 1} - \pi_x E_x) \geq \\
\pi_{q^*} (\pi_{q^* + 1} E_{q^* + 1} - \pi_x E_x) + \pi_{q^* + 1} (\pi_{q^*} E_{q^*} - \pi_y E_y)
\]
But since the NIDO assumptions imply \(\pi_x < \pi_{q^* + 1}\) and \(\pi_y < \pi_{q^*}\), this cannot hold. Then the cycles \((q^*, q^* + 1)\) and \((x, y)\) must result in higher ex-ante welfare than \((q^*, x)\) and \((q^* + 1, y)\).

The proof of Theorem 4 then follows directly from Lemmas 4 and 5.

5 Maximising ex-ante welfare

So far the ex-ante welfare of \(a\) when the set of learning cycles \(o(a)\) is stable has only been discussed in relation to the maximum ex-ante welfare achievable in equilibrium. However, it is not immediately apparent how this compares to the maximum ex-ante welfare that could be achieved overall.

5.1 IDO and IDO+

For some environments, the difference between the ex-ante welfare when the learning cycles are stable and the maximum ex-ante welfare achievable overall is either identical or almost negligible. This includes the IDO environment but it also applies to a broader environment, here called IDO+ where objects are not necessarily identical.

IDO+: As in NIDO, agents are ex-ante identical but objects are not so \(f_k^i = f_k\) for all \(i \in N, k \in K\) such that \(i \neq k\) and \(\omega_k^i\) is an iid draw from \(f_k\). Also, as in NIDO, since \(f_k\) is non-identical across objects, let \(E_k := E(\omega_k^i \mid \omega_k^i > 0)\). However, as in IDO, let \(\pi := \pi(\omega_k^i > 0)\) for all \(i \in N, k \in K\). Then all objects are equally likely to be acceptable but may differ in their expected value conditional on them being acceptable. Without loss of generality I assume that agent \(i\) is endowed with the \(i^{th}\) best object such that \(E_1 \geq E_2 \geq \ldots \geq E_n\).
In this environment, as in IDO, the set of stable learning cycles is not necessarily unique. Theorem 5 shows that despite this, all stable learning cycles yield the same ex-ante welfare and that this is either identical to, or very ‘close’ to the maximum ex-ante welfare overall.

**Theorem 5.** Let \( \bar{a} \in \arg \max_{a \in A} W(a) \) and \( o(a^+) \) be any set of stable learning cycles in IDO+. Then either \( W(\bar{a}) = W(a^+) \) or \( W(\bar{a}) \) and \( W(a^+) \) differ only in the expected utility of agents \( n, n-1 \) and \( n-2 \).

**Lemma 6.** The ex-ante welfare of any \( a \in A^+ \) is a stable equilibrium, but it may not be the unique stable equilibrium.

**Proof.** Suppose \( E_1 > E_2 > \ldots > E_n \). Then the proof for Lemma 4 holds and \( o(a^*) \), where \( a^* \in A^+ \), is the unique stable equilibrium. Now suppose there exists at least one pair of agents, \( i, j \in N \) such that \( E_i = E_j \). Group the agents using the following steps:

**Step 1:** Let \( i^* = \min N \). Define \( \beta^1 := \{ i \mid i \in N \text{ and } E_i = E_{i^*} \} \). If \( |\beta^1| = N \) then end, if not continue to step 2.

Then in general at step \( r \):

**Step r:** Let \( i^* = \min N \setminus (\bigcup_{j=1}^{r-1} \beta^j) \). Define \( \beta^r := \{ i \mid i \in N \text{ and } E_i = E_{i^*} \} \).

If \( \bigcup_{j=1}^r \beta^j = N \) then end, if not continue to step \( r + 1 \).

This must terminate at some step since the number of agents is finite. Once terminated, agents will be separated into groups where every agent in each particular group is endowed with an object of the same ex-ante value. Note that if \( |\beta^1| = N \) then this is equivalent to the IDO environment. The bijection \( p : N \to N \) is such that if \( p(i) \in \beta^j \) if \( i \in \beta^j \) for all \( i \in N \). Let \( P = \{p_1, \ldots, p_x\} \) be the set of all permutations of \( p \). Define \( O := \{(p_1(1), p_1(2)), (p_1(3), p_1(4)), \ldots, (p_1(k), p_1(k')) \mid p_1 \in P \} \) where if \( n \) is odd, \( k = n - 1 \) and \( k' = n \), if \( n \) is even, \( k = n - 2 \) and \( k' = n - 1 \). Then \( O \) consists of sets of cycles that differ only in so much as agents endowed with objects of identical ex-ante value may have switched places in the cycles (an example is shown in Figure 7). The set of learning decisions that generate these cycles is \( A^+ := \{a \mid o(a) \in O \} \).

Then a set of learning cycles \( o(a) \) is stable if and only if \( a \in A^+ \). This can be seen by repeating the proof for Lemma 4, replacing the set \( A^+ \) with \( A^+ \). The key difference between \( A^+ \) and \( A^+ \) is that if \( a \in A^+ \) there is a unique set of stable learning cycles, \( o(a) \) but, if \( a \in A^+ \), then there are \( |O| \) sets of stable learning cycles.
Figure 7: Stable learning cycles in IDO+. If $E_1 > E_2 = E_3 > E_4$ then both the cycles shown in (a) and (b) are stable.

**Lemma 7.** The ex-ante welfare of any $a^+ \in A^+$ is $W(a^+) = \sum_{i=1}^k E_i$, where $k = n$ if $n$ is even and $k = n - 1$ if $n$ is odd.

**Proof.** If $n$ is even then all agents are in 2-cycles and $U_i(a^+) = \pi^2 E_{a_i}$ for all $i \in N$. Since all agents are in 2-cycles, no two agents are learning about the same object; $a_i \neq a_j$ for all $i \neq j$. Then $W(a^+) = \sum_{i=1}^n U_i(a^+) = \pi^2 \sum_{i=1}^n E_i$. If $n$ is odd then one agent is not in a cycle and all other agents are in 2-cycles. Suppose the agent not in a cycle is $n$. Then $W(a^+) = \pi^2 \sum_{i=1}^{n-1} E_i$. Now suppose the agent not in a cycle is $b \neq n$. Then $W(a^+) = \pi^2 \sum_{i \in N \backslash \{b\}} E_i$. Since, by the construction of $A^+$, $E_n = E_b$, $\pi^2 \sum_{i \in N \backslash \{b\}} E_i = \pi^2 \sum_{i=1}^{n-1} E_i$. \qed

**Lemma 8.** If $\bar{a} \in \text{arg max}_{a \in A} W(a)$ then:

$$W(a) = \begin{cases} \pi^2 \sum_{i=1}^n E_i & \text{if } n \text{ is even} \\ \pi^2 \sum_{i=1}^{n-1} E_i & \text{if } n \text{ is odd and } \pi \leq \gamma \\ \pi^2 \sum_{i=1}^{n-3} E_i + \pi^3 (E_n + E_{n-1} + E_{n-2}) & \text{if } n \text{ is odd and } \pi > \gamma \end{cases}$$

where $\gamma = \pi > \frac{E_{n-2} + E_{n-1} + E_n}{E_{n-2} + E_{n-1} + E_n}$.

**Proof.** If an agent is in a 2-cycle then $U_i(a) = \pi^2 E_i$. If an agent is not in a cycle then
$U_i(a) = 0 < \pi^2 E_i$ and if in an $m$-cycle such that $m > 2$, $U_i(a) = p_i m E_i < \pi^2 E_i$. Then if $\bar{a} \in \arg \max_{a \in A} W(a)$ and $n$ is even, all agents must be in 2-cycles and so $\bar{a}_i \neq \bar{a}_j$ for all $i \neq j$, $i, j \in N$. If $n$ is odd, then it is not possible for all agents to be in 2-cycles. Since for any agent not in a cycle, $U_i(a) = 0$, to maximise ex-ante welfare either $n - 1$ agents are in 2-cycles and one agent is in no cycle or $n - 3$ agents are in 2-cycles and three agents are in a single 3-cycle. Which set of cycles is ex-ante welfare dominating depends on the value of $\pi$. Fix some $N = \{i, j, k\}$ such that $E_i > E_j > E_k$. A 3-cycle between these agents will ex-ante welfare dominate a 2-cycle if $\pi^3 E_i + \pi^3 E_j + \pi^3 E_k > \pi^2 E_i + \pi^2 E_j$, which holds only if $\pi > \frac{E_i + E_j}{E_i + E_j + E_k}$.

Then it is possible for a set of learning cycles containing $n - 3$ 2-cycles and a single 3-cycle to ex-ante welfare dominate a set of $n - 1$ 2-cycles. If this is the case then the maximum ex-ante welfare can be achieved by forming the 3-cycle between agents in $X = \{n - 2, n - 1, n\}$. To see this, under $\bar{a}$ let the 3-cycle form between all members of $X$ and under some other $a' \in A$ let a 3-cycle form between some $S \subset N$ such that $S \neq X$ and $|S| = 3$. In both $\bar{a}$ and $a'$, let all other agents not in $X$ or $S$ respectively, be in 2-cycles. If $W(a') > W(\bar{a})$, then,

$$\pi^2 \sum_{i \in N \setminus S} E_i + \pi^3 \sum_{i \in S} E_i > \pi^2 \sum_{i \in N \setminus X} E_i + \pi^3 \sum_{i \in X} E_i$$

$$\Rightarrow \pi > \frac{\sum_{i \in N \setminus X} E_i - \sum_{i \in N \setminus S} E_i}{\sum_{i \in S} E_i - \sum_{i \in X} E_i}$$

Since $\pi \in (0, 1)$:

$$\sum_{i \in S} E_i - \sum_{i \in X} E_i > \sum_{i \in N \setminus X} E_i - \sum_{i \in N \setminus S} E_i$$

But this implies, $\sum_{i \in N} E_i > \sum_{i \in N} E_i$ and so $W(a') \leq W(\bar{a})$.

Then if $\pi > \frac{E_{n-2} + E_{n-1}}{E_{n-2} + E_{n-1} + E_n}$, $W(\bar{a}) = \pi^2 \sum_{i=1}^{n-3} E_i + \pi^3 (E_n + E_{n-1} + E_{n-2})$. If $\pi \leq \frac{E_{n-2} + E_{n-1} + E_n}{E_{n-2} + E_{n-1} + E_n}$, $W(\bar{a}) = \pi^2 \sum_{i=1}^{n-1} E_i$.

The proof of Theorem 5 follows directly from Lemmas 6, 7 and 8. If $n$ is even then any stable equilibrium also maximises ex-ante welfare; $W(\bar{a}) = W(a^+)$. If $n$ is odd then $W(\bar{a}) \geq W(a^+)$ but the difference is at most the difference in utilities of agents $n$, $n - 1$ and $n - 2$; $\pi^2 (E_{n-2} + E_{n-1}) - \pi^3 (E_{n-2} + E_{n-1} + E_n)$.
5.2 NIDO

The comparison with the ex-ante welfare maximising $a$ is less clear in the NIDO environment, as it depends on the relative values of $\pi_k^i$ and $E_k^i$ for every $i \in N$, $k \in K$. If the values of $\pi_k^i$ and $E_k^i$ are particularly low in relation to other agents then it may no longer be the case that maximising the number of 2-cycles and minimising the number of 3-cycles is no longer optimal for ex-ante welfare.

Example 2: Suppose $N = \{1, 2, 3, 4, 5, 6\}$ and the values of $\pi_k^i$ and $E_k^i$ are as follows:

<table>
<thead>
<tr>
<th>$k$</th>
<th>$E_k^i$</th>
<th>$\pi_k^i$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>6</td>
<td>0.99</td>
</tr>
<tr>
<td>2</td>
<td>5</td>
<td>0.98</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>0.97</td>
</tr>
<tr>
<td>4</td>
<td>3</td>
<td>0.1</td>
</tr>
<tr>
<td>5</td>
<td>2</td>
<td>0.01</td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td>0.001</td>
</tr>
</tbody>
</table>

Let $o(a') = \{(1,2), (3,4), (5,6)\}$ and $o(a'') = \{(1,2,3), (4,5,6)\}$. Then $o(a')$ contains three 2-cycles and $o(a'')$ two 3-cycles. Following the proof of Lemma 5, $W(a')$ is the maximum welfare that can be achieved by any group of 2-cycles. The ex-ante welfare achieved under each set of cycles is:

\[
W(a') = (0.99 \times 0.98)\cdot(6 + 5) + (0.97 \times 0.1)\cdot(4 + 3) + (0.01 \times 0.001)\cdot(2 + 1) = 11.35
\]

\[
W(a'') = (0.99 \times 0.98 \times 0.97)\cdot(6 + 5 + 4) + (0.1 \times 0.01 \times 0.001)\cdot(3 + 2 + 1) = 14.12
\]

Then $W(a') < W(a'')$ and so the set of two 3-cycles ex-ante welfare dominates. But this is not always the case. Suppose instead the values of $\pi_k^i$ and $E_k^i$ are as follows:

<table>
<thead>
<tr>
<th>$k$</th>
<th>$E_k^i$</th>
<th>$\pi_k^i$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>6</td>
<td>0.9</td>
</tr>
<tr>
<td>2</td>
<td>5</td>
<td>0.8</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>0.7</td>
</tr>
<tr>
<td>4</td>
<td>3</td>
<td>0.6</td>
</tr>
<tr>
<td>5</td>
<td>2</td>
<td>0.5</td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td>0.4</td>
</tr>
</tbody>
</table>

Then $W(a') = 11.46 > W(a'') = 8.28$ and now the set of three 2-cycles ex-ante welfare dominates.
6 Remarks

Remark 1. If there exists a pair \( k \) and \( k' \) such that \( \pi_k > \pi_{k'} \) and \( E_{k'} > E_k \) then Theorem 4 need not hold.

Suppose the objects with the highest \( E_k \) have the lowest \( \pi_k \). That is, the probability an object will be of greater value than an agent’s own endowment is very low but if it is greater, then the expected value is very high. For example:

<table>
<thead>
<tr>
<th>( k )</th>
<th>( E_k )</th>
<th>( \pi_k )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>500</td>
<td>0.2</td>
</tr>
<tr>
<td>2</td>
<td>200</td>
<td>0.4</td>
</tr>
<tr>
<td>3</td>
<td>100</td>
<td>0.6</td>
</tr>
<tr>
<td>4</td>
<td>50</td>
<td>0.8</td>
</tr>
</tbody>
</table>

The highest expected utility agents 1 and 2 can achieve is \( U_1 = \pi_1 \pi_2 E_2 \) and \( U_2 = \pi_1 \pi_2 E_1 \) respectively. So for a set of learning cycles to be stable, it must contain the cycle \((1, 2)\). Since there are only 4 agents, it is easily shown that the unique stable set of learning cycles is \( o(a^*) = \{(1, 2), (3, 4)\} \). The ex-ante welfare for these cycles is \( W(a^*) = (0.2 \times 0.4)(500 + 200) + (0.6 \times 0.8)(100 + 50) = 128 \). However, if under \( a' \), \( o(a') = \{(1, 4), (2, 3)\} \) then \( W(a') = 160 > W(a^*) \). This contrasts with the conclusion of Theorem 4, since the unique set of stable learning cycles does not then coincide with the set of equilibrium learning cycles which maximise ex-ante welfare.

Remark 2. In IDO+, as with the set of stable learning cycles, the ex-ante welfare maximising equilibrium is also not unique.

The equilibrium with the highest ex-ante welfare must maximise the number of 2-cycles. However, if \( \pi_k = \pi^* \) for all \( k \in N \) then any set of 2-cycles which contains all \( k \in N \) if \( n \) is even (or all \( k \in N \setminus \{n\} \) if \( n \) is odd) is ex-ante welfare maximising. For any \( a \) which maximises the number of 2-cycles:

\[
W(a) = \begin{cases} 
\pi^2 \sum_{k \in N} E_k & \text{if } n \text{ is even} \\
\pi^2 \sum_{k \in N \setminus \{n\}} E_k & \text{if } n \text{ is odd}
\end{cases}
\]

\(^4\)I would like to thank Maris Goldmanis for providing this example.
7 Conclusion

If information is required in order to enable trade in a housing problem then it impacts the nature of any exchange that may eventually take place. The risk that an agent may find the object they investigate to be unacceptable means that large cycles cannot exist in equilibrium. This applies in all the environments discussed here, when the objects are both ex-ante identical and non-identical. Small, two agent cycles then arise endogenously as a result of learning which contrasts with applications where the restriction to two agent cycles is exogenously applied (Roth et. al (2004), (2005)). The endogenous information acquisition also has a notable impact on the relationship between stability and ex-ante welfare. In all environments covered here, any set of learning cycles which is stable will also be a set of learning cycles which maximises ex-ante welfare in equilibrium. When objects are ex-ante non-identical the number of sets of cycles that this applies to is significantly reduced. There is a unique set of stable learning cycles and a unique set of ex-ante welfare maximising, equilibrium learning cycles and the two sets perfectly coincide. Furthermore, if those objects differ only in $E_k$ (their potential value if found acceptable) then the set of stable learning cycles also maximises or come very ‘close’ to maximising ex-ante welfare overall. It will differ in at most the utility of the three agents endowed with the worst three objects and even then only if the number of agents is odd and the probability of finding an object acceptable is high. Since this holds irrespective of the number of agents in the problem, the relative difference in ex-ante welfare between the stable and ex-ante welfare maximising sets of learning cycles is either zero or negligible when the number of agents is large. Then the societal goal of maximising ex-ante welfare is not incompatible with the aim of taking individual incentives into account.

References


