# Essays in Matching and Information Acquisition

Submitted by

Annika Johnson

for the degree of Doctor of Philosophy in Economics of Royal Holloway, University of London in 2018.

# Declaration of Authorship

I, Annika Rhiannon Lindsay Johnson, hereby declare that this thesis and the work presented in it is entirely my own. Where I have consulted the work of others, that is clearly stated.

20th September 2018

## 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' endowments must also be introduced. For the instance in which each agent has only the resources to learn about one other object, I show how agents' decisions over what to learn about restricts the size of the trading cycles. Large cycles are risky and so no cycle containing more than two agents can exist in equilibrium. I then give the conditions under which stability and ex-ante welfare maximisation are mutually compatible objectives. If objects are 'well ranked' in the sense that the objects of highest value are also more likely to be acceptable, then any profile of agents' learning decisions which is stable is also an ex-ante welfare maximising equilibrium. Introducing a time dimension which allows agents to choose not only what to learn about, but when, does not rule out equilibria in which all agents learn quickly and at the same time. The same learning pattern as observed when agents are forced to make the decision simultaneously, remains an equilibrium when this restriction is removed. Even offering the agents the opportunity to learn carefully, one by one, making decisions with the most information they can does not prevent the rush to learn at the same time as others in equilibrium. The information acquisition problem is by no means unique to unilateral matches and so I also consider the particular allocation mechanism used for university entry in the UK. The combination of allowing applications to be submitted to only two institutions and students only being able to acquire information on their grades after submission results in poorly assortative allocations where the best students are unable to attend higher-ranked instiutions.

# Acknowledgements

This work would not have been possible without support, guidance, encouragement and cooperation from a great many people. Its creation, development and funding from the Economic and Social Research Council have benefited greatly from the comments of Dan Anderberg, Sourav Bhattacharya, Manolis Galenianos, Maris Goldmanis, Michael Mandler, Klaus Ritzberger, Juan Pablo Rud, Rob Sauer, Andrew Seltzer, Ija Trapeznikova and Jonathan Wadsworth, amongst all other academic and professional services staff in the Department of Economics at Royal Holloway. Special mention must also go to fellow candidates Florentyna Farghly and Nathan Woolley. I am indebted to the patience of Sophie Bade whose support has changed my life's direction so completely.

I am also grateful for the many conference attendees who have taken time to read or discuss this work including Subir Bose, Eddie Dekel, Anne Go, Stephen Jagau, Jinpeng Ma, Foivos Savva, Andre Stenzel and Christopher Wallace and audience members at the 2016 Africa Meeting of the Econometric Society, the 29th International Conference in Game Theory at Stony Brook, the 4th International PhD Conference at the University of Leicester, the 2018 RES Annual Conference and the 2018 RES Junior Symposium.

I could not have reached this point without every opportunity provided by my parents, Lyall and Mark Johnson and cannot underestimate the value of the support from all the Johnson, Lambrick and Thornthwaite families. Without the influence of illustrative bargaining pie created by Perry Johnson or the example set by the family's leading academic, Dr David Lambrick this process may not even have begun. My grandparents' wisdom has been vital, particularly given how hard I know both Jeanette Lambrick and Pauline Johnson fought for their education. When unsure of one's place in the world, my grandfather Lawrie Lambrick was also able to provide a homing beacon. He died before I could complete this work, but it was his advice I took on the use of 'him' and 'her' as I consider his life's work left him far better placed to reach a thoughtful conclusion. Finally, I owe everything to David Thornthwaite, without whom this journey would not have started, continued, or survived long enough to reach all that follows.

# Contents

# Introduction

Matching markets, such as college admissions, human kidney exchange, graduate job assignment and house allocation have long been the subject of investigation. In such markets, indivisible goods and services must often be allocated without monetary transfers. Two of the most famous mechanisms designed to allocate goods or assign places in such circumstances are Gale's Top Trading Cycles (Shapley and Scarf (1974)) and Deferred Acceptance (Gale and Shapley (1962)). These mechanisms use agents' preferences to create a match with desirable properties such as Pareto optimality, strategyproofness, individual rationality and stability. However, the requirement that agents always know their preferences over all options available to them can be an onerous one as their ability to acquire that information may easily be limited in some way, either through high costs, restricted resources or institutional constraints. I examine the impact of an information acquisition requirement on 'high stakes' markets such as kidney exchange and within the UK's higher education system to determine how such requirements may affect the outcome that is realised.

Chapters 1 and 2 explore models of 'high stakes' unilateral matching markets with endogenous information acquisition. Agents are fully informed about their own endowments: a patient in need of a kidney exchange will already know she is incompatible with her willing living donor and a council tenant will be living in her current property and know the local area intimately. However, although some cursory information may be available, each agent knows far less about the other endowments available for exchange and is unwilling to give up her own without knowing she is getting something better in return. This is particularly clear in the case of living donor kidney exchange, since a doctor would not consent to transplant an organ without first ensuring compatibility. For the case where each agent is able to investigate just one other object (for example, due to a back up at the lab or prohibitively high costs), I analyse the impact of the endogenous information acquisition on the exchanges which can take place under Gale's Top Trading Cycles. I identify conditions under which the outcome of any stable learning profile is the same (or negligibly close to) the best society can achieve were the process to be managed by a designer able to dictate actions to each agent.

Chapter 2 builds on a version of the model in Chapter 1, but where agents are not required to make their decision simultaneously. Instead, agents are free to choose not only which object they investigate but when to do so. This means the equilibrium outcomes begin to differ from that which a social planner might wish to impose. Fully sequential learning, where agents learn one after the other, prevents wasteful learning as few agents as possible are forced to investigate an object without first knowing the preferences of its owner. However, fully sequential learning can also create situations where individuals lose out and by the time they come to choose which object to investigate, there are no more partners to exchange with. The threat of being left alone means that more agents learn simultaneously than is socially desirable.

Chapter 3 turns to a two-sided market where students are assigned places at university. In the UK, the University and College Admissions Service (UCAS) requires students to begin their application process not only before learning their exam results but before sitting their exams. Universities base their acceptance decisions on the students' grades but students are able to acquire this information only after submitting a preference to UCAS. In addition, students are limited to submitting their application to only two universities. Despite the UCAS system having features otherwise in common with the student-proposing deferred acceptance mechanism, this combination of factors can have severe consequences for the assortativity of the resulting match, where the highest-ranked institutions are not attended by the best students.

CHAPTER 1:

# Top Trading Cycles in Endogenous Information Acquisition

## 1.1 Introduction

Consider a unilateral matching problem in which each agent is endowed with one object that they are free to exchange with other such agents. Common examples of such problems include living donor kidney exchange, council house exchange programmes and college dorm swapping policies. In each of these cases the agent is fully informed about her own endowment: a patient will already know she is incompatible with her donor, a council tenant will be living in her current property and the college student will know both her dorm and roommates. However, though some cursory information may be available, each agent knows far less about the other endowments available for exchange. This is more problematic for some scenarios than others. The college student may be willing to take a gamble on a new room but the council tenant is unlikely to want to uproot her life for a property she has not viewed in an area she does not know. For the patient in need of a new kidney, the consequences of a transplant with an incompatible organ are deadly. In such cases, and whether it be due to prohibitively high costs or limited resources, resistance to exchange without complete information only becomes more acute as the agent's ability to investigate diminishes. But if agents will only exchange for objects they know to be compatible then they must carefully consider which objects they want to investigate; if they cannot learn everything then they do not want to waste opportunities on dead ends. In such cases, one must take into account the organisation of agents' investigations alongside the design of the exchange process itself. This paper explores such a model in which agents must acquire information about objects before committing to an exchange. The importance of including learning in the model in this way is most easily demonstrated within the context of kidney exchange.

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 third case in which each patient can find only an incompatible living donor. If the mechanism used to exchange these donor organs is to be individually rational, Pareto optimal and strategyproof then it must be Gale's Top Trading Cycles (GTT) (Shapley and Scarf (1974)) since it is the unique mechanism satisfying these properties (Ma (1994), Roth (1982)). However, before any exchange can take place the organs must be tested; a kidney can never be safely transplanted without first confirming its compatibility with the patient. This means concentrating solely on the exchange process may not be sufficient to ensure an optimal outcome. 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. A model which incorporates the need to learn before exchange may reduce the incidence of failed matches.

This paper explores a model in which each of a finite set of agents is endowed with a single indivisible object. Each agent knows about their own endowment but has little information on anyone else's object; they know the probability they will find any given object to be acceptable, that is, to have a higher value than their own endowment and the expected value conditional on that being the case. Agents are only willing to exchange for objects they are certain to strictly prefer to their endowment. In this sense, the model is closer to the 'high-stakes' kidney example above than the 'low-stakes' dorm swap. In order to ascertain whether they strictly prefer an object or not, each agent is able to simultaneously investigate one object (other than their own endowment). A learning profile details which object each agent investigates and upon learning the outcome of these investigations, objects can be exchanged. Since agents will only exchange for objects certain to be strictly preferred to their own endowment, agents can only exchange for investigated objects and so the pattern of investigations affects which trades may eventually take place.

The question is then over how the investigations and exchanges should be organised. When information is known, the obvious candidate is GTT since it is the unique individually rational, Pareto optimal and strategyproof mechanism, and so one option is to naïvely continue using GTT with incomplete information. In this case, since GTT is strategyproof, each agent need only consider which object to investigate. The consequences of this decision, however, have significant impact on trade in equilibrium. A trading cycle can occur only if every agent in that cycle learns she prefers the object she has investigated to her own endowment. The more agents involved in the potential trading cycle, the higher the risk at least one test, and consequently the trade itself, will fail. Potential trading cycles such as this are known here as learning cycles. Large learning cycles create an incentive for at least one agent to 'short circuit' the cycle and create a smaller one in which she can trade for an object of weakly greater expected value but with strictly higher probability of exchange. As a result, in equilibrium only small learning cycles of two agents exist and so any given trading cycle under GTT will involve at most two agents. This is in contrast with the kidney exchange literature, where due to the constraints of hospitals, exchange is often exogenously restricted to two agent-donor pairs (Roth et al. (2004), (2005)). In this paper, since the two-agent cycles arise endogenously, such a restriction would leave the results unaffected.

There is little reason, however, to assume that the naïve GTT approach will work well in an incomplete information environment and so other organisational structures should be considered. Many alternatives exists and so here I examine two extremes with respect to the role of the designer. At one extreme, designers are able to dictate which test each agent performs and which exchanges take place. In this case, any learning profile can be enforced and so learning cycles (and in turn, trading cycles) may consist of any number of agents and their endowments. At the other extreme, the designer has no input and agents are left to coordinate amongst themselves. In this case I consider the set of stable learning profiles. The stable set has an important feature in common with the set of equilibria in the naïve GTT approach; learning cycles contain only two agents. If not, then any two agents with the best two endowments in any given learning cycle will be strictly better off by investigating each other's respective endowments. Unlike the equilibria of the naïve GTT approach, if a learning profile is stable then the maximum number of agents possible must be in learning cycles. It is also important which agents are in which cycle; the two agents with the best endowments must be paired together, otherwise they both have the incentive to deviate. In the same way, the agents with the next two best endowments must also be paired together and so on. When no two objects are the same, this means a unique set of learning cycles forms under any stable learning profile.

When all investigations and exchanges can be dictated to agents, the size of the learning cycles is uninhibited by the same threat of deviation, but nevertheless large learning cycles can be undesirable. If the designer's goal is to maximise ex-ante welfare, then large learning cycles still carry unnecessarily high risk. For this reason, any learning cycle arising under an ex-ante welfare maximising learning profile will consist of at most three agents. Which learning cycle each particular agent is in and whether that learning cycle consists of two or three agents depends on the given set of objects but, for a class of objects known here as 'strictly well ranked,' there is a clear structure to the learning cycles. Objects are strictly well ranked if they are ordered in the same way whether by the probability they will be found acceptable or by their potential value. When objects are strictly well ranked then agents must be grouped in cycles in descending order of the potential value of their endowments in order to maximise ex-ante welfare. This is similar to the structure of the unique set of stable learning cycles discussed above, with the exception that some of those learning cycles may contain three agents. If objects are strictly well ranked, however, then three-agent cycles will occur only if there are significant differences in the probability of objects being acceptable. Such differences divide endowments into groups of 'good' and 'junk' objects. The designer will create three-agents cycles to avoid condemning an agent with a good endowment to a learning cycle with one or more junk objects that jeopardise the probability of exchange. If there are no junk objects, then these three-agent cycles are not required and it is the set of two-agent learning cycles which maximises ex-ante welfare. In this case, strategic concerns do not compromise the welfare goal; any stable learning profile yields the highest level of ex-ante welfare possible. Since this is the highest ex-ante welfare that can be reached through two-agent cycles alone, the stable learning profile also compares favourably with the equilibria of the naïve GTT approach. When objects are strictly well ranked, and regardless of the presence of junk objects, no equilibrium can generate ex-ante welfare which exceeds that of the stable learning profile.

The literature on endogenous information acquisition within the matching field is not extensive. Bade (2015) shows that when learning is costly, serial dictatorship is the unique ex-ante Pareto optimal, strategyproof and nonbossy 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 compromising individual rationality. In the kidney exchange literature, Dickerson et. al (2013) use random graph models to 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.

### 1.2 Environment

A finite set of agents $N = \{1, ...n\}$ is such that each agent $i$ is endowed with object $i$ in the set $K = \{k_1, ..., k_n\}$. An agent values his own endowment at zero but does not know his value for any other object, only that 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 each agent prefers to keep his own endowment. More formally, there exists a state space $\Omega$ consisting of profiles of values $\omega = (\omega^i_k)_{i\in N, k\in K}$, where $\omega^i_k$ is the value of object $k$ to agent $i$ in state $\omega$. Since agents value their own endowment at zero, $\omega^i_i = 0$ for all $\omega \in \Omega$. For all other objects, $\omega^i_k$ is an independent draw from some distribution $f^i_k$ with some support not containing zero[1]

[1] This is to prevent indifferences in what follows.

and such that $\mathbb{E}(\omega_k^i) < 0$ for all $i \in N, k \in K$. This means firstly, Agent $i$ knows the probability that an object $k \neq i$ will be **acceptable**, $\pi_k^i = \pi(\{\omega \mid \omega_k^i > 0\})$, which is the probability that $\omega_k^i$ is higher than his own endowment. Secondly, it means Agent $i$ also knows the **conditional value** of the object $E_k^i = \mathbb{E}(\omega_k^i | \omega_k^i > 0)$, which is the expected value of $k$ conditional on the object being acceptable. For the majority of the paper, unless otherwise specified, I assume agents are ex-ante identical while objects may differ. In this case, $f_k^i = f_k$ for all $i \in N, k \in K$ such that $i \neq k$ and $\omega_k^i$ is an iid draw from $f_k$. Let $E_k := E(\omega_k^i \mid \omega_k^i > 0)$ and $\pi_k := \pi(\omega_k^i > 0)$ for all $i \in N, k \in K$. Two objects $j, k \in K$ are considered to be **ex-ante identical** if $\pi_k E_k = \pi_j E_j$. For simplicity, when at least two objects are not ex-ante identical and without loss of generality, I assume $\pi_1 E_1 \geq \pi_2 E_2 \geq ... \geq \pi_n E_n$. In this way, Agent 1 is always endowed with the best object, Agent 2 is endowed with either the best or second best object and so on until Agent $n$ who is endowed with the worst object.

Each agent $i$ learns the value $\omega_k^i$ of one object $k \neq i$. A **learning profile** is a strategy profile[2] $a = (a_i)_{i \in N}$ such that $a \in A = \times_{i \in N} A_i$. 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. The set of learning cycles that forms under $a$ is $o(a) = \{o_1(a), ...o_\nu(a)\}$. For any $j \in \{1, ...\nu\}$, let $\phi_j(a) = \{k_1, ...k_m\}$ be the set of agents in the learning cycle $o_j(a) = (k_1, ...k_m)$. The set of all agents in learning cycles under $a$ is $C(a)$ and $B(a) = N \setminus C(a)$ is the set of agents not in learning cycles under $a$. The set of all agents in $m$-cycles under $a$ is $C_m(a)$ and $B_m(a) = N \setminus C_m(a)$.

Although the probability an an agent will find Object $k$ acceptable ($\pi_k$) and the conditional value of Object $k$ ($E_k$) are drawn from the same distribution ($f_k$) for all agents $i \neq k$, given the state and each agent having chosen which object to investigate, each

[2]For the moment, attention is restricted to pure strategies, where each agent investigates a particular object with certainty. Relaxing this and allowing for mixed strategies can, under some circumstances, affect the results in this chapter. These effects are explored in Appendix D.

agent may have different (ex-post) preferences over the set of objects. Agent $i$'s transitive ex-post preference relation under any given learning profile $a$ and set of object values $\omega$ is $R_i(a,\omega)$, where $kR_i(a,\omega)k'$ if Agent $i$ weakly prefers $k$ to $k'$. If $kR_i(a,\omega)k'$ but not $k'R_i(a,\omega)k$ then it is denoted $kP_i(a,\omega)k'$. The ex-post preference profile is $R(a,\omega) = (R_i(a,\omega))_{i\in N}$ and the set of all possible ex-post preference profiles is $\mathcal{R}$. If $i$ chooses to learn about object $k$ then $kR_i(a,\omega)i$ if $\omega_k^i > 0$. 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 $kR_i(a,\omega)i$. There may, however, be many untested objects $k' \neq i$ such that $iR_i(a,\omega)k'$ and it is possible that an agent is indifferent between two such objects. A matching is a bijection $\mu : N \to K$. The set of all matchings is $\mathcal{M}$. Under any given learning profile $a$ and set of object values $\omega$, a matching is individually rational if $\mu(i)R_i(a,\omega)i$ for all $i \in N$ and a matching $\mu'$ Pareto dominates $\mu$ if $\mu'(i)R_i(a,\omega)\mu(i)$ for all $i \in N$ and $\mu'(i^*)P_{i^*}(a,\omega)\mu(i^*)$ for at least one $i^* \in N$. If a matching is not Pareto dominated then it is Pareto optimal. A mechanism, $M : \mathcal{R} \to \mathcal{M}$, is individually rational and Pareto optimal if it always results in an individually rational and Pareto optimal matching. A mechanism is strategyproof if $M(R(a,\omega))(i)R_i(a,\omega)M(R_i'(a,\omega), R_{-i}(a,\omega))(i)$ for all $i \in N, R_i'(a,\omega)$ so that no agent $i$, whose truthful preferences are $R_i$, can be matched with an object they strictly prefer to $M(R(a,\omega))(i)$ by stating some alternative preference profile $R_i' \neq R_i$.

## 1.3 Solution Concepts

Given the environment, the questions remains as to how both investigations and exchanges should be organised. There are many ways in which this task could be executed and so I focus on three approaches. The first focusses on the Gale's Top Trading Cycles ($GTT$) mechanism. Without the need to learn about other objects, $GTT$ would be the sole candidate since it is the unique individually rational, Pareto optimal and strategyproof mechanism. With the additional information acquisition stage, however, simply declaring $GTT$ as the mechanism to be used in the matching stage may affect agents' decision over which objects to investigate. To this end, I analyse the range of equilibria which may arise when agents first choose which objects to investigate and then, having determined the results of those tests, arematched via $GTT$.

To establish the extent to which the equilibria which arise under $GTT$ are desirable, I compare these results to two further approaches. These two approaches capture the range of possible outcomes with respect to input from a designer (or dictator). Under the second approach I focus on the set of stable learning profiles. These are the learning profiles which arise when agents are permitted to decide on both which object they choose to learn about and which objects they would like to exchange for, without input from a designer. Under the third approach I move to the opposite end of the spectrum and analyse the best possible outcomes achievable when a dictator is able to determine both the investigation and match.

**Approach 1: Equilibrium**

Under this approach, agents first choose $a_i$ and then, having learned the value of their investigated objects, declare their preferences $R(a, \omega)$ and are matched via the Gale's Top Trading Cycles mechanism, $GTT : \mathcal{R} \to \mathcal{M}$ which works as follows:

> **Step $r$:** Each unmatched agent $i$ points at his most preferred object according to $R_i(a, \omega)$ 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.

$GTT$ 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 $iP_i(a,\omega)k$ and $iP_i(a,\omega)k'$. Therefore, $i$ will always point at and be matched to his own object before needing to choose between objects to which he is indifferent. Even in this domain, $GTT$ remains the unique individually rational, Pareto optimal and strategyproof mechanism.[3]

$GTT$'s property of strategyproofness plays an important role here. Once agents have completed their investigations, it is a weakly dominant strategy for an agent to truthfully report her preferences; an agent cannot induce a better outcome for herself by misrepresenting her true preferences. Having reached the matching stage of the problem, the equilibrium is determined through weak dominance. Given the use of the strategyproof $GTT$ in this stage, agents' expected utilities are determined largely by the learning profile arising in the first 'learning' stage where each agent chooses which object to investigate.

Agent $i$'s expected utility firstly depends on the probability $i$ is matched with $a_i$: $\pi(\{\omega \mid GTT(R(a,\omega))(i) = a_i\})$. In order for $i$ to be matched to $a_i$, it not only needs to be that $i$ finds $a_i$ acceptable but that a trading cycle can form between a group of $S$ agents such that $i \in S$. For this to be the case, every agent in $S$ must investigate the endowment of a different agent in $S$ and it must that $\omega^j_{a_j} > 0$ for all $j \in S$. Secondly, $i$'s expected utility depends on the conditional value of the object agent $i$ chooses to learn about, $E^i_{a_i}$. Agent $i$'s expected utility is the product of these two terms:

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

Agent $i$ will choose $a_i$ to maximise her ex-ante expected utility. Since agents simulta-

[3]The proof in Bade (2019) applies to the domain $\mathcal{R}$. The proof in Ma (1994) is not applicable since it uses the full domain of preferences.

neously make their decision over which object to investigate and agents are assumed to follow their weakly dominant strategy of revealing their true preferences in the matching stage (when using the strategyproof $GTT$ mechanism) a learning profile $a$ is a Bayes-Nash equilibrium if $U_i(a) \geq U_i(a'_i, a_{-i})$ for all $i \in N$.

**Approach 2: Stability**

The second approach examines what is possible when there is minimal (no) input from a designer and agents are able to act strategically at both the learning and matching stage of the problem. Comparison of the outcomes under this approach and the equilibrium approach discussed above is simplified since any differences can only occur as a result of what happens in the learning stage and not the matching stage. This is because, having completed all investigations and learned the value of their investigated objects, any mechanism, $M$, used by the agents which is individually rational and Pareto optimal must yield the same match as would arise under $GTT$: $M(R(a,\omega)) = GTT(R(a,\omega))$[4]. The properties of individual rationality and Pareto optimality are important to retain in this approach since agents cannot be forced into matches by some authority and by the same token, cannot be prevented from seeking Pareto improvements which make some subset of agents strictly better off. Since $M(R(a,\omega)) = GTT(R(a,\omega))$ and $GTT$ is strategyproof, $M$ must also be strategyproof and so it is a weakly dominant strategy for any agent to truthfully report her preferences.

Since the matching stage is identical under both Approaches 1 and 2, any difference must occur in the learning stage. In Approach 1 there was a Bayes-Nash Equilibrium if no single agent could benefit by deviating from her learning decision. In Approach 2, I focus on learning profiles where no subset of agents can benefit by deviating from their learning decisions: A learning profile $a$ is **stable** if under $a$ there is no subset $S$ of $m$ agents who are not in a given $m$-cycle but all agents in $S$ would have a strictly higher

[4]A proof is provided in Appendix C

expected utility if they were. If this holds for all subsets of $m$ agents such that $m = 2$ then the learning profile is **pairwise stable**. Since $M(R(a, \omega)) = GTT(R(a, \omega))$, Agent $i$'s expected utility under profile $a$, $U_i(a)$ can be expressed in the same way as detailed in Approach 1.

**Approach 3: Ex-ante Welfare Maximisation**

At the opposite end of the scale to the stability approach (where decisions are decentralised and agents are able to decide both which object to investigate and match with), in this approach I examine the learning profiles which are most socially desirable as measured by **ex-ante welfare**, $W(a) = \sum_{i \in N} U_i(a)$. A dictator is able to decide which investigation each agent should complete. For example, a centralised health service may decide which crossmatch blood tests to perform between potential patients and donors. The dictator is able to determine tests and to decide exchanges subject to the minimal requirements of individual rationality (ie. the dictator cannot compel a individual to give up their kidney against their will) and Pareto optimality (the dictator cannot deny transplants which do not disadvantage other patients). In this sense, the dictator seeks the learning profile $a^*$ such that $a^* \in \arg\max_{a \in A} W(a)$. As with the two previous approaches, since the matching stage is individually rational and Pareto optimal, it must be that $M(R(a^*, \omega)) = GTT(R(a^*, \omega))$. Since $GTT$ is strategyproof, it is a weakly dominant strategy for each agent to report her true preferences given the information acquired in the learning stage. This means that the utility of each individual Agent $i$ under the learning profile $a^*$ can be calculated as detailed in Approach 1: $U_i(a^*) = \pi(\{\omega \mid GTT(R(a^*, \omega))(i) = a_i^*\}) \cdot E^i_{a_i^*}$.

Given some preference profile, all three approaches will result in an identical matching between agents and objects. However, the three approaches can differ drastically in terms of permissible learning profiles. The next three sections characterise these learning profiles and compare the ex-ante welfare achievable under each approach.

## 1.4 Properties of Stable Learning Profiles

Depending on the set of agents and their endowments, the characteristics of learning profiles under each of the three approaches may differ. In this section I characterise the set of stable learning profiles as it provides a useful starting point when also considering any equilibrium or ex-ante welfare maximising learning profile. Section 1.3 introduced the definitions of stability and pairwise stability. Before continuing, note that since pairwise stability implies stability, it is sufficient to focus on pairwise deviations in the discussion which follows.

**Lemma 1.** *If $a$ is pairwise stable then $a$ is stable*

*Proof.* Suppose $a$ is pairwise stable but not stable. Then there exists an $a'$ under which an $m$-cycle, $(b, c, ..., d)$ forms between some $S = \{b, c, ..., d\}$ such that $U_k(a'_S, a_{-S}) > U_k(a)$ for all $k \in S$. Since $a$ is pairwise stable, $|S| > 2$. W.l.o.g let $b < d < k$ for all $k \in S \setminus \{b, d\}$. Since $U_k(a'_S, a_{-S}) > U_k(a)$ for all $k \in S$, $U_b(a'_S, a_{-S}) = \prod_{k \in S} \pi_k E_c > U_b(a)$ 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 $U_b(a''_{\{b,d\}}, a'_{S \setminus \{b,d\}}, a_{-S}) = \pi_b \pi_d E_d$ and $U_d(a''_{\{b,d\}}, a'_{S \setminus \{b,d\}}, a_{-S}) = \pi_b \pi_d E_b$. 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 $a$ is pairwise not stable, which is a contradiction. □

### 1.4.1 The bi-cycle set, $A^S$

In order to characterise the set of all stable learning profiles, I first characterise a set of learning profile $A^S$ as the set of learning profiles which meet conditions **I**, **II** and **III** below. These conditions restrict learning profiles to those which contain only 2-cycles and also limits the range of agents each agent can be in a learning cycle with. The size of $A^S$ varies depending on the values of $\pi_i$ and $E_i$ for each object but after introducing the set $A^S$, I introduce the set $A^\circ \subset A^S$ which produces a fixed set of learning cycles and exists for any given set of objects. In the next section, I then show that a learning profile is stable if and only if it is contained in the set $A^S$.

Define the set $A^S$ such that a learning profile, $a$ is in the set $A^S$ if and only if:

**I**: If $n$ is even then all agents are in 2-cycles. If $n$ is odd then all agents are in 2-cycles except for some $i^*$ such that $\pi_{i^*}E_{i^*} = \pi_n E_n$.

and for any pair of agents $i, j$ such that $\pi_i E_i > \pi_j E_j$ and $(i, j) \in o(a)$:

**II**: If there are two agents $i', j'$ such that $(i', j') \in o(a)$ and $\pi_i E_i = \pi_{i'} E_{i'}$ then $\pi_{j'} E_{j'} \geq \pi_i E_i$.

**III**: There is no agent $j^*$ such that $\pi_i E_i > \pi_{j^*} E_{j^*} > \pi_j E_j$.

These three conditions mean that $A^S$ may contain many different learning profiles that each generate different sets of learning cycles. The number of learning profiles in $A^S$ depends in part on the the number of objects which are ex-ante identical. Let $\gamma = \{\gamma^1, ...\gamma^{\bar{r}}\}$ form a partition on $N$ such that for any $i \in \gamma^r$ and $i' \in \gamma^{r'}$, $\pi_i E_i > \pi_{i'} E_{i'}$ if and only if $r < r'$. Then all the objects in any given $\gamma^t$ are ex-ante identical. The following example illustrates some of the learning profiles and their associated learning cycles that are and are not in $A^S$ for a given set of agents and objects.

**Example 1.** Let $N = \{1, ...8\}$ and $\pi_1 E_1 > \pi_2 E_2 > \pi_3 E_3 = \pi_4 E_4 > \pi_5 E_5 = \pi_6 E_6 = \pi_7 E_7 > \pi_8 E_8$. For this set of objects, $A^S = \{a^1, a^2, a^3\}$ as illustrated in Figure 1. In contrast, Figure 2 shows three learning profiles $a^4, a^5, a^6 \notin A^S$. Clearly $a^4 \notin A^S$ since it contains two $m$-cycles such that $m > 2$, violating condition **I**. Under $a^5$, $\pi_3 E_3 = \pi_4 E_4$ and $\pi_3 E_3 > \pi_5 E_5 = \pi_6 E_6$, violating **II**. Finally, $(1, 3) \in o(a^6)$ but $\pi_1 E_1 > \pi_2 E_2 > \pi_3 E_3$ and so **III** does not hold. Also note that $a^4$, $a^5$ and $a^6$ each only violate one of the three conditions.

(a)

(b)

(c)

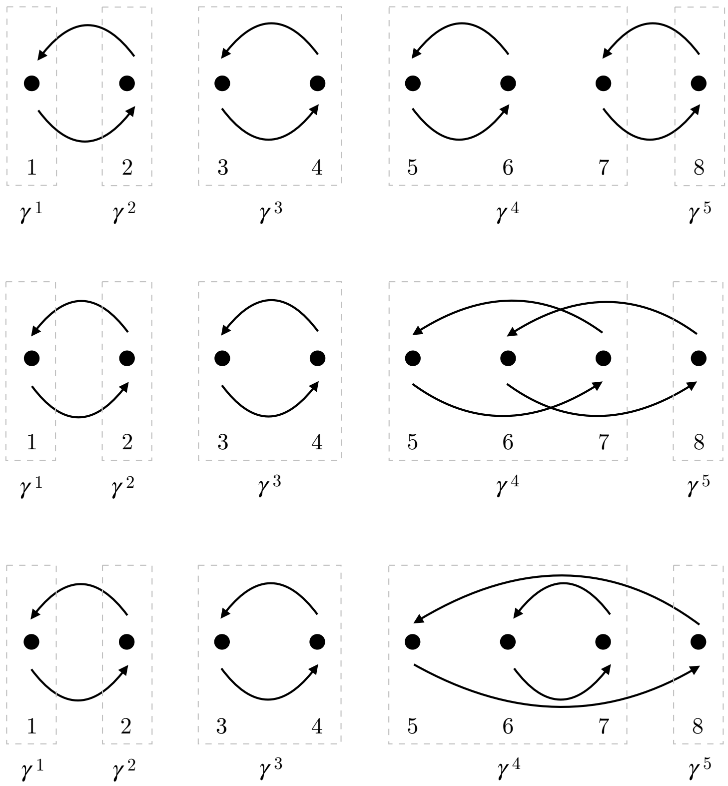



Figure 1: The set of learning cycles generated under the three learning allocations in $A^S$: (a) $a^1$ and (b) $a^2$ and (c) $a^3$.

The size of $A^S$ will vary depending on the number of different values of $\pi_i E_i$ and the number of agents with each of those values. There is, however, a set of learning allocations which is always in $A^S$.

**Observation 1.** *Define* $A^\circ := \{a \mid a_i = i - 1 \text{ } if\, i \in N \text{ } even \text{ }, a_i = i + 1 \text{ } if\, i \in N \setminus \{n\} \text{ } odd \}$, *then* $A^\circ \subseteq A^S$.

An example of a learning profile in the set $A^\circ$ is shown in Figure 1(a). There are eight agents, all in in 2-cycles so it meets condition **I**. Since every even numbered agent $i$ is in a 2-cycle with $i - 1$ it also meets conditions **II** and **III**. The learning profiles in

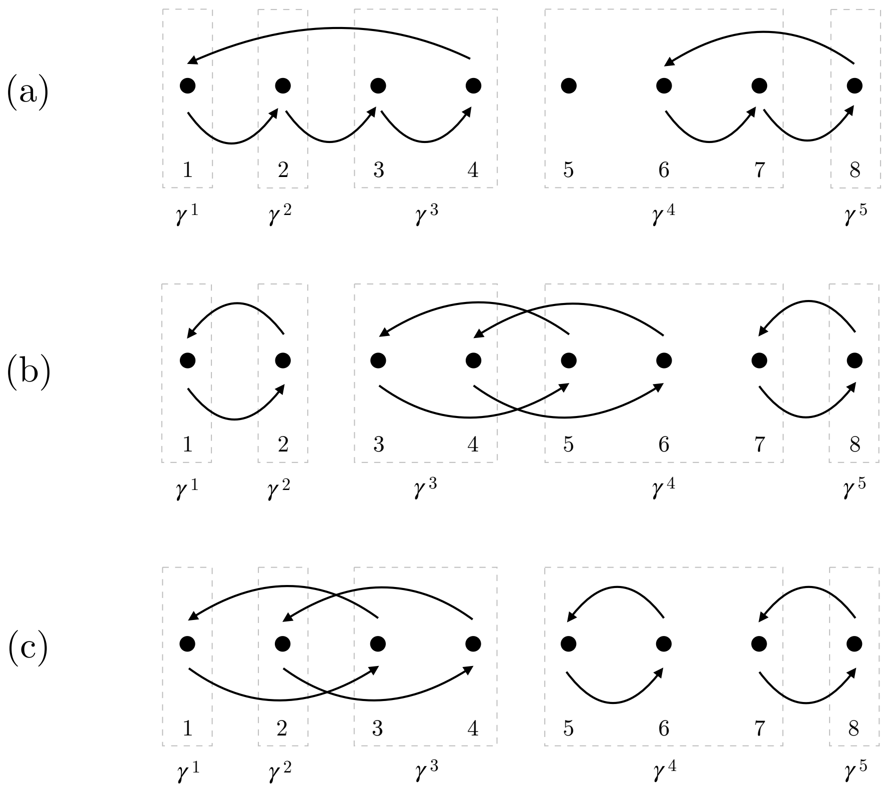



Figure 2: The set of learning cycles generated under learning allocations not in $A^S$: (a) $a^4$ and (b) $a^5$ and (c) $a^6$.

$A^\circ$ can be constructed for any set of agents and objects and since any $a \in A^\circ$ satisfies conditions **I**, **II** and **III**, $A^\circ \subseteq A^S$. Since the number of agents is even, the learning profile illustrated in Figure 1(a) is in fact the only learning profile in $A^\circ$ for this set of objects. The definition of $A^\circ$ prescribes $a_i$ for all $i \in N$ if $n$ is even and for all $i \in N \setminus \{n\}$ otherwise so $|A^\circ| = 1$ if $n$ is even. If $n$ is odd the learning profiles in $A^\circ$ can differ only in $a_n$ and so $|A^\circ| = n - 1$ if $n$ is odd. However, as shown in Figure 3, even though $A^\circ$ may contain multiple learning profiles, the learning cycles arising from those profiles is unique.

**Observation 2.** *For any given set of objects, the set of learning cycles $o(a)$ that can form under any $a \in A^\circ$ is unique.*

If $n$ is even then Observation 2 follows from the fact that $|A^{\circ}| = 1$. If $n$ is odd then since there is no $i \in N$ such that $a_i = n$ under any $a \in A^{\circ}$, $n$ is never in a learning cycle under any $a \in A^{\circ}$. By the definition of $A^{\circ}$ , $a_i$ is constant across all $a \in A^{\circ}$ for all $i \in N \setminus \{n\}$. As such, the set of learning cycles which forms is:

$$o^{\circ} = o(a) = \{(1, 2), (3, 4), ...(k, k + 1)\}$$

for all $a \in A^{\circ}$, where $k = n - 1$ if $n$ is even and $k = n - 2$ otherwise.

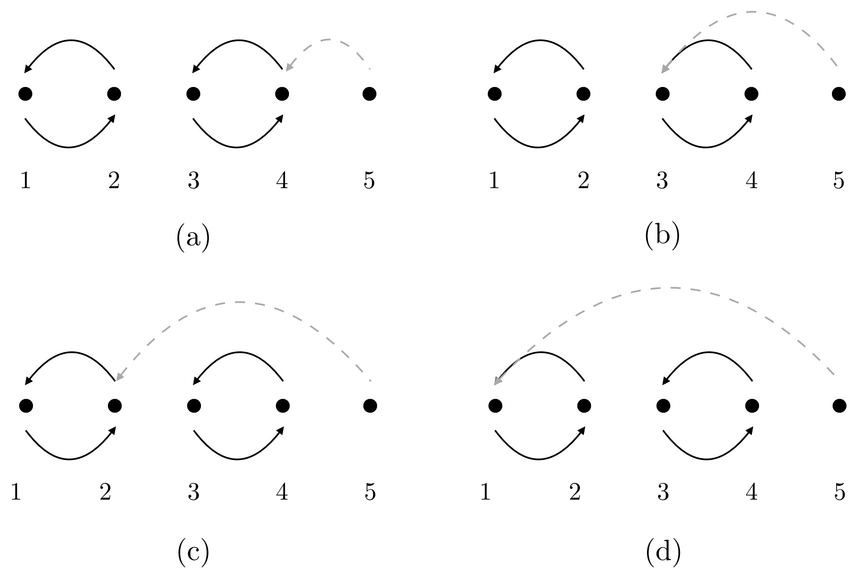



Figure 3: Learning profiles in $A^{\circ}$ when $n = 5$

In general, $A^S$ will contain learning profiles other than those in $A^{\circ}$. Figure 1 shows three sets of learning cycles for the same agents, all of which result from learning profiles in $A^S$ but only (a) is generated by a learning profile in $A^{\circ}$. The only difference between the three sets of learning cycles in Figure 1 is that agents 5, 6 and 7 have exchanged places which is possible because $\pi_5 E_5 = \pi_6 E_6 = \pi_7 E_7$. If there are no such equalities and $\pi_1 E_1 > ... > \pi_n E_n$ then the set $A^S$ will only contain the learning profiles in $A^{\circ}$.

### 1.4.2 The stable set, $A^S$

The stable set of learning profiles is characterised by conditions **I**, **II** and **III**. Lemmas 2 and 3 show that a learning profile is stable if and only if $a \in A^S$.

**Lemma 2.** *If $a \in A^S$ then $a$ is stable.*

*Proof.* If any $a \in A^S$ is not stable then by Lemma 1 there is some pair of agents $c, c' \in N$ such that $(c, c') \notin o(a)$ and for $a'_c = c'$, $a_c = c'$, $U_i(a) < U_i(a'_{\{c,c'\}}, a_{-\{c,c'\}})$ for $i \in \{c, c'\}$. By **I** at most one agent is not in a 2-cycle under $a$ so w.l.o.g let $(c, b) \in o(a)$, where $b \in N$. Then $U_c(a) = \pi_c \pi_b E_b < \pi_c \pi_{c'} E_{c'} = U_c(a'_{\{c,c'\}}, a_{-\{c,c'\}})$, which implies $\pi_b E_b < \pi_{c'} E_{c'}$.

First suppose $\pi_c E_c < \pi_{c'} E_{c'}$. Since $\pi_{c'} E_{c'} \neq \min_{i \in N} \pi_i E_i$, by **I** there is some $b' \in N$ such that $(c', b') \in o(a)$. Since $U_{c'}(a) = \pi_{c'} \pi_{b'} E_{b'} < \pi_{c'} \pi_c E_c = U_{c'}(a'_{\{c,c'\}}, a_{-\{c,c'\}})$, $\pi_{b'} E_{b'} < \pi_c E_c$. But then $\pi_{b'} E_{b'} < \pi_c E_c < \pi_{c'} E_{c'}$ and since $(c', b') \in o(a)$, $a$ violates **III** and so $a \notin A^S$.

Now suppose $\pi_c E_c \geq \pi_{c'} E_{c'}$. Then $\pi_b E_b < \pi_{c'} E_{c'} \leq \pi_c E_c$. By **III**, since $(c, b) \in o(a)$ it must be that $\pi_b E_b < \pi_{c'} E_{c'} = \pi_c E_c$. Since $\pi_{c'} E_{c'} \neq \min_{i \in N} \pi_i E_i$, by **I** $(c', b') \in o(a)$. Since $\pi_{c'} E_{c'} = \pi_c E_c$, **II** implies $\pi_{b'} E_{b'} \geq \pi_c E_c = \pi_{c'} E_{c'} > \pi_b E_b$. Then $U_{c'}(a) = \pi_{c'} \pi_{b'} E_{b'} \geq \pi_{c'} \pi_c E_c = U_{c'}(a'_{\{c,c'\}}, a_{-\{c,c'\}})$ and so no pair of agents $c$, $c'$ exists such that $a \in A^S$ is not stable. $\square$

**Lemma 3.** *A learning allocation $a$ is stable only if $a \in A^S$.*

*Proof.* Fix some $a^S \in A$ such that $a^S$ is stable but $a^S \notin A^S$ because it violates one or more of **I**, **II** or **III**.

Suppose $a^S \notin \arg\max_{a \in A} |C_2(a)|$ so that $o(a^S)$ does not contain the maximum possible number of 2-cycles that can form between agents, violating **I**. Then $|B_2(a^S)| \geq 2$. Let $b, b' \in B_2(a^S)$ be such that $b = \max B_2(a^S)$ and $b' = \max B_2(a^S) \setminus \{b\}$. Since $b, b' \in B_2(a^S)$, $b$ and $b'$ are either not in a cycle or in an m-cycle such that $m > 2$. In either case $U_i(a^S) < \pi_b \pi_{b'} E_{b'}$ for $i \in \{b, b'\}$. But for $a'_b = b'$ and $a'_{b'} = b$, $U_i(a'_{\{b,b'\}}, a^S_{-\{b,b'\}}) = \pi_b \pi_{b'} E_{a'_i}$ for $i \in \{b, b'\}$ and so $a^S$ is not stable.

So it must be that if $a^S \in A^S$ then $a^S \in \arg\max_{a \in A} |C_2(a)|$. If the number of agents $n$ is even then all agents will be in 2-cycles and $B_2(a^S) = \emptyset$. If $n$ is odd then there is one agent $b^*$ not in any cycle such that $\{b^*\} = B_2(a^S)$. Suppose $\pi_{b^*} E_{b^*} > \pi_n E_n$. Then there exists some cycle $(c, c') \in o(a^S)$ such that $\pi_{c'} E_{c'} = \pi_n E_n$. Under this set of cycles, $U_{b^*}(a^S) = 0$ and $U_c(a^S) = \pi_c \pi_{c'} E_{c'}$. But for $a'_{b^*} = c$ and $a'_c = b^*$, $U_i(a'_{\{b^*,c\}}, a^S_- \{b^*, c\}) = \pi_{b^*} \pi_c E_{a'_i} > U_i(a^S)$ for $i \in \{b^*, c\}$ and so $a^S$ is not stable.

Now suppose there are two cycles $(i, j), (i', j') \in o(a)$ such that $\pi_i E_i > \pi_j E_j$ and $\pi_i E_i = \pi_{i'} E_{i'} > \pi_{j'} E_{j'}$, so that $a^S$ violates **II**. Then $U_i(a^S) = \pi_i \pi_j E_j$ and $U_{i'}(a^S) = \pi_{i'} \pi_{j'} E_{j'}$. But since $\pi_i E_i > \pi_j E_j$ and $\pi_i E_i = \pi_{i'} E_{i'} > \pi_{j'} E_{j'}$, for $a'_i = i'$ and $a'_{i'} = i$, $U_i(a'_{\{i,i'\}}, a^S_{-\{i,i'\}}) = \pi_i \pi_{i'} E_{i'} > U_i(a^S)$ and $U_{i'}(a'_{\{i,i'\}}, a^S_{-\{i,i'\}}) = \pi_i \pi_{i'} E_i > U_{i'}(a^S)$ and so $a^S$ is not stable.

Now suppose $a^S$ violates **III** and there is some $j^*$ such that $\pi_i E_i > \pi_{j^*} E_{j^*} > \pi_j E_j$. Since $\pi_{j^*} E_{j^*} > \pi_j E_j$, by **I** there is some $(j^*, k) \in o(a^S)$. Then $U_i(a^S) = \pi_i \pi_j E_j$, $U_{j^*} = \pi_{j^*} \pi_k E_k$ and $U_k = \pi_{j^*} \pi_k E_{j^*}$. If $\pi_i E_i > \pi_k E_k$ then for $a'_i = j^*$ and $a'_{j^*} = i$, $U_i(a'_{\{i,j^*\}}, a^S_{-\{i,j^*\}}) = \pi_i \pi_{j^*} E_{j^*} > U_i(a^S)$ and $U_{j^*}(a'_{\{i,j^*\}}, a^S_{-\{i,j^*\}}) = \pi_i \pi_{j^*} E_i > U_{j^*}(a^S)$ and so $a^S$ is not stable. If $\pi_i E_i \leq \pi_k E_k$ then for $a''_i = k$ and $a''_k = i$, $U_i(a''_{\{i,k\}}, a^S_{-\{i,k\}}) = \pi_i \pi_k E_k > U_i(a^S)$ and $U_k(a''_{\{i,k\}}, a^S_{-\{i,k\}}) = \pi_i \pi_k E_i > U_k(a^S)$ and $a^S$ is not stable. □

### 1.5 Properties of equilibrium learning profiles

Having characterised the set of stable learning profiles in the previous section, the set of learning profiles considered can now be expanded to cover all those which can result in equilibrium. The set of equilibrium learning profiles $A^e$, shares one important characteristic with the set of stable learning profiles $A^S$; every learning cycle that forms under any $a \in A^e$ is a 2-cycle. However, unlike stable learning profiles where the number of 2-cycles is always maximised, the number of 2-cycles varies over different equilibria anywhere between one and $\frac{n}{2}$ if $n$ is even or $\frac{n-1}{2}$ if $n$ is odd. Some examples of equilibria are shown in Figure 4. This is not to say that any learning profile $a$ which only produces 2-cycles is an equilibrium. The set of equilibria is characterised in Lemma 4 below. It shows that not only is cycle size important, but whether or not a learning profile is an equilibrium also depends on which objects the agents in $B_2(a)$ are investigating. Every agent in $B_2(a)$ must be learning about the endowment of an agent in $C_2(a)$ who is in a 2-cycle; there can be no 'chains' of agents not in learning cycles such as that shown in Figure 6(a). It also matters which agents are in $B_2(a)$ and $C_2(a)$. If agents with endowments of sufficiently high potential value and probability of being acceptable are not in 2-cycles then there may be an incentive for an agent already in a 2-cycle to create an alternative 2-cycle or 3-cycle as shown in Figure 7.

**Lemma 4.** *A learning profile $a$ is an equilibrium if and only if all learning cycles are 2-cycles and for any agent $b \in B_2(a)$ not in a learning cycle, there is some $c \in C_2(a)$ such that $a_b = c$ , $\pi_b E_b \leq \pi_{a_c} E_{a_c}$ and $\pi_c E_c \geq \pi_c \pi_b E_b$.*

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

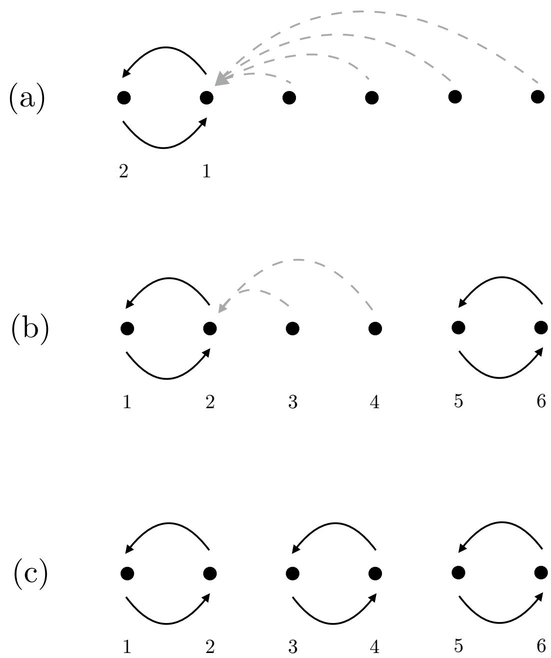



Figure 4: Examples of equilibria

and $\pi_{k^*} E_{k^*} \leq \pi_i E_i$ for all $i \in S$ [5]. This is illustrated in figure 5(a). For agent $k'$:

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

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

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

This is shown in figure 5(b). Since $k' \in S \setminus \{k^*\}$, $\pi_{k'} > \prod_{i \in S \setminus k^*} \pi_i$. But then since $\pi_{k^*} E_{k^*} \leq \pi_k E_k$, $U_{k'}(a) < U_{k'}(a'_{k'}, a_{-k'})$ and $a$ cannot be an equilibrium.

Then if $a$ is an equilibrium all learning cycles are 2-cycles which means no $i \in B_2(a)$ is in a learning cycle and so $U_i(a) = 0$ for all $i \in B_2(a)$. Suppose there is some $b \in B_2(a)$ such that $a_b = b' \in B(a)$ as in Figure 6(a). Then for some $a'$ such that $a'_{b'} = b$,

[5]If $|S| = 3$ then $k^* = k''$

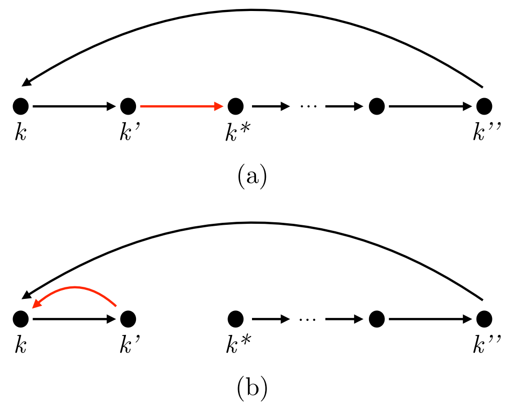



Figure 5: (a) The $m$-cycle $(k, k', k^*, ..., k'')$ (b) The alternative 2-cycle $(k, k')$

$(b, b') \in o(a'_{b'}, a_{-b'})$ as shown in Figure 6(b) and $U_{b'}(a'_{b'}, a_{-b'}) = \pi_{b'}\pi_b E_b > 0 = U_{b'}(a)$. But then $a$ cannot be an equilibrium and so it must be that $a_i \in C_2(a)$ for all $i \in B_2(a)$.

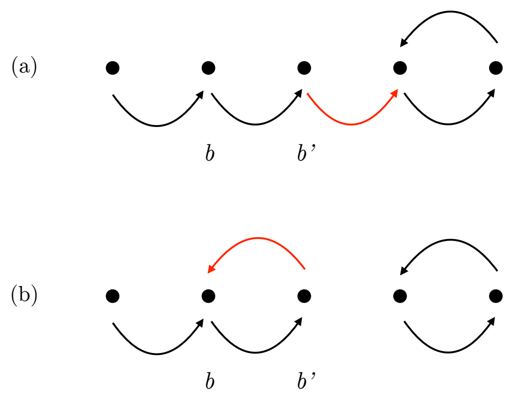



Figure 6: If $a$ is an equilibrium then $a_i \in C(a)$ for all $i \in B(a)$

I next show that if all learning cycles are 2-cycles and $a_i \in C_2(a)$ for all $i \in B_2(a)$, then for any $b \in B_2(a)$ such that $a_b = c \in C_2(a)$, $\pi_b E_b \leq \pi_{a_c} E_{a_c}$ and $\pi_c E_c \geq \pi_c \pi_b E_b$. This means that no agent in a 2-cycle is able to attain greater utility by creating an alternative 2-cycle or 3-cycle with any agent in $B_2(a)$. To see this, suppose that for some $b \in B_2(a)$ such that $a_b = c \in C_2(a)$, either $\pi_b E_b > \pi_{a_c} E_{a_c}$ or $\pi_c E_c < \pi_c \pi_b E_b$

or both. Under $a$ (as in Figure 7(a)), $U_c(a) = \pi_c \pi_{a_c} E_{a_c}$ and $U_{a_c}(a) = \pi_c E_c$. For some $a'$ let $a'_c = a'_{a_c} = b$. Under $(a'_c, a_{-c})$, $(b, c) \in o(a'_c, a_{-c})$ (as in Figure 7(b)) and $U_c(a'_c, a_{-c}) = \pi_c \pi_b E_b$. If $\pi_b E_b > \pi_{a_c} E_{a_c}$ then $U_c(a) < U_c(a'_c, a_{-c})$ and $a$ is not an equilibrium. Under $(a'_{a_c}, a_{-a_c})$, $(b, c, a_c) \in o(a'_{a_c}, a_{-a_c})$ (as in Figure 7(c)) and $U_{a_c}(a'_{a_c}, a_{-a_c}) = \pi_{a_c} \pi_c \pi_b E_b$. If $\pi_c E_c < \pi_c \pi_b E_b$ then $U_{a_c}(a) < U_{a_c}(a'_{a_c}, a_{-a_c})$ and again, $a$ is not an equilibrium.

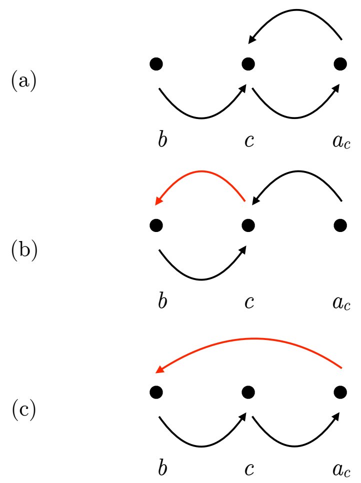



Figure 7: If $a$ is an equilibrium then no $i \in C(a)$ can attain greater utility by forming a cycle with any $i \in B(a)$

It remains to be shown that $a$ is an equilibrium if $o(a)$ contains only 2-cycles and for every agent $b \in B_2(a)$ there is some $c \in C_2(a)$ such that $a_b = c$ , $\pi_b E_b \leq \pi_{a_c} E_{a_c}$ and $\pi_c E_c \geq \pi_c \pi_b E_b$. Since all agents in $C_2(a)$ are in 2-cycles, $U_i(a'_i, a_{-i}) = 0$ for all $a'_i \in C_2(a) \setminus \{a_i\}$, $i \in N$. Now consider all $a'_i \in B_2(a)$ for all $i \in B_2(a)$. Since $a_i \in C_2(a)$ for all $i \in B_2(a)$, no $i \in B_2(a)$ can form a cycle by learning about the endowment of any agent in $B_2(a)$ and so $U_i(a'_i, a_{-i}) = 0$ for all $a'_i \in B_2(a)$, $i \in B_2(a)$. Since no $i \in B_2(a)$ is in a learning cycle under $a$, $U_i(a) = 0 = U_i(a'_i, a_{-i})$ for all $a'_i \in A_i$, $i \in B_2(a)$. Now consider all $a'_i \in B_2(a)$ for all $i \in C_2(a)$. Fix some $c^* \in C_2(a)$ and $a'_{c^*} = b^* \in B_2(a)$. Let $(c^*, c') \in o(a)$ (as in Figure 8(a) and (b)) so that $U_{c*}(a) = \pi_{c^*} \pi_{c'} E_{c'}$. Under $(a'_{c^*}, a_{-c^*})$, $c^*$ is either in a learning cycle or not. If $c^*$ is not in a learning cycle then $U_{c^*}(a'_{c^*}, a_{-c^*}) = 0 < U_{c*}(a)$. If $c^*$ is in a learning cycle then it is either a 2-cycle $(c^*, b^*)$ if $a_{b^*} = c^*$ (as in Figure 8(c)) or a 3-cycle $(b^*, c', c^*)$ if $a_{b^*} = c'$ (as in Figure 8(d)). Since $a_i \in C_2(a)$ for all $i \in B_2(a)$, $c^*$ cannot be in an $m$-cycle such that $m > 3$ under $(a'_{c^*}, a_{-c^*})$. If $(c^*, b^*) \in (a'_{c^*}, a_{-c^*})$ then $U_c^*(a'_{c^*}, a_{-c^*}) = \pi_{c^*}\pi_{b^*}E_{b^*}$. Since $a_{b^*} = c^*$ and $(c^*, c') \in o(a)$, $\pi_{b^*}E_{b^*} \leq \pi_{c'}E_{c'}$ and so $U_{c^*}(a) \geq U_{c^*}(a'_{c^*}, a_{-c^*})$. If $(b^*, c', c^*) \in (a'_{c^*}, a_{-c^*})$ then $U_c^*(a'_{c^*}, a_{-c^*}) = \pi_{c^*}\pi_{c'}\pi_{b^*}E_{b^*}$. Since $a_{b^*} = c'$ and $(c^*, c') \in o(a)$, $\pi_{c'}E_{c'} \geq \pi_{c'}\pi_{b^*}E_{b^*}$ and so $U_{c^*}(a) \geq U_{c^*}(a'_{c^*}, a_{-c^*})$. Then $U_i(a) \geq U_i(a'_i, a_{-i})$ for all $a'_i \in A_i$, $i \in N$ and so $a$ is an equilibrium. $\square$

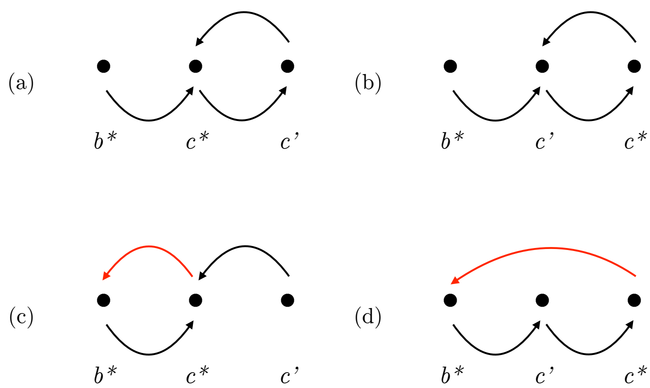



Figure 8: (a) $a_{b^*} = c^*$, (b) $a_{b^*} = c'$, (c) the 2-cycle $(c^*, b^*)$ and (d) the 3-cycle $(b^*, c', c^*)$

Lemma 4 shows that many learning profiles which generate the maximum number of 2-cycles are also equilibria. This includes the stable set, $A^S$.

**Lemma 5.** *Any stable learning profile is an equilibrium and so $A^S \subseteq A^e$, where $A^e$ is the set of equilibrium learning profiles.*

*Proof.* If $n$ is even then by **I** all agents are in 2-cycles and Lemma 4 holds. If $n$ is odd then by **I** all agents are in 2-cycles with the exception of some $i^*$ such that $\pi_{i^*}E_{i^*} = \min_{i \in N} \pi_i E_i$. Then $\pi_{i^*}E_{i^*} \leq \pi_i E_i$ for all $i \in N$ and $\pi_i\pi_j\pi_{i^*}E_{i^*} < \pi_i\pi_j E_j$ for all $i, j \in N$

and so Lemma 4 holds. □

## 1.6 Ex-ante welfare

In Section 1.3 three approaches to the learning and matching problem were introduced. Sections 1.4 and 1.5 characterised the set of stable and equilibrium learning profile respectively but did not explore the different effects on welfare. This section compares the ex-ante welfare across those stable and equilibrium learning profiles and considers the conditions under which a stable learning profile yields the highest ex-ante welfare possible (in line with Approach 3 described in Section 1.3). To more easily distinguish between different levels of ex-ante welfare, in what follows let the maximum possible ex-ante welfare for any given set of objects be $W^* = \max_{a \in A} W(a)$. Focussing on what occurs in equilibrium, let the maximum and minimum possible ex-ante welfare of any equilibrium learning profile be $W^E = \max_{a \in A^e} W(a)$ and $W^e = \min_{a \in A^e} W(a)$ respectively. The set of equilibrium learning profiles that yield $W^E$ is $A^E = \{a \mid a \in \arg\max_{a \in A^e} W(a)\}$. A learning profile is an **ex-ante welfare maximising equilibrium** if and only if $a \in A^E$.

The results presented here are strongly reliant on a number of crucial assumptions. The first is that the distribution from which each Agent $i$ draws their value for Object $k \neq i$ varies across objects but not across agents so that any Agent $i$ will observe the same distribution for any given Object $k \neq i$. This can be thought of as each agent having access to the same public information on the object but further investigation is required in order to determine more personalised information. The effects of relaxing this assumption are explored in Section 1.6.4. Secondly, the assumption that all agents learn simultaneously is critical to the results in this section. Agents are time constrained and do not have the capacity to wait and observe others' actions over time. This assumption is restrictive but it is to some extent relaxed and treated as a dynamic problem in Chapter

2. Thirdly, each agent has only a single test and is unwilling to exchange for an untested object. This means attention is restricted to those 'high-stakes' exchanges discussed in Section 1.1. Expanding the learning capacity adds new dimensions to the problem and an example illustrating some of the possible effects to consider all illustrated in the conclusion (Section 1.7). This section will also add one further assumption, that objects are 'well ranked,' in order to compare ex-ante welfare maximisation across approaches. The concept of well ranked objects applies when objects can be ranked in similar ways regardless of whether one focusses on the acceptability or conditional value of an object. It is explained further below, but the consequences of relaxing this assumption are given in Section 1.6.3.

### 1.6.1 Well ranked objects

Objects are well ranked if their order is the same when ranked by either their acceptability or conditional values. Agent 1 is endowed with an object which not only has the highest probability of being acceptable ($\pi_1 \geq \pi_1$ for all $i \in N$) but also the highest conditional value ($E_1 \geq E_n$ for all $i \in N$), while Agent 2 is endowed with an object which has the same or next highest values of $\pi_i$ and $E_i$ and so on. More formally, objects are **well ranked** if $\pi_1 \geq \pi_2 \geq ...\pi_n$ and $E_1 \geq E_2 \geq ...E_n$ and **strictly well ranked** if $\pi_1 > \pi_2 > ...\pi_n$ and $E_1 > E_2 > ...E_n$. Whether objects are well ranked or strictly well ranked $\pi_1 E_1 \geq \pi_2 E_2 \geq ...\pi_n E_n$ holds and so all the results on stability and equilibrium in Lemmas 1 to 5 and Observations 1 and 2 hold.

When objects are well ranked the relationship between the ex-ante welfare of all stable learning profiles, $W^S$, and the highest ex-ante welfare possible in equilibrium $W^E$ is precise and straightforward: $W^S = W^E$.

**Theorem 1.** *If objects are well ranked then the ex-ante welfare of any stable learning* *profile, $W^S$, is equal to that of any ex-ante welfare maximising equilibrium, $W^E$.*

The full proof of Theorem 1 can be found in Appendix A. The proof provided here in Lemma 6 is only for the case of $n = 4$ strictly well ranked objects but it is sufficient to illustrate the key steps of the full proof in Appendix A. Restricting attention to only four objects makes the proof more tractable, but the fact that objects are strictly well ranked and the fact that $n$ is even also somewhat simplify the proof. When objects are strictly well ranked, the set of stable learning profiles $A^S$ is identical to the set of ex-ante welfare maximising learning profiles $A^E$. As such, the conditions that define $A^S$ play a central role in the proof of Lemma 6. When objects are strictly well ranked, $\pi_1 E_1 > \pi_2 E_2 > ...\pi_n E_n$ and so **II** trivially holds for any set of strictly well ranked objects. For this reason, Lemma 6 refers only to conditions **I** and **III**. Restricting $n$ to an even number means that the part of condition **I** referring to odd values of $n$ is also trivially satisfied.

**Lemma 6.** *When objects are strictly well ranked and $n = 4$, $a$ is an ex-ante welfare maximising equilibrium if and only if $a \in A^S$.*

*Proof.* Let $a^* \in A^e$. If $a^* \notin A^S$ then it must violate at least one of conditions **I** and **III**.

**Condition I:** If $a^*$ is an equilibrium then by Lemma 4, $o(a^*)$ can only contain 2-cycles and if it violates **I** then $o(a^*)$ contains only a single 2-cycle. Say $\{(1, 2)\} = o(a^*)$ so that $W(a^*) = \pi_1\pi_2(E_1 + E_2)$. Now suppose that under $a^\circ$, agents 3 and 4 form a second 2-cycle (as shown in Figure 9(b)) so that $o(a^\circ) = \{(1, 2), (3, 4)\}$. Then $W(a^\circ) = \pi_1\pi_2(E_1 + E_2) + \pi_3\pi_4(E_3 + E_4)$. By Lemma 4, $a^\circ$ is an equilibrium and since $W(a^\circ) > W(a^*)$, $a^* \notin A^E$.

Then $a^*$ must satisfy condition **I** which implies there are two 2-cycles in $o(a^*)$. When $n = 4$ only the three learning profiles $a^\circ \in A^\circ$, $a'$ and $a''$ illustrated in Figure 10 can

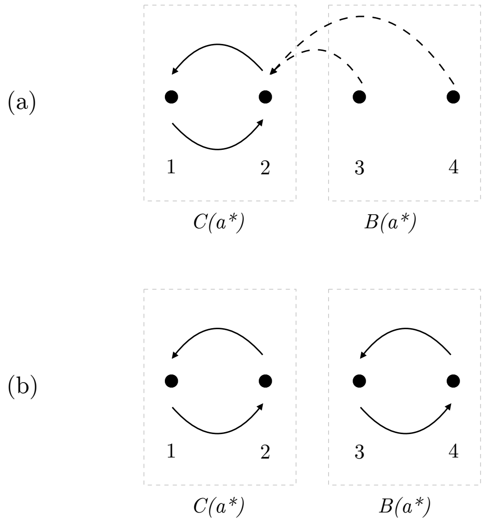



Figure 9: Two learning profiles: (a) A single 2-cycle under $a^*$ (b) Two 2-cycles under $(a'_{\{b,b'\}}, a^*_{-\{b,b'\}})$

generate sets of two 2-cycles and so $a^* \in \{a^\circ, a', a''\}$. Note that since all agents are in 2-cycles in all three learning profiles, by Lemma 4, $a^\circ$, $a'$ and $a''$ are all equilibria.

**Condition III:** Since $a^*$ satisfies **I** and $a^* \notin A^*$, $a^*$ must violate **III**. Of the three learning profiles $a^\circ$, $a'$ and $a''$, only $a^\circ$ satisfies **III** and so $a \in \{a', a''\}$. Consider the ex-ante welfare under $a^\circ$ and $a'$. If $a^* \in A^E$ and $a^* = a'$ then since $a^\circ$ is an equilibrium, ex-ante welfare under $a'$ must be at least as great as under $a^\circ$:

$$W(a') \geq W(a^\circ)$$

$$\pi_1\pi_3(E_1 + E_3) + \pi_2\pi_4(E_2 + E_4) \geq \pi_1\pi_2(E_1 + E_2) + \pi_3\pi_4(E_3 + E_4)$$

$$\pi_4(\pi_2 E_2 - \pi_3 E_3) + \pi_3(\pi_1 E_1 - \pi_4 E_4) \geq \pi_1(\pi_2 E_2 - \pi_3 E_3) + \pi_2(\pi_1 E_1 - \pi_4 E_4) \quad (1)$$

However, since objects are well ranked, $\pi_1 > \pi_2 > \pi_3 > \pi_4$ and so (1) is a contradiction. Then $W(a') < W(a^\circ)$. The same argument can be applied mutatis mutandis to $a^\circ$ and $a''$. Since $a^* \in \{a', a''\}$ and $W(a^\circ)$ is strictly greater than both $W(a')$ and $W(a'')$, $a^* \notin A^E$.

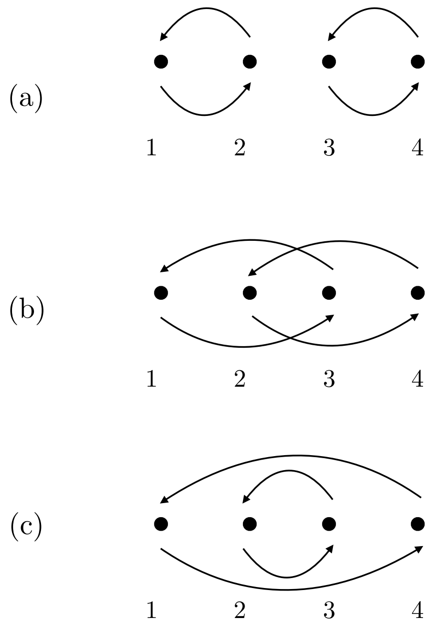



Figure 10: Three learning profiles generating two 2-cycles: (a) $a^{\circ}$ (b) $a'$ (c) $a''$.

Then if $a^*$ is an ex-ante welfare maximising equilibrium it must satisfy both **I** and **II** and so $a^* \in A^S$. Since $A^S = \{a^*\}$ and $a^* \in A^e$, $a^*$ must be the ex-ante welfare maximising equilibrium. □

Though Lemma 6 only covers the case where $n = 4$ and objects are strictly well ranked, the proof in appendix A which applies to any set of well ranked objects follows a similar pattern. The key difference is that when objects are well ranked (rather than strictly well ranked) the set of stable learning profiles, $A^S$, is only a subset of the set of ex-ante welfare maximising equilibria, $A^E$. This is demonstrated in Example 4.

**Example 4:** Four agents have a set of well ranked objects where $\pi_1 = \pi_2 = \pi_3 = \pi_4$ and $E_1 > E_2 > E_3 > E_4$. Consider again the three profiles $a^{\circ}$, $a'$, and $a''$ shown in Figure 10. Ex-ante welfare for all three profiles is:

$$W(a^{\circ}) = W(a') = W(a'') = \pi_1^2(E_1 + E_2 + E_3 + E_4)$$

Since $\pi_1 E_1 > \pi_2 E_2 > \pi_3 E_3 > \pi_4 E_4$ but $(1,3) \in o(a')$ and $(1,4) \in o(a'')$, both $a'$ and $a''$ violate condition **III**. Since $a^\circ$ does not violate **III**, $a^\circ$ is the sole learning profile in $A^S = \{a^\circ\}$. But since $W(a^\circ) = W(a') = W(a'')$, $A^S \subsetneq A^E$.

In Example 4 the probability an object will be acceptable is identical across objects and it is this feature which means that learning profiles which are not stable can be ex-ante welfare maximising equilibria. For this reason, the full proof of Theorem 1 (in Appendix A) utilises the set $A^* \supseteq A^S$ to show that there may be a large set of ex-ante welfare maximising equilibria but it always includes all stable learning profiles when objects are well ranked. Since by Observation 1, $A^\circ \subseteq A^S$, the ex-ante welfare of any stable learning profile can be expressed as $W^E = W^S = \sum_{i \in N, i \text{ even}} \pi_{i-1}\pi_i(E_{i-1} + E_i)$.

To complete the picture of the relative position of $W^S$, Theorem 2 compares the ex-ante welfare of all stable learning profiles with that of the worst equilibria, $W^e$ and the highest ex-ante welfare that can be achieved over all learning profiles, $W^*$. To show that $W^S$ is greater than the ex-ante welfare of some equilibria is not a complex task since in contrast to stable learning profiles which maximise the number of 2-cycles, Lemma 4 has already shown that equilibria can exist even when there are many agents and only a single 2-cycle. Whether a learning profile can yield a higher ex-ante welfare than $W^E$ depends on the exact values of $\pi_i$ and $E_i$ for each object. As Lemma 8 will show, sometimes these values are such that $m$-cycles other than 2-cycles can result in a higher level of ex-ante welfare.

**Theorem 2.** *When objects are well ranked, the ex-ante welfare of any stable learning profile, $W^S$, is at least as great as that of any equilibrium and may be less than the maximum possible ex-ante welfare, $W^*$.*

The proof of Theorem 2 is via Lemma 7 and Lemma 8.

**Lemma 7.** *If objects are well ranked, the ex-ante welfare of any stable learning profile, $W^S$ is at least as great as that of the equilibrium with the lowest ex-ante welfare, $W^e$ and strictly greater if there are more than three agents.*

*Proof.* Fix some $a^\circ \in A^\circ$. By Observation 1 and Lemma 2, $A^\circ \subseteq A^S$. Then $W(a^\circ) = W^S$[6]. If $n > 3$ then $\{(1,2),(3,4)\} \subseteq o^\circ$. By Lemma 4, for the same set of agents there is some equilibrium $a^e \in A^e$ such that $\{(1,2)\} = o(a^e)$ so that no agent in $N \setminus \{1,2\}$ is in a learning cycle. Since $(3,4) \in o^\circ$, $U_i(A^\circ) > 0$ for $i \in \{3,4\}$. Then $W(A^\circ) = W^S > W(a^e)$ and since $W^e = \min_{a \in A^E} W(a)$, $W(a^e) \geq W^e$. □

**Lemma 8.** *The maximum possible ex-ante welfare over all learning profiles, $W^*$ may exceed that of any equilibrium so that $W^* > W^E$.*

The proof of Lemma 8 is via Example 5, which demonstrates that the restriction to 2-cycles in equilibrium noted in Lemma 4 can also restrict ex-ante welfare. In Example 5, larger $m$-cycles result in higher ex-ante welfare.

**Example 5.** Suppose $N = \{1,2,3,4,5,6\}$ with the following values of $\pi_i$ and $E_i$:

| $i$ | $\pi_i$ | $E_i$ | $\pi_i E_i$ |
|---|---|---|---|
| 1 | 0.99 | 6 | 5.94 |
| 2 | 0.98 | 5 | 4.9 |
| 3 | 0.97 | 4 | 3.88 |
| 4 | 0.1 | 3 | 0.3 |
| 5 | 0.01 | 2 | 0.02 |
| 6 | 0.001 | 1 | 0.001 |

Consider the learning cycles that form under two learning profiles $a^E$ and $a'$: $o(a^E) = \{(1,2),(3,4),(5,6)\}$ and $o(a') = \{(1,2,3),(4,5,6)\}$. Note that $a^E \in A^\circ$ so by Observation 1, Lemma 2 and Theorem 1, $a^E \in A^E$.

Since $o(a')$ contains two 3-cycles, by Lemma 4 it cannot be an equilibrium. For the given

[6] In Lemma 6 this holds trivially since $A^S$ is a singleton. Lemmas 12 and 13 in Appendix A prove $W^E = W^S = W(a)$ for all $a \in A^S$ and any set of well ranked objects.

values:

$$W(a^E) = (0.99 * 0.98)(6 + 5) + (0.97 * 0.1)(4 + 3) + (0.01 * 0.001)(2 + 1) = 11.35$$

$$W(a') = (0.99 * 0.98 * 0.97)(6 + 5 + 4) + (0.1 * 0.01 * 0.001)(3 + 2 + 1) = 14.12$$

Since $a^E \in A^E$, $W(a^E) = W^E < W(a') \leq W^*$.

Example 5 does not imply that a set of 3-cycles will always ex-ante welfare dominate a set of 2-cycles. It occurs in Example 5 because the value of $\pi_i$ for agents 4, 5 and 6 is so low relative to agents 1, 2 and 3. In the three cycle $(1, 2, 3) \in o(a')$, there is a very high probability that all three agents will find the object they are learning about acceptable. Since $\pi_4 = 0.1$ it is far less likely that that the learning cycle $(3, 4) \in o(a^E)$ will result in an exchange. Then under $a^E$ there is a high probability that only agents 1 and 2 will exchange objects whilst the remaining four agents will keep their own endowments and as such, $W(a^E) < W(a')$.

### 1.6.2 When stable is best

Example 5 demonstrates that in order to maximise ex-ante welfare, it is sometimes necessary to implement larger learning cycles than the 2-cycles to which all stable learning profiles are restricted. It is, however, significant that the learning cycles in Example 5 are 3-cycles. Although there may be many 3-cycles under a learning profile which maximises ex-ante welfare, as Lemma 9 shows, there can never be an $m$-cycle containing more than three agents. A large $m$-cycle (as in Figure 11(a)) can always be broken into at least one 2-cycle and one other $m$-cycle (as in Figure 11(b)) to increase ex-ante welfare.

**Lemma 9.** ***Condition IV***$^c$*: If* $W(a) = W^*$*, then* $o(a)$ *contains no* $m$*-cycles such that* $m > 3$.

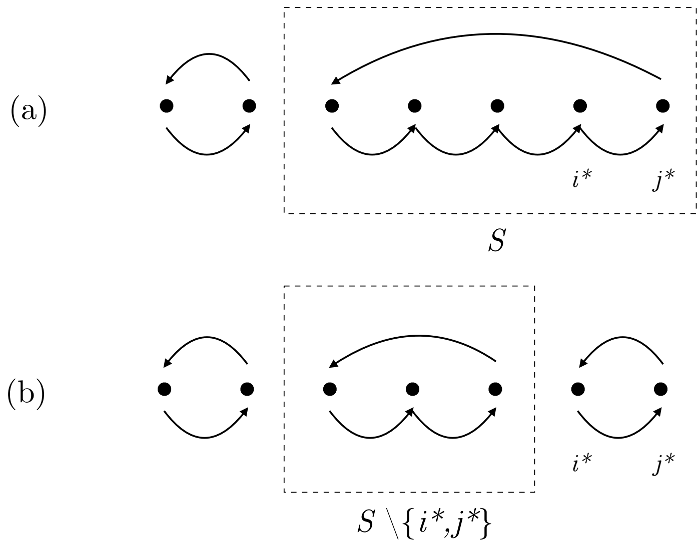



Figure 11: Ex-ante welfare and large $m$-cycles. (a) An $m$-cycle between agents in $S$, (b) The $m$-cycle broken into two smaller cycles.

*Proof.* Suppose $o(a)$ contains an $m$-cycle between $S \subset N$ agents such that $m > 3$ (as in Figure 11(a)). The sum of the expected utilities of all agents in $S$ is:

$$\sum_{i \in S} U_i(a) = \prod_{i \in S} \pi_i \sum_{i \in S} E_i$$

Under $a'$, let a 2-cycle form between agents $i^*, j^* \in S$ and an $(m-2)$-cycle between all agents in $S \setminus \{i^*, j^*\}$ (as in Figure 11(b)) so that the sum of expected utility of agents in $S$ is:

$$\sum_{i \in S} U_i(a'_S, a_{-S}) = \pi_{i^*}\pi_{j^*}(E_{i^*} + E_{j^*}) + \prod_{i \in S \setminus \{i^*, j^*\}} \pi_i \sum_{i \in S \setminus \{i^*, j^*\}} E_i$$

Since $\pi_{i^*} E_{j^*}, \prod_{i \in S \setminus \{i^*, j^*\}} \pi_i > \prod_{i \in S} \pi_i$, the sum of agents utility is higher under the two cycles: $\sum_{i \in S} U_i(a'_S, a_{-S}) > \sum_{i \in S} U_i(a)$. All $i \in N \setminus S$ are in the same cycles under $a$ and $(a'_S, a_{-S})$ and so $\sum_{i \in N \setminus S} U_i(a) = \sum_{i \in N \setminus S} U_i(a'_S, a_{-S})$. Then $\sum_{i \in N} U_i(a'_S, a_{-S}) >$

$\sum_{i \in N} U_i(a)$ and so $W(a'_S, a_{-S}) > W(a)$. $\square$

It is also significant, in Example 5, that there is a large gap in the probability an object will be acceptable between the best and worst three objects. Such large gaps in the value of $\pi_i$ are necessary in order for 3-cycles to feature in any ex-ante welfare maximising equilibrium. However, when the differences in the probability two objects being acceptable is sufficiently small, sets of 2-cycles produce higher ex-ante welfare. Objects are **closely well ranked** if objects are well ranked and $\pi_i \leq \frac{\pi_{i+3}}{\pi_{i+1}}$ for all $i \in N$ for which there exists an agent $i+3 \in N$. When $n$ is even and objects are closely well ranked, maximum ex-ante welfare can be achieved by learning profiles containing only 2-cycles. This restriction to 2-cycles means that any stable learning profile yields not only the highest ex-ante welfare of any equilibrium but also the highest ex-ante welfare of any equilibrium. The case where $n$ is odd is similar, with the exception that when the value of $\pi_n$ is high enough, it can be better to form a three cycle between three of the worst agents rather than leave one agent not in any learning cycle at all.

**Theorem 3.** *When objects are closely well ranked, if $n$ is even then $W^S = W^*$ and if $n$ is odd then either $W^* = W^S$ or $W^*$ and $W^S$ differ only in the sum expected utilities of three agents with the three worst endowments so that $W^S - W^{E*} = \pi_{n-2}\pi_{n-1}(\pi_n(E_{n-2} + E_{n-1} + E_n) + E_{n-2} + E_{n-1})$.*

The proof of Theorem 3 is given in Appendix B. It utilises a set of conditions similar to **I**, **II** and **III** but allows for 3-cycles as well as 2-cycles. Two extra conditions are also required. The first was given in Lemma 9 and restricts all learning cycles to 2-cycles and 3-cycles. The second, given in Lemma 21 restricts the number of 3-cycles to at most one. Collectively, these conditions show that when objects are closely well ranked, stable learning profiles yield the maximum (or very close to the maximum) ex-ante welfare.

### 1.6.3 Ex-ante welfare without the well ranked assumption

If objects are not well ranked then whilst the characterisations of stable and equilibrium learning profiles hold, the results regarding ex-ante welfare do not. It is still the case ex-ante welfare of any stable learning profile lies between that of the worst and best equilibria ($W^e$ and $W^E$) but a key difference is that the ex-ante welfare of any stable learning profile may be strictly less than $W^E$.

**Theorem 4.** *For any given set of objects, the ex-ante welfare of any stable learning profile is weakly greater than the lowest ex-ante welfare of any equilibrium, $W^e$ and may be exceeded by the highest ex-ante welfare of any equilibrium, $W^E$.*

*Proof.* By Lemma 5, $A^S \subseteq A^e$ and so $W^e \leq W(a^S) \leq W^E$ for all $a^S \in A^S$. As when objects are well ranked, if $n > 3$ then $W^e < W(a^S)$ for all $a^S \in A^S$. To see this, fix some $a^S \in A^S$. By conditions **I**, **II** and **III** on $A^S$, there is some cycle $(c_1, c_2) \in o(a^S)$ such that $\pi_{c_1} E_{c_1} = \pi_1 E_1$ and $\pi_{c_2} E_{c_2} = \pi_2 E_2$. By condition **I**, $|o(a^S)| > 2$ and $|C_2(a^S)| \geq 4$. By Lemma 4, there is also some $a^e \in A^e$ such that $\{(c_1, c_2)\} = o(a^e)$. Since $U_i(a^S) > 0$ for all agents in learning cycles and $|C_2(a^S)| \geq 4$, $W(a^S) > W(a^e) \geq W^e$.

The fact that the ex-ante welfare of an equilibrium can exceed that of a stable learning profile (in contrast to when objects are well ranked) is demonstrated through Example 6.

**Example 6.** Suppose $N = \{1, 2, 3, 4\}$ with the following values of $\pi_i$ and $E_i$[7]:

| $i$ | $E_i$ | $\pi_i$ | $\pi_i E_i$ |
|---|---|---|---|
| 1 | 500 | 0.2 | 100 |
| 2 | 200 | 0.4 | 80 |
| 3 | 100 | 0.6 | 60 |
| 4 | 50 | 0.8 | 40 |

[7] I would like to thank Maris Goldmanis for providing this example

Note that since $E_i$ and $\pi_i$ are inversely related, the objects are not well ranked. Consider the learning cycles that form under two learning profiles $a^S$ and $a^E$: $o(a^S) = \{(1,2),(3,4)\}$ and $o(a^E) \in \{(1,3),(2,4)\}$. Note that $a^S \in A^{\circ}$ so by Observation 1 and Lemma 2, $a^S \in A^S$. By Lemma 4, $a^E \in A^e$. Since $(1,3) \in o(a^E)$ and $\pi_1 E_1 > \pi_2 E_2 > \pi_3 E_3$, $a^E \notin A^S$ as it violates condition III. For this set of agents:

$$W(a^S) = (0.2 * 0.4)(500 + 200) + (0.6 * 0.8)(100 + 50) = 128$$

$$W(a^E) = (0.2 * 0.6)(500 + 100) + (0.4 * 0.8)(200 + 50) = 152$$

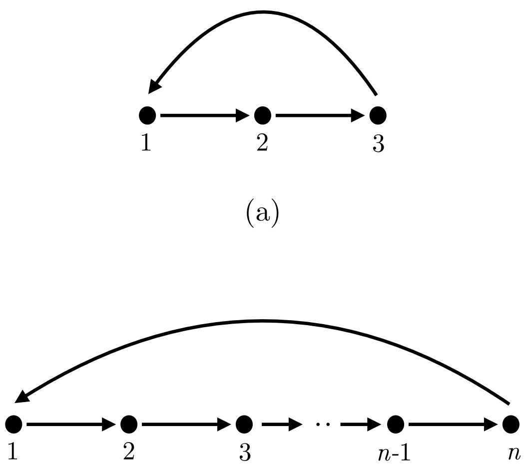
Then $W(a^S) < W(a^E) \leq W^E$. □

### 1.6.4 Non-identical ex-ante object values

It has been assumed throughout that all agents have the same ex-ante value for any given object other than their own endowments. Without this assumption, the characterisations of the set of stable learning profiles and the set of equilibria do not hold. In particular, if $f_k^i$ is non-identical across both agents and objects then, as the following example shows, learning cycles larger than 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 she ex-ante prefers agent 2's endowment to 3's since $E_2^1 = 10 > E_3^1 = 1$ and $\pi_2^1 = 0.9 > \pi_3^1 = 0.1$. Similarly, agent 2 ex-ante prefers 3's endowment and 3 ex-ante prefers 1's: $E_3^2 = E_1^3 = 10 > E_1^2 = E_2^3 = 1$ and $\pi_3^2 = \pi_1^3 = 0.9 > \pi_1^2 = \pi_2^3 = 0.1$. Then $a_1 = 2, a_2 = 3$ and $a_3 = 1$ is an equilibrium (as shown in Figure 12(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 $i \in N$, $a_i' \in A_i$, $a$ is an equilibrium. Since there are only three agents, $U_i(a) > U_i(a_i', a_{-i})$ for all $i \in N$, $a_i' \in A_i$ also implies $a$ is stable. Furthermore

$\sum_{i \in N} U_i(a) = 21.87 = \max_{a \in A} W(a)$.




(a)

(b)

Figure 12: (a) A 3-cycle between all agents. (b) An $n$-cycle between all agents.

Example 1 demonstrates that not only can 3-cycles exist in both equilibria and stable learning profiles, but that such learning profiles can also be ex-ante-welfare maximising. This result is not restricted to 3-cycles. As the number of agents increases, so too does the maximum possible cycle size. In fact, Example 3 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 3:** For some $N = \{1, ..., n\}$, 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 12(b). Let the ex-ante values of the objects be such that $E^i_{i+1} > E^i_{i-1} > E^i_j$ 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, $E^1_2 > E^1_n > E^1_j$ and $\pi^1_2 > \pi^1_n > \pi^1_j$ for all $j \in N \setminus \{2, n\}$ and for agent $n$, $E^n_1 > E^n_{n-1} > E^n_j$ and $\pi^n_1 > \pi^n_{n-1} > \pi^n_j$ for all

$j \in N \setminus \{1, n-1\}$. Then $U_i(a) = E^i_{a_i} \prod_{i=1}^n \pi^i_{a_i}$ and $U_i(a'_i, a_{-i}) \leq E^i_{i-1}\pi^i_{i-1}\pi^{i-1}_i$ for all $a'_i \in A_i$, $i \in N \setminus \{1\}$ and $U_1(a'_1, a_{-1}) \leq E^1_n \pi^1_n \pi^n_1$ for all $a'_1 \in A_1$. So for a sufficiently large $E^n_1$ and $E^i_{i+1}$ for all $i \in N \setminus \{n\}$ and a sufficiently small $E^1_n$ and $E^i_{i-1}$ for all $i \in N \setminus \{1\}$, $U_i(a) > U_i(a'_i, a_{-i})$ for all $a'_i \in A_i$ and so $a$ is an equilibrium. Also by making $E^n_1$ and $E^i_{i+1}$ for all $i \in N \setminus \{n\}$ sufficiently large and $E^1_n$ and $E^i_{i-1}$ for all $i \in N \setminus \{1\}$ sufficiently small, $a$ can be made both stable and the ex-ante welfare maximising learning profile.

## 1.7 Conclusion

When the ability to acquire information is limited or prohibitively costly, it can inhibit the function of the matching process. Mechanisms such as Gale's Top Trading Cycles despite possessing compelling properties such as individually rationality, Pareto optimality and strategyproofness may not deliver the best or maximum number of matches when information acquisition is endogenous. When each agent's ability to learn is restricted to just one other object then the need to design the learning and matching process is sometimes limited. When objects are well ranked and sufficiently similar (or closely well ranked) then any stable learning profile yields the maximum ex-ante welfare possible. The impact of relaxing some of the key assumptions has been examined in Section 1.6 but all the analysis presented here applies only to the case when learning is so limited that each agent is only able to acquire information about a single object. To illustrate some of the consequences of relaxing this assumption consider the agents and learning profiles illustrated in Figure 13. Suppose that not only is Agent 1 endowed with the best object but also the ability to investigate two other objects while the remaining agents can each investigate only one. In Figure 13(a) both Agents 2 and 3 investigate Agent 1's endowment. Despite Agent 1 owning the best object, the chance Agent 2 will be able to exchange for it depends in part on the outcome of Agent 1's investigation into Agent 3's endowment and vice versa for Agent 3. If instead, Agents 2 and 3 investigate each other's respective endowments (as in Figure 13(b)) then the probability they are able to trade depends only on the probability each of the objects are acceptable. In contrast with the results presented in prior sections, stable learning profiles exist which do not maximise ex-ante welfare and leave the agents with the best endowments and greatest testing ability unable to trade. Increasing the learning capacity will require further understanding of the trade off between the probability of acquiring an object and its value.

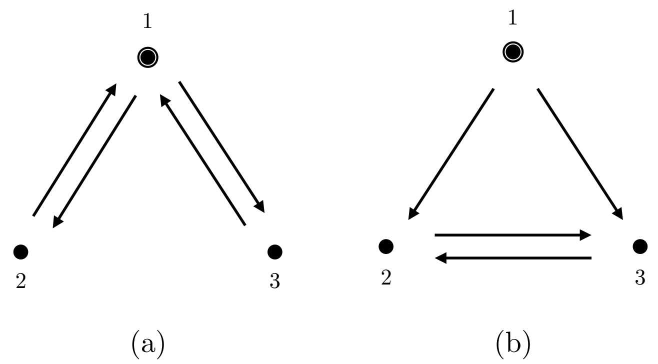



Figure 13: Learning cycles when $i^*$ can test two objects

CHAPTER 2:

# All Together Now?

## Information Acquisition over Time with Unilateral Matching

### 2.1 Introduction

Performing scientific tests and investigations, seeking out compatible new work schedules, or moving house can all involve substantial costs. So in one-sided matching markets, where each agent looks to exchange their endowment, high research costs can impact the information an agent acquires. In situations where the costs are so prohibitively high that each agent can only complete one investigation, the choice over what to learn is a critical one. This decision, however, may not necessarily be taken immediately; if the agent has a window of time in which to form her preferences then she can choose exactly when she investigates. For example, consider a council house exchange programme[8] and a set of families living in council homes trying to use such a scheme to move in the summer school holiday break. Since investigating new potential homes is a costly, time-consuming exercise, each family has only one weekend to completely dedicate to viewing a property, researching local schools and transport options before submitting the paperwork in time to formalise the exchange. Families cannot afford to get the decision wrong and move to an environment where they find themselves worse off. If this is the case, then the family must decide not only which property to investigate, but when. This leaves the family with two competing problems: Firstly, if Family A searches for a house and commits to viewing it too quickly, although they may find a house they prefer, they will not know which other families might prefer the Family A home and therefore which exchanges are possible. In short, they may waste their viewing weekend on a property they can never attain. Secondly, if Family A waits to see whether there will be other families who like their home and which families they are, Family A may miss out on the opportunity to exchange at the end of the summer altogether. It is therefore important for Family A to consider not only which property they learn about, but when they do so. Restricting attention to situations where there are only sufficient resources (time, money, opportunity etc.) to investigate one other property may seem extreme but it need not be an unrealistic assumption since those relying on social housing schemes may be more likely to be resource constrained.

This paper examines a model in which a finite set of agents, all with ex-ante identical endowments, are each equipped with the resources to investigate one other agent's endowment. Each agent knows the value of their own endowment but wants to be certain she prefers any object with which she commits to exchange. Agents may choose both which object to learn about and when to conduct their investigation from amongst a

[8] In the UK, many local authorities, such as Brighton and Hove City Council or the Royal Borough of Kensington and Chelsea run their own 'mutual exchange schemes' but privately run national schemes such as 'HomeSwapper' and 'Exchange Locata' also exist.

finite set of time periods (known here as learning periods). At the end of each learning period agents declare their up to date preferences so that other agents know both which tests were performed in that period and whether or not those tests were successful. At the conclusion of the final learning period, agents' endowments are exchanged using their ex-post preferences in Gale's Top Trading Cycles, the unique individually rational, Pareto optimal and strategyproof mechanism for this preference domain (as discussed in the introduction to Chapter 1). Under Gale's Top Trading Cycles, agents are only able to exchange for objects strictly preferred to their own and so the decision over what to learn about and when affects which exchanges can potentially be realised.

The previous paper (Chapter 1) showed that when agents and their objects are ex-ante identical, the ex-ante welfare maximising equilibrium also attained (or came very close to) the highest ex-ante welfare possible within the model. This 'fully simultaneous' equilibrium is not prohibited in the model explored here, agents are free to all learn together in the same period if they choose to do so. However, affording agents the opportunity to learn at their choice of learning period presents other options. When all agents learn simultaneously, it might result in many wasted investigations. If Agents 1 and 2 both test each other's respective endowments at the same time, and either one of those tests fails then both agents will be left unable to exchange. If instead Agent 2 waits to see the results of Agent 1's test then, if Agent 1's test fails, she has the option to test another object altogether. In this sense, the fewer tests performed in each period, the more other agents can benefit from the information gathered in each test. This makes 'fully sequential' learning an attractive candidate learning pattern for agents.

Working against the 'slow and steady' nature of fully sequential learning, however, is the pressure to not be left alone. Fully sequential learning requires, by definition, some agent to be the last to learn. If that agent has already had their endowment successfully tested by another agent then there is no problem, but the risk the last agent takes is that by the time they conduct their test, either all other agents have already identified other agents to exchange with or expended their tests and had them fail. In such a situation, the last agent will be left without a potential exchange partner, no matter which object she tests. This possibility creates an incentive for that 'last agent' to change her mind and instead, conduct her test in an earlier time period. For example, if there are ten agents and each Agent $i$ tests in the $i$th learning period then Agent 10 is the last to conduct her test. If instead she tests, say, Agent 9's endowment in the first period then, if the test is successful, she creates an incentive for Agent 9 to test the object belonging to 10. This pressure, to not be left until last, prevents fully sequential learning from arising in any equilibrium and causes multiple agents to choose to test in the same period as others, despite the information lost in doing so. This problem is similar to the well documented 'unravelling' phenomenon often observed in matching markets, particularly those involved in the allocation of new graduates to their first professional position (Roth and Xing, (1984)). In those two-sided matching markets, firms (hospitals/law firms/sports teams etc.) looking to hire new workers are often incentivised to make offers to new potential workers earlier and earlier in their education which can lead to 'explosive' offers made as early as possible in the process before all the information required which may affect the quality of the match (eg. medical training) is acquired.

This paper clearly builds on Chapter 1 and the literature discussed there. The most relevant of those remain Bade (2015) and Harless and Manjunath (2018), on the topic of endogenous information acquisition within matching problems. Bade (2015) examines a two-sided matching problem, and the superior role of serial dictatorship when agents have a choice over whether or not to acquire information. This does not apply to the model discussed here since agents are unwilling to risk exchange with objects which they have not investigated and the cost prohibits the number of investigations, not the extensive choice over whether to investigate at all. The two sided, school assignment problem discussed in Harless and Manjunath (2018) is presented as a static learning problem where agents simultaneously choose which single institution to investigate and focus on welfare properties highlights the advantages of Gale's Top Trading Cycles. Whilst more uncommon in the matching field, sequential information acquisition can be found in other areas of mechanism design. In voting models, for example, Gershkov and Szentes (2009) determine the optimal voting mechanism, in the face of costly information acquisition, is shown to be sequential and this makes obeying the central planner's instructions optimal. In the model discussed here, although fully sequential learning may be optimal from an ex-ante welfare perspective, it cannot be enforced by a central planner due to the 'last' agent's incentive to expend her test in an earlier period.

Dynamic mechanism design is explored in a variety of contexts, but largely concentrates on a changing set of agents as they arrive and depart. Parkes (2007), for example, defines an efficient mechanism appropriate for online auction environments where the set of customers morphs over time. In matching, Bloch and Cantala (2013) and Kurino (2014) use an overlapping generations models within two-sided assignment problems. Ünver (2010) proposes mechanisms for use in kidney exchange problems where the set of available patients and donors evolves over time, although their preferences do not. Bade (2017) proposes mechanisms for use in shift exchange problems with an infinite number of agents who arrive over time and, though their preferences do not change, they cannot be determined simultaneously. In contrast, the model discussed here applies to a fixed set of agents whose preferences can change over time and are determined before an exchange is executed.

This paper begins by defining the model in Section 2.2. Section 2.3 discusses the simultaneous learning equilibria that still persist, despite the option to learn over time. Section 2.4 demonstrates why exchanges can only happen between at most two agents in equilibrium and uses this to explain why the fully sequential learning pattern cannot occur in equilibrium. Section 2.5 concludes.

## 2.2 Model

### 2.2.1 Agents, Objects, Values

A set of agents $N = \{1, ...n\}$ is each endowed with an object in $K = \{1, ...n\}$ where agent $i$ is endowed with object $i$ and $n$ is even. The value of object $k$ to agent $i$ is $\omega^i_k$. Each agent values his own object at 0 so if $i = k$ then $\omega^i_k = 0$. Each agent knows she will find an object belonging to another agent to be either 'good' with value $\bar{v}$ or 'bad' with value $\underline{v}$, where $\underline{v} < 0 < \bar{v}$; if $i \neq k$ then $\omega^i_k \in \{\underline{v}, \bar{v}\}$. The ex-ante value of any Object $k$ to Agent $i$ is such that it is less than $i$'s endowment: $\mathbb{E}(\omega^i_k) = p\bar{v} + (1-p)\underline{v} < 0$. The vector $\omega^i := (\omega^i_1, ...\omega^i_n)$ gives the value of each object $k \in K$ to agent $i$ and $\omega := (\omega^1, ...\omega^n)$ is the vector of values of each object to each agent. The set of all possible such vectors $\omega$ is the state space $\Omega$. The state $\omega \in \Omega$ is determined by the chance player $c$ and is drawn from a uniform distribution over $\Omega$. The probability $i$ values some object $k \neq i$ as 'good' under any $\omega \in \Omega$ is $p = \pi(\{\omega \mid \omega^i_k = \bar{v}\})$. Since all states occur with equal probability, at any given state agents' values for objects are iid; in each state the value of one object $k$ to Agent $i$ conveys no information to another Agent $j \neq i$ about the value of any object to $j$.

### 2.2.2 Learning and Preference Declaration

Over time, agents are able to learn about exactly one other object. As discussed in the introduction to this chapter within the context of social housing, focus is restricted to those individuals who are so resource constrained, either in terms of time, money or opportunity that they are only able to take the opportunity to learn once. There is a set of $T = \{t_0, t_1, ...t_{2n}\}$ time periods which begin with nature, $c$, in $t_0$ drawing $\omega \in \Omega$ from a uniform distribution and then alternate between learning and preference declaration periods. Period $t_r$ is a learning period if $r$ is odd and a preference declaration period if $r$ is even. In every period $t_r$ where $r \neq 0$, all agents must take some action and so the player correspondence, $\mathcal{P} : N \to H$, is as follows: $\mathcal{P}(\emptyset) = c$ and $\mathcal{P}(h) = N$ for all $h \in H \setminus \{\emptyset\}$, where $H$ is the set of histories. A history is a sequence of vectors of agents' actions, where each vector $a^m$ details the actions taken by players who move after the history $(a^q)_{q=1}^{m-1}$. The length of history $h$ is $l(h)$ and since there are $2n+1$ time periods, $h$ is terminal if and only if $l(h) = 2n$. Since all agents move after every nonterminal history (with the exception of the initial history), all histories of length one or more will be a state followed by a sequence of $1 \times n$ vectors. The history $h'$ is a subhistory of $h = (a^q)_{q=1}^{m}$ if $h' = (a^q)_{q=1}^{m'}$, where $m' < m$. Agents are unaware of the state (and therefore the values of any object other than their own endowment) selected by $c$, so all $h \in H$ such that $\mathcal{P}(h) = N$ and $l(h) = 1$ are in the same information set. Agents know that $c$ draws $\omega$ from a uniform distribution over $\Omega$. Agents do observe each other's actions (other than $c$'s) and so all remaining information sets are singletons. Agent $i$'s strategy is $s_i$ and $S_i$ is the set of all possible strategies for Agent $i$. A strategy profile is $s = (s_i)_{i \in N}$ and the set of all strategy profiles is $S = \times_{i \in N} S_i$. The set of actions available to each $i \in N$ under the strategy profile $s$ at any $h \in H \setminus \{\emptyset\}$ is $A_i(s, h)$ and $a_i(s, h)$ is the action taken by $i$ at $h$ under $s$. The vector of all actions taken by all $i \in N$ at any $h \in H \setminus \{\emptyset\}$ under the strategy profile $s$ is $a(s, h)$.

Any $h \in H$ where $l(h)$ is odd coincides with a learning period. Each agent is equipped with a single test which can be used on any object and in the agent's choice of learning period. Then, in each learning period an agent can either choose to test some object $k \in K \setminus \{i\}$ or choose option $x$, which is the option not to test an object at all. Since each agent has only a single test, the actions available to Agent $i$ after any given history $h$ depends on Agent $i$'s actions at any subhistory of $h$, as determined by $s$: $A_i(s, h) = \{1, ...n, x\} \setminus \{i\}$ if there is no subhistory $h^k$ of $h$ such that $l(h^k)$ is odd and $a_i(s, h^k) \neq x$, and $A_i(s,h) = x$ otherwise.

Any $h \in H$ where $l(h)$ is even coincides with a preference declaration period. In such a period, agents report their preferences over objects. Agent $i$'s transitive preference relation at $h$ under the strategy profile $s$ is $R_i(s,h)$ where $kR_i(s,h)k'$ means agent $i$ weakly prefers object $k$ to $k'$ and $kP_i(s,h)k'$ implies $kR_i(s,h)k'$ but not $k'R_i(s,h)k$. Since $\mathbb{E}(\omega_k^i) < 0$ for all $i \neq k$, note that before an agent has learned the value of any object, $iP_i(s,h)k$ for all $k \in K \setminus \{i\}$ and after an agent has learned the value of an object there is at most one object such that $iP_i(s,h)k$ for all $k \in K \setminus \{i\}$. A preference profile at $h$ under the strategy profile $s$ is $R(s,h) = (R_i(s,h))_{i \in N}$ and the set of all preference profiles is $\mathcal{R}$. A **preference chain** is a vector of agents $(i_1, ..., i_m)$ such that $i_{k+1}R_{i_k}(s,h)i_k$ for all $k < m$. A **preference cycle** is a preference chain $o_b(s,h) = (i_1, ..., i_m)$ such that $i_{k+1}R_{i_k}(s,h)i_k$ for all $k < m$ and $i_1R_{i_m}(s,h)i_m$. The set of agents in the preference cycle $o_b(s,h)$ is $\phi_b(s,h)$.

### 2.2.3 Exchange and Equilibrium

The set of terminal histories is $Z \subset H$, where $z \in Z$ if and only if $l(z) = 2n$. When a terminal history $z \in Z$ is reached, objects are exchanged according to agents' preferences at $z$, $R(s,z)$. A matching is a bijection $\mu : N \to K$ and the set of all matchings is $\mathcal{M}$. Under any given strategy profile $s$ and terminal history $z$, a matching is individually rational if $\mu(i)R_i(s,z)i$ for all $i \in N$ and a matching $\mu'$ Pareto dominates $\mu$ if $\mu'(i)R_i(s,z)\mu(i)$ for all $i \in N$ and $\mu'(i^*)P_{i^*}(s,z)\mu(i^*)$ for at least one $i^* \in N$. If a matching is not Pareto dominated then it is Pareto optimal. A mechanism, $M : \mathcal{R} \to \mathcal{M}$ is individually rational and Pareto optimal if it always results in an individually rational and Pareto optimal matching. A mechanism is strategyproof if $M(R(s,z))(i)R_i(s,z)M(R'_i(s,z), R_{-i}(s,z))(i)$ for all $i \in N$, $r'_i(s,z)$. The mechanism used to match agents to objects is Gale's Top Trading Cycles, $GTT : \mathcal{R} \to \mathcal{M}$, as it is the unique individually rational, Pareto optimal and strategyproof mechanism.[9] $GTT$ is executed as follows:

> **Step $r$:** Each unmatched agent $i$ points at his most preferred object under $R_i(s, z)$ from amongst all unmatched agents. 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.

Agent $i$ may be indifferent between two objects $k$ and $k'$, however, since the indifference occurs only between objects for which $i$ strictly prefers her own endowment, it does not affect the mechanism's function; if $i$ strictly prefers her own endowment she will always point at and be matched to her own object before having to choose between $k$ and $k'$.

Agent $i$'s expected utility under any given strategy profile $s$, $U_i(s)$ depends on the probability an agent is matched with a given object and the value of that object. The probability an agent is matched with a given object is affected only by the decision over which object to investigate and not by which preferences to report. Since $GTT$ is used to decide the matching, and $GTT$ is strategyproof[10], it is a weakly dominant strategy for all agents to truthfully report there preferences as at $z$. It is also a weakly dominant strategy for agents to report their preferences truthfully in any preference declaration period. Agents know which tests have been performed in each period so an agent cannot deceive others to her own advantage by misrepresenting her true preferences if she has yet to complete the test. If she has performed her test and found she prefers her own endowment to the object she investigated then it is a weakly dominant strategy to report her true preferences since the object she prefers ex-post is her own. If she has performed her test and found she prefers the tested object to her own endowment then, since all other objects are ex-ante identical, stating she prefers the tested object makes it both

[9]The proof in Bade (2019) applies to the domain $\mathcal{R}$.
[10]As in Chapter 1, the proof in Bade (2019) also applies to the domain $\mathcal{R}$ discussed here.

weakly less attractive for other agents to test the same object and weakly more beneficial for the owner of the tested object (or the owner of the object at the end of the preference chain) to reciprocate the test.

Since agents report their preferences truthfully, agents are either matched with their own endowments (valued at 0) or a good object (valued at $\bar{v}$). Given the strategy profile $s$, the probability a terminal history $z \in Z$ is reached under which $i$ is matched with a good object is $\pi(\{z \mid GTT(R(s,z))(i)P_i(s,z)i\})$. So, given the strategy profile $s$, Agent $i$'s expected utility is:

$$U_i(s) = \pi(\{z \mid GTT(R(s,z))(i)P_i(s,z)i\}) \cdot \bar{v}$$

Let $H|_h$ be the set of sequences $h'$ of vectors of agents' actions for which $(h, h') \in H$. Given Agent $i$'s strategy $s_i$, let $s_i|_h$ be the strategy $i$ follows at each $h' \in H|_h$ and $s|_h$ the profile of strategies followed by all $i \in N$ at each $h' \in H|_h$. A strategy profile $s$ is a subgame perfect equilibrium if $U_i(s|_h) \geq U_i((s'_i, s_{-i})|_h)$ for all $s'_i \in S_i$, $i \in N$, $h \in H \setminus (Z \cup \{\emptyset\})$.

## 2.3 Learning Together and Equilibrium

Giving agents the option to choose not only which object to investigate, but also when to conduct that investigation affords them the opportunity to wait and see others' results before expending their own test: All else being equal, it is less risky for Agent $i$ to test the endowment of an Agent $j$ who is known to prefer $i$ to $j$ than an Agent $j'$ who has yet to expend her test and form ex-post preferences. In this sense, reducing the number of periods in which agents conduct tests simultaneously can lead to fewer wasted tests and more informative investigations. However, despite the possible advantages of fully sequential testing, a 'fully simultaneous' equilibrium remains in which all agents conduct their tests in the same period. Figures 14 and 15 show one such equilibrium. Both figures show which objects each agent investigates in each of the $n$ learning periods. In Figure 14 (a), every agent learns about a different object in period $t_1$. Figure 14 (b) illustrates how when any one of those agents tries to investigate an alternative object in that same period (in this case the endowment of Agent 2), then that agent cannot possibly be included in any later preference cycle. Figure 15 shows a similar effect when Agent 2 tries to investigate in a later learning period $t_7$; since all agents have expended their test in the first learning period, the only agent that Agent 2 can match with is Agent 1 no matter the period in which Agent 2 expends her test. Theorem 1 shows that there is a fully simultaneous equilibrium for any number of agents and objects, $n$.

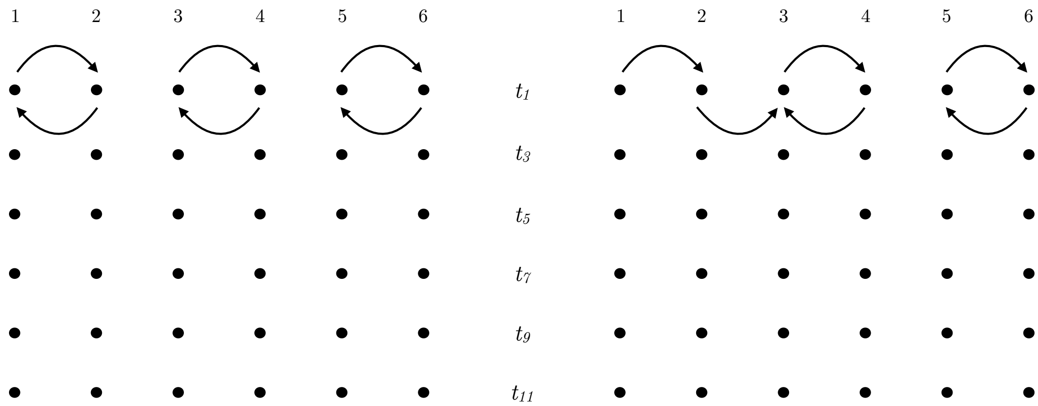
**Theorem 5.** *There is an equilibrium in which all agents learn in the same learning period.*

*Proof.* Let $S^\circ \subset S$ be a set of strategy profiles where every agent learns about a different object in the first learning period and in such a way that learning cycles contain no more than two agents:

$$S^\circ = \{s \mid a_i(s,h) = i+1 \text{ for } i \text{ odd }, a_i(s,h) = i-1 \text{ for } i \text{ even }, l(h) = 1\}$$

Since each agent only has one test to expend and all agents truthfully declare their preferences, $U_i(s^\circ) = p^2\bar{v}$ for all $i \in N$, $s^\circ \in S^\circ$. Now consider some $s'_i \in S_i$ such that $s'_i \neq s^\circ_i$, for any $s^\circ \in S^\circ$. Then, if $l(h) = 1$, either $a_i((s'_i, s^\circ_{-i}), h) \in N \setminus \{a_i(s^\circ, h), i\}$ (so that $i$ investigates a different object in the same period) or $a_i((s'_i, s^\circ_{-i}), h) = x$ (so that $i$ investigates an object in a different period).

First suppose $a_i((s'_i, s^\circ_{-i}), h) \in N\setminus\{a_i(s^\circ, h), i\}$ (Figure 14 shows one example where $i = 2$ and $a_2((s'_2, s^\circ_{-2}), h) = 3$). If $i$ chooses to learn about some object other than $a_i(s^\circ, h)$




(a) (b)

Figure 14: *Fully simultaneous equilibrium - Deviation within period.* Objects investigated in each learning period under the strategy profiles (a) $s^\circ$ and (b) $(s'_2, s^\circ_{-2})$.

at $h$ then, regardless of whether $i$ values $a_i((s'_i, s^\circ_{-i}), h)$ as either good ($\bar{v}$) or bad ($\underline{v}$), there can be no preference cycle containing both $i$ and another agent $j \neq a_i(s^\circ, h)$ at any history. Since agents reveal their preferences truthfully, $i$ will be matched with her own endowment. This means $U_i(s'_i, s^\circ_{-i}) = 0 < p^2\bar{v}$.

Now suppose $a_i((s'_i, s^\circ_{-i}), h) = x$ (Figure 15(b) shows Agent 2 choosing to instead investigate 4 in $t_7$). Under $(s'_i, s^\circ_{-i})$, all $j \in N \setminus \{i\}$ expend their test in the first learning period and so there is no $(h, h') \in H$ and no $j \in N \setminus \{a_i(s^\circ, h)\}$ such that $iP_j(s^\circ, (h, h'))j$. Since agents reveal their preferences truthfully, $i$ can never be matched with any $j \in N \setminus \{i, a_i(s^\circ, h)\}$ and so under any strategy profile $(s_i, s^\circ_{-i})$ where $s_i \in S_i$, $a_i(s^\circ, h)$ is the only object Agent $i$ can match with (besides $i$'s own endowment): $U_i(s'_i, s^\circ_{-i}) = 0$ for all

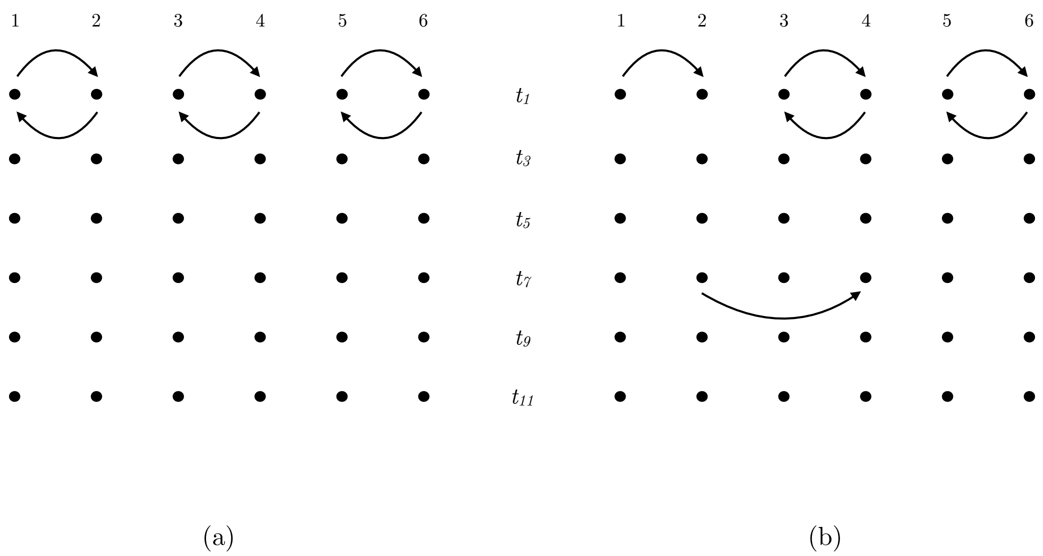



Figure 15: *Fully simultaneous equilibrium - Deviation across periods.* Objects investigated in each learning period under the strategy profiles (a) $s^\circ$ and (b) $(s_2', s_{-2}^\circ)$.

$a_i((s_i', s_{-i}^\circ, \cdot) = j$. Regardless of the history at which $i$ tests $a_i(s^\circ, h)$, $U_i(s_i', s_{-i}^\circ) = p^2\bar{v}$ and so $s^\circ$ is an equilibrium. □

## 2.4 Learning Apart and Equilibrium

When the opportunity to learn over time exists, from an ex-ante welfare perspective it is wasteful for all learning to take place simultaneously in the same period; ex-ante welfare (the sum of all agents' expected utilities) is higher if Agent $i$ can wait to first see whether another agent's test of $i$'s endowment is successful or not. But can the rush to learn together be halted? Do equilibria exist where agents patiently wait their turn in order to take full advantage of the information on offer? While agents may learn over more than one period in equilibrium, Theorem 6 shows that there is no 'fully sequential' equilibrium in which only one agent plans to learn in each period. Before turning to Theorem 6, it is first necessary to establish the limits of preference cycle size in equilibrium.

### 2.4.1 Preference Cycles in Equilibrium

It is not a coincidence the proof of Theorem 5 utilised an equilibrium in which each preference cycle contained exactly two agents; Lemma 10 shows that in equilibrium no preference cycle contains more than two agents.

**Lemma 10.** *In equilibrium, preference cycles are comprised of at most two agents.*

The proof of Lemma 10 can be found in Appendix E, but the key features of the argument are illustrated here for the case where $n = 4$, in Example 1. Though in this example a single large preference cycle is considered, Lemma 10 applies to all incidences of 'large' preference cycles comprised of three or more agents.

**Example 4:** Let $n = 4$ and $s$ be a strategy profile under which a preference cycle forms between all four agents. The following four cases illustrate the argument as to why $s$ cannot be an equilibrium.

*Case 1: Two or more agents learn in the same period, completing a preference cycle in the next period.*

Suppose that under the strategy profile $s$ two agents both learn in the same period and this results in a preference cycle in the next period. Figure 16(a) shows Agents 1 and 3 both learning in period $t_7$. The preference cycle which results in $t_8$ is shown in the box. In order for the preference cycle shown in Figure 16(a) to form, both Agent 1's test of 2 and Agent 3's test of 4 must be successful and so $U_3(s|_h) = p^2\bar{v}$, where $h$ is the history which coincides with the learning period $t_7$ illustrated in Figure 16(a).

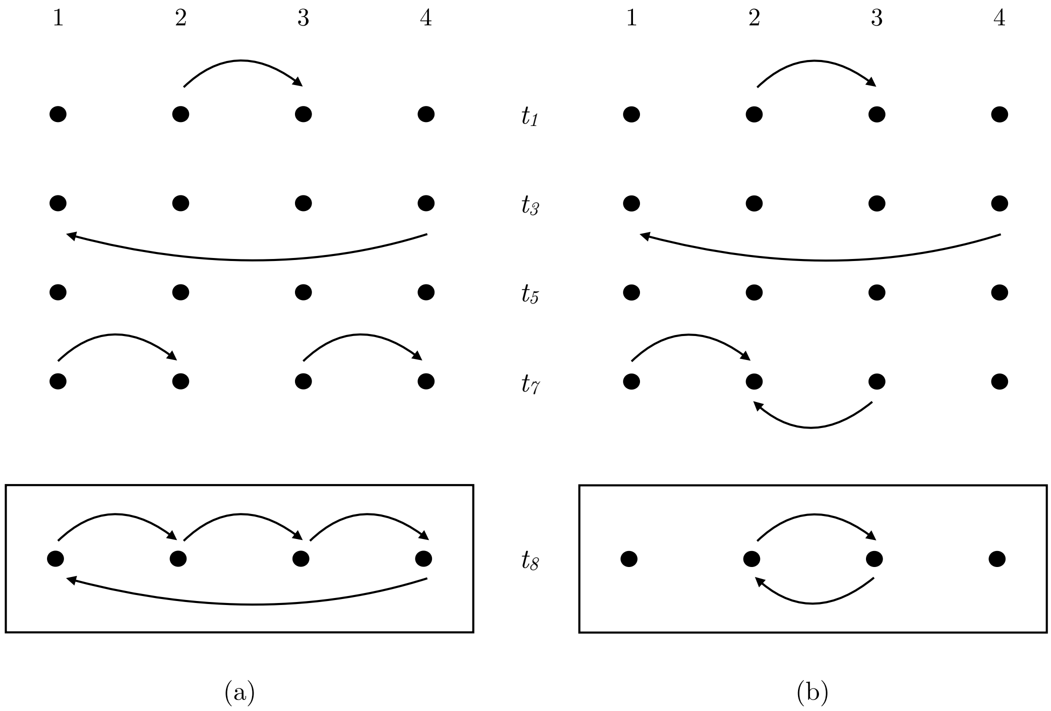



Figure 16: *Example 4 - Case 1.* Objects investigated in each learning period under the strategy profiles (a) $s$ and (b) $(s_3', s_{-3})$. The box shows the preference cycle that results in the preference declaration period, $t_8$.

If the preference cycle shown in Figure 16(a) forms at $t_8$, it must be that Agent 2 found Agent 3's endowment to be good: $w_3^2 = \bar{v}$. Now consider $s_3'$ such that in $t_7$ Agent 3 tests Agent 2's endowment as in Figure 16(b). In order to be in a preference cycle, Agent 3 only needs her single test of Object 2 to be successful. Then $U_3((s_3', s_{-3})|_h) = p\bar{v} > U_3(s|_h)$ and so $s$ is not an equilibrium.

*Case 2: Two or more agents learn in the same period, extending a preference chain but not a preference cycle in the next period.*

Suppose under $s$ two agents both learn in the same period and this 'extends' a preference chain but does not create a preference cycle in the next period. Figure 17(a) shows agents 2 and 3 both learning in period $t_5$, extending the preference chain started by Agent 1 in $t_1$. The preference cycle shown in the box is not complete at $t_6$ because Agent 4 has not yet expended her test. In order for the preference cycle to form, every test conducted by agents 2, 3 and 4 must be successful. Then $U_3(s|_h) = p^3\bar{v}$, where $h$ coincides with $t_5$.

If the preference cycle shown in Figure 17(a) forms at $t_8$, it must be that Agent 2 found Agent 3's endowment to be good at $t_5$: $w_3^2 = \bar{v}$. Now consider $s_3'$ such that in $t_5$ Agent 3 instead tests Agent 2's endowment. In order to be in a preference cycle, Agent 3 only needs two tests (those shown in $t_5$ of Figure 17(b)) to be successful. Then $U_3((s_3', s_{-3})|_h) = p^2\bar{v} > U_3(s|_h)$ and so $s$ is not an equilibrium.

*Case 3: Two or more agents learn in the same period, but neither extend a preference chain, nor create a preference cycle in the next period.*

Suppose two or more agents learn in the same period, but neither extend a preference chain, nor create a preference cycle in the next period. Figure 18(a) shows one such example where Agents 1 and 3 both investigate different objects in $t_1$. If the large preference cycle is to be realised in $t_8$ then at some time period after $t_1$, the 'gaps' in the preference cycle must be completed: Agent 2 and Agent 4 must also complete their investigations. Consider Agent 2, the penultimate agent to expend her test. Since she is the penultimate agent she extends the preference chain started by Agent 1 in $t_1$ but since Agent 4 is yet to expend her test, the preference cycle is also not yet complete. In this case $U_2(s|_h) \leq p^2\bar{v}$, where $h$ coincides with $t_3$. This leaves Agent 2 in a similar situation to that described in Case 2: For $s_2'$ such that $a_2((s_2', s_{-2}), h) = 1$, $U_2((s_2', s_{-2})|_h) = p\bar{v} > U_2(s|_h)$ (as shown in Figure 18(b)).

*Case 4: One agent learns in each period*

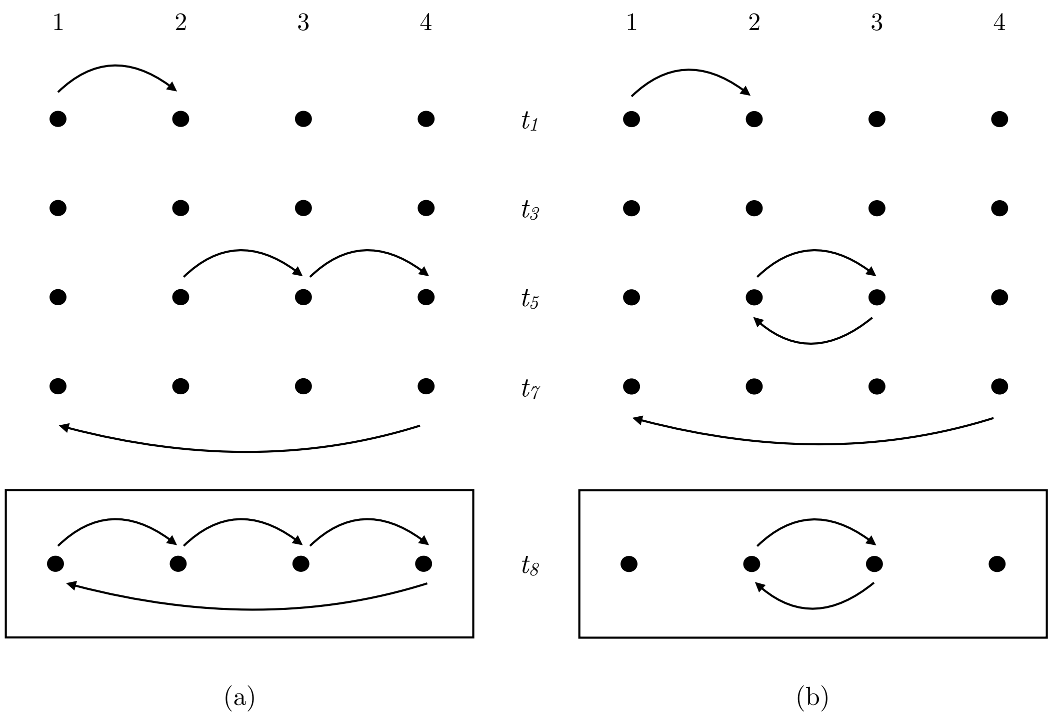



Figure 17: *Example 4 - Case 2.* Objects investigated in each learning period under the strategy profiles (a) $s$ and (b) $(s'_3, s_{-3})$. In period $t_5$, Agents 2 and 3 extend the preference chain started by Agent 1 in $t_1$. The box shows the preference cycle that results in the preference declaration period, $t_8$.

If exactly one agent learns in each period then, in order for the large preference cycle to form, there must be one or more occasions where an agent extends an existing preference chain but does not complete the preference cycle. Any such agent has a profitable deviation from $s$. Figure 19(a) shows an example where only one agent learns in each learning period. The penultimate agent to expend her test is Agent 3 and in doing so she creates a preference chain with Agent 1 but does not complete the cycle as Agent 4 has yet to expend her test. So $U_2(s|_h) \leq p^2\bar{v}$, where $h$ coincides with $t_5$. However, if Agent 2 instead chooses to learn about Agent 1's endowment then since at $t_5$ it is already known that Agent 1 strictly prefers Agent 2's endowment, for $s'_2$ such that $a_2((s'_2, s_{-2}), h) = 1$,

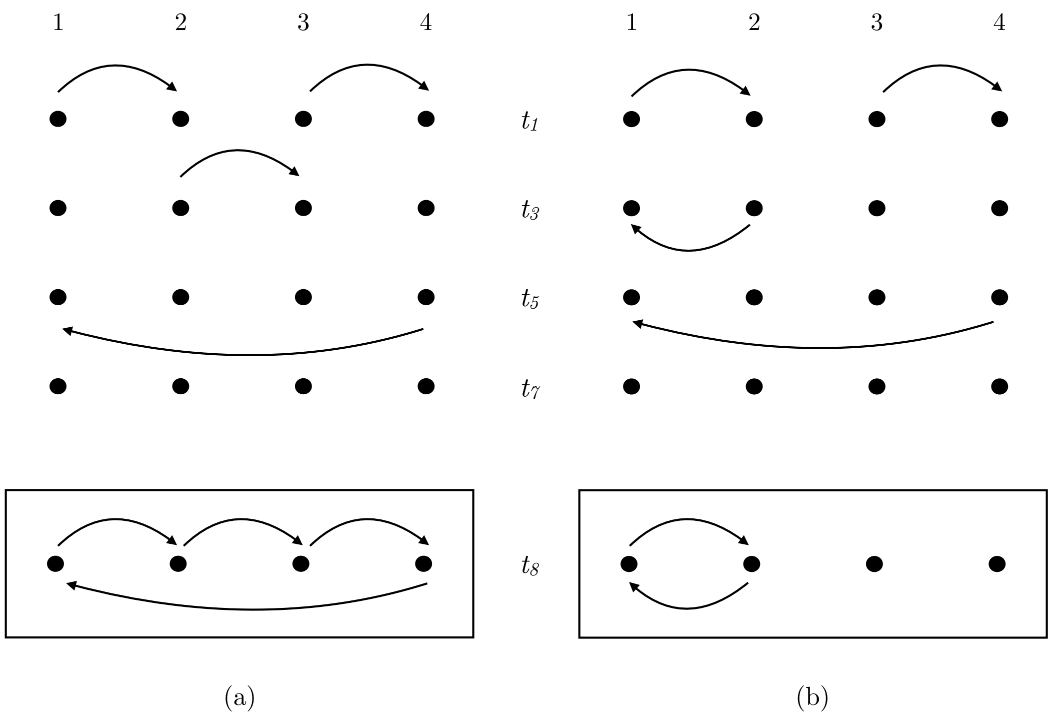



Figure 18: *Example 4 - Case 3.* Objects investigated in each learning period under the strategy profiles (a) $s$ and (b) $(s'_2, s_{-2})$. In period $t_3$, Agent 2 has an incentive to create a smaller preference cycle rather than completing the three-agent preference chain. The box shows the preference cycle that results in the preference declaration period, $t_8$.

$U_2((s'_2, s_{-2})|_h) = p\bar{v} > U_2(s|_h)$ (as shown in Figure 19(b)).

Chapter 1, Section 1.5 discussed a much broader range of objects than are currently being considered; in Chapter 1 objects were not necessarily ex-ante identical but nevertheless the restriction on cycle size is similar in both the simultaneous (Chapter 1) and dynamic (Chapter 2) setting. The fact that objects are ex-ante identical is not driving the conclusion of Example 4; similar statements could be made about the dynamic setting even if the ex-ante value of objects differs. If we assume (as in Chapter 1) that agent $i$ is endowed with the $i$th best endowment and that objects are well ranked (as in Section

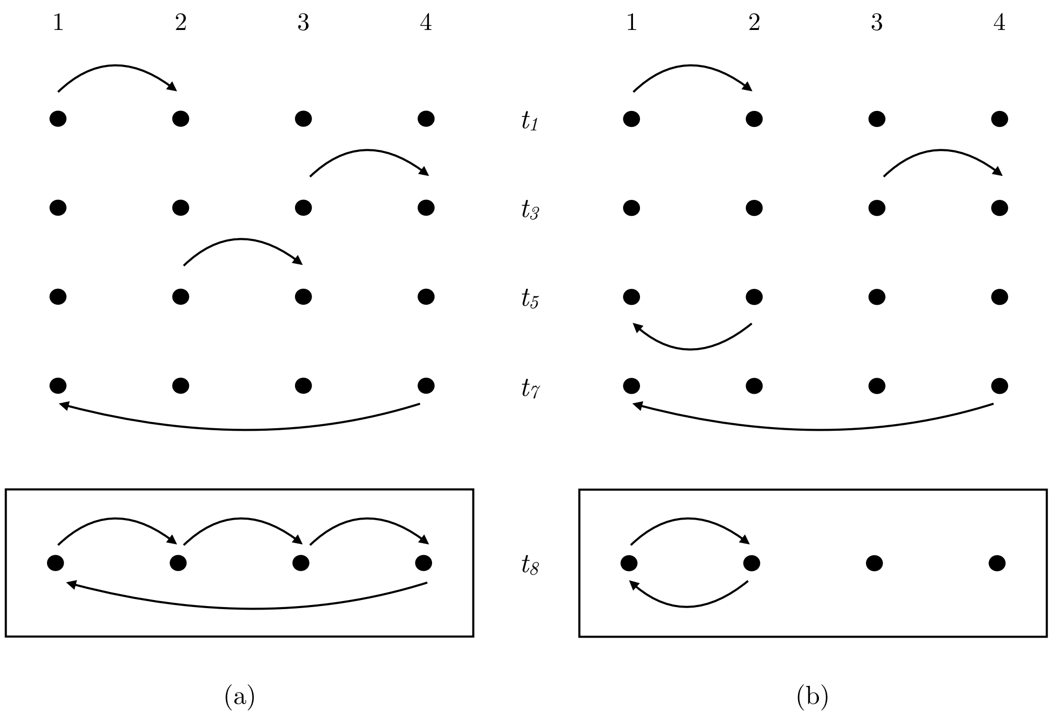



Figure 19: *Example 4 - Case 4*. Objects investigated in each learning period under the strategy profiles (a) $s$ and (b) $(s_2', s_{-2})$. In period $t_5$, Agent 2 has an incentive to create a smaller preference cycle. The box shows the preference cycle that results in the preference declaration period, $t_8$.

1.6.1) then it leaves Agents 1 and 2 in very influential positions. If they do not learn about each other's respective endowments then it gives the agents who the objects they do investigate the incentive to reciprocate the investigation and create a 2-cycle. If either Agent 1 investigates Agent 2's endowment or Agent 2 investigates Agent 1's endowment then the owner of the investigated object is incentivised to reciprocate the investigation in a later period. If, for example, Agent 2 test Agent 1's object in the first period and the test proves unsuccessful then Agents 1 and 3 become the two best remaining objects and similar logic applies as for Agents 1 and 2. All the deviations illustrated in Figures 31 to 19 would also be true in the strictly well ranked environment.

### 2.4.2 Fully Sequential Learning in Equilibrium

Lemma 10 dictates that preference cycles contain at most two agents in equilibrium, but it does not determine the period in which each preference cycle first forms. Whilst Theroem 5 proved that a fully simultaneous equilibrium does exist in which agents all learn in the same period, it is not necessarily the most desirable outcome as agents' are unable to use information learned by other agents to inform their own choice of investigation. A 'fully sequential' learning pattern would allow agents to learn slowly and steadily, thereby avoiding wasting tests on objects with which they know they will be unable to exchange. Whilst for the majority of agents this leads to higher expected utility than under fully simultaneous learning, it does create one victim. The final agent to execute her test may be left in the unfortunate position of having no potential partners with which to trade. Figure 20(a) shows how the failed tests of Agents 1, 2 and 3 (indicated by the dashed arrows) led to a preference cycle forming between Agents 4 and 5. If Agent 6 conducts his test in $t_{11}$ then he cannot hope to exchange for any object at all. However, if Agent 6 instead chooses to test in the first period, $t_1$ then it is a weakly dominant strategy for Agent 5 to investigate Agent 6's endowment as shown in Figure 20(b). Theorem 6 provides a more formal argument.

**Theorem 6.** *There is no equilibrium in which only one agent plans to learn in each period: If $s$ is an equilibrium and $n \geq 2$ then there is at least one learning period in which multiple agents learn.*

*Proof.* Let $s$ be an equilibrium such that in each of the $m$ learning periods, only one agent performs a test on another object. Let $i^*$ be the last agent to expend her test, which she does at $h^*$, then $A_j(s, h^*) = \{x\}$ for all $j \in N \setminus \{i^*\}$. Since, by Lemma 10 there are no large cycles in equilibrium, $U_{i^*}(s) \leq p^2\bar{v}$. Since each agent finds another object to be 'good' only with probability $p < 1$, under $s$ it is possible that all tests conducted

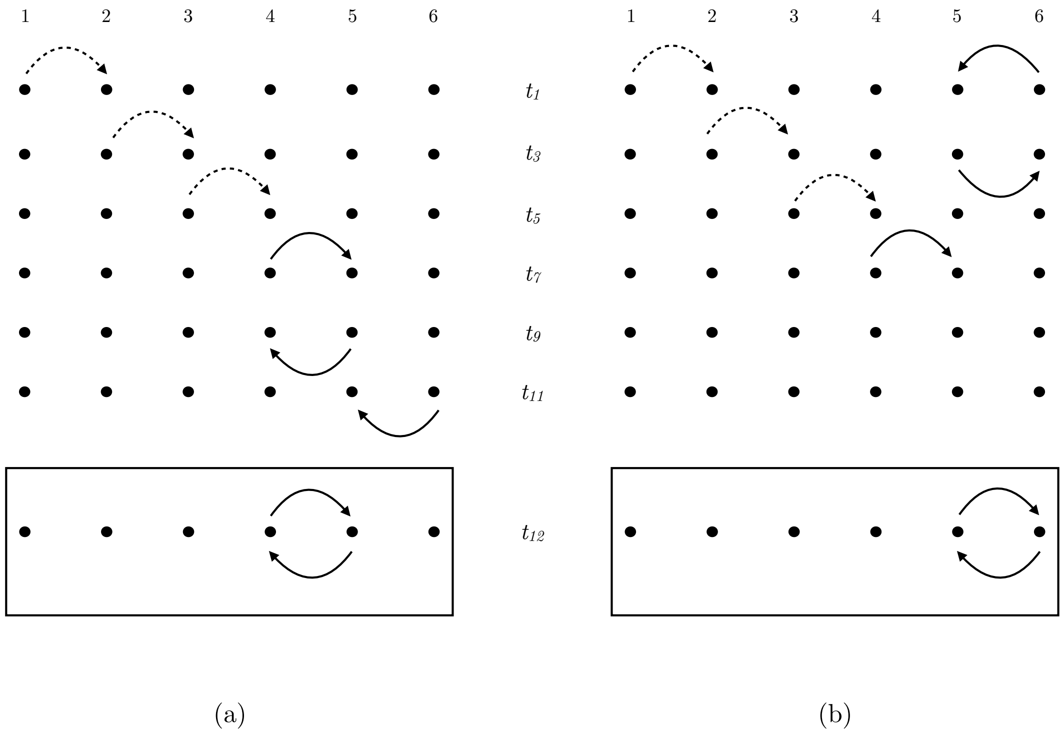



Figure 20: *Problems in fully sequential learning.* Solid arrows indicate successful tests and dashed arrows indicate failed tests. The box shows the preference cycles in the preference declaration period $t_{12}$.

by other agents fail if $\omega^j_{a_j(s,\cdot)} = \underline{v}$ for all $j \in N \setminus \{i^*\}$, in which case $i^*$ will not be able to form a preference cycle with any agent other than herself and so $U_{i^*}(s) < p^2\bar{v}$.

Suppose instead that under $s'_{i^*}$, $a_{i^*}(s'_{i^*}, h^1) = j^*$ where $h^1$ occurs in the first learning period and there is no $j \in N$ such that $a_j(s, h^1) = j^*$. Suppose $\omega^{i^*}_{j^*} = \bar{v}$ and consider $j^*$'s decision at $h^1$. Since $i^*$ is the only agent to investigate $j^*$ at $h^1$, if $a_{j^*}(\hat{s}_{j^*}, h^1) = i^*$ for some $\hat{s}_{j^*} \in S_{j^*}$ then $U_{j^*}(\hat{s}|_{h^1}) = p\bar{v}$. If $a_{j^*}(\hat{s}, h^1) \in (N \setminus \{i^*, j^*\}) \cup \{x\}$ then $U_{j^*}(\hat{s}|_{h^1}) \leq p^2\bar{v} < p\bar{v}$. Since $s$ is an equilibrium, it must be that $a_{j^*}(s_{j^*}, h^1) = i^*$. But then $U_{i^*}(s'_i, s_{-i}) = p^2\bar{v} > U_{i^*}(s)$. $\square$

## 2.5 Conclusion

Giving agents the opportunity to learn over a certain time period does not necessarily reduce wasteful learning. The threat of being left until last, with no option to exchange, prevents agents from taking full advantage of the information which can be gathered through sequential learning. This results in agents learning together in the same time periods and allows fully simultaneous learning as a possible equilibrium. Since the model presented here did not use discount factors, the pressure to 'learn together' is not linked to the need to learn fast, before their utility of an object is eroded. The fact that agents are all ex-ante identical, however, is likely playing a critical role. If one agent was thought to have an object with a much higher ex-ante value than others, agents may find it advantageous to wait and see if exchange with such an agent is possible. Different ex-ante values would also play a role whenever an agent is fortunate enough to have more than one agent interested in her endowment. In the model presented, a tie breaking rule is needed to determine how an agent chooses between two identical agents, but if objects are ex-ante differentiable then the agent's choice over which object to learn about is perhaps more clear. When each agent has the capacity for only one test, this would not necessarily prevent the fully simultaneous learning taking place in equilibrium, but it may affect the number of periods in which such learning can take place.

CHAPTER 3:

# UCAS: Prizes for Some

## Short Lists, Unknown Grade and Assortative Matching

### 3.1 Introduction

The Universities and College Admissions Service (UCAS) runs the UK's centralised university application process. With the exception of a select few private courses, entry to all UK institutions is conducted through UCAS; In 2017, UCAS processed over 2.5 million applications on behalf of almost 700000 students[11]. The design of such a process is critical to ensuring an efficient match, with the 'correct' students being placed at each institution. Students may have individual preferences but assortative matching is

[11] See https://www.ucas.com/about-us/who-we-are [accessed 01/11/18]

desirable with the students achieving the highest grades attending the most prestigious institutions. The 'gold standard' in college admissions, deferred acceptance, has been a staple mechanism of matching markets since the field's nascent paper of Gale and Shapley (1962). Deferred acceptance is appealing because it produces stable outcomes and because it makes telling the truth about one's preferences a dominant strategy (Dubins and Freedman (1981) and Roth (1982b)). The conditions which ensure these features are realised include firstly, that both students and colleges know their preferences over all the members of the opposite set and secondly that they are able to report those preferences. Neither of these features is present in the UCAS system. Students applying through UCAS must do so before they know their school exam results. This information is acquired only after their application has been submitted but before the universities make their decisions. UCAS also limits students to reporting their preferences over only two colleges (known as 'firm' and 'insurance' choices). Although it remains a dominant strategy to tell the truth about their preferences over the two chosen colleges, the combination of unknown grades and short preference lists prevents students from reporting honestly about which two colleges are their most preferred. As a result, students must be strategic about which two options they choose and this, in turn affects the assortativity of the final assignment, where the most able students are not able to access the high performing institutions.

This paper explores a model where a set of students are assigned to a set of colleges[12] through a student-proposing deferred acceptance style mechanism that differs from the canonical Gale-Shapley mechanism in two key aspects. Firstly, the colleges' preferences are determined by a students grades and while this information is known when the assignment is made, students do not know their grades at the time of application. Students learn their grades only after submitting preferences over colleges. Secondly, students

[12] I use the terms 'college' and 'university' interchangeably throughout to mean a higher education, degree awarding institution

must choose to express their preference over only two colleges by selecting a first 'firm' choice and a second 'insurance' choice to be used in case they are unsuccessful in securing a place at the first. To see that this corresponds to the UCAS system used in the UK, see Figure 21 which shows the decision faced by a UK student in the 2016/17 application cycle. Students may invite conditional offers from up to five universities. These offers will only be honoured if a students meets the conditions on A level grades (For example, in Figure 21 the student must attain three A grades at A level[13] to attend King's College London and one A and two B grades at A level to attend the University of East Anglia. However, from amongst any offers they receive from those five universities, they must choose one firm and one insurance choice (see the 'your reply' column in Figure 21). Furthermore, they must do so not only before they know which grades they will receive but before they have even sat the majority of their exams.

| university/college | course | starting | decision | | your reply |
|---|---|---|---|---|---|
| **University of Southampton**<br>campus: entry point: 1 | **QW36** | **01-09-16** | **Conditional** | **ABB** | **Insurance** |
| **University of Leeds**<br>campus: entry point: 1 | **6T3X** | **01-09-16** | **Conditional** | **AAB** | **Firm** |
| **King's College London (University of London)**<br>campus: entry point: 1 | **Q3P3** | **01-09-16** | **Conditional** | **AAA** | **Declined** |
| **Nottingham Trent University**<br>campus: entry point: 1 | **QP33** | **01-09-16** | **Conditional** | **280 points** | **Declined** |
| **University of East Anglia**<br>campus: entry point: 1 | **QW36** | **01-09-16** | **Conditional** | **ABB** | **Declined** |

Figure 21: *A student's UCAS in the 2016/17 application cycle.*

[13]The final exams to be taken at the high school level for the majority of 18 year olds are known as Advanced Levels or, more commonly, A Levels.

Using a model with both unknown grades and short preference lists permits exploration of how this affects the final allocation of students to universities. Theorem 7 demonstrates that unknown grades and short lists need not necessarily lead to an ex-ante undesirable outcome. When various parameters such as college capacity and students' preferences are set at the right level, the UCAS style system can produce the same outcome as would be expected under the student-proposing deferred acceptance mechanism. However, this result is very sensitive to the parametric assumptions. When students become more risk averse, the assortative matching can deteriorate, with the better universities missing out on some of the better students and some students remaining unassigned altogether. Increasing student capacity can benefit the best students and colleges but leave a large proportion of students not assigned to any college. Such effects are concerning as it not only prevents students life outcomes (Belfield et al. (2018)) and universities' ability to plan but if the strategic decision requires expert information then the strategic complexity may impact low income groups who are already at a disadvantage (Jerrim (2013)).

The college admissions problem has been considered from a number of different angles, an overview can be found in Pathak (2011). Short lists are a concern in a number of matching markets. Cseh et al. (2016) find mechanisms for use in the roommates problem with short lists and Immorlica and Mahdian (2005) focus on stability in the marriage problem when one side of the market must submit short preferences lists and find that under such circumstances, agents are unlikely to have more than one stable partner. Beyhaghi et al. (2017) look at the effects of short lists on doctors' choices and social welfare in the National Resident Matching Program (NRMP). They focus on where individuals apply and the extent to which 'safe' options are chosen and find that the Nash equilibrium outcome is not drastically different to the optimal one. Unlike the UCAS style system discussed here, however, all agents (both doctors and hospitals) have all the required information before an application is processed. In addition, social welfare is considered as the sum of agents utilities from their assignments and are not concerned with the assortative quality. Chade et al. (2014) look at the portfolio choice problem students face when they are uncertain about a college preferences and the number of applications is a choice limited by cost rather than the system itself. They find the portfolio choice problem can cause sorting to fail.

Closely related to the unknown grades problem is the early admissions problem in the US college market. Avery and Levin (2010) focus on the student's response to the possibility of early action and the positive benefits of allowing students to signal their preferences. The combination of the unknown grades and short list problems are known to cause adverse effects in school matches. Ajayi and Sidibe (2017) analyse the complex school match in Ghana and suggest changes to improve their measure of student welfare. Common with UCAS, the Ghana school match used both short lists and students were unaware of their priorities at each school at the time of application. The authors recommend increasing the length of the preference lists as well as informing students of their test scores in order to improve welfare. Such solutions are difficult to apply directly to UCAS due to the fact that students submit their rankings before even sitting their exams. Rectifying this would involve more than a change to the application mechanism, but a full restructure of the final year of secondary school. Extending the list would be simple to implement but it also creates high levels of uncertainty for universities. In the Ghanaian school match problem, each school can expect a number of students within a given range and the teaching needs will be uniform across the intake. By contrast, universities using the UCAS system need to be able to predict whether a course will be 100% or 10% full and plan resources accordingly. Given the match takes places often with six weeks of the start of teaching, it is difficult to acquire the correct resources and hire the correct staff at the eleventh hour.

The education system UCAS operates within does pose many seemingly immovable obstacles (such as short lists and unknown grades) but the concern over the lack of assortativity in the match is also a concern from the point of view of widening participation. Even when a strategyproof mechanism is used, students sometimes do not truthfully report their preferences. Artemov, Che and He (2017) show that when students believe there is little chance of being accepted to the best colleges then those colleges are likely to be omitted from their lists. This becomes concerning when it may not only be the academically less able students adopting such behaviour but when students from more disadvantaged backgrounds are more likely to apply this strategy. Chen and Pereyra (2018) find that students from low socio-economic backgrounds are more likely to decide to deliberately misreport their preferences and not list the most aspirational schools. Assessing the possible outcomes of the UCAS mechanism may then be key in determining how to aid students most likely to be disadvantaged by the strategic complexity.

Section 3.2 introduces a simple model which can be used to analyse the problems of short lists and unknown grades with a UCAS type mechanism. Section 3.3 then uses this model to tentatively explore some of the factors which influence the outocme of the mechanism such as university capacity and quality of education at a given institution and presents some conditions under which the ex-ante performance of UCAS is the same as the ex-ante outcome of deferred acceptance.

## 3.2 Model

A continuum of students $N = [0, 1]$, are to be assigned to one of four college options in the finite set $X = \{A, B, C, \varnothing\}$, where $A$, $B$ and $C$ are all colleges and being assigned to the 'null' college, $\varnothing$, is equivalent to not being assigned to a college. All students agree college $A$ is the best and college $C$ the worst and so if $\succ_i$ is $i$'s transitive preference relation, $A \succ_i B \succ_i C$ for all $i \in N$. The utility any student $i$ receives from attending college $x$ is $U_i(x)$, where $U(A) > U(B) > U(C) > U(\varnothing)$. Each student receives a grade from the set $G = \{1, 2, 3\}$, where colleges regard 1 to be the best grade and 3 the worst. The probability any student receives a grade $g \in G$ is $P(g)$. The grades are drawn from a uniform distribution and so $P(1) = P(2) = P(3)$. Each college $x \in X$ has capacity $q_x$. Since all colleges agree on which grade is the best, all colleges strictly prefer grade 1 students to grade 2 students and grade 2 students to grade 3 students but are indifferent between students with the same grade. Each college $x$ prefers to be full to capacity $q_x$ than not and so prefers a student with any grade to having spare capacity.

A matching is a function $\mu : N \rightarrow X$ and the set of all possible matchings is $\mathcal{M}$. The match is conducted through the following three-stage process:

**Stage 1 - Application**

Students submits a college application prior to receiving their grades. Each student lists two colleges on the application: a firm choice and an insurance choice. Each agent $i \in N$ chooses an application strategy $s_i$ from from the set $S = \{AB, AC, BC\}$, where $s_i = x_1x_2$ means $x_1$ is $i$'s firm choice and $x_2$ is $i$'s insurance choice college.

**Stage 2 - Grades**

Students learn their grades. Student $i$'s grade will be visible to any college receiving her application in the assignment stage.

**Stage 3 - Assignment**

- **Step 1:** Students applications are sent to their firm choice college. Each college considers all the applications it receives. If the number of applications to college $x$ exceeds $q_x$ then $x$ tentatively accepts its most preferred $q_x$ applicants. If the number of applications to college $x$ is $q_x$ or less then $x$ tentatively accepts all applicants. If all students are tentatively accepted to a college then the process terminates, otherwise continue to Step 2.

In general, at step $k$:

- **Step k:** The applications of any student rejected from their insurance choice college in Step $k-1$ are accepted by the null college $\varnothing$. The applications of any students rejected from their firm choice college in Step $k-1$ are sent to their insurance choice college. Each college considers all the new applications it receives in Step $k$ alongside the applications of the students it tentatively accepted in the Step $k-1$. If the number of applications to college $x$ exceeds $q_x$ then $x$ tentatively accepts its most preferred $q_x$ applicants. If the number of applications to college $x$ is $q_x$ or less then $x$ tentatively accepts all applicants. If all students are tentatively accepted to a college then the process terminates, otherwise continue to Step $k+1$.

The strategy profile $s = (s_i)_{i \in N}$ details the strategy chosen by each student. The set of all possible strategy profiles is $\mathcal{S}$. Stage 3 above describes the assignment mechanism $M : \mathcal{S} \to \mathcal{M}$. The mechanism $M$ is closely related to the canonical student-proposing deferred acceptance mechanism, (Gale and Shapley (1962)). Any difference here between $M$ and deferred acceptance is driven by the limit on the number of colleges to which a student can apply; under $M$ a student is only permitted to send their application to two colleges whereas under deferred acceptance, a student would be permitted to send their application to all colleges the student prefers to being left unassigned at the null college. This means that whilst the number of tentative matches and steps may be very large under deferred acceptance, the assignment mechanism $M$ will terminate after at most four steps.

Under a given strategy profile $s$, the mass of students adopting strategy $s_i \in S$ is $m(s_i)$, where $m(AB) + m(AC) + m(BC) = 1$. Note that the set $S$ implies no student will be able to choose a strategy where they list a college they prefer less than their insurance choice as a firm choice. Such a strategy could never be advantageous to the student choosing it since in a student-proposing deferred acceptance mechanisms it is a dominant strategy for students to be truthful about their order of preferences (Dubins and Freedman (1981) and Roth (1982)). So while students may carefully consider which two colleges they should send their application to, it is assumed they will always send it first to their most preferred college of the two.

The college $x$ a student is assigned to depends on both the mass of students adopting each strategy and the grade they each receive. Since all students are ex-ante identical, the probability student $i$ is assigned to college $x$ given their choice of strategy $s_i$ is $P(x \mid s_i)$. The probability student $i$ is assigned to college $x$ given their choice of strategy $s_i$ and having received grade $g$ is $P(x \mid s_i, g)$. These probabilities determine the expected utility $u_i(s_i)$ of an agent who adopts some strategy $s_i \in S$. Agent $i$'s expected utility of strategy $s_i = x_1x_2 \in S$ is:

$$u_i(x_1x_2) = P(x_1 \mid x_1x_2)U(x_1) + P(x_2 \mid x_1x_2)U(x_2) \tag{2}$$

Since preferences over colleges and their utilities are the same for all agents, $u_i(x_1x_2) = u(x_1x_2)$ for all $i \in N$.

## 3.3 Comparative Statics

Using the model in Section 3.2, it is possible to explore factors affecting both students application strategies and the assortativity of resulting match. In particular, I focus on two factors relevant to the changing UK higher education system in 2018. Firstly, I examine the possible impact on assortativity when one university is considered to be of disproportionately higher quality than the remainder and secondly, the possible impact on assortativity when universities expand their capacity. Before turning to these questions, however, it is necessary to first establish what happens in the model at equilibrium.

### 3.3.1 Equilibrium

In equilibrium it must be that the case that the expected utility of each application strategy is the same and that the total number of students choosing each strategy is equal to the total mass of students:

$$u(AB) = u(AC) = u(BC) \tag{3}$$

$$m(AB) + m(AC) + m(BC) = 1 \tag{4}$$

Determining the values of $u(AB)$, $u(AC)$ and $u(BC)$ is a complex task and so it is necessary to make some assumptions about how agents might behave in equilibrium. This has the benefit of allowing us to examine what sort of behaviour might arise in equilibrium but it is important to state that this is only one equilibrium out of many given that the following assumptions may not hold:

**(i)** Firstly, it is assumed that the top ranked college is filled to capacity with students. If places were otherwise left unfilled, some students would be able to benefit by changing their strategy. College $A$, therefore, is assumed to receive at least $q_A$ applications and its capacity will exhausted by students receiving a grade 1.

**(ii)** Since the top ranking institution is always able to take the very best students (those with grade 1), if there are any grade 1 students who are not admitted to College $A$ then these will be most preferred by College $B$. Since grading systems are designed to discriminate between ability, it's assumed that the number of students receiving top grades does not exceed the capacity of the top (College $A$) and middle (College $B$) ranking colleges. Then, college $B$ must accept all grade 1 students choosing the strategy $BC$ and all grade 1 students with strategy $AB$ who are not accepted to $A$. The remainder of of $B$'s capacity is exhausted by grade 2 students with strategies $AB$ and $BC$.

**(iii)** Lastly, it is assumed that there is sufficient provision for all Grade 1 and 2 students (with strategies $AC$ and $BC$) who are not accepted to $A$ or $B$ to be accepted to College $C$.

Under these assumptions, it is possible to find an expression for the expected utility of each strategy. For example, as stated in assumption (i), college $A$'s capacity is exhausted by grade 1 applicants so:

$$P(A \mid AB, 1) = \frac{q_A}{P(1)[m(AB) + m(AC)]} \tag{5}$$

$$P(A \mid AC, 1) = P(A \mid AB, 1) \tag{6}$$

The capacity remaining at $A$ for grade 2 and 3 students is $q_{A,2} = q_{A,3} = 0$ (where $q_{x,g}$ is the capacity at college $x$ for students with grade $g$ or below). Since by (ii), $B$'s capacity is not exhausted by grade 1 students, the number of grade 1 students at $B$ depends on the number rejected from $A$:

$$P(B \mid BC, 1) = 1 \tag{7}$$

$$P(B \mid AB, 1) = 1 - P(A \mid AB, 1) \tag{8}$$

The capacity remaining at $B$ for students receiving other grades is:

$$q_{B,2} = q_B - P(1)[m(AB)(1 - P(A \mid AB, 1)) + m(BC)] \tag{9}$$

By (iii), College $C$'s capacity exceeds the student population and since by (ii), all grade 1 students with the strategy $BC$ will be guaranteed a place at $B$, the only grade 1 students assigned to $C$ are those with the strategy $AC$ who are not assigned to $A$.

$$P(C \mid AC, 1) = 1 - P(A \mid AB, 1) \tag{10}$$

$$P(C \mid BC, 1) = 0 \tag{11}$$

Similar expressions for the probabilities of Grade 2 and 3 students can be found in Appendix F.1 and they can then be used to express the expected utility of each strategy as in Equation 2. These expected utilities ($u(A,B)$, $u(A,C)$, $u(B,C)$) can be found in Appendix F.2 and the mass of students choosing each strategy as a function of the utility of each college and the capacity of each college is in Appendix F.3.

Normalising $U(C)$ to 1, yields the following student masses applying using each strategy in equilibrium:

$$\begin{aligned} m(AC) = [&U(A) - 1 - U(A)U(B)q_A + U(B)^2 q_A - U(B)q_A + q_A - U(A)U(B)q_B \\ &+ U(B)^2 q_B - U(B)q_B + q_B]/[U(A) - U(B)] \end{aligned} \tag{12}$$

$$m(AB) = \frac{U(B)[q_A - m(AC)(q_a + q_B)] + m(AC)(1 - m(AC)) - q_A(1 - m(AC))}{(q_A + q_B)U(B) - 1 - m(AC)} \tag{13}$$

$$m(BC) = 1 - m(AC) - m(AB) \tag{14}$$

In order to compare the effects of changing utilities and and capacities in the following sections, I use this solution in the following example:

**Example 1:** Suppose college capacity is such that $q_x = \frac{1}{4}$ for all $x \in X$, $P(g) = \frac{1}{3}$ for all $g \in G$ and $U(A) = 3, U(B) = 2, U(C) = 1$ and $U(\varnothing) = 0$. Then, by equations 12, 13 and

14, $m(AB) = \frac{1}{4}$, $m(AC) = \frac{1}{2}$ and $m(BC) = \frac{1}{4}$ and $u(AB) = u(AC) = u(BC) = \frac{2}{3}$.

The ex-ante proportion of students receiving each grade expected to be assigned to each institution is shown in Figure 22. The left most bar represents the $\frac{1}{3}$ mass of students who are all expected to achieve a Grade 1. It shows that of those Grade 1 students, a mass of 0.25 students will go to College $A$ while the remaining mass of $\frac{1}{12}$ students are assigned to College $B$.

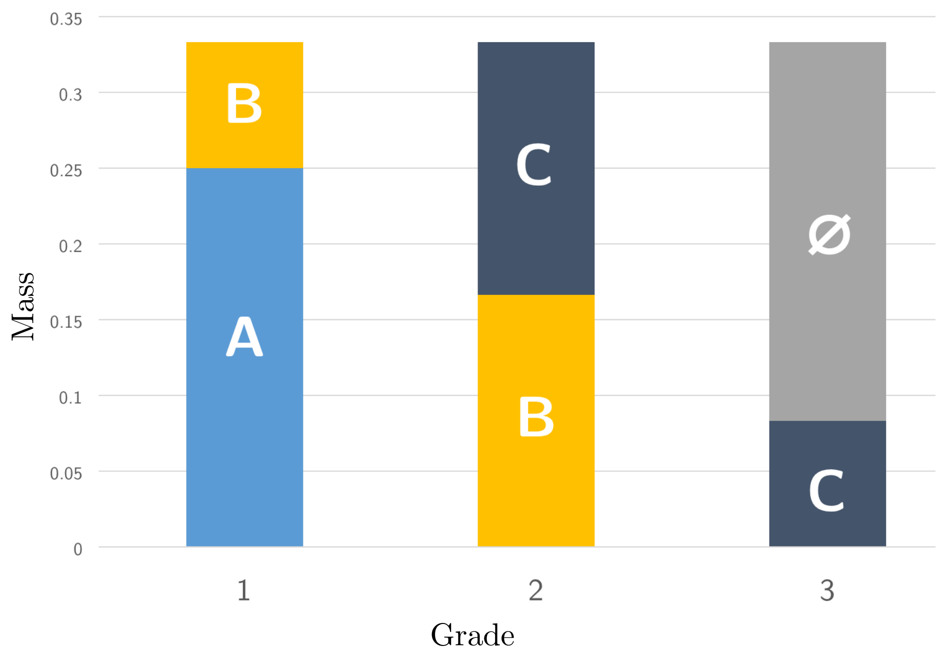



Figure 22: *Ex-ante student assignment in Example 1.* Proportion of students assigned to each college $A$, $B$, $C$ or $\varnothing$ by grade.

Example 1 and Figure 22 demonstrate an important result: despite the perceived problems associated with short lists and unknown grades, from an ex-ante perspective, the mechanism $M$ may perform just as well as student-proposing deferred acceptance. The distribution of students illustrated in Figure 22 is also the expected (ex-ante) outcome of deferred acceptance.

**Theorem 7.** *Ex-ante, the mechanism $M$ can result in the same matching as the student-proposing deferred acceptance mechanism.*

To see that the student-proposing deferred acceptance mechanism ($DA$) would yield the same result, recall that under $DA$, any student $i$ can express their complete, strict preference relation $\succ_i$: $A \succ_i B \succ_i C$. Therefore, even though students do not know their grades prior to application, their application can be considered by all colleges if necessary. This means, in Step 1 of $DA$, college $A$ accepts only grade 1 students, up to $q_A = 0.25$. The remainder of grade 1 students go to $B$ in Step 2 of DA with the remainder of $B$'s capacity being filled by grade 2 students. In Step 3, College $C$ is filled with the remaining grade 2 students and some grade 3 students, rejecting those in excess of capacity $q_c$.

### 3.3.2 Assortative Matching and $U(A)$

Of course, Theorem 7 relies heavily on the parametric assumptions in Example 1 and simply because $M$ works well in some circumstances does not suggest it will work equally well in others. For example, consider the case where one college is considered disproportionately better than others. Such cases are not uncommon, Belfield et al. (2018) find that men from the most selective Russell Group universities earn up to 50% more than those from other institutions within the same group. Such conditions can affect sorting within the match.

**Example 2:** Let all parameters be the same as in Example 1, with the exception that $U(A) = 3.5$. This increases the number of students who apply to $A$ through either the $AB$ or $AC$ strategy. Figure 23 shows that as $U(A)$ increases to 3.5, $m(AB)$ approaches $\frac{2}{3}$ and $m(AC)$ approaches $\frac{1}{3}$ while $m(BC)$ decreases to zero.

Example 2 yields the allocations represented in Figure 24(a), with the outcome from

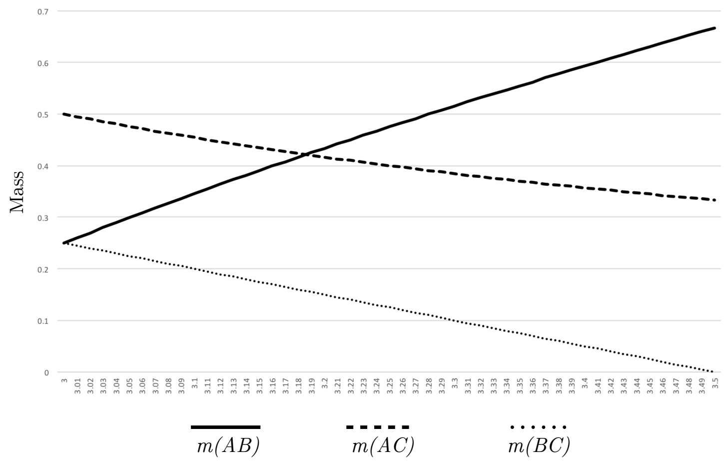



Figure 23: *Change in strategy as $U(A)$ increases from 3.0 to 3.5*

Example 1 shown in panel (b) for comparison. If the 'optimal' match is considered to be that achieved under $DA$, then the parameters in Example 2 take us to a less socially desirable outcome. College $A$ is still filled to capacity with grade 1 students, however the talent of some grade 1 students is now wasted as they are assigned to college $C$. Some grade 2 students are also left unassigned. The winner in this situation, however, is College $C$ who accept some grade 1 students it would not otherwise have recruited.

#### 3.3.3 Assortative Matching and $q_x$

Increased college capacity is another feature of the UK higher education market. In 2015/16 government caps on student numbers were removed (Hillman (2014)) allowing all universities to expand their capacity. Large changes in capacity at all institutions can also have dramatic effects on assortativity in matching, as shown by Example 3.

**Example 3:** Let all parameters be the same as in Example 1, with the exception that $q_x = \frac{1}{3}$ for all $x \in X$. As the capacity of each institution increases to $\frac{1}{3}$, $m(AB)$ increases to 0.99 and $m(AC)$ decreases to 0.01, while $m(BC)$ decreases to zero.

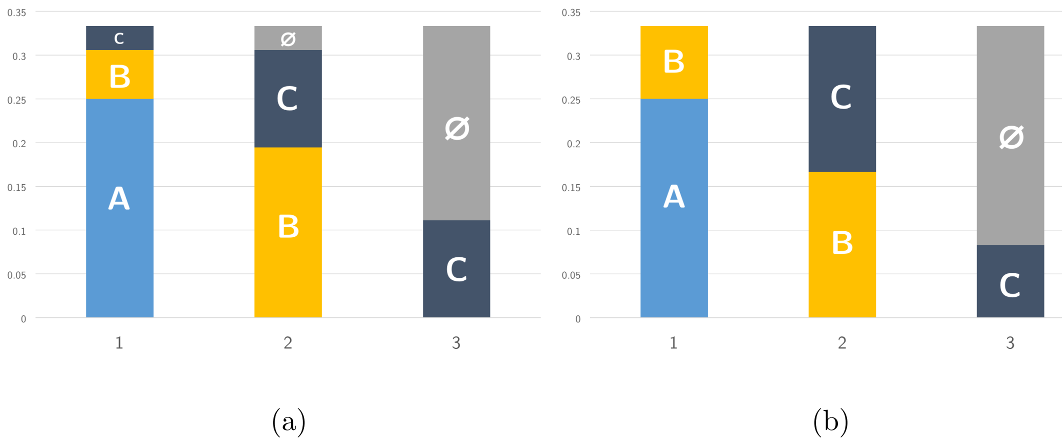



Figure 24: *Increasing* $U(A)$ (a) Ex-ante matching when $U(A) = 3.5$. (b) Ex-ante matching when $U(A) = 3.0$.

Figure 25 shows how the increased capacity affects assortativity in matching. As in all previous example, College $A$ still fills with grade 1 students, but with $A$'s increased capacity, $B$ is almost exclusively filled with grade 2 students; $B$ can no longer take the best students who are rejected by $A$. College $C$ also suffers as the number of student using a strategy featuring $C$ approaches to zero. Grade 3 students also suffer the most in this example as they become increasingly unassigned.

### 3.4 Conclusion

The results in Section 3.3 are clearly very sensitive to the parametric assumptions, but the examples illustrate how different groups of both colleges and students can be severely

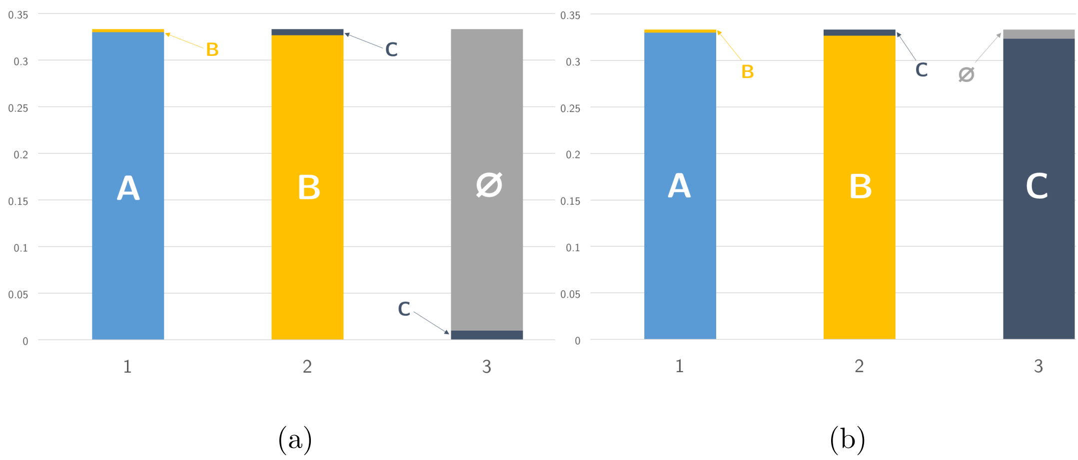



Figure 25: *Increasing* $q_x$ *to approaching* $\frac{1}{3}$ (a) Ex-ante matching under $M$ (b) Matching under $DA$.

disadvantaged by introducing the twin features of short listing and unknown grades into an otherwise appealing mechanism. In the UK's UCAS system, students are required to cope with both of these hurdles as they must both narrow their options to a first 'firm' choice and a second 'insurance' choice. If the pre-results application is to persist then this suggests further study should be conducted into increasing the number of options any given student may choose. The argument for providing a less strategic entry process is even more pertinent when considering widening participation. The inability to master the UCAS admissions process is one feature likely to block entry to high status universities for those from disadvantaged background (Jerrim (2013)). From a theoretic stance, one important feature absent from the model discussed above is heterogenous groups of students. Students may differ in their levels of risk aversion or in the accuracy of their information about programme entry. Many other policies to help those from low-participation schools exist; costly tutoring and enrichment programmes for such students may be partial solutions but reform in the admissions system may be more efficient and can be made easily accessible to all. Further work is needed to establish the extent to which reform in the admissions system might do more to allow fair access to all institutions.

# Appendix

## A Well ranked objects: Stability and Equilibrium

In Section 1.6, the proof of Theorem 1 was given only for the case where objects are strictly well ranked. Lemmas 11 to 13 provide the proof of Theorem 1 for any set of well ranked objects. In contrast to the strictly well ranked case (as in Lemma 6), when objects are well ranked the set of stable learning profiles is only a subset of ex-ante welfare maximising equilibria. For this reason, the proof of Theorem 1 utilises a wider set of learning profiles, $A^*$ in order to identify the ex-ante welfare maximising equilibria in $A^E$. A learning profile $a$ is an element of $A^*$ if and only if it meets the following three conditions:

**I**: If $n$ is even then all agents are in 2-cycles. If $n$ is odd then all agents are in 2-cycles except for some $i^*$ such that $\pi_{i^*}E_{i^*} = \pi_n E_n$.

and for any pair of agents $i, j$ such that $\pi_i > \pi_j$ and $(i, j) \in o(a)$:

**II**$^*$: If there are two agents $i', j'$ such that $(i', j') \in o(a)$ and $\pi_i = \pi_{i'}$ then $\pi_{j'} \geq \pi_i$.

**III**$^*$: There is no agent $j^*$ such that either $\pi_i = \pi_{j^*} > \pi_j$ and $E_i > E_{j^*} \geq E_j$ or $\pi_i > \pi_{j^*} \geq \pi_j$ and $E_i \geq E_{j^*} > E_j$.

The set $A^*$ has some similarities with the set $A^S$ in Section 1.4. The difference is that conditions **II**$^*$ and **III**$^*$ refer separately to the probability an object is acceptable, $\pi_i$, and its conditional expected value, $E_i$. However, as shown in Lemma 15 when objects are strictly well ranked, the sets $A^S$ and $A^*$ are identical.

**Lemma 11.** *When objects are well ranked, if $a$ is an equilibrium and not an element of $A^*$ then $a$ is also not in the set of ex-ante welfare maximising equilibria, $A^E$.*

*Proof.* Fix some equilibrium learning profile $a^E \in A^e$ such that $a^E \notin A^*$. Since $a^E \notin A^*$ it must violate one or more of **I**, **II**$^*$ and **III**$^*$. I consider each of these conditions in turn and show that in every case a learning profile violating a condition cannot be an ex-ante welfare maximising equilibrium since it is possible to construct an alternative learning profile which is also an equilibrium and yields higher ex-ante welfare.

If $a^E$ violates **I** then either $a^E \notin \arg\max_{a \in A} |C_2(a)|$ or $a^E \in \arg\max_{a \in A} |C_2(a)|$ and there is some $\{i^*\} = B_2(a^E)$ such that $\pi_{i^*} E_{i^*} \neq \min_{i \in N} \pi_i E_i$. First consider the case where $a^E \notin \arg\max_{a \in A} |C_2(a)|$, then there are two agents $b, b' \in B_2(a^E)$ as in Figure 26(a). By Lemma 4 none of the agents in $B_2(a^E)$ are in learning cycles and so $U_b(a^E) = 0$ and $U_{b'}(a^E) = 0$. Under $a'$ let $a'_b = b'$ and $a'_{b'} = b$. Since $a^E \in A^e$, $a_i \notin \{b, b'\}$ for all $i \in N$ and so $(a'_{\{b,b'\}}, a_{-\{b,b'\}})$, as in Figure 26(b), is also an equilibrium. Since $(b, b') \in o(a'_{\{b,b'\}}, a_{-\{b,b'\}})$, $U_b(a'_{\{b,b'\}}, a_{-\{b,b'\}}) > 0$ and $U_{b'}(a'_{\{b,b'\}}, a_{-\{b,b'\}}) > 0$. Since $o(a'_{\{b,b'\}}, a_{-\{b,b'\}}) = o(a^E) \cup \{(b, b')\}$, $W(a^E) < W(a'_{\{b,b'\}}, a_{-\{b,b'\}})$. Then $W(a^E) < W^E$ and $a^E \notin A^E$.

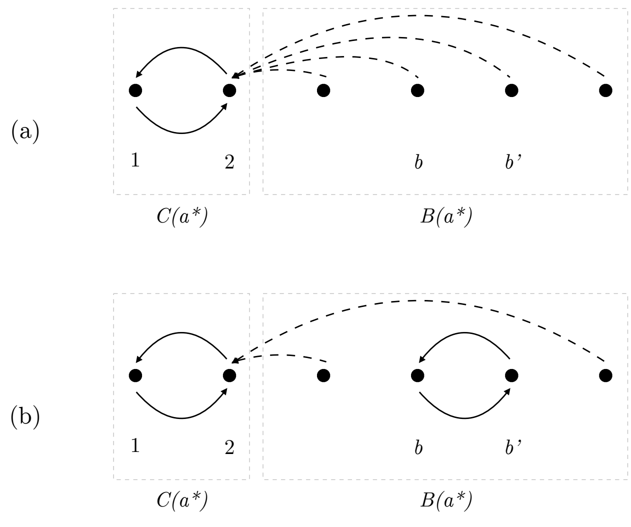



Figure 26: Two learning profiles: (a) $a^E$, (b) $(a'_{\{b,b'\}}, a_{-\{b,b'\}})$. Ex-ante welfare can be increased by pairing agents not in learning cycles in 2-cycles

Next consider the case that $a^E \in \arg\max_{a\in A} |C_2(a)|$ and there is some $\{i^*\} = B_2(a^E)$ such that $\pi_{i^*}E_{i^*} \neq \min_{i\in N} \pi_i E_i$, then there is some $(j, n) \in o(a^E)$ such that $\pi_n E_n = \min_{i\in N} \pi_i E_i$. An example is shown in Figure 27(a). Under $a^E$, $U_{i^*}(a^E) = 0$, $U_j(a^E) = \pi_j \pi_n E_n$ and $U_n(a^E) = \pi_n \pi_j E_j$. Let $a'_j = i^*$ and $a'_{i^*} = j$. Since $|o(a^E)| = |o(a'_{\{i^*,j\}}, a_{-\{i^*,j\}})|$ and $\{n\} = B(a'_{\{i^*,j\}}, a_{-\{i^*,j\}})$, by Lemma 4, $(a'_{\{i^*,j\}}, a_{-\{i^*,j\}})$ is an equilibrium, as shown in Figure 27(b). Under the learning profile $(a'_{\{i^*,j\}}, a_{-\{i^*,j\}})$, $U_{i^*}(a'_{\{i^*,j\}}, a_{-\{i^*,j\}}) = \pi_{i^*}\pi_j E_j$, $U_j(a'_{\{i^*,j\}}, a_{-\{i^*,j\}}) = \pi_j \pi_{i^*} E_{i^*}$ and $U_n(a'_{\{i^*,j\}}, a_{-\{i^*,j\}}) = 0$. Since $\pi_{i^*}E_{i^*} \neq \min_{i\in N} \pi_i E_i$, $\pi_j \pi_n E_n < \pi_j \pi_{i^*} E_{i^*}$. Since objects are well ranked, $\pi_n \leq \pi_i$ for all $i \in N$ so $\pi_n \pi_j E_j \leq \pi_{i^*}\pi_j E_j$. Then, $\sum_{i\in\{i^*,j,n\}} U_i(a^E) < \sum_{i\in\{i^*,j,n\}} U_i(a'_{\{i^*,j\}}, a_{-\{i^*,j\}})$. Since every $i \in N \setminus \{i^*, j, n\}$ is in the same learning cycle under $a^E$ and $(a'_{\{i^*,j\}}, a_{-\{i^*,j\}})$, $W(a^E) < W(a'_{\{i^*,j\}}, a_{-\{i^*,j\}})$ and so $a^E \notin A^E$.

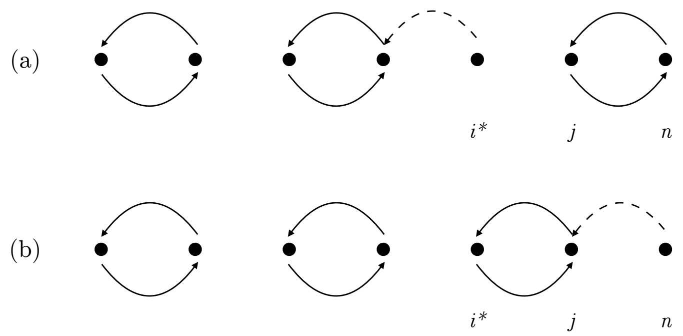



Figure 27: Two learning profiles: (a) $a^E$, (b) $(a'_{\{i^*,j,n\}}, a_{-\{i^*,j,n\}})$. If $n$ is odd then then agent $n$ is not in a 2-cycle.

If $a^E$ violates **II**$^*$ then there are two cycles $(i, j), (i', j') \in o(a^E)$ such that $\pi_i = \pi_{i'} > \pi_j$ and $\pi_i > \pi_{j'}$ as shown in Figure 28(a). Let $a'_i = i'$, $a'_{i'} = i$, $a'_j = j'$, $a'_{j'} = j$ and $S = \{i, i', j, j'\}$. If $a^E \in A^E$ then $a^E$ must meet **I** and since $C_2(a^E) = C_2(a'_S, a^E_{-S})$, by Theorem 4, $(a'_S, a^E_{-S})$ is also an equilibrium as shown in Figure 28(b). If $a \in A^E$ then

$W(a^E) \geq W(a'_S, a^E_{-S})$. Since all $i \in S$ are in 2-cycles with only other agents in $S$ under both learning profiles, $W(a^E)$ and $W(a'_S, a^E_{-S})$ differ only in the sum of the expected utilities of the four agents in $S$. To see $\sum_{i \in S} U_i(a'_S, a^E_{-S}) > \sum_{i \in S} U_i(a^E)$ first recall $\pi_{i'} > \pi_j$ and $\pi_i > \pi_{j'}$. Then the following must hold:

$$\pi_j(\pi_i E_i - \pi_{j'} E_{j'}) + \pi_{j'}(\pi_{i'} E_{i'} - \pi_j E_j) < \pi_{i'}(\pi_i E_i - \pi_{j'} E_{j'}) + \pi_i(\pi_{i'} E_{i'} - \pi_j E_j)$$
$$\Rightarrow \pi_i \pi_j (E_i + E_j) + \pi_{i'} \pi_{j'} (E_{i'} + E_{j'}) < \pi_i \pi_{i'} (E_i + E_{i'}) + \pi_j \pi_{j'} (E_j + E_{j'})$$
$$\Rightarrow \sum_{i \in S} U_i(a^E) < \sum_{i \in S} U_i(a'_S, a^E_{-S})$$

Since $W(a'_S, a^E_{-S}) > W(a^E)$, $a^E$ cannot be an ex-ante welfare maximising equilibrium if $a^E$ violates **II**$^*$.

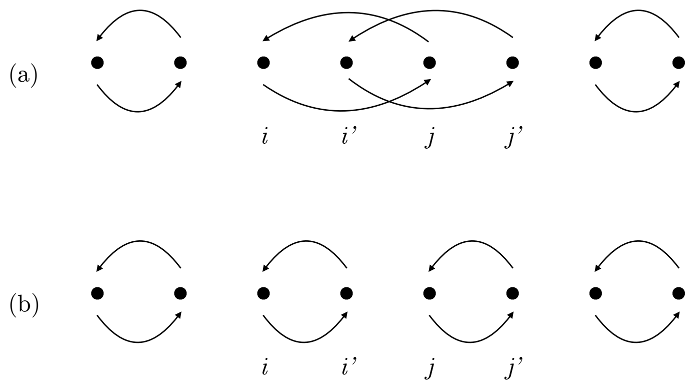



Figure 28: Two learning profiles: (a) $a^E$, (b) $(a'_S, a_{-S})$. Two agents with equally acceptable objects cannot both be in 2-cycles with agents endowed with objects less likely to be acceptable.

If $a^E$ violates **III**$^*$ then for some $(i, j) \in o(a^E)$ there is a $j^*$ such that either $\pi_i = \pi_{j^*} > \pi_j$ and $E_i > E_{j^*} \geq E_j$ or $\pi_i > \pi_{j^*} \geq \pi_j$ and $E_i \geq E_{j^*} > E_j$. If $a^E \in A^E$ then **I** must hold for $a^E$. Since $\pi_{j^*} E_{j^*} > \pi_j E_j$, there is some $(j^*, k) \in o(a)$. An example of such learning cycles is shown in Figure 29(a). If $k > j^*$ then $\pi_k \leq \pi_{j^*}$ and so let $a'_i = j^*$, $a'_{j^*} = i$, $a'_k = j$ and $a'_j = k$ as shown in Figure 29(b). Let $S = \{i, j, j^*, k\}$. Since **I** holds for $a^E$ and $C(a^E) = C(a'_S, a^E_{-S})$, by Lemma 4, $(a'_S, a^E_{-S})$ is also an equilibrium and an example is shown in Figure 29(b). If $a \in A^{E*}$ then $W(a^E) \geq W(a'_S, a^E_{-S})$. Since every $i \in N \setminus S$ is in the same learning cycle under $a^E$ and $(a'_S, a^E_{-S})$, $W(a^E)$ and $W(a'_S, a^E_{-S})$ can differ only in the utilities of the four agents in $S$. Since either $\pi_i = \pi_{j^*} \geq \pi_k$ and $\pi_{j^*} > \pi_j$ or $\pi_i > \pi_{j^*} \geq \pi_k$ and $\pi_{j^*} \geq \pi_k$ the following must hold:

$$\pi_j(\pi_i E_i - \pi_k E_k) + \pi_k(\pi_{j^*} E_{j^*} - \pi_j E_j) < \pi_{j^*}(\pi_i E_i - \pi_k E_k) + \pi_i(\pi_{j^*} E_{j^*} - \pi_j E_j)$$
$$\Rightarrow \pi_i \pi_j (E_i + E_j) + \pi_{j^*} \pi_k (E_{j^*} + E_k) < \pi_i \pi_{j^*} (E_i + E_{j^*}) + \pi_j \pi_k (E_j + E_k)$$
$$\sum_{i \in S} U_i(a^E) < \sum_{i \in S} U_i(a'_S, a^E_{-S})$$

Then $W(a^E) < W(a'_S, a^E_{-S})$ and so $a \notin A^E$. If $k < j^*$ then $\pi_k \geq \pi_{j^*}$, $a'_i = k$, $a'_k = i$, $a'_{j^*} = j$ and $a'_j = j^*$, an example of which is shown in Figure 30(b). The above argument for $k > j^*$ can then be applied to the case where $k < j^*$ mutatis mutandis. □

**Lemma 12.** *When objects are well ranked, all learning profiles in $A^*$ yield the same ex-ante welfare and so $A^*$ is the set of ex-ante welfare maximising equilibria $A^E$.*

*Proof.* Let $\beta = \{\beta^1, ...\beta^{\bar{r}}\}$ be a partition on $N$ such that for any $i \in \beta^t$ and $i' \in \beta^{t'}$, $\pi_i > \pi_{i'}$ if and only if $t < t'$. Ex-ante welfare of any learning profile $a$ can then be expressed as:

$$W(a) = \sum_{i=1}^{n} U_i(a) = \sum_{i=1}^{\bar{r}} \sum_{i \in \beta^r} U_i(a)$$

To define $\sum_{i \in \beta^r} U_i(a)$ for any given $\beta^r$, consider each of the following cases. In all cases $i' \in \arg\max_{i \in \beta^r} \pi_i E_i$ and $i'' \in \arg\min_{i \in \beta^r} \pi_i E_i$ . If $\beta^{r-1}$ exists then $i^- \in \arg\min_{i \in \beta^{r-1}} \pi_i E_i$ and if $\beta^{r+1}$ exists then $i^+ \in \arg\max_{i \in \beta^{r+1}} \pi_i E_i$. Examples of such agents are shown in Figure 31.

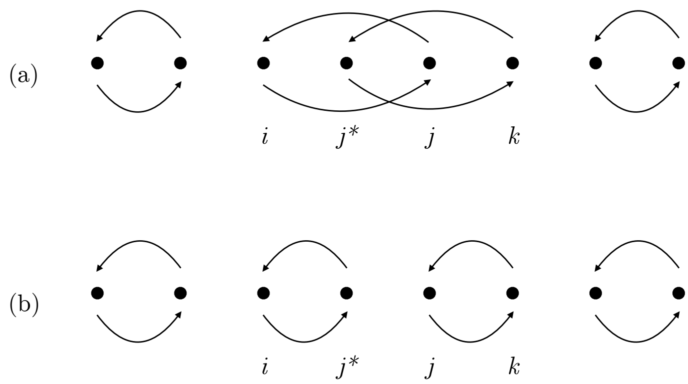



Figure 29: Two learning profiles for the case where $k > j^*$: (a) $a^E$, (b) $(a'_S, a_{-S})$. A learning profile $a$ cannot violate **III**$^*$ if $a$ is a welfare maximising equilibrium.

*Case 1: Either* $r = 1$ *or* $|\cup_{i=1}^{r-1} \beta^r|$ *even.*

*Case 1a:* $|\beta^r|$ *even.* If $|\beta^r|$ even then **I**, **II**$^*$ and **III**$^*$ imply that no $i \in \beta^r$ is in a learning cycle with any $i \in N \setminus \beta^r$. By **I** all agents in $\beta^r$ must be in 2-cycles with other agents in $\beta^r$ as shown in Figure 31(a) Then $\sum_{i\in\beta^r} U_i(a) = \sum_{i\in\beta^r} U_i(a) = \pi_i^2 E_i$.

*Case 1b:* $|\beta^r|$ *odd.* If $|\beta^r|$ odd then **I**, **II**$^*$ and **III**$^*$ imply no $i \in \beta^r$ is in a learning cycle with any $i \in \beta^{r'}$ where $r' < r$. By **I** all agents in $\beta^r \setminus \{i''\}$ must be in 2-cycles with other agents in $\beta^r \setminus \{i''\}$. Again by **I**, if $r = \bar{r}$ then $i''$ is not in a learning cycle (as shown in Figure 31(b)). Then $\sum_{i\in\beta^r} U_i(a) = \sum_{i\in\beta^r\setminus\{i''\}} \pi_i^2 E_i$. If $r < \bar{r}$ then **I**, **II**$^*$ and **III**$^*$ imply $(i'', i^+) \in o(a)$ and so $\sum_{i\in\beta^r} U_i(a) = \sum_{i\in\beta^r\setminus\{i''\}} \pi_i^2 E_i + \pi_{i''}\pi_{i^+} E_{i^+}$ (as shown in Figure 31(c)).

*Case 2:* $|\cup_{i=1}^{r-1} \beta^r|$ *odd.*

*Case 2a:* $|\beta^r|$ *even.* If $|\beta^r|$ even then **I**, **II**$^*$ and **III**$^*$ imply that $(i^-, i') \in o(a)$ and

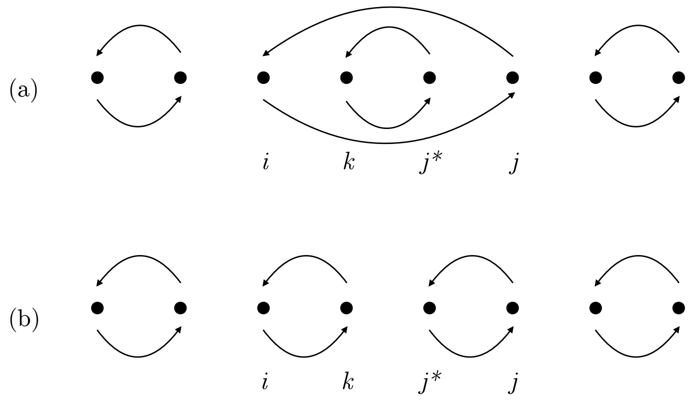



Figure 30: Two learning profiles for the case where $j^* > k$: (a) $a^E$, (b) $(a'_S, a_{-S})$. A learning profile $a$ cannot violate **III**$^*$ if $a$ is a welfare maximising equilibrium.

all $i \in \beta^r \setminus \{i', i''\}$ are in 2-cycles with other agents in $\beta^r \setminus \{i', i''\}$. If $r = \bar{r}$ then by **I**, $i''$ is not in a learning cycle (as shown in Figure 32(a)) and so $\sum_{i \in \beta^r} U_i(a) = \sum_{i \in \beta^r \setminus \{i', i''\}} \pi_i^2 E_i + \pi_{i'} \pi_{i^-} E_{i^-}$. If $r < \bar{r}$ then **I**, **II**$^*$ and **III**$^*$ imply $(i'', i^+) \in o(a)$ (as shown in Figure 32(b)) and so $\sum_{i \in \beta^r} U_i(a) = \sum_{i \in \beta^r \setminus \{i', i''\}} \pi_i^2 E_i + \pi_{i'} \pi_{i^-} E_{i^-} + \pi_{i''} \pi_{i^+} E_{i^+}$.

*Case 2b:* $|\beta^r|$ *odd.* If $|\beta^r|$ odd then **I**, **II**$^*$ and **III**$^*$ imply that $(i^-, i') \in o(a)$ and since $|\beta^r \setminus \{i'\}|$ is even, all $i \in \beta^r \setminus \{i'\}$ are in 2-cycles with other agents in $\beta^r \setminus \{i'\}$ (as shown in Figure 32(c)). Then $\sum_{i \in \beta^r} U_i(a) = \sum_{i \in \beta^r \setminus \{i'\}} \pi_i^2 E_i + \pi_{i'} \pi_{i^-} E_{i^-}$.

Then when $a \in A^*$, $\sum_{i \in \beta^r} U_i(a)$, and in turn $\sum_{i=1}^n U_i(a)$, depends only on the the number of agents in each $\beta^r$. Since the number of agents does not change with the learning profile, all $a \in A^*$ must yield the same ex-ante welfare. By Lemma 4, $A^e \neq \emptyset$ and so $A^E$ must be nonempty. By Lemma 11 there is no $a \in A^E$ that is not also an element of $A^*$. Since all $a \in A^*$ yield the same ex-ante welfare and by Lemma 4 all $a \in A^*$ are equilibria, $A^* = A^E$. □

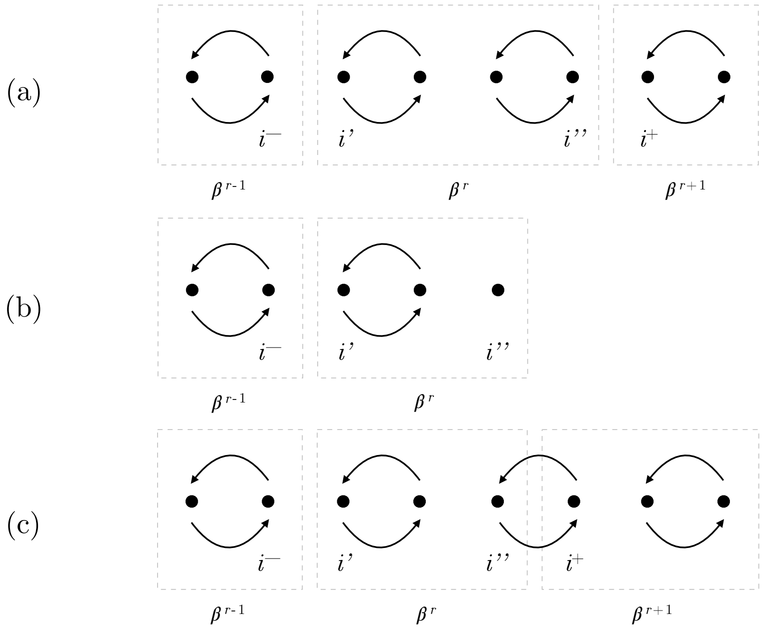



Figure 31: Learning cycles for agents in $\beta^r$ when $|\bigcup_{i=1}^{r-1} \beta^r|$ is even.

**Lemma 13.** *If objects are well ranked then any stable learning profile is an ex-ante welfare maximising equilibrium.*

*Proof.* Let $a \in A^S$ and $a \notin A^*$. Then $a$ violates at least one of conditions $\mathbf{II}^*$ and $\mathbf{III}^*$. If $a$ violates $\mathbf{II}^*$ then there are two pairs of agents $i, j$ and $i', j'$ such that $(i, j), (i', j') \in o(a)$, $\pi_i > \pi_j$ and $\pi_i = \pi_{i'} > \pi_{j'}$. Since objects are well ranked and $\pi_i > \pi_{j'}$, $E_i \geq E_{j'}$ and so $\pi_i E_i > \pi_{j'} E_{j'}$. Now consider the relationship between $E_i$ and $E_{i'}$.

*Case 1:* $E_i = E_{i'}$. If $E_i = E_{i'}$ then $\pi_i E_i = \pi_{i'} E_{i'}$ and by condition **III**, $\pi_i E_i \leq \pi_{j'} E_{j'}$, which is a contradiction.

*Case 2:* $E_i > E_{i'}$. If $E_i > E_{i'}$ then since $\pi_i = \pi_{i'}$, $\pi_i E_i > \pi_{i'} E_{i'}$. Since $\pi_i = \pi_{i'}$ and $\pi_i > \pi_j$, $\pi_{i'} > \pi_j$ which, since objects are well ranked, implies $E_{i'} \geq E_j$ and so $\pi_{i'} E_{i'} > \pi_j E_j$. Then $\pi_i E_i > \pi_{i'} E_{i'} > \pi_j E_j$ and since $(i, j) \in o(a)$, this violates condition **III** implying $a \notin A^*$.

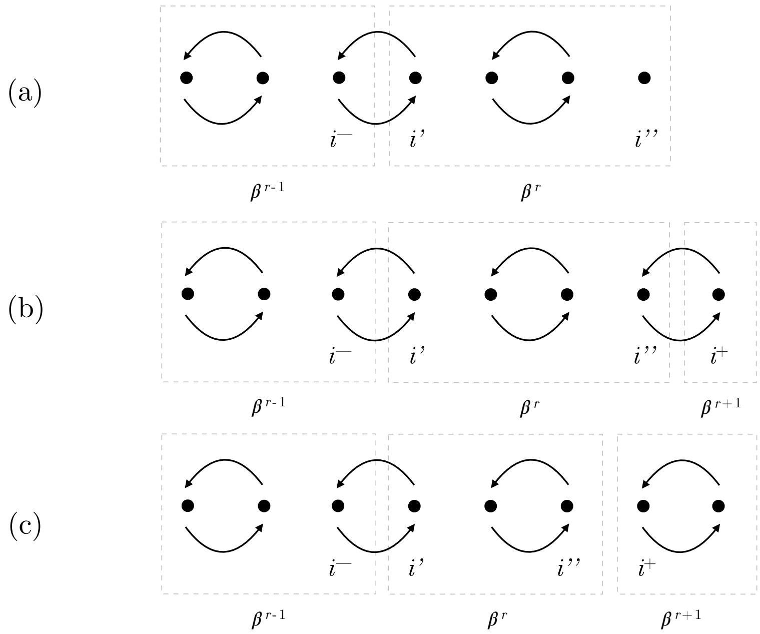



Figure 32: Learning cycles for agents in $\beta^r$ when $|\bigcup_{i=1}^{r-1}\beta^r|$ is odd.

*Case 3:* $E_i < E_{i'}$. If $E_{i'} > E_i$ then since $\pi_i = \pi_{i'}$, $\pi_{i'}E_{i'} > \pi_i E_i$. Since objects are well ranked and $\pi_i > \pi_{j'}$, $E_i \geq E_{j'}$ which implies $\pi_i E_i > \pi_{j'}E_{j'}$. Then $\pi_{i'}E_{i'} > \pi_i E_i > \pi_{j'}E_{j'}$ and since $(i', j') \in o(a)$, this violates condition **III** implying $a \notin A^S$.

If $a$ violates **III**$^*$ then there exists a three agents $i, j, j^* \in N$ such that $(i, j) \in o(a)$ and either $\pi_i = \pi_{j^*} > \pi_j$ and $E_i > E_{j^*} \geq E_j$ or $\pi_i > \pi_{j^*} \geq \pi_j$ and $E_i \geq E_{j^*} > E_j$. In either case, $\pi_i E_i > \pi_{j^*}E_{j^*} > \pi_j E_j$ and since $(i, j) \in o(a)$, this violates condition **III** implying $a \notin A^S$.

So it must be that if $a \in A^S$ then $a \in A^*$ and since by Lemma 12 $A^* = A^E$, $a$ is an ex-ante welfare maximising equilibrium. □

**Lemma 14.** *When objects are well ranked, the maximum ex-ante welfare that can be achieved in equilibrium is:*

$$W^E = \sum_{i \in N, i \text{ even}} \pi_{i-1}\pi_i(E_{i-1} + E_i)$$

*Proof.* By Observation 1 $A^\circ \subseteq A^S$ and by Lemmas 2 and 13, $A^S \subseteq A^E$. Following 2, the set of learning cycles formed under any $a^\circ \in A^\circ$ is $o^\circ$ and so the maximum ex-ante welfare that can be achieved in equilibrium is:

$$W^E = W(a^\circ) = \sum_{i \in N, i \text{ even}} \pi_{i-1}\pi_i(E_{i-1} + E_i)$$

for all $a^\circ \in A^\circ \subseteq A^E$. □

Lemma 6 demonstrated that when objects are strictly well ranked and $n = 4$, a learning profile is an ex-ante welfare maximising equilibrium if and only if it is stable. This is not a coincidence and holds for any set of strictly well ranked objects.

**Lemma 15.** *If objects are strictly well ranked then a learning profile is an ex-ante welfare maximising equilibrium if and only if it is stable.*

*Proof.* When objects are well ranked, $A^S$ is characterised by conditions **I** and **III** and $A^{**}$ is characterised by conditions **I** and **III**$^*$. Condition **III**$^*$ states that for any pair of agents $i, j$ such that $\pi_i > \pi_j$ and $(i, j) \in o(a)$ there exists no $j^*$ such that either (i) $\pi_i = \pi_{j^*} > \pi_j$ and $E_i > E_j \geq E_j$ or (ii) $\pi_i > \pi_{j^*} \geq \pi_j$ and $E_i \geq E_j > E_j$. Since objects are strictly well ranked, only (ii) is applicable and implies $\pi_i E_i > \pi_{j^*} E_{j^*} > \pi_j E_j$. Then **III**$^*$ implies **III** and so $A^* \subseteq A^S$. By Lemma 12, $A^* = A^E$ and so $A^E \subseteq A^S$. Then by Lemma 13, $A^S \subseteq A^E$ and so $A^S = A^{E*}$. □

# B Closely well ranked objects and ex-ante welfare

When objects are closely well ranked then, as stated in Theorem 3, stable learning profiles yield either the maximum or close to the maximum ex-ante welfare over all learning profiles. In addition to Lemma 9, the proof is given here via Lemmas 16 to 24. Conditions $\mathbf{I}^c$, $\mathbf{II}^c$, $\mathbf{III}^c$ are closely related to **I**, **II** and **III** but allow for 3-cycles in addition to 2-cycles.

**Lemma 16.** ***Condition*** $\boldsymbol{I^c}$*: If* $a \in \arg\max_{a\in A} W(a)$ *then either all agents are in learning cycles or all agents are in learning cycles with the exception of some* $i^*$ *such that* $\pi_{i^*} = \min_{i\in N} \pi_i$ *and* $E_{i^*} = \min_{i\in N} E_i$

*Proof.* Let $a \in \arg\max_{a\in A} W(a)$ and $B(a) > 1$. Since no agent in $B(a)$ is in a learning cycle, $U_i(a) = 0$ for all $i \in B(a)$. Let $(a'_{B(a)}, a_{-B(a)})$ be such that $o(a'_{B(a)}, a_{-B(a)})$ contains a $|B(a)|$-cycle between all agents in $B(a)$. Then $U_i(a'_{B(a)}, a_{-B(a)}) > 0$ for all $i \in B(a)$. Since no agent in $B(a)$ is in a learning cycle under $a$, all agents in $C(a)$ are in the same learning cycle under both $a$ and $(a'_{B(a)}, a_{-B(a)})$ and so $\sum_{i\in C(a)} U_i(a) = \sum_{i\in C(a)} U_i(a'_{B(a)}, a_{-B(a)})$. Then $\sum_{i\in N} U_i(a) < \sum_{i\in N} U_i(a'_{B(a)}, a_{-B(a)})$ and $a \notin \arg\max_{a\in A} W(a)$. So it must be that $|B(a)| \leq 1$.

Now suppose $B(a) = \{i^*\}$ and $\pi_{i^*} \neq \min_{i\in N} \pi_i$. Since objects are well ranked and $\pi_{i^*} > \pi_n$ it must be that $E_{i^*} \geq E_n$. Since $B(a) = \{i^*\}$, $n \in C(a)$ and $n$ is in some $m$-cycle consisting of all agents in $T \subset N$, where $n \in T$. Under this $|T|$-cycle:

$$\sum_{i\in T\cup i^*} U_i(a) = \pi_n \prod_{i\in T\setminus\{n\}} \pi_i \left(E_n + \sum_{i\in T\setminus\{n\}} E_i\right) \tag{15}$$

Let $(a'_{T\cup\{i^*\}}, a_{-T\cup\{i^*\}})$ be such that $n$ is not in a learning cycle and $o(a'_{T\cup\{i^*\}}, a_{-T\cup\{i^*\}})$ contains a single $|T|$-cycle between all agents in $(T\setminus\{n\})\cup\{i^*\}$. Under this new learning profile:

$$\sum_{i\in T\cup i^*} U_i(a'_{T\cup\{i^*\}}, a_{-T\cup\{i^*\}}) = \pi_{i^*} \prod_{i\in T\setminus\{i^*\}} \pi_i \left(E_{i^*} + \sum_{i\in T\setminus\{i^*\}} E_i\right) \tag{16}$$

Since $T \setminus \{n\} = T \setminus \{i^*\}$, $\pi_{i^*} > \pi_n$ and $E_{i^*} \geq E_n$:

$$\sum_{i \in T \cup i^*} U_i(a) < \sum_{i \in T \cup i^*} U_i(a'_{T\cup\{i^*\}}, a_{-T\cup\{i^*\}}) \tag{17}$$

Since no agent in $T \cup \{i^*\}$ is in a learning cycle with any agent in $C(a) \setminus (T \cup \{i^*\})$, all agents in $C(a) \setminus (T \cup \{i^*\})$ are in the same learning cycles under both $a$ and $(a'_{T\cup\{i^*\}}, a_{-T\cup\{i^*\}})$ and so for these agents:

$$\sum_{i \in C(a)\setminus(T\cup\{i^*\})} U_i(a) = \sum_{i \in C(a)\setminus(T\cup\{i^*\})} U_i(a'_{T\cup\{i^*\}}, a_{-T\cup\{i^*\}}) \tag{18}$$

But then $a \notin \arg\max_{a \in A} W(a)$ since $\sum_{i \in N} U_i(a) > \sum_{i \in N} U_i(a'_{T\cup\{i^*\}}, a_{-T\cup\{i^*\}})$. So it must be that $\pi_{i^*} = \min_{i \in N} \pi_i$.

Now suppose $E_{i^*} > E_n$. Repeating the argument above, if $\pi_{i^*} = \min_{i \in N} \pi_i$ and $E_{i^*} > E_n$ then the inequality in (16) and the equality in (17) hold, which again implies that $a \notin \arg\max_{a \in A} W(a)$. So it must be that $E_{i^*} = \min_{i \in N} E_i$.

□

**Lemma 17.** *If $a \in \arg\max_{a \in A} W(a)$ such that $o(a)$ contains a learning cycle between all agents in $S \subset N$ then for all $i, j \in S$ such that $\pi_i > \pi_j$:*

> ***Condition II$^c$:*** *If there is another set of agents $S'$ in a cycle under $a$ then for any $i', j' \in S$ such that $\pi_i = \pi_{i'}$ it must be that $\pi_{j'} \geq \pi_i$.*
>
> ***Condition III$^c$:*** *There is no agent $j^* \in N$ such that either (i) $\pi_i = \pi_{j^*} > \pi_j$ and $E_i > E_{j^*} \geq E_j$ or (ii) $\pi_i > \pi_{j^*} \geq \pi_j$ and $E_i \geq E_{j^*} > E_j$.*

*Proof.* Let $a \in \arg\max_{a \in A} W(a)$ such that under $a$ a learning cycle forms between all agents in $S \subset N$ and for some $i, j \in S$, $\pi_i > \pi_j$. By Lemma 9 the agents in $S$ are in either a 2-cycle or a 3-cycle.

Suppose $a$ violates $\mathbf{II}^c$. Then there is also another set of agents $S'$ in a cycle under $a$ such that for some $i', j' \in S$, $\pi_i = \pi_{i'}$ and $\pi_j < \pi_i$. By Lemma 9 the agents in $S'$ are also in either a 2-cycle or a 3-cycle. Let $a'$ be such that $o(a')$ contains a learning cycle between all agents in $(S \cup i') \setminus \{j\}$ and a learning cycle between all agents in $(S' \cup j) \setminus \{i^*\}$ (as shown in Figures 33, 34 and 35). Since no agent is in $N \setminus (S \cup S')$ is in a learning cycle with any agent in $S \cup S'$, all agents in $N \setminus (S \cup S')$ are in the same learning cycles under $a$ and $(a'_{(S \cup S')}, a_{-\{(S \cup S')\}})$. Then:

$$\sum_{k \in N \setminus (S \cup S')} U_k(a) = \sum_{k \in N \setminus (S \cup S')} U_k(a'_{(S \cup S')}, a_{-\{(S \cup S')\}})$$

.

**Case 1:** $S, S' \subset C_3(a)$

If all agents in $S$ and $S'$ are in a 3-cycles under $a$ (as shown in Figure 33) then $a \notin \arg\max_{a \in A} W(a)$ as $\sum_{k \in (S \cup S')} U_k(a) < \sum_{k \in (S \cup S')} U_k(a'_{(S \cup S')}, a_{-\{(S \cup S')\}})$. To see this, note that since $\pi_i = \pi_{i'} > \pi_j$, $E_{i'} \geq E_j$ and so the following inequality must hold:

$$E_{i'}(\pi_i \pi_k - \pi_{j'} \pi_{k'}) \geq E_j(\pi_i \pi_k - \pi_{j'} \pi_{k'})$$

Adding $\pi_i \pi_k (E_i + E_k) - \pi_{j'} \pi_{k'} (E_{j'} + E_{k'})$ to both sides of the inequality:

$$\pi_i \pi_k (E_i + E_{i'} + E_k) - \pi_{j'} \pi_{k'} (E_{i'} + E_{j'} + E_{k'}) \geq \pi_i \pi_k (E_i + E_j + E_k) - \pi j' \pi k' (E_j + E_{j'} + E_{k'})$$

Since $\pi_{i'} > \pi_j > 0$:

$$\pi_i \pi_{i'} \pi_k (E_i + E_{i'} + E_k) + \pi_j \pi_{j'} \pi_{k'} (E_j + E_{j'} + E_{k'}) > \pi_i \pi_j \pi_k (E_i + E_j + E_k) + \pi_{i'} \pi_{j'} \pi_{k'} (E_{i'} + E_{j'} + E_{k'})$$

$$\sum_{k \in (S \cup S')} U_k(a'_{(S \cup S')}, a_{-\{(S \cup S')\}}) > \sum_{k \in (S \cup S')} U_k(a)$$

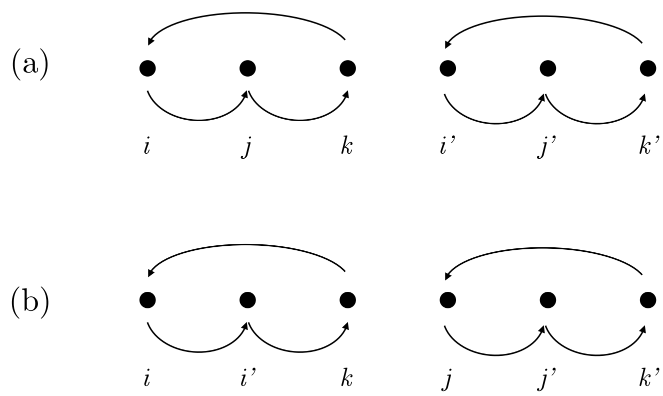



Figure 33: Violating condition $\mathbf{II}^c$ - Case 1. (a) Two learning cycles under $a$ between $S$ and $S'$. (b) Two learning cycles under $(a'_{(S\cup S')}, a_{-\{(S\cup S')\}})$ between agents in $(S \cup i') \setminus \{j\}$ and agents in $(S' \cup j) \setminus \{i^*\}$.

**Case 2:** $S \subset C_3(a)$ and $S' \subset C_2(a)$

If the agents in $S$ are in a 3-cycle and those in $S'$ are in a 2-cycle under $a$ (as shown in Figure 34) then $a \notin \arg\max_{a\in A} W(a)$ as $\sum_{k\in(S\cup S')} U_k(a) < \sum_{k\in(S\cup S')} U_k(a'_{(S\cup S')}, a_{-\{(S\cup S')\}})$. To see this, note that since $\pi_i = \pi_{i'} > \pi_j$, $E_{i'} \geq E_j$ and so the following inequality must hold:

$$E_{i'}(\pi_i\pi_k - \pi_{j'}) \geq E_j(\pi_i\pi_k - \pi_{j'})$$

Adding $\pi_i\pi_k(E_i + E_k) - \pi_{j'}E_{j'}$ to both sides of the inequality:

$$\pi_i\pi_k(E_i + E_{i'} + E_k) - \pi_{j'}(E_{i'} + E_{j'}) \geq \pi_i\pi_k(E_i + E_j + E_k) - \pi j'(E_j + E_{j'})$$

Since $\pi_{i'} > \pi_j > 0$:

$$\pi_i \pi_{i'} \pi_k (E_i + E_{i'} + E_k) + \pi_j \pi_{j'} (E_j + E_{j'}) > \pi_i \pi_j \pi_k (E_i + E_j + E_k) + \pi_{i'} \pi_{j'} (E_{i'} + E_{j'})$$
$$\sum_{k \in (S \cup S')} U_k(a'_{(S \cup S')}, a_{-\{(S \cup S')\}}) > \sum_{k \in (S \cup S')} U_k(a)$$

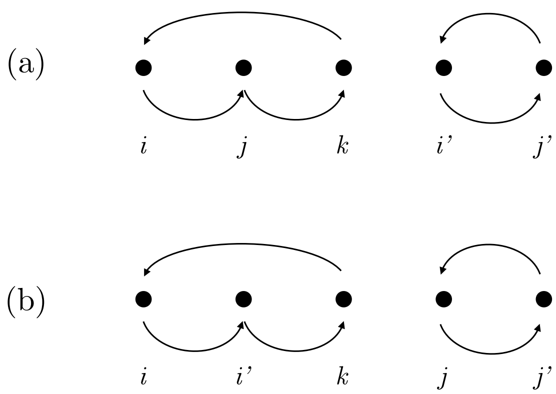



Figure 34: Violating condition $\mathbf{II}^c$ - Case 2. (a) Two learning cycles under $a$ between $S$ and $S'$. (b) Two learning cycles under $(a'_{(S \cup S')}, a_{-\{(S \cup S')\}})$ between agents in $(S \cup i') \setminus \{j\}$ and agents in $(S' \cup j) \setminus \{i^*\}$.

**Case 3:** $S \subset C_2(a)$ and $S' \subset C_3(a)$

If the agents in $S$ are in a 2-cycle and those in $S'$ are in a 3-cycle under $a$ (as shown in Figure 35) then $a \notin \arg\max_{a \in A} W(a)$ as $\sum_{k \in (S \cup S')} U_k(a) < \sum_{k \in (S \cup S')} U_k(a'_{(S \cup S')}, a_{-\{(S \cup S')\}})$. To see this, note that since $\pi_i = \pi_{i'} > \pi_j$, $E_{i'} \geq E_j$ and so the following inequality must hold:

$$E_{i'}(\pi_i - \pi_{j'} \pi_{k'}) \geq E_j(\pi_i - \pi_{j'} \pi_{k'})$$

Adding $\pi_i E_i - \pi_{j'}\pi_{k'}(E_{j'} + E_{k'})$ to both sides of the inequality:

$$\pi_i(E_i + E_{i'}) - \pi_{j'}\pi_{k'}(E_{i'} + E_{j'} + E_{k'}) \geq \pi_i(E_i + E_j) - \pi j'\pi k'(E_j + E_{j'} + E_{k'})$$

Since $\pi_{i'} > \pi_j > 0$:

$$\pi_i\pi_{i'}(E_i + E_{i'}) + \pi_j\pi_{j'}\pi_{k'}(E_j + E_{j'} + E_{k'}) > \pi_i\pi_j(E_i + E_j) + \pi_{i'}\pi_{j'}\pi_{k'}(E_{i'} + E_{j'} + E_{k'})$$
$$\sum_{k\in(S\cup S')} U_k(a'_{(S\cup S')}, a_{-\{(S\cup S')\}}) > \sum_{k\in(S\cup S')} U_k(a)$$

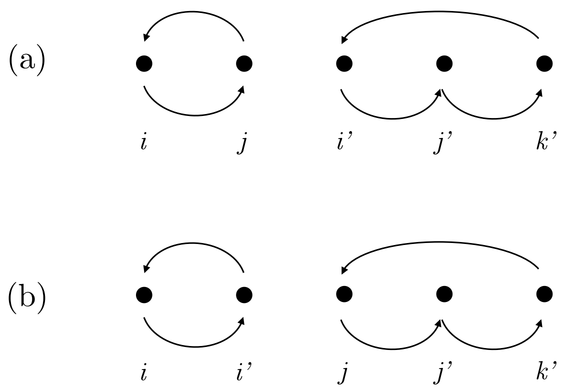



Figure 35: Violating condition $\mathbf{II}^c$ - Case 3. (a) Two learning cycles under $a$ between $S$ and $S'$. (b) Two learning cycles under $(a'_{(S\cup S')}, a_{-\{(S\cup S')\}})$ between agents in $(S \cup i') \setminus \{j\}$ and agents in $(S' \cup j) \setminus \{i^*\}$.

**Case 4:** $S, S' \subset C_2(a)$

If $S, S' \in C_2(a)$ then the proof in Lemma 11 regarding condition $\mathbf{II}^*$ applies and again $\sum_{k\in(S\cup S')} U_k(a'_{(S\cup S')}, a_{-\{(S\cup S')\}}) > \sum_{k\in(S\cup S')} U_k(a)$.

Since $\sum_{k\in(S\cup S')} U_k(a'_{(S\cup S')}, a_{-\{(S\cup S')\}}) > \sum_{k\in(S\cup S')} U_k(a)$ holds for all cases, $W(a) <$

$W(a'_{(S\cup S')}, a_{-\{(S\cup S')\}})$ and so $a \notin \arg\max_{a\in A} W(a)$ if it violates $\mathbf{II}^c$.

Now suppose $a$ violates $\mathbf{III}^c$ and there is some $j^* \in N$ such that either (i) $\pi_i = \pi_{j^*} > \pi_j$ and $E_i > E_{j^*} \geq E_j$ or (ii) $\pi_i > \pi_{j^*} \geq \pi_j$ and $E_i \geq E_{j^*} > E_j$. Since either $\pi_{j^*} > \pi_j$ or $E_{j^*} > E_j$, by Lemmas 9 and 16, $j^*$ must be in either a 2-cycle or a 3-cycle under $a$. But then the arguments for Cases 1 to 4 above can be applied mutatis mutandis for each of conditions (i) and (ii), replacing $j'$ with $j^*$ . □

**Lemma 18.** *For a set of agents $S = \{i_1, i_2, i_3, i_4, i_5\}$ with well ranked objects, where $\pi_{i_1} \geq \pi_{i_2} \geq \pi_{i_3} \geq \pi_{i_4} \geq \pi_{i_5}$, if two learning profiles $a$ and $a'$ are such that $(i_1, i_2), (i_3, i_4, i_5) \in o(a)$ and $(i_1, i_2, i_3), (i_4, i_5) \in o(a')$ then $\sum_{k\in S} U_k(a) \geq \sum_{k\in S} U_k(a')$.*

*Proof.* W.l.o.g let $S = \{1, 2, 3, 4, 5\}$. Since $\pi_3 \in (0, 1)$:

$$E_1 + E_2 - E_4 - E_5 \geq \pi_3(E_1 + E_2 - E_4 - E_5)$$
$$E_1 + E_2 - E_4 - E_5 \geq \pi_3(E_1 + E_2 + E_3) - \pi_3(E_3 + E_4 + E_5)$$
$$E_1 + E_2 - \pi_3(E_1 + E_2 + E_3) \geq E_4 + E_5 - \pi_3(E_3 + E_4 + E_5)$$

Since $1 \geq \pi_1 \geq \pi_2 \geq \pi_4 \geq \pi_5 > 0$, $\frac{\pi_4\pi_5}{\pi_1\pi_2} \leq 1$:

$$E_1 + E_2 - \pi_3(E_1 + E_2 + E_3) \geq E_4 + E_5 - \pi_3(E_3 + E_4 + E_5)\frac{\pi_4\pi_5}{\pi_1\pi_2}$$
$$\pi_1\pi_2(E_1 + E_2) + \pi_3\pi_4\pi_5(E_3 + E_4 + E_5) \geq \pi_1\pi_2\pi_3(E_1 + E_2 + E_3) + \pi_4\pi_5(E_4 + E_5)$$
$$\sum_{k\in S} U_k(a) \geq \sum_{k\in S} U_k(a')$$

□

**Lemma 19.** *For a set of agents $S = \{i_1, i_2, i_3, i_4, i_5, i_6\}$ with closely well ranked objects, where $\pi_{i_1} \geq \pi_{i_2} \geq \pi_{i_3} \geq \pi_{i_4} \geq \pi_{i_5} \geq \pi_{i_6}$, if two learning profiles $a$ and $a'$ are such that $(i_1, i_2), (i_3, i_4), (i_5, i_6) \in o(a)$ and $(i_1, i_2, i_3), (i_4, i_5, i_6) \in o(a')$ then $\sum_{k\in S} U_k(a) >$*

$\sum_{k \in S} U_k(a')$.

*Proof.* W.l.o.g let $S = \{1, 2, 3, 4, 5, 6\}$ and consider $\sum_{k \in S} U_k(a)$ and $\sum_{k \in S} U_k(a')$:

$$\sum_{k \in S} U_k(a) = \pi_1\pi_2(E_1 + E_2) + \pi_3\pi_4(E_3 + E_4) + \pi_5\pi_6(E_5 + E_6)$$

$$\sum_{k \in S} U_k(a') = \pi_1\pi_2\pi_3(E_1 + E_2 + E_3) + \pi_4\pi_5\pi_6(E_4 + E_5 + E_6)$$

Since $\pi_i \in (0, 1)$ for all $i \in S$, $(\pi_1\pi_2\pi_3)E_1 < (\pi_1\pi_2)E_1$, $(\pi_1\pi_2\pi_3)E_2 < (\pi_1\pi_2)E_2$, $(\pi_4\pi_5\pi_6)E_5 < (\pi_5\pi_6)E_5$ and $(\pi_4\pi_5\pi_6)E_6 < (\pi_5\pi_6)E_6$ . Since objects are also well ranked $\pi_3 > \pi_5\pi_6$ and so $(\pi_4\pi_5\pi_6)E_4 < (\pi_3\pi_4)E_4$. Finally, since objects are close in acceptability, $\pi_1 \leq \frac{\pi_4}{\pi_2}$ and so $(\pi_1\pi_2\pi_3)E_3 \leq (\pi_3\pi_4)E_3$. Then $\sum_{k \in S} U_k(a) > \sum_{k \in S} U_k(a')$. □

**Lemma 20.** *For a set of agents $S = \{i_1, i_2, i_3, i_4\}$ with closely well ranked objects, where $\pi_{i_1} \geq \pi_{i_2} \geq \pi_{i_3} \geq \pi_{i_4}$, if two learning profiles $a$ and $a'$ are such that $(i_1, i_2), (i_3, i_4) \in o(a)$ and $(i_1, i_2, i_3) \in o(a')$ then $\sum_{k \in S} U_k(a) > \sum_{k \in S} U_k(a')$.*

*Proof.* W.l.o.g let $S = \{1, 2, 3, 4\}$ and consider $\sum_{k \in S} U_k(a)$ and $\sum_{k \in S} U_k(a')$:

$$\sum_{k \in S} U_k(a) = \pi_1\pi_2(E_1 + E_2) + \pi_3\pi_4(E_3 + E_4)$$

$$\sum_{k \in S} U_k(a') = \pi_1\pi_2\pi_3(E_1 + E_2 + E_3)$$

Since $\pi_i \in (0, 1)$ for all $i \in S$, $(\pi_1\pi_2\pi_3)E_1 < (\pi_1\pi_2)E_1$, $(\pi_1\pi_2\pi_3)E_2 < (\pi_1\pi_2)E_2$ and $(\pi_3\pi_4)E_4 > 0$ . Since objects are close in acceptability, $\pi_1 \leq \frac{\pi_4}{\pi_2}$ and so $(\pi_1\pi_2\pi_3)E_3 \leq (\pi_3\pi_4)E_3$. Then $\sum_{k \in S} U_k(a) > \sum_{k \in S} U_k(a')$. □

**Lemma 21.** ***Condition $V^c$:*** *If $a \in \arg\max_{a \in A} W(a)$, and objects are closely well ranked then there is at most one 3-cycle in $o(a)$.*

*Proof.* Let $a \in \arg\max_{a \in A} W(a)$. By Lemma 9 $o(a)$ contains only 2-cycles and 3-cycles.

If $n \leq 5$ then there is at most one 3-cycle in $o(a)$. Now consider sets of agents and objects where $n > 5$. Since $a \in \arg\max_{a \in A} W(a)$ by Lemma 17 it must meet conditions $\mathbf{II}^c$ and $\mathbf{III}^c$ so let $o(a) = \{o_1(a), ..., o_\nu(a)\}$ where if $i \in \phi_j(a)$ and $i' \in \phi_{j'}(a)$ and $j < j'$ then $\pi_i \geq \pi_{i'}$ and $E_i \geq E_{i'}$. Suppose $o(a)$ contains 2 or more 3-cycles. Let $o_k(a)$ and $o_{k'}(a)$ be two of those 3-cycles such that if there is some $o_j(a)$ such that $k < j < k'$ then $o_j(a)$ is a 2-cycle. To see that the presence of two 3-cycles in $o(a)$ implies $a \notin \arg\max_{a \in A} W(a)$, let $a^0 = a^1 = a$ and use the following algorithm:

> **Step q:** If $k' = k + q$ then since both $o_{k+q-1}(a^q)$ and $o_{k+q}(a^{q-1})$ are 3-cycles, by Lemma 19, $a^q \notin \arg\max_{a \in A} W(a)$ so end. If $k' \neq k + q$ then by the definition of $k$ and $k'$, $o_{k+q}(a^{q-1})$ is a 2-cycle and $o_{k+q-1}(a^q)$ is a 3-cycle. Let $\phi_{k+q-1}(a^q) = \{q_1, q_2, q_3\}$ and $\phi_{k+q}(a^{q-1}) = \{q_4, q_5\}$ where $S^q = \{q_1, q_2, q_3, q_4, q_5\}$ and since $a^q$ meets $\mathbf{II}^c$ and $\mathbf{III}^c$, for any $q_i, q_{i'} \in S^q$ if $i < i'$ then $\pi_i \geq \pi_{i'}$ and $E_i \geq E_{i'}$. Let $a^{q+1}$ be such that $\phi_{k+q-1}(a^{q+1}) = \{q_1, q_2\}$, $\phi_{k+q}(a^{q+1}) = \{q_3, q_4, q_5\}$ and all $i \in N \setminus S^q$ are in the same learning cycles under both $a^q$ and $a^{q+1}$. Then by Lemma 18, $W(a^q) \leq W(a^{q+1})$. Note $a^{q+1}$ still meets $\mathbf{II}^c$ and $\mathbf{III}^c$. Continue to step $q + 1$.

Since the number of learning cycles in $o(a)$ is finite, the algorithm must terminate at some step $q$, where $a^q \notin \arg\max_{a \in A} W(a)$. Since by the construction of $a^q$, $W(a) \leq W(a^q)$, $a \notin \arg\max_{a \in A} W(a)$ and so $a$ cannot contain more than one 3-cycle. $\square$

**Lemma 22.** *If there exists some $a \in \arg\max W(a)$ and $|C_3(a)| = 3$ then $W(a) = W(a^*)$ for any $a^* \in A$ satisfying conditions $\mathbf{I}^c$ to $\mathbf{V}^c$ where $o(a^*)$ contains a single 3-cycle, $o_\nu(a^*)$ such that $\pi_i \leq \pi_j$ and $E_i \leq E_j$ for all $i \in \phi_\nu(a^*)$, $j \in C(a^*) \setminus \phi_\nu(a^*)$.*

*Proof.* Let $a, a^* \in A$ be two learning profiles each containing exactly one 3-cycle and satisfying conditions $\mathbf{I}^c$, $\mathbf{II}^c$, $\mathbf{III}^c$, $\mathbf{IV}^c$ and $\mathbf{V}^c$. Since they both satisfy $\mathbf{I}^c$, $\mathbf{II}^c$ and $\mathbf{III}^c$, let $o(a) = \{o_1(a), ...o_\nu(a)\}$ and $o(a^*) = \{o_1(a^*), ...o_\nu(a^*)\}$ where if $i \in \phi_j(a')$ and $i' \in \phi_{j'}(a')$ and $j < j'$ then $\pi_i \geq \pi_{i'}$ and $E_i \geq E_{i'}$ for all $a' \in \{a, a^*\}$. Since $a$ and $a^*$

satisfy $\mathbf{I}^c$, $\mathbf{IV}^c$ and $\mathbf{V}^c$, $|o(a)| = |o(a^*)|$ so $\nu = \nu'$. In $o(a^*)$ let $o_\nu(a^*)$ be the 3-cycle. Since $a^*$ meets $\mathbf{I}^c$, $\mathbf{II}^c$, $\mathbf{III}^c$, $\pi_i \leq \pi_j$ and $E_i \leq E_j$ for all $i \in \phi_\nu(a^*)$, $j \in C(a^*) \setminus \phi_\nu(a^*)$. In $o(a)$ let $o_k(a)$ be the 3-cycle such that $\pi_i \geq \pi_j$ and $E_i \geq E_j$ for all $i \in \phi_\nu(a)$, $j \in \phi_\nu(a^*)$ and either $\pi_{i^*} > \pi_j$ or $E_{i^*} > E_j$ or both for some $i^* \in \phi_k(a)$ and for all $j \in \phi_\nu(a^*)$. Then $k \neq \nu$. Since $o(a)$ contains only one 3-cycle, any $o_{k'}(a)$ where $k \neq k'$ is a 2-cycle. Now let $a = a^1$ and apply the following algorithm:

> **Step q:** Since $o_k(a^q)$ is a 3-cycle and $o_{k+1}(a^q)$ is a 2-cycle, let $\phi_k(a') = \{q_1, q_2, q_3\}$, $\phi_{k+1}(a^q) = \{q_4, q_5\}$ and $S^q = \{q_1, q_2, q_3, q_4, q_5\}$. Since $a^q$ meets $\mathbf{II}^c$ and $\mathbf{III}^c$, for any $q_i, q_{i'} \in S^q$ if $i < i'$ then $\pi_i \geq \pi_{i'}$ and $E_i \geq E_{i'}$. Let $a^{q+1}$ be such that $\phi_k(a^{q+1}) = \{q_1, q_2\}$ and $\phi_{k+1}(a^{q+1}) = \{q_3, q_4, q_5\}$ and all $i \in N \setminus S^q$ are in the same cycles under both $a^q$ and $a^q + 1$. Note that $\mathbf{II}^c$ and $\mathbf{III}^c$ continue to hold for $a^{q+1}$. Then by Lemma 18, $W(a) \leq W(a^{q+1})$. If $k + 1 = \nu$ then end, otherwise continue to step $q + 1$.

Since $o(a)$ is finite, the algorithm terminates at some step $t$. At step $t$, $o_\nu(a^t)$ is a 3-cycle. Since $a^t$ still satisfies $\mathbf{II}^c$ and $\mathbf{III}^c$, $\pi_i \leq \pi_j$ and $E_i \leq E_j$ for all $i \in \phi_\nu(a^t)$, $t \in N \setminus C(a^t)$. If $B(a^*) = \{i\}$ for some $i \in N$ then by $\mathbf{I}^c$, $\mathbf{II}^c$, $\mathbf{III}^c$, $\mathbf{IV}^c$ and $\mathbf{V}^c$, $B(a) = B(a^t) = \{i'\}$ for some $i' \in N$ and by $\mathbf{I}^c$, $\pi_i = \pi_{i'}$ and $E_i = E_{i'}$. Then $\sum_{i \in o_\nu(a^t)} U_i(a) = \sum_{i \in o_\nu(a^*)} U_i(a)$. W.l.o.g let $o_\nu(a^t) = o_\nu(a^*)$. All $i \in C(a^t) \setminus \phi_\nu(a^t)$ and $j \in C(a^*) \setminus \phi_\nu(a^*)$ are in 2-cycles. Since when all agents are in 2-cycles $\mathbf{II}^c$ implies $\mathbf{II}^*$ and $\mathbf{III}^c$ implies $\mathbf{III}^*$, by Lemma 12, $\sum_{i \in C(a^t)} U_i(a^t) = \sum_{i \in C(a^*)} U_i(a^*)$. Then $W(a^t) = W(a^*)$ and by construction of $a^t$, $W(a) \leq W(a^t) = W(a^*)$. □

**Lemma 23.** *If $n$ is even and objects are closely well ranked then $W^* = W^S$.*

*Proof.* Let $n$ be even and fix some $a \in A$ satisfying $\mathbf{I}^c$, $\mathbf{II}^c$, $\mathbf{III}^c$, $\mathbf{IV}^c$ and $\mathbf{V}^c$. Suppose $o(a)$ contains the 3-cycle $o_\nu(a)$ where $\pi_i \leq \pi_j$ and $E_i \leq E_j$ for all $i \in \phi_\nu(a)$, $j \in C(a) \setminus \phi_\nu(a)$. By Lemma 22, if some $a' \in \arg\max_{a \in A} W(a)$ such that $C_3(a') \neq \emptyset$

does exist then $W(a) = W(a')$ and so $a \in \arg\max_{a \in A} W(a)$. Since $a$ satisfies $\mathbf{IV}^c$ and $\mathbf{V}^c$ all other cycles in $o(a)$ must be 2-cycles. Since $n$ is even this implies there is some $i^* \in B(a)$. By $\mathbf{I}^c$, $\pi_{i^*} \leq \pi_i$ and $E_{i^*} \leq E_i$ for all $i \in N$. Since $a$ satisfies $\mathbf{II}^c$ and $\mathbf{III}^c$, all $i \in \phi_\nu(a) \cup \{i^*\}$ must be closely well ranked. But then by Lemma 20, $a \notin argmax_{a \in A} W(a)$. Then all cycle in $o(a)$ are 2-cycles.

Since when all cycles are 2-cycles, $\mathbf{II}^c$ implies $\mathbf{II}^*$ and $\mathbf{III}^c$ implies $\mathbf{III}^*$, $a \in A^*$ and so by Lemma 12, $W^* = W^S$. □

**Lemma 24.** *If $n$ is odd and objects are closely well ranked then either $W^* = W^S$ or $W^*$ and $W^S$ differ only in the sum expected utilities of three agents with the three worst endowments so that $W^S - W^{E*} = \pi_{n-2}\pi_{n-1}(\pi_n(E_{n-2} + E_{n-1} + E_n) + E_{n-2} + E_{n-1})$.*

*Proof.* Suppose there is no $a \in \arg\max_{a \in A} W(a)$ such that $C_3(a) = \emptyset$. Since all $a \in \arg\max_{a \in A} W(a)$ must satisfy $\mathbf{I}^c$, $\mathbf{II}^c$, $\mathbf{III}^c$, $\mathbf{IV}^c$ and $\mathbf{V}^c$, for each $a \in \arg\max_{a \in A} W(a)$ there exists some $i^* \in B(a)$ such that $\pi_{i^*} \leq \pi_i$ and $E_{i^*} \leq E_i$ for all $i \in N$, while all $i \in N \setminus \{i^*\}$ are in 2-cycles. Since when all cycles are 2-cycles, $\mathbf{II}^c$ implies $\mathbf{II}^*$ and $\mathbf{III}^c$ implies $\mathbf{III}^*$, $a \in A^*$ and so by Lemma 12, $W^* = W^S$.

Suppose there is some $a \in \arg\max_{a \in A} W(a)$ such that $C_3(a) \neq \emptyset$. Since $a$ satisfies $\mathbf{I}^c$, $\mathbf{II}^c$, $\mathbf{III}^c$, $\mathbf{IV}^c$ and $\mathbf{V}^c$, by Lemma 22:

$$W^* = \sum_{i \in N \setminus \{n-2, n-1, n\}, i \text{ even}} \pi_{i-1}\pi_i(E_{i-1} + E_i) + \pi_{n-2}\pi_{n-1}\pi_n(E_{n-2} + E_{n-1} + E_n)$$

By Lemma 14, $W^S = \sum_{i \in N, i \text{ even}} \pi_{i-1}\pi_i(E_{i-1} + E_i)$ and so $W^S - W^* = \pi_{n-2}\pi_{n-1}(\pi_n(E_{n-2} + E_{n-1} + E_n) + E_{n-2} + E_{n-1})$. □

# C Outcomes of an individually rational, Pareto optimal mechanism

**Lemma 25.** *If $M$ is an individually rational and Pareto optimal mechanism then $M(R(a,\omega)) = GTT(R(a,\omega))$.*

*Proof.* Let $M$ be an individually rational and Pareto optimal mechanism and suppose $M(R(a,\omega)) \neq GTT(R(a,\omega))$. Since a different matching is produced under $M$ and $GTT$, there is some $i^* \in N$ for which $M(R(a,\omega))(i^*) \neq GTT(R(a,\omega))(i^*)$. Since there is one test per agent, either (i) there is a single $j^* \in N$ such that $j^* P_{i^*}(a,\omega) i^*$ or (ii) $i^* P_{i^*}(a,\omega) i$ for all $i \in N \setminus \{i^*\}$. If (ii) holds then since $GTT$ is individually rational, $GTT(R(a,\omega))(i^*) = i^*$ and so $M(R(a,\omega))(i^*) \neq i^*$. But then $M$ is not individually rational and so it must be that $P_{i^*}$ is as described in (i).

Since $M$ and $GTT$ are individually rational, $M(R(a,\omega))(i^*), GTT(R(a,\omega))(i^*) \in \{i^*, j^*\}$. If $GTT(R(a,\omega))(i^*) = j^*$ then $M(R(a,\omega))(i^*) = i^*$. Since each agent has only one test, there is a vector of agents $(i_1, ... i_{\bar{t}})$ such that $i^*, j^* \in \{i_1, ... i_{\bar{t}}\}$, $i_{t+1} P_{i_{\bar{t}}}(a,\omega) j$ for all $j \in N \setminus \{i_{t+1}\}$, $t \in \{1, ..., \bar{t}-1\}$ and $i_1 P_{i_{\bar{t}}}(a,\omega) j$ for all $j \in N \setminus \{i_{\bar{t}}\}$. Since $M(R(a,\omega))(i^*) = i^*$ and $M$ is individually rational, can only be matched with their own endowment under $M$: $M(R(a,\omega))(j) = j$ for all $j \in \{i_1, ..., i_{\bar{t}}\}$. Then, under $M(R(a,\omega))$ there is a Pareto improvement possible where each $j \in \{i_1, ..., i_{\bar{t}}\}$ is matched with $GTT(R(a,\omega))(j)$ and all $i \in N \setminus \{i_1, ..., i_{\bar{t}}\}$ are matched with $M(R(a,\omega))(i)$. Since $M$ is Pareto optimal, it cannot be that $GTT(R(a,\omega))(i^*) = j^*$ and $M(R(a,\omega))(i^*) = i^*$.

If $GTT(R(a,\omega))(i^*) = i^*$ then $M(R(a,\omega))(i^*) = j^*$. Since $M$ is individually rational, there must be some vector of agents $(i_1, ... i_{\bar{t}})$ such that $i^*, j^* \in \{i_1, ... i_{\bar{t}}\}$, $M(R(a,\omega))(i_t) = i_{t+1}$ for all $t \in \{1, ..., \bar{t}-1\}$ and $M(R(a,\omega))(i_{\bar{t}}) = i_1$. Let $P^*(a,\omega)$ be such that $i_{i+1} P^*_{i_t}(a,\omega) j$ for all $j \in N \setminus \{i_{t+1}\}$, $t \in \{1, ..., \bar{t}-1\}$ and $i_1 P^*_{i_{\bar{t}}}(a,\omega) j$ for all $j \in N \setminus \{i_{\bar{t}}\}$. If $P^*_i(a,\omega) = R_i(a,\omega)$ for all $i \in \{i_1, ... i_{\bar{t}}\}$ then in Step 1 of $GTT$, all

$i \in \{i_1, ...i_{\bar{t}}\}$ point at and are matched to their most preferred object under $P^*(a, \omega)$ and so $GTT(R(a, \omega))(i^*) = j^*$. Since $GTT(R(a, \omega))(i^*) = i^*$, it must be that $P_i^*(a, \omega) \neq R_i(a, \omega)$ for all $i \in \{i_1, ...i_{\bar{t}}\}$. Since each agent prefers at most one object to their own endowment, there is some $j' \in \{i_1, ..., i_{\bar{t}}\}$ matched under $M(R(a, \omega))$ with an object not strictly preferred to their own endowment. Since there is an exchagne between all $i \in \{i_1, ..., i_{\bar{t}}\}$, $M(R(a, \omega))(j') \neq j'$ and so $j' P_{j'} M(R(a, \omega))(j')$ and $M$ is not individually rational. □

# D Mixed Strategies

Chapter 1 focusses on only pure strategies for all agents and compares the maximum ex-ante welfare achievable under both equilibrium ($W^E$) and stable learning profiles ($W^S$). Theorem 2 shows that when agents are restricted to using pure strategies and objects are well ranked $W^E = W^S$. Allowing for mixed strategies changes this relationship so that the stable learning profile may yield a lower ex-ante welfare than that of the ex-ante welfare maximising equilibrium. This can be seen via the following three agent example in which all agents and objects are ex-ante identical. It shows firstly that a mixed strategy equilibrium exists in which no agent is using a pure strategy, secondly that the ex-ante welfare is greater than can be achieved under pure strategies and lastly that it exceeds the ex-ante welfare of any stable learning profile for this set of agents.

**Example 1 (Part a: Equilibrium):** Let $N = \{1, 2, 3\}$, $\pi_1 = \pi_2 = \pi_3$ and $E_1 = E_2 = E_3$ so that all agents and their endowments are ex-ante identical. Let $a$ be such that each agent learns about one of the two other agents' endowments with a probability of 0.5. Since each agent must learn about one other object, whichever endowment each agent actually investigates, at least one learning cycle must form. That learning cycle will either be a 2-cycle or a 3-cycle. Let $Pr(o_j \mid a)$ be the probability the learning cycle

$o_j$ occurs under the mixed strategy learning profile $a$. Under this learning profile, there is some probability Agent 1 will be in either the 2-cycle (1,2) or (1,3). If this is the case then because all agents endowments are ex-ante identical, the expected utility of such a 2-cycle for a Agent 1 is $\pi_1^2 E_1$. However there is also a probability that Agent 1 will be in one of the possible 3-cycles, (1,2,3) or (1,3,2). Agent 1's expected utility of either of these cycles is $\pi_1^3 E_1$. Given the probability of any given learning cycle occurring, Agent 1's expected utility can be expressed as:

$$U_1(a) = Pr((1,2) \mid a)\pi_1^2 E_1 + Pr((1,3) \mid a)\pi_1^2 E_1 + Pr((1,2,3) \mid a)\pi_1^3 E_1 + Pr((1,3,2) \mid a)\pi_1^3 E_1$$

If Agent 1 chooses to alter her strategy by increasing the probability she investigates the endowment of Agent 2 then the probability of the the learning cycle (1,2) increases and the probability of (1,3) occurring decreases. Similarly, the probability of (1,2,3) occurring increases and decreases for (1,3,2). However, since the endowments of Agents 2 and 3 are ex-ante identical for Agent 1, this does not increase Agent 1's expected utility. This can be seen in the expression for Agent 1's utility. Let $a'$ be such that Agent 1 chooses to investigate Agent 2's endowment with probability $\alpha$ and Agent 3's endowment with probability $(1 - \alpha)$ and Agents 2 and 3 choose to investigate each of the other agent's endowments with probability 0.5 (as under $a$):

$$\begin{aligned} U_1(a') &= \frac{\alpha}{2}\pi_1^2 E_1 + \frac{(1-\alpha)}{2}\pi_1^2 E_1 + \frac{\alpha}{4}\pi_1^3 E_1 + \frac{(1-\alpha)}{4}\pi_1^3 E_1 \\ &= \frac{\pi_1^2 E_1}{2} + \frac{\pi_1^3 E_1}{4} \end{aligned}$$

Since given the mixed strategies of Agents 2 and 3, Agent 1's utility does not depend on $\alpha$, Agent 1 cannot achieve a higher expected utility than she does under $a$. Since all agents and their endowments are ex-ante identical, the same argument can be made mutatis mutandis for Agents 2 and 3 and so $a$ is an equilibrium.

**Example 1 (Part b: Ex-ante Welfare):** Under the learning profile $a$ (described in Part a of this example), each of the five possible learning cycles can arise with some positive probability: the 2-cycles (1,2), (1,3) and (2,3) each occur with probability 0.25 and the 3-cycles (1,2,3) and (1,3,2) each occur with probability 0.125. Since all agents and their endowments are ex-ante identical, conditional on a 2-cycle occurring, ex-ante welfare is $2\pi_1^2 E_1$ whereas conditional on a 3-cycle occurring, ex-ante welfare is $3\pi_1^3 E_1$. Depending on the probability an object is acceptable $\pi_1$, the ex-ante welfare of a 3-cycle can exceed that of a 2-cycle:

$$2\pi_1^2 E_1 < 3\pi_1^3 E_1$$
$$\frac{2}{3} < \pi_1$$

Since under pure strategies a 3-cycle cannot occur in equilibrium, ex-ante welfare under pure strategies is $2\pi_1^2 E_1$. However, since under the mixed strategy learning profile $a$, 3-cycles occur with some positive probability, $W(a)$ is greater than ex-ante welfare under only pure strategies if $\pi_1 > \frac{2}{3}$.

**Example 1 (Part c: Stability):** It is the possibility of a 3-cycle occurring which means that the learning profile $a$ cannot be stable. This is because simply being in a learning cycle is no guarantee to an agent in the cycle that they will be able to exchange their endowment; each agent in the cycle needs their test to be successful such that all agents in the cycle prefer their investigated object to their own endowments. The larger the learning cycle, the lower the probability of an exchange taking place. Under $a$, the probability a 3-cycle results in an exchange for any given agent is $\pi_1^3$, while the probability a 2-cycle results in an exchange for either agent in the 2-cycle is $\pi_1^2 > \pi_1^3$. Since $E_1 = E_2 = E_3$ this means that an agent has a higher expected utility if they are in any 2-cycle rather than one of the 3-cycles. This means that $a$ cannot be stable since Agents 1 and 2 have a higher expected utility under any $a''$ where $a_1'' = 2$ and $a_2'' = 3$

than under $a$. If, for this set of agents and endowments, under a stable learning profile there can be no possibility of a 3-cycle occurring then the ex-ante welfare of any stable learning profile is $2\pi_1^2 E_1$. Part b of the example demonstrated that when $\pi_1 > \frac{2}{3}$, this is less than the ex-ante welfare of the mixed strategy equilibrium learning profile $a$.

## E Size of Preference Cycles in Equilibrium

*Proof of Lemma 10:*

*Proof.* Suppose there is some preference cycle $o_1(h') = (i_1, ..., i_m)$ such that $m \geq 3$ at $h'$. Then there is a set of subhistories of $h'$, $\{h^1, ...h^m\}$, such that $a_{i_k}(s_{i_k}, h^k) = i_{k+1}$ for $k \in \{1, ..., m\}$ and $a_{i_m}(s_{i_m}, h^m) = i_1$.

*Case 1: Two or more agents learn in the same period and complete a preference cycle in the revelation period immediately following.*

First consider the case where two or more agents learn at the same history (a learning period) and complete the preference cycle $o(h')$ in the next period. Then there is some $h^k = h^{k'}$ such that $k \neq k'$ and since $o_1(h')$ forms in the next period, $h' = (h^k, a(s, h^k))$. Let $Y^k$ be the set of agents which comprise the cycle $o_1(h')$ who all expend their test at $h^k$: $Y^k := \{i \mid i \in \phi_1(h') \text{ and } a_i(s_i, h^k) \neq x\}$. Then $i_k, i_{k'} \in Y^k$.

Since the preference cycle $o_1(h')$ is complete at $h'$ and $|Y^k| \geq 2$, $U_i(s\ |_{h^k}) \leq p^2\bar{v}$ for all $i \in Y^k$. If $Y^k = \phi_1(h')$ then $|Y^k| \geq 3$ and $U_i(s|_{h^k}) < p^2\bar{v}$ for all $i \in Y^k$ but for $s'_{i_k}$ such that $a_{i_k}(s'_{i_k}, h^k) = i$, $U_{i_2}(s|_{h^k}) = p^2\bar{v} > U_{i_k}(s|_{h^k})$.

If $Y^k \neq \phi_1(h')$ then, since the preference cycle $o_1(h')$ exists at $h'$, there is some $i^* \in Y^k$ such that at $h^*$, $a_{j^*}(s_{j^*}, h^*) = i^*$ where $j^* \in \phi_1(h')$ and $h^*$ is a subhistory of $h^k$. Since $|Y^k| \geq 2$, $U_i(s\ |_{h^k}) \leq p^2\bar{v}$ for all $i \in Y^k$ but for $s'_{i^*}$ such that $a_{i^*}(s'_{i^*}, h^k) = j^*$, a learning cycle forms between two agents and so $U_{i^*}((s'_{i^*}, s_{-i^*})|_{h^k}) = p\bar{v} > U_{i^*}(s|_{h^k})$.

*Case 2: Two or more agents learn in the same period and extend a preference chain, but do not complete a preference cycle, in the revelation period immediately following.*

Now consider the case where two or more agents learn at the same history (in a learning period) but do not complete the preference cycle $o(h')$ in the next preference revelation period. That is, there is some $h^k = h^{k'}$ such that $k \neq k'$ and $h'' = (h^k, a(s, h^k)) \neq h'$ but $(h^k, a(s, h^k))$ is a subhistory of $h'$. Let $Y^k := \{i \mid i \in \phi_1(h') \text{ and } a_i(s_i, h^k) \neq x\}$.

Suppose there is some subset of agents $Y^j \subset Y^k$ such that $|Y^j| \geq 2$ and all members of $Y^j$ form a single preference chain at $h''$. Since the preference cycle $o_1(h')$ does not exist at $h''$ and $|Y^j| \geq 2$, further tests by other agents in $\phi_1(h')$ must be successful before the preference cycle $o_1(h')$ is complete and so $U_i(s|_{h^k}) < p^2\bar{v}$ for all $i \in Y^j$. But since the members of $Y^j$ form a preference chain, there are two agents $j^*, j' \in Y^j$ such that $a_{j'}(s_{j'}, h^k) = j^*$ and so for $s'_{j^*}$ such that $a_{j^*}(s'_{j^*}, h^k) = j'$, $U_{j^*}((s'_{j^*}, s_{-j^*})|_{h^k}) = p^2\bar{v} > U_{j^*}(s|_{h^k})$ and $s$ is not an equilibrium.

*Case 3: Two or more agents learn in the same period but neither extend a preference chain or complete a preference cycle in the revelation period immediately following.*

There is a gap at $h$ if between any two agents $i_r$ and $i_{r''}$, who both expend their tests at $h$, there is another agent $i_{r'}$ who comes between $i_r$ and $i_{r''}$ in the preference cycle $o_1(h')$ but $i_{r'}$ only expends her test in a later period. That is, there is a **gap** at $h$ if there are two agents $i_r, i_{r''} \in o_1(h')$ such that $a_{i_r}(s_{i_r}, h), a_{i_{r''}}(s_{i_{r''}}, h) \neq x$ but $a_{i_{r'}}(s_{i_{r'}}, h) = x$ and $A_{i_{r'}}(s_{i_{r'}}, h) = (N \setminus \{i_{r'}\} \cup \{x\}$ for all $i_{r'} \in o_1(h')$ such that $r < r' < r''$.[14] Let $G(h)$ be the set of gaps that exist at $h$. Since the preference cycle $o_1(h')$ occurs at $h'$ there must be some subhistory $(h, h^g)$ of $h'$ at which $G(h) \neq G(h, h^g)$ so that the gaps which existed at $h$ no longer exist at $(h, h^g)$. Let $\mathcal{H}$ be the set of histories $(h, h^g)$ such that $l(h, h^g) < l(h')$ and $G(h) \neq G(h, h^g)$. If at one such $h \in \mathcal{H}$, two or more gaps cease to exist for the first time at the same period then either Case 1 or Case 2 above applies

[14]where $r, r'$ and $r''$ are modulo $m$

and $s$ is not an equilibrium. If only one gap ceases to exist in each learning period in every history $h \in \mathcal{H}$ then consider the history at which the penultimate gap ceases to exist as a result of $a_{i_{r'}}(s_{i_{r'}}, h)$ then Case 2 applies.

*Case 4: Only one agent learns in each period*

If at most one agent in $\phi_1(h')$ learns in each period then there is no $i_k, i_{k'} \in \phi_1(h')$ such that $h^k = h^{k'}$. W.l.o.g let $(h^1, ..., h^m, a(s, h^m)) = h'$. Consider agent $i_{m-1}$. When agent $i_{m-1}$ investigates $a_{i_{m-1}}(s_{i_{m-1}}, h^{m-1}) = m$, agent $i_m$ has not yet expended her test and so $U_{i_{m-1}}(s|_{h^{m-1}}) \leq p^2\bar{v}$. However, for $s'_{i_{m-1}}(s'_{i_{m-1}}, h^{m-1}) = m - 2$, since at some subhistory of $h^{m-1}$, $i_{m-1}$'s endowment has already been investigated by $i_{m-2}$, $U_{i_{m-1}}((s'_{i_{m-1}}, s_{-i_{m-1}})|_{h^{m-1}}) = p\bar{v}$. $\square$

# F Expected utility calculation

In addition to the probabilities of grade 1 students attending each college $x$ given each strategy, the probabilities can also be determined for grade 2 and grade 3 students to calculate expected utility.

## F.1 Probabilities

*For students with grade 2:*

$$P(A \mid AB, 2) = 0 \tag{19}$$

$$P(A \mid AC, 2) = 0 \tag{20}$$

$$P(B \mid BC, 2) = \frac{q_{B2}}{P(2)[m(AB) + m(BC)]} \tag{21}$$

$$P(B \mid AB, 2) = P(B \mid BC, 2) \tag{22}$$

$$P(C \mid BC, 2) = 1 - P(B \mid BC, 2) \tag{23}$$

$$P(C \mid AC, 2) = 1 \tag{24}$$

*Capacities for grade 3 students:*

$$q_{A3} = 0 \tag{25}$$

$$q_{B3} = 0 \tag{26}$$

*($q_C$ is irrelevant since it's capacity is never exhuasted.)*

*For students with grade 3:*

$$P(A \mid AB, 3) = 0 \tag{27}$$

$$P(A \mid AC, 3) = 0 \tag{28}$$

$$P(B \mid AB, 3) = 0 \tag{29}$$

$$P(B \mid BC, 3) = 0 \tag{30}$$

$$P(C \mid BC, 3) = 1 \tag{31}$$

$$P(C \mid AC, 3) = 1 \tag{32}$$

### F.2 Expected utilities

$$u(AB) = P(A \mid AB) \cdot U(A) + P(B \mid AB) \cdot U(B) \tag{33}$$

$$= U(A)P(A \mid AB, 1)P(1) + U(B)[P(B \mid AB, 1)P(1) + P(B \mid AB, 2)P(2)] \tag{34}$$

$$= U(A)\frac{q_A}{m(AB) + m(AC)} + U(B)[P(1)(1 - \frac{q_A}{P(1)[m(AB) + m(AC)]} + \frac{q_{B2}}{m(AB) + m(BC)}] \tag{35}$$

$$u(AC) = P(A \mid AC) \cdot U(A) + P(C \mid AC) \cdot U(C) \tag{36}$$

$$= U(A)P(A \mid AC, 1)P(1) + U(C)[P(C \mid AC, 1)P(1) + P(C \mid AC, 2)P(2) + P(C \mid AC, 3)P(3)] \tag{37}$$

$$= U(A)\frac{q_A}{m(AB) + m(AC)} + U(C)(P(1)(1 - \frac{q_A}{P(1)[m(AB) + m(AC)]}) + P(2) + P(3)) \tag{38}$$

$$
\begin{aligned}
u(AC) &= P(B \mid BC) \cdot U(B) + P(C \mid BC) \cdot U(C) && (39)\\
&= U(B)[P(B \mid BC,1)P(1) + P(B \mid BC,2)P(2)] + U(C)[P(C \mid BC,2)P(2) + P(C \mid BC,3)P(3)] && (40)\\
&= U(B)(P(1) + \frac{q_{B2}}{m(AB)+m(BC)}) + U(C)(P(2)(1 - \frac{q_{B2}}{P(2)(m(AB)+m(BC)})) + P(3)) && (41)
\end{aligned}
$$

### F.3 Equilibrium Solution

*Solving the equations for $u(AB)$, $u(AC)$ and $u(BC)$ in Section F.2, in combination with equations 3 and 4 yields:*

$$
\begin{aligned}
m(AC) = [&U(A)U(C) - U(C)^2 - U(A)U(B)q_A + U(B)^2 q_A - U(B)U(C)q_A + U(C)^2 q_A - U(A)U(B)q_B \\
&+ U(B)^2 q_B - U(B)U(C)q_B + U(C)^2 q_B]/[U(C)(U(A) - U(B))]
\end{aligned}
$$

$$
m(AB) = \frac{U(B)[q_A - m(AC)(q_a + q_B)] + U(C)[m(AC)(1 - m(AC)) - q_A(1 - m(AC))]}{(q_A + q_B)U(B) - U(C)(1 - m(AC))}
$$

$$
m(BC) == 1 - m(AC) - m(AB)
$$

*Or, as shown in Section 3.3.1, if normalising U(C) to 1:*

$$
\begin{aligned}
m(AC) = [&U(A) - 1 - U(A)U(B)q_A + U(B)^2 q_A - U(B)q_A + q_A - U(A)U(B)q_B \\
&+ U(B)^2 q_B - U(B)q_B + q_B]/[U(A) - U(B)]
\end{aligned}
$$

$$m(AB) = \frac{U(B)[q_A - m(AC)(q_a + q_B)] + m(AC)(1 - m(AC)) - q_A(1 - m(AC))}{(q_A + q_B)U(B) - 1 - m(AC)}$$

$$m(BC) == 1 - m(AC) - m(AB)$$

# References

**Ajayi, K., and Sidibe, M. (2015).** An empirical analysis of school choice under uncertainty. Boston University Working paper, http://people. bu. edu/kajayi/research. html.

**Artemov, G., Che, Y. K., and He, Y. (2017).** Strategic ?Mistakes?: Implications for Market Design Research. mimeo.

**Avery, C., and Levin, J. (2010).** Early admissions at selective colleges. American Economic Review, 100(5), 2125-56.

**Bade, S. (2015).** Serial dictatorship: The unique optimal allocation rule when information is endogenous. Theoretical Economics, 10(2), 385-410.

**Bade, S. (2017).** Shift Exchange Problems. mimeo.

**Bade, S. (2019).** Matching with single-peaked preferences. Journal of Economic Theory, 180, 81-99.

**Belfied, C., Britton, J., Bushca, F., Dearden, L., Dickson, M., van der Erve, L., Sibieta, L., Vignoles, A., Walker, I and Zhu, Y. (2018)** The relative labour market returns to different degrees, IFS Research report.

**Beyhaghi, H., Saban, D., and Tardos, E. (2017).** Effect of selfish choices in deferred acceptance with short lists. arXiv preprint arXiv:1701.00849.

**Bloch, F., and Cantala, D. (2013).** Markovian assignment rules. Social Choice and Welfare, 40(1), 1-25.

**Chade, H., Lewis, G., and Smith, L. (2014).** Student portfolios and the college admissions problem. Review of Economic Studies, 81(3), 971-1002.

**Chen, L., and Pereyra, J. (2018).** Self-selection in school choice. Available at SSRN 2880182.

**Cseh, ., Irving, R. W., and Manlove, D. F. (2016, September).** The Stable Roommates problem with short lists. In International Symposium on Algorithmic Game Theory (pp. 207-219). Springer, Berlin, Heidelberg.

**Dickerson, J. P., Procaccia, A. D., and Sandholm, T. (2013, June).** Failure-aware kidney exchange. In Proceedings of the fourteenth ACM conference on Electronic commerce (pp. 323-340). ACM.

**Dubins, L. E., and Freedman, D. A. (1981).** Machiavelli and the Gale-Shapley algorithm. The American Mathematical Monthly, 88(7), 485-494.

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

**Gershkov, A., and Szentes, B. (2009).** Optimal voting schemes with costly information acquisition.

**Harless, P., and Manjunath, V. (2018).** Learning matters: Reappraising object allocation rules when agents strategically investigate. International Economic Review, 59(2), 557-592.

**Immorlica, N., and Mahdian, M. (2005, January).** Marriage, honesty, and stability. In Proceedings of the sixteenth annual ACM-SIAM symposium on Discrete algorithms (pp. 53-62). Society for Industrial and Applied Mathematics.

**Jerrim, J. (2013)** Family background and access to high status institutions. The Sutton Trust.

**Kurino, M. (2014).** House allocation with overlapping generations. American Economic Journal: Microeconomics, 6(1), 258-89.

**Ma, J. (1994).** Strategy-proofness and the strict core in a market with indivisibilities. International Journal of Game Theory, 23(1), 75-83.

**Parkes, D. C. (2007).** Online mechanisms.

**Roth, A. E., and Xing, X. (1994).** Jumping the gun: Imperfections and institutions related to the timing of market transactions. The American Economic Review, 992-1044.

**Roth, A. E. (1982).** Incentive compatibility in a market with indivisible goods. Economics letters, 9(2), 127-132.

**Roth, A. E. (1982b).** The economics of matching: Stability and incentives. Mathematics of operations research, 7(4), 617-628.

**Roth, A. E., Snmez, T., and Ünver, M. U. (2004).** Kidney exchange. The Quarterly Journal of Economics, 119(2), 457-488.

**Roth, A. E., Snmez, T., and Ünver, M. U. (2005).** Pairwise kidney exchange. Journal of Economic theory, 125(2), 151-188.

**Shapley, L., and Scarf, H. (1974).** On cores and indivisibility. Journal of mathematical economics, 1(1), 23-37.

**Ünver, M. U. (2010).** Dynamic kidney exchange. The Review of Economic Studies, 77(1), 372-414.