# Markets, Bargaining, and Networks with Heterogeneous Agents

Arnold Polanski
University of East Anglia

Fernando Vega-Redondo
European University Institute

February 10, 2013

**Abstract**

The paper proposes an intertemporal model of bargaining among heterogeneous buyers and sellers placed on a bipartite network. First, it characterizes conditions on the network under which its trading restrictions are inessential and the outcome is arbitrage-free. Instead, when the system is segmented in different trading components, we show how these come about and how prices are determined in each of them. Second, we turn to the issue of network endogeneity, focusing on those networks that are Pairwise Stable. Such networks are shown to always exist and be arbitrage-free. In the latter respect, therefore, they satisfy one of the key properties displayed by frictionless markets. We identify, however, a sharp contrast regarding another key feature: Pairwise-Stable networks are generically inefficient if the matching process is genuinely decentralized. This uncovers a fundamental incompatibility between individual incentives and social welfare in endogenous trading networks. We explain that such incompatibility is not only due to buyer/seller heterogeneity but is also caused by the incentives underlying network formation in a trading context.

*Keywords*: Bargaining, Networks, Markets, Pairwise Stability, Arbitrage, Competition, Matching
JEL classif. codes: D41, D61, D85, C78.

## 1 Introduction

A central question in economics is the extent to which decentralized mechanisms can achieve an efficient allocation of resources. The paradigmatic approach to the problem is based on the notion of competitive equilibrium where (if only one good is traded) all transactions are assumed to take place at some (common) competitive equilibrium price (CEP) and the market clears. But this formulation of the issue still begs the question of how such uniform trading conditions are to be implemented when agents interact in a truly non-coordinated manner. The main point we make in this paper is that, if the underlying interaction mechanism is truly decentralized (both in terms of the prevailing pattern of interaction and the matching rules operating on it), there is a fundamental generic impossibility of guaranteeing intertemporal efficiency, even as the discount factor approaches unity.

To be precise, we consider a theoretical framework with the following features:

(a) Buyers and sellers meet in pairs at every point in time and bargain over their instantaneous bilateral surplus. If they strike a deal, they leave the economy and are replaced by two other agents with the same characteristics.

(b) Buyers and sellers have heterogeneous valuations and costs and their matching possibilities are restricted to lie in a bipartite (buyer-seller) trading network.

(c) The trading network is strategically (pairwise) stable,[1] with links being infinitesimally costly and all maximal matchings arising with some *ex ante* positive probability.

The features spelled out in (a) are standard in many of the models studied in the literature on bargaining and markets – see , in particular, the seminal papers by Rubinstein and Wolinsky (1985, 1990), Fraja and Sákovics (2001) and Gale (1987). On the other hand, the features postulated in (b) – heterogeneity and a specific trading structure – are not novel either, but have generally been considered in a separate manner. Thus, for example, Gale (1987) allows for agent heterogeneity but in a context with *random* "unstructured" matching, while Manea (2011) studies how bargaining among *homogeneous* agents is affected by a specific trading network.[2] In contrast, our analysis and its main insights crucially depend on the *interplay* of agent heterogeneity and a network of trading possibilities that, as stated in (c), is not just exogenously given but endogenously shaped by agents' strategic incentives.[3]

In a theoretical framework displaying features (a)-(c), we shall plainly ask: can efficiency be at all guaranteed? The hope here would be that, giving to the agents the possibility of establishing their own links, they could successfully tackle – if linking costs are low – the different channels through which inefficiency can arise. One such channel, of course, is that there might not be enough links for all efficiency gains to be exhausted – e.g. if two agents who would enjoy the highest gains from trade are not connected, a welfare loss cannot be avoided. But this, however, is never a problem in a Pairwise-Stable Network (PSN). For,

[1] The notion of pairwise-stable network, by now standard in the literature, was introduced by Jackson and Wolinsky (1996). Informally, it requires that no single agent could profit from unilaterally destroying an existing link and that no pair of agents could benefit from forming a non-existing link between them. See below for a formal definition.

[2] Other related papers are Corominas-Bosch (2004) and Polanski (2007). Both extend the bargaining framework of Rubinstein and Wolinsky (1985) to allow for any general trading structure, exogenously given. They differ from Manea (2011) (as well as from our present approach) in that they contemplate a bargaining protocol with some extent of centralization in either resource allocation or matching.

[3] The independent work of Kappas (2012) also studies the implications of heterogeneity on bargaining in a specific trading network, but one that is *exogenously* given. Another key difference with our approach – and, for that matter, with most of the literature – is that inter-agent matching requires some central coordination procedure and hence cannot be regarded as decentralized.

as it turns out, a PSN is always *arbitrage-free* in the following sense: it induces the same terms of trade as that obtained under no prior restrictions on trade (i.e. when sellers and buyers are connected by complete bipartite network).

An opposite problem is that the network might display too many links. This can cause an inefficiency because, if buyer-seller matching is genuinely decentralized, some trading opportunities may arise that the agents involved may prefer to discard even if, from a welfare-maximizing viewpoint, they should profit from them. Here the trouble is that some agents may refuse a deal in the hope of meeting an even better opportunity later on, thus leading to the (irreversible) loss of some current surplus. In this case, however, we may hope that the requirement of pairwise stability could also help. If agents anticipate that a particular link will never be used, they should not be ready to incur the cost of maintaining it. This indeed happens to be the case, as no PSN can involve inactive (or redundant) links.

Thus, in a certain sense, pairwise stability ensures that the network neither has too many nor too few links. Can one then conclude that every PSN guarantees efficiency? As advanced, the answer is negative, and it is so in a strong sense. For we show that, generically, every PSN (which always exists) is bound to produce an inefficient outcome if the matching mechanism is decentralized. To understand better the nature of this conclusion, some elaboration on our notions of efficiency and decentralization is in order.

First, efficiency requires that no available surplus be squandered. In effect, given that distinct trading opportunities arise (and then vanish) in every period, it demands, in particular, that no pair of agents ever refuse to trade when they can profitably do so. Second, decentralization is identified with the idea that agents are matched every period in an uncoordinated manner. In practical terms, this is simply embodied by the condition that, given the prevailing network, all links that have some positive surplus associated to them have some positive probability of belonging to some matching outcome. Intuitively, the first requirement of efficiency demands that the network only includes high-surplus links. As it turns out, however, this cannot be done in a way that satisfies pairwise stability and effective decentralized matching. And it is precisely such impossibility that in the end precludes the attainment of efficient outcomes.

The fact that agent heterogeneity in bargaining can lead to inefficiencies was originally highlighted by Gale (1987). Our approach sheds new light on the phenomenon from a complementary viewpoint. We show that, contrary to what might have been conjectured (see our former discussion), allowing the trading network to be formed endogenously does not mitigate the severity of the problem. From this perspective, we can say that Gale's conclusion is robust to the endogenization of the trading network. For, as we shall see, even though a Pairwise-Stable network is always sparser than the complete network studied by Gale, this by itself does not overcome the inefficiency. Indeed, we show that the nature of the problem is not just the aforementioned heterogeneity. Even when agents on the two sides of the market are homogeneous (in which case, Gale's model would produce efficient outcomes), it is the architecture of the endogenous social network that precludes the attainment of efficiency. We find, therefore, that inefficiencies also arise as a consequence of how agents choose to establish their trading connections.

The rest of the paper is organized as follows. In Section 2 we describe the theoretical framework and the key notions used throughout our analysis: decentralized matching, pairwise stability, freedom from arbitrage, and efficiency. In Section 3, we introduce the equilibrium notion – "trading equilibrium" – used to study the intertemporal bargaining game played by our population. Section 4 focuses on the benchmark case where the subpopulation of buyers is connected to that of sellers through a complete bipartite network. In this context, we characterize the bargaining equilibrium, identify some of its interesting properties, and compare it to the competitive (Walrasian) outcome. Section 5 then turns to a general setup where the network can display any possible architecture. Here, we characterize Arbitrage-Free networks, discuss the related notion of trading component, and show how average payoffs in any such component depend on the relative size of each side of the market. Section 6 then studies the central issue of network formation and obtains our main results: a Pairwise Stable networks always exist and are arbitrage-free but, generically, they are inefficient. Section 7 concludes with a summary and a discussion of future research. For the sake of a smooth discussion, all formal proofs are relegated to an Appendix. This includes the characterization of equilibrium in a general completely connected K-sided market, which may be of independent interest.

# 2 Theoretical framework

Our setup consists of a set of sellers and buyers, each of them (seller or buyer) connected to a certain subset of agents of the other side of the market. These connections are formalized through a bipartite network $G = \{S, B, L\}$ where $S$ is the set of sellers, $B$ is the set of buyers, and $L \subseteq \{\{s, b\} : s \in S, b \in B\}$ stands for the set of links that connect sellers and buyers. Since, in our context, links are undirected, we shall often use the notation $sb$ to denote a link between the seller $s$ and the buyer $b$. We assume that every seller $s \in S$ can produce at most one unit of the good being traded and incurs a (production or opportunity) cost $c_s$ in doing so. On the other hand, each buyer $b \in B$ is supposed to care for just one unit of the good and have an idiosyncratic valuation $v_b$ for it.

Trading opportunities arise every period $t = 1, 2, ...$ for every pair of connected agents. These opportunities then materialize (in general, just a subset of them), as determined by the following trading game.

**The trading game**
At every $t$, two different steps take place in sequence.

1. *First*, a matching $m$ is chosen that is consistent with the network $G$. To be precise, a *matching* is defined as a collection of seller-buyer pairs, $m = \{s_1b_1 , s_2b_2 ..., s_qb_q\}$, where every particular seller or buyer is included at most once, i.e. $s_i \neq s_j$ and

$b_i \neq b_j$ for every $i, j = 1, 2, ..., q$ $(i \neq j)$. The condition of network consistency simply embodies the requirement that every $s_i b_i \in m$ is also a link in $G$, i.e. $m \subset L$. Thus, denote by $\mathcal{M}_G$ the set of matchings consistent with the network $G$. Then, we assume that there is a probability distribution function (p.d.f.) $\varphi_G : \mathcal{M}_G \to [0, 1]$. according to which a matching is selected every period in a stochastically independent manner.

2. *Second*, for every pair $sb \in m$, one of the two agents is selected at random with equal probability to make a proposal $p$ on the price at which trade can be conducted.
   (a) If this proposal is accepted, the good is transferred and the price paid. The buyer $b$ earns $v_b - p$ and the seller $p - c_s$. These two agents then leave the economy and are replaced by another buyer and seller with the same characteristics who occupy the same network positions next period, $t + 1$.
   (b) If the proposal is refused, then agents remain active in the same network position (i.e. with the same kind of connections) and participate in the new bargaining round taking place at $t + 1$.

Given the prevailing network $G$ and the set of costs $\{c_s\}_{s \in S}$ and valuations $\{v_b\}_{b \in B}$, the process continues as described above for all $t = 1, 2, ....$ Every agent active at any given $t$ discounts the instantaneous payoffs that might be obtained at some future $t'$ with the factor $\delta^{t'-t}$, where the discount rate $\delta < 1$. This defines an intertemporal game, which is assumed played under complete information on all relevant details of the situation (i.e. the payoffs of all agents, the prevailing network, etc.) As usual in the literature, we shall be focusing our analysis on the Stationary Subgame Perfect Equilibria (SSPE) of the game, i.e. Subgame Perfect Equilibria where players' behavior at any given period is independent of previous history.

Our analysis will revolve around four key notions:

1. *Decentralization* of the matching mechanism
2. *Arbitrage-free* trading pattern
3. *Strategic stability* of the social network
4. *Intertemporal efficiency* of bargaining

Next we explain and motivate each of them, as adapted to our set-up.

**Decentralized matching**
A key assumption of our model is that the matching process, through which buyers and sellers are paired and bargain, operates in a *decentralized* manner. More specifically, our notion of decentralization rules out that matching is governed by a *global* (or centralized) procedure that, for example, may block certain pairs of *connected* agents from being matched

so that other alternative (welfare-improving) trades can be realized. Before elaborating further on this idea, we spell it formally.

Given the underlying network $G = \{S, B, L\}$, recall that $\varphi_G$ stands for the distribution function determining the probability $\varphi_G(m)$ with which any particular matching $m \in \mathcal{M}_G$ is selected in every trading period. The matching mechanism at work is fully captured by the function $\varphi_G$. It is in the spirit of our approach that we should allow for a matching procedure that, short of operating in a centralized manner, is as effective as possible (see, however, Remark 1 for a weakening of this idea). This will be taken to imply the following two features:

(i) Any matching realized is maximal, given the prevailing trading network. Or, in other words, no pair of agents can be left unmatched if they are potential partners in the trading network.

(ii) Every maximal matching can occur with some positive probability.

To fix ideas, consider as a simple illustration a matching procedure that, in some random order, matches pairs of agents who are connected in the trading network but are yet unmatched. One may think of this procedure as being implemented, for example, by agents meeting potential (connected) partners at random while "roaming" through a common location – "the market place." And, as soon as they meet a potential partner, they leave for the day to undertake one round of bargaining. We argue that, in an intuitive sense, this procedure can be conceived as "decentralized." Clearly, it will satisfy conditions (i) and (ii) above if meeting is fully random and it proceeds for long enough to make sure that every pair of unmatched partners eventually meet. More formally, the concept can be formulated as follows.

DEFINITION 1 *Given the prevailing network $G = \{S, B, L\}$, let the set of maximal matchings consistent with $G$ be given by $\mathcal{M}_G^* \equiv \{m \in \mathcal{M}_G : \nexists m' \neq m \text{ s.t. } m \subset m'\}$. The matching mechanism embodied by the p.d.f. $\varphi_G$ is said to be* decentralized *if the support of $\varphi_G$ coincides with $\mathcal{M}_G^*$, i.e.*

$$\forall m \in \mathcal{M}_G,\ m \in \mathcal{M}_G^* \ \Leftrightarrow\ \varphi_G(m) > 0. \tag{1}$$

The assumption that the underlying matching mechanism is decentralized will be maintained throughout the paper. To clarify further its implications, it may be useful to highlight, by way of contrast, a certain property that is ruled out by our notion of decentralization. Consider a mechanism that ensures that all matching outcomes occurring with positive probability should display a maximum cardinality, i.e. should maximize the *number* of individuals matched, provided it is consistent with the prevailing trading network. In fact, this property could be achieved in a decentralized manner if the underlying trading network were complete or segmented in buyer-seller dyads. But, in general, if the procedure qualifies as decentralized in the sense proposed, it will fail to deliver such a property. As a trivial illustration, suppose we only have two sellers and two buyers where one seller and one buyer have two network partners, while the other two agents have only one. Then, if the seller and the buyer with two connections are matched with each other, only one pair is formed while two pairs are obviously possible. This latter possibility, however, cannot be enforced with probability one under decentralization, as this notion is presented in the above definition.

REMARK 1 One of our main results in this paper (Proposition 7) establishes the generic inefficiency of decentralized matching. From this viewpoint, contemplating a demanding notion of decentralization such as that reflected by Definition 1 only strengthens that result. But it is worth stressing that our whole analysis remains unaffected if we consider instead a weaker concept that does not presume that all matching outcomes are maximal. The key requirement is that all maximal matchings do enjoy positive probability, with other matching outcomes (even if quite limited and hence inefficient) can also occur with substantial probability. In terms of the specific example used above to illustrate features (i)-(ii), one could assume, for example, that the underlying circumstances are such that it cannot be ensured that, with probability one, every unmatched feasible pair meet before the end of the period (say, because not enough time always elapses). Then, strict subsets of maximal matchings (as well as these maximal matchings) would enjoy positive probability. A weaker reformulation of Definition 1 that is consistent with this state of affairs would be to replace (1) by

$$\forall \mathsf{m} \in \mathcal{M}_G,\ \mathsf{m} \in \mathcal{M}_G^* \ \Rightarrow\ \varphi_G(\mathsf{m}) > 0, \tag{2}$$

which allows for the possibility that non-maximal matchings have positive probability.

**Arbitrage-free networks**

The elimination of arbitrage opportunities is often associated to the operation of frictionless allocation mechanisms and the corresponding attainment of an efficient use of resources. It is useful, therefore, to understand the conditions under which, in our context, arbitrage-free trading networks arise.

To provide an analysis of the problem that highlights the role of the network, we focus our attention throughout on a context where agents are extremely patient (or, analogously, on a scenario where the time period is very short). Formally, this amounts to focusing on the limit case where $\delta \to 1$. In this context, we introduce the following notation. Given a seller-buyer network $G = \{S, B, L\}$, denote by $\mathbf{p}(G) \equiv [p_{ij}(G)]_{ij \in L}$ the limit prices at which trade is conducted in each of the links of $G$ at a SSPE, when the discount rate $\delta \to 1$.[4] If the corresponding probability that trade occurs at some link $ij \in L$ is zero, we simply make $p_{ij} = \emptyset$. The following definition captures the traditional idea of "no arbitrage," particularized to our context.

[4] We shall argue in Section 3 that such prices are uniquely well-defined for any given trading network $G$ when the matching mechanism is decentralized, i.e. $\varphi_G$ satisfies Definition 1.

DEFINITION 2 *A network $G$ is* Arbitrage-Free *(AF) if the price vector $\mathbf{p}(G) \equiv [p_{ij}(G)]_{ij\in L}$ obtained at any SSPE as $\delta \to 1$ satisfies that, for all $ij$, $k\ell \in L$, if $p_{ij}(G) \neq \emptyset$ and $p_{k\ell}(G) \neq \emptyset$, then $p_{ij}(G) = p_{k\ell}(G)$.*

Thus, arbitrage-free networks are those where all buyers and sellers who actually trade do so at the same price. Clearly, a forceful way of attaining such a no-arbitrage outcome is through the complete bipartite network $G^c = \{S, B, L^c\}$ with $L^c = \{\{i, j\} : i \in S, j \in B\}$, provided the matching mechanism is decentralized (cf. Definition 1). For, in this case, if there were some arbitrage opportunity, there would also be some pair of trading agents who could profit from it and would be ready to wait until they meet (for we focus on the case where $\delta \to 1$ and hence they are "infinitely patient"). But, in general, as we shall see, significantly less connected networks may be equally effective in eliminating *de facto* all arbitrage opportunities, in the sense of Definition 2.

**Strategically stable networks**

Given the impact that the underlying network $G$ has on the matching possibilities, it is obviously important to understand which networks are consistent with agents' incentives to connect. In this respect, we shall adopt an approach based on what is one of the most common notions used in the literature to capture the idea of strategic stability in the formation of social networks: pairwise stability, as originally proposed by Jackson and Wolinsky (1996). First, we need to specify what benefits and costs agents anticipate from their alternative linking decisions.

For the sake of simplicity, we adopt a particularly stark formulation in this respect. Benefits, on the one hand, are quite naturally identified with the expected discounted limit payoffs that an agent foresees from any given network $G$ under the bargaining process described. Linking costs, on the other hand, are assumed to be positive but infinitesimally small – hence the cost of any link is always less important than any effect it may have on the bargaining surplus. Based on such specification of costs and benefits, the notion of pairwise stability simply captures the idea that each link (existing or potential) should be considered independently and formed or maintained if, and only if, the two agents involved profit from it and hence agree.

To proceed formally, we now introduce some convenient notation. First, given any trading network $G$, denote by $\mathbf{x}(G) \equiv [x_i(G)]_{i\in S\cup B}$ the (unique) payoff vector obtained at any SSPE by each agent $i$ from the bargaining game played on any given network $G$ when the discount factor $\delta \to 1$.[5] Next, for any (existing or potential) link $ij \in S \times B$, we write $G \oplus ij$ to represent the network $G' = \{S\,, B\,, L'\}$ where $L' = L \cup ij$, while $G \ominus ij$ stands for the network $G'' = \{S\,, B\,, L''\}$ where $L'' = L \setminus ij$. We may then define pairwise stability on networks as follows.

[5] Again, as throughout the paper, we focus on the limit case of arbitrarily patient players in order to highlight the network role in the analysis. Concerning equilibrium uniqueness (under decentralized matching), see Section 3.

Definition 3 *A network $G$ is* pairwise stable *if $\forall i \in S$, $\forall j \in B$, the following two conditions hold.*

(i) $x_\ell(G) > x_\ell(G \ominus ij), \quad \ell = i, j$ .

(ii) $x_\ell(G \oplus ij) > x_\ell(G) \Rightarrow x_{\ell'}(G \oplus ij) \leq x_{\ell'}(G), \quad \ell, \ell' \in \{i, j\}, \ell \neq \ell'$.

In a PS network, therefore, two stability conditions hold. First, from (i), all agents want to keep their links. For, if any agent were to sever one of her existing links, this would entail for her a loss of bargaining surplus that could not be compensated by the (infinitesimal) linking cost thus being saved. The second requirement, embodied by (ii), expresses the idea that no pair of agents can both benefit from creating a link. For, if one of them were to profit from the new link, her partner would experience a payoff loss (e.g., because her bargaining surplus would not increase but her linking cost would do so).

**Efficient networks**

Finally, we discuss the issue of trading efficiency, as applied to our context. A first point to note is that, since *fresh* trading opportunities arise every period, efficiency requires that, also in every period, there is *immediate* agreement between every pair of seller $s$ and buyer $b$ who are matched and satisfy that $v_b - c_s > 0$. For otherwise, the failure to trade under these circumstances would entail an *irreversible* waste of trading surplus.

But, of course, the previous no-disagreement requirement is just a necessary condition for efficiency. In general, efficiency also demands that the matching mechanism pairs those agents who can earn the highest bilateral surplus. The key concern here is that any mechanism that matches agents who can earn only relatively low surplus will typically impose a high *social* opportunity cost, by blocking one of the agents involved from entering more profitable arrangements.

Combining the former considerations, we can heuristically conceive a trading network $G$ as efficient when it avoids both delay and inefficient matching patterns. Making this idea precise requires the following notation. Consider any decentralized matching mechanism, as embodied by a corresponding p.d.f. $\varphi_G$. As usual, we want to focus on the (unique) SSPE outcome achieved in the limit $\delta \to 1$. Here we are interested in singling out the probabilities $\left\{\zeta_G^{ij}\right\}_{ij \in L}$ that agreement takes place between each pair of connected agents $i$ and $j$ if the matching selected by the (decentralized) mechanism includes it. Finally, assume for convenience that the prevailing costs and valuations, $\{c_s\}_{s \in S}$ and $\{v_b\}_{b \in B}$, are indexed so that $c_s \leq c_{s+1}$ for all $s = 1, 2, ..., \#S - 1$, while $v_b \geq v_{b+1}$ for all $b = 1, 2, ..., \#B - 1$, where $\#(\cdot)$ stands for the cardinality of the set in question. Then we may concisely define our notion of efficiency as follows:

Definition 4 *A network $G$ is* efficient *if for all $m \in \mathcal{M}_G^*$:*

$$\sum_{ij \in m} \zeta_{ij}(v_j - c_i) = \sum_{k=1}^{\min\{\#S, \#B\}} \max\{v_k - c_k\, ,\, 0\}\,. \tag{3}$$

In essence, the above definition declares a trading network efficient if it guarantees, at equilibrium, that social (utilitarian) welfare is maximized. To be sure, it may be worth emphasizing that any inefficiency resulting from the violation of (3) has significant intertemporal effects on total welfare. To see this, note that, in our stationary context, any bargaining frictions that prevent current surplus from being maximized in *every period* will mount into an arbitrarily large loss over time. And, indeed, this will be the case even if, from the perspective of the population currently involved in bargaining, those frictions only mean some insignificant delay – say, because agents are arbitrarily patient. The total cost induced by such a delay on the whole of "society" will become arbitrarily large in terms of forgone surplus. More specifically, for any fixed point in time, far enough into the future, the effect on the number of agents who have enjoyed some positive utility will be arbitrarily large. From a utilitarian viewpoint, therefore, the resulting welfare loss will be substantial.

# 3 Trading equilibrium

Within the intertemporal trading game described above for a given network $G$, agents are assumed to play a Subgame Perfect Equilibrium in stationary strategies, i.e. what we have labelled a SSPE. In this context, a stationary strategy $\sigma_i$ for any given agent $i$ (a seller or a buyer) embodies two distinct components. First, it must include, for every $j$ such that $ij \in L$, a proposed price $p_{ij}$ at which $i$ offers to trade with $j$ when the link $ij$ is chosen by the matching mechanism and $i$ is the proposer. Thus, in total, agent $i$ must have a vector of such proposals $p_i \equiv (p_{ij})_{ij \in L}$. On the other hand, also for every one of his links $ij$, agent $i$ must have a function $\psi_{ij} : \mathbb{R} \to \{Y, N\}$ that specifies what price proposals from $j$ he will accept ($Y$) or not ($N$). So, overall, agent $i$'s behavior as a responder is given by a set of functions $\psi \equiv (\psi_{ij})_{ij \in L}$. Bringing together the set of proposals and response functions $\sigma_i = (p_i, \psi_i)$ for every agent $i \in S \cup B$ we have stationary strategy profile $\sigma \equiv (\sigma_i)_{i \in S \cup B}$ that represents a well-defined description of (contingent) play for the trading game taking place on the given network $G$. Any such profile induces a unique corresponding payoff vector, which we shall denote by $\mathbf{x}(\sigma; G, \delta) \equiv [x_i(\sigma; G, \delta)]_{i \in S \cup B}$, hence making explicit its dependence on the strategy profile $\sigma$ being played by agents, as well as on the underlying network $G$ and the discount factor $\delta$.

Given the probability distribution $\varphi_G$ formalizing the matching mechanism operating on network $G$, recall that $\pi_G^{ij}$ stands for the marginal probability on any given pair $\{i, j\}$ induced by $\varphi_G$. Correspondingly, let $v^{ij}$ stand for the surplus that can be jointly earned by $i$ and $j$. Naturally, this surplus is zero if the agents are not connected in the trading network $G$. Instead, if they are connected and, say, $i \in B$ and $j \in S$, we define $v^{ij} \equiv \max\{v_i - c_j, 0\}$. With the former notation in place, the requirement that any SSPE $\sigma^*$ should be *intertemporally self-consistent* implies that the induced payoffs $\mathbf{x}(\sigma^*; G, \delta)$ must satisfy, for every $i \in N \equiv S \cup B$, the following Bellman-like conditions:

$$x_i = \sum_{j \in N} \pi^{ij} \left( \frac{1}{2} \max\{v^{ij} - \delta x_j, \delta x_i\} + \frac{1}{2} \delta x_i \right) + \left( 1 - \sum_{j \in N} \pi^{ij} \right) \delta x_i \tag{4}$$

where, for notational clarity, we suppress the arguments for each $x_i \,(= x_i(\sigma^* \,; G, \delta))$. These conditions simply state that, at equilibrium, the expected payoff that any given agent $i$ obtains at any general date $t$ must be precisely equal to the one he expects to obtain if he and his partner react optimally at that period (striking a deal if, and only if, there is an intertemporal non-negative surplus to share) and anticipates the same payoff if he is still active the following period.

Polanski and Lazarova (2011) showed that the system defined in (4) has a single solution for any $\delta < 1$, which in turn defines the unique payoff vector achieved at any SSPE $\sigma^*$ of the trading game. Since this payoff vector is independent of the specific SSPE, we shall omit the equilibrium strategy profile and simply denote the induced payoff profile by $\mathbf{x}^*(G, \delta) \equiv [x_i^*(G, \delta)]_{i \in S \cup B}$. Then, in order to highlight the role of the network architecture on the outcome we shall focus on the scenario where agents are arbitrarily patient (or equivalently, the trading periods occur very fast). Formally, this will be captured by making $\delta \to 1$, thus considering the limit payoff vector $\mathbf{x}^*(G) \equiv [x_i^*(G)]_{i \in S \cup B}$ given by

$$\mathbf{x}^*(G) \equiv \lim_{\delta \to 1} \mathbf{x}^*(G, \delta)$$

which is well defined and we shall refer to as the *Limit Bargaining Outcome* (LBO). We start in the next section by studying the problem in the useful benchmark case where the underlying network $G$ is complete.

# 4 The complete trading network

We start by analyzing a context where all buyers are connected to all sellers, a setup that was already studied by Gale (1987). This context will serve as a helpful term of comparison to understand the implications of richer and more realistic (also endogenous) network topologies, where the structure of interaction has a key influence on the outcome. Let $G^c = \{S, B, L^c\}$ denote the complete bipartite network with $L^c = \{ij : i \in S, j \in B\}$. To focus on non-trivial cases, we shall assume henceforth (without making it explicit) that some non-detrimental trade can take place, since for some seller $s$ and buyer $b$ we have $v_b - c_s \geq 0$. Under this condition, the first conclusion that is established about the LBO is that all buyers and sellers (connected by a complete bipartite network) who *actually* exchange the good must do so at the same *implicit price*. More precisely, the statement is formulated by the following proposition.

PROPOSITION 1 *Let $\mathbf{x}(G^c) \equiv [x_i(G^c)]_{i \in S \cup B}$ be the LBO obtained for the complete bipartite network. Then there exists some $p \in \mathbb{R}_+$ such that for any seller $s$ and buyer $b$ who actually trade, $x_s(G^c) = p - c_s$ and $x_b(G^c) = v_b - p$.*

**Proof**: See Appendix A.1.

Let the price $p$ referred to in Proposition 1 be called the Limit Bargaining Price (LBP). The following result establishes that it is well defined (i.e. exists and is unique) and also specifies a condition that implicitly determines it. To state this result precisely, we introduce the following notation. For any set of sellers $S' \subseteq S$ and any set of buyers $B' \subseteq B$, denote by $C(S')$ the set of costs and by $V(B')$ the set of valuations of agents in $S'$ and $B'$, respectively. Furthermore, given any price $p$, let $[C(S')]^{\leqslant p}$ stand for the costs in $C(S')$ that do not exceed $p$, while $[V(B')]^{\geqslant p}$ represents the valuations in $V(B')$ that are no lower than $p$. Finally, for any set of costs and/or valuations $X$, the notation $\langle X \rangle$ indicate the average value of the elements in this set.

PROPOSITION 2 *If buyers and sellers are connected by a complete bipartite network, there is a unique LBP $p(G^c)$ obtained as the solution to the following equation in $p$:*

$$p = \left\langle [C(S)]^{\leqslant p} \cup [V(B)]^{\geqslant p} \right\rangle . \tag{5}$$

**Proof**: See Appendix A.1.

The fact that, under complete bipartite connectivity and infinitesimal costs of delay, our bargaining setup induces a unique price begs the following question. Is the outcome equivalent to that which would result if buyer-seller "unrestricted connectivity" was implemented through a standard competitive (i.e. Walrasian) market devoid of all frictions? The answer is generically negative, as we now explain.

In the present context, the Competitive Equilibrium Price (CEP) can be defined as the price $p^w$ that satisfies the Walrasian market clearing condition:

$$\#\left([C(S)]^{\leqslant p^w}\right) = \#\left([V(B)]^{\geqslant p^w}\right) \tag{6}$$

In general, we cannot expect that $p^w = p(G^c)$, other than for (non-generic) "coincidence" in the prevailing costs $\{c_s\}_{s \in S}$ and valuations $\{v_b\}_{b \in B}$. And, *a priori*, the discrepancy between the LBP resulting from bargaining and the CEP determined in a Walrasian market could go in either direction. To see this, simply note that a solution $\hat{p}$ to equation (5) must satisfy:

$$\hat{p} = \frac{\#\left([C(S)]^{\leqslant \hat{p}}\right)}{\#\left([C(S)]^{\leqslant \hat{p}}\right) + \#\left([V(B)]^{\geqslant \hat{p}}\right)} \left\langle [C(S)]^{\leqslant \hat{p}} \right\rangle + \frac{\#\left([V(B)]^{\geqslant \hat{p}}\right)}{\#\left([C(S)]^{\leqslant \hat{p}}\right) + \#\left([V(B)]^{\geqslant \hat{p}}\right)} \left\langle [V(B)]^{\geqslant \hat{p}} \right\rangle$$

or, equivalently,

$$\#\left([C(S)]^{\leqslant \hat{p}}\right) \left[\hat{p} - \left\langle [C(S)]^{\leqslant \hat{p}} \right\rangle\right] \; = \; \#\left([V(B)]^{\geqslant \hat{p}}\right) \left[\left\langle [V(B)]^{\geqslant \hat{p}} \right\rangle - \hat{p}\right]$$

which, in view of (6) implies that

$$p^w = p(G^c) \quad \Longleftrightarrow \quad \sum_{c_s \in [C(S)]^{\leqslant p(G^c)}} (p(G^c) - c_s) = \sum_{v_b \in [V(B)]^{\geqslant p(G^c)}} (v_b - p(G^c)).$$

That is, the LBP and CEP coincide if, and only if, the total surplus earned by sellers at a limit SSPE exactly equals that of buyers.

To illustrate matters further, suppose that we have a very large set of sellers and buyers, whose respective costs and valuations can be suitably described (or approximated) by continuous distributions with respective cumulative distribution functions $F^s$ and $F^b$ and support on the interval $[0, 1]$. In this context, the LBP $p(G^c)$ can be obtained by reformulating the equilibrium condition (5) as follows:

$$p = \frac{1}{1 - F^b(p) + F^s(p)} \left[ \int_0^p c dF^s(c) + \int_p^1 v dF^b(v) \right].$$

Suppose, for simplicity, that both distributions are the same, i.e. $F^s = F^b = F$. Then, the LBP is just given by the average of $F$,

$$p(G^c) = \int_0^1 x dF(x).$$

On the other hand, the CEP $p^w$ is determined from the "market-clearing" condition:

$$\int_0^{p^w} dF(c) = \int_{p^w}^1 dF(v) \quad \Longleftrightarrow \quad F(p^w) = 1 - F(p^w),$$

which simply implies that

$$p^w = F^{-1}(1/2).$$

To fix ideas, consider the parametric context where $F(x) = x^\alpha$ for some $\alpha > 0$. The case with $\alpha = 2$ is considered in Figure **??**, which depicts inverse supply and demand functions constructed in the standard manner. The supply function is given by $S(p) = F(p)$ (hence $S^{-1}(q) = \sqrt{q}$) while the demand function is $D(p) = 1 - F(p)$ (and therefore $D^{-1}(q) = \sqrt{1-q}$). Thus, for this specific case, we readily find that $p(G^c) = 2/3$. Graphically, this corresponds to the price $\hat{p}$ such that the area above the straight horizontal line $p = \hat{p}$ that is limited by the inverse demand curve is precisely equal to the area below that line that is limited by the inverse supply curve. On the other hand, we have $p^w = 1/\sqrt{2}$. Consequently, in this case, we have $p^w > p(G^c)$, i.e. the CEP is higher than the LBP. It is immediate to see that the opposite would apply if $\alpha < 1$. The only (knife-edge) case where both prices coincide is $\alpha = 1$, i.e., when the distribution of costs and valuations is uniform on the interval $[0, 1]$.

CAPTION FIGURE 1:

A context with a common distribution of costs and valuations is given by the c.d.f. $F(x) = x^2$. The inverse supply is given by $S^{-1}(q) = \sqrt{q}$, and the inverse demand by $D^{-1}(q) = \sqrt{1-q}$. The LBP is $p(G^c) = 2/3$, graphically determined by the unique horizontal price line where the areas comprised between this line and the inverse supply and inverse demand functions (dark and light greyed, respectively) are equal. The CEP, given by the intersection of the supply and demand curves, is $p^w = 1/2$. Hence, in this case, the CEP is higher than the LBP.

Intuitively, the generic discrepancy between the bargaining outcome and the Walrasian allocation derives from the fact that, even as agents become arbitrarily patient, there are still gains to be obtained at equilibrium from striking a deal with some partners and not others, depending on their costs or valuations. It is, therefore, a consequence of the heterogeneity displayed by buyers and/or sellers. Such a discrepancy, in particular, would not arise if all sellers had the same costs and all buyers had the same valuation.

The fact that the trading price at an equilibrium of the bargaining process generally differs from the competitive price has important implications in our case. Specifically, it is the essential *strategic* reason why inefficient delay materializes when matching is decentralized. To understand this point better, let us briefly illustrate that, in general, decentralized matching *per se* is not the key factor at work. Suppose, for example, that the distributions of costs and valuations are such that every (maximal) matching allows for immediate and efficient trade. As a variation of the previous example, this would occur if, say, the distribution $F^s$ of sellers' costs $c$ over $[0, 1]$ is given by

$$F^s(c) = \begin{cases} (2c)^\alpha & if \quad c \leq 1/2, \\ 1 & if \quad c > 1/2 \end{cases}$$

and the distribution $F^b$ of buyers' valuations $v$ over $[0, 1]$ is given by

$$F^b(v) = \begin{cases} 0 & if \quad v \leq 1/2, \\ [2(v - 1/2)]^\alpha & if \quad v > 1/2 \end{cases}$$

Then, for every maximal matching, if trade were conducted at the CEP $p^w = 1/2$ (which is independent of $\alpha$), the trade would be maximized through behavior that is individually rational (in the sense that no agent's payoff would fall below her reservation value of zero). But this is typically not strategically stable, i.e. it is not an SSPE of the bargaining game. Again, if $\alpha \neq 1$, the LBP $p(\phi_{G_c})$ is different from the CEP: lower or higher, depending on whether $\alpha$ is higher or lower than 1. This results in strategic-induced impasse that leads some pairs to refuse a viable agreement and thus generate wasteful delay.

# 5 Bargaining in general networks: freedom from arbitrage and trading components

Now we extend the analysis to the case where the trading network may be incomplete and, therefore, the structure of trading possibilities may have an interesting impact on the bargaining outcome. One of the key issues here is to what extent the network is sufficiently connected to eliminate all arbitrage possibilities. The question, therefore, is when will there be a single price at which all trades are conducted, thus reproducing the situation that would occur if the bipartite trading network were complete.

Naturally, a necessary condition for such arbitrage-free conditions to prevail is, generically, that the trading network is connected, i.e. there must exist a network path connecting every two agents. But how much more than that, if any, is really needed? A precise answer is spelled out in the following result, which characterizes Arbitrage-Free (AF) networks, as this notion was formulated in Definition 2. (We remind the reader that here, as throughout the paper, we maintain that assumption that the underlying matching mechanism is decentralized in the sense of Definition 1.)

First, we introduce some convenient notation. Given non-empty subsets of buyers $B' \subseteq B$ and sellers $S' \subseteq S$, denote by $P(B', S')$ the LBP that would be obtained through our bargaining model if these were completely connected, i.e. in the complete bipartite network $G' = \{S', B', L'\}$ where $L' = \{\{i, j\} : i \in S',\ j \in B'\}$. That is, in analogy with (5), the price $P(B', S')$ is obtained as a solution to the following equation:

$$P(B', S') = \left\langle [C(S')]^{\leqslant P(B',S')} \cup [V(B')]^{\geqslant P(B',S')} \right\rangle . \tag{7}$$

In addition, for any subset of agents $X$, we shall have $N_G(X)$ stand for the set of other agents who are collectively connected to those in $X$ in the network $G$, i.e. $N_G(X) \equiv \{j \in N : \exists i \in X \quad s.t. \quad ij \in L\}$. With this notation in place, we can now state the indicated characterization result. For notational simplicity (and without loss of generality), it applies to networks $G$ where every node has some connection and every link can be conceivably active.

Proposition 3 *Consider a network $G = \{S,\ B,\ L\}$ where $N_i(G) \neq \emptyset$ for all $i \in S \cup B$ and $v_i > c_j$ for all $i \in B$, and $j \in S$ such that $ij \in L$. Let $p^* \equiv P(B, S)$. Then, the network $G$ is Arbitrage-Free if, and only if,*

$$\begin{aligned} &\forall B' \subseteq B, B' \neq \emptyset \quad P(B', N_G(B')) \leq p^*, \\ &\forall S' \subseteq S, B' \neq \emptyset \quad P(N_G(S'), S') \geq p^*. \end{aligned} \tag{8}$$

**Proof**: See Appendix A.2.

Intuitively, Proposition 3 characterizes an AF network as a pattern of trading possibilities such that

(a) concerning buyers, those subsets of them who display relatively high valuations are connected to enough low-cost sellers that they cannot be forced into accepting prices that are higher than what they could get without any connectivity restrictions (i.e. under the complete bipartite network);

(b) symmetrically for sellers, those subsets of them who enjoy relatively low costs are connected to enough high-valuation buyers that they cannot be forced into accepting prices that are lower that they could enjoy without any connectivity restrictions.

Condition (8) is reminiscent of that contemplated by the celebrated *Marriage Theorem* (see Hall (1935) or Chartrand (1985)) in its characterization of the bipartite networks that admit a perfect matching (i.e. a matching where every node in one part is associated to exactly one other node in the other part). Indeed, this similarity is no coincidence. As we now discuss, the aforementioned theorem can be seen as a particularization of Proposition 3 to binary networks where all links are uniformly "weighted," i.e. they all provide an identical surplus to the seller and buyer involved.

Thus consider the case where every buyer $b \in B$ has an identical valuation that is normalized to $v_b = 1$ and every seller $s$ incurs an identical costs that is normalized to $c_s = 0$. In this case, Condition (8) becomes,

$$\begin{aligned} \forall B' \subseteq B, \quad & \frac{\#(B')}{\#(B') + \#(N_G(B'))} \leq \frac{\#(B)}{\#(B) + \#(S)}, \\ \forall S' \subseteq S, \quad & \frac{\#(N_G(S'))}{\#(S') + \#(N_G(S'))} \geq \frac{\#(B)}{\#(B) + \#(S)}, \end{aligned}$$

or, equivalently,

$$\begin{aligned} \forall B' \subseteq B, \quad & \frac{\#(B')}{\#(N_G(B'))} \leq \frac{\#(B)}{\#(S)}, \\ \forall S' \subseteq S, \quad & \frac{\#(S')}{\#(N_G(S'))} \leq \frac{\#(S)}{\#(B)}. \end{aligned}$$

Then, provided $\#S = \#B$ (which must obviously hold for a perfect matching to exist), the former conditions become:

$$\begin{aligned} \forall B' \subseteq B, \quad & \#(B') \leq \#(N_G(B')) \\ \forall S' \subseteq S, \quad & \#(S') \leq \#(N_G(S')). \end{aligned}$$

That is, it is required that every subset of agents on one side of the market is connected to a set on the other side that is at least as numerous. This is precisely the condition established by Hall's Marriage Theorem as necessary and sufficient for a perfect matching.

The notion of arbitrage (or freedom thereof) is intimately associated to that of trading component. Prior to elaborating on this relationship, we provide a formal definition of this new concept.

Definition 5 *Given a trading network $G$, let $\Gamma \subset G$ be the subnetwork*[6] *consisting of the links where trade takes place with positive probability at a limit SSPE. A trading component is a connected*[7] *subnetwork of* $\Gamma$.

[6] A subnetwork of $G$ is defined as a network defined on a subset of agents $S \cup B$ whose links are also a subset of those in $G$.

[7] A(sub)network is said to be connected if there is a network path joining every two nodes in it.

Since agents are assumed arbitrarily patient, all possible trades that can be undertaken by any given agent at equilibrium must be conducted at the same price. This in turn implies that any two agents belonging to the same trading component must be trading at the same price. Within any such component, therefore, there is no room for arbitrage. So our notion of freedom from arbitrage is guaranteed if the bargaining equilibrium induces a single trading component.

In contrast, if trade is segmented in more than one component, the underlying trading network will generically fail to be arbitrage-free. That is, only by chance may one expect that all trades in the economy should take place at the same price. Within each trading component, the prevailing price will depend on the whole profile of costs and valuations of the sellers and buyers included in it. But this component-specific price cannot depend on its internal linking structure. (Note, in particular, that it must be the same as if the component had its two parts completely connected.) In the same vein, a further interesting observation is that, in a trading component, the average surplus earned by the agents in each side of the market only depends, in relative terms, on their relative numbers. Formally, this is the conclusion stated in the following result.

PROPOSITION 4 *Consider a trading component including the sellers in $S'$ and the buyers in $B'$. Then, the LBP $p$ prevailing at a SSPE within that component satisfies:*

$$\frac{\langle V(B')\rangle - p}{p - \langle C(S')\rangle} = \frac{\#(S')}{\#(B')},$$

*which implies that the payoff obtained on average by buyers relative to that of sellers is equal to the relative number of the former to the latter.*

**Proof**: See Appendix A.2.

In general, however, the interplay between the profile of costs and valuations on the one hand, and the network topology on the other, can lead to rich considerations in the determination of the size and form of trading components. To illustrate matters, consider the bipartite network depicted in Figure **??**.

CAPTION FIG. 2 A simple seller-buyer network.

First, assume, as before, that all links in this network display a uniform weight – say, equal to one, with all buyers displaying a unit valuation and all sellers a zero cost. In this case, the network topology alone dictates the bargaining outcome, which is easily seen to yield three separate trading components:[8]

[8] Interestingly, the partition of the network in trading components parallels the Gallai-Edmonds decomposition that has been used by the bargaining literature (see Corominas-Bosch (2004) or Polanski (2007)). We shall not elaborate on this approach here, since it is peripheral to our main concerns in this paper. It will suffice to say that any binary bipartite network can be *uniquely partitioned* into subnetworks of the three types illustrated in the example.

- G1, which is of a balanced type and includes an equal number of buyers and sellers.
- G2, where the buyers are relatively scarce, their number lower than that of sellers.
- G3, where the sellers are relatively scarce, their number lower than that of buyers.

In fact, by relying on Proposition 4, we can easily determine the limit equilibrium payoffs in this case. In G1, all agents earn a common payoff equal to $1/2$; in G2, the single buyer earns $2/3$ while each of the two sellers earns $1/3$; in G3, the role of buyers and sellers in G2 is reversed.

As emphasized, however, the previous sharp segmentation in separate components must be expected to depend on sellers' costs and buyers' valuations. Consider, for example, a trading network with the same architecture as in Figure **??**, and only the valuation of consumer 8 changes to $v_8 = 10$ (all other costs and valuations are kept as in the binary case). Then, it is immediate to see that we obtain four trading components. One consists of the former $G_2$ enlarged with sellers 2 and 6, i.e. it consists of the agents in the set $\{2,3,4,5,8\}$. This leaves buyers 9 and 10 as isolated singleton components $\{9\}$ and $\{10\}$ – which means that they do not trade – and a residual component consisting of the set $\{1,6,7\}$.

Possibly a more interesting, somewhat polar, illustration arises if we modify only the costs of sellers 3 and 4, raising them to $c_3 = c_4 = 1/4$, and the valuations of buyers 9 and 10, lowering them to $v_9 = v_{10} = 3/4$. Then, it is straightforward to check that the candidate for arbitrage-free price is $p^* = 1/2$ and, relative to this price, condition 8 is satisfied for all non-empty subsets of buyers $B' \subseteq \{1,2,3,4,5\}$ and sellers $S' \subseteq \{6,7,8,9,10\}$. Hence the full trading network is arbitrage-free and trade will be conducted at equilibrium at the price $p^* = 1/2$.

# 6 Pairwise stability, freedom from arbitrage, and efficiency

So far in the analysis, the trading network has been taken as fixed. But, as explained, the main objective of the paper is to understand what are the implications of letting the network itself respond to agents' incentives. These incentives are captured by the notion of pairwise stability introduced in Definition 3. As a first crucial step in the analysis, we establish that, in our context, this notion is never void, i.e. some Pairwise-Stable (PS) network always exists.

PROPOSITION 5 *Given any profile of sellers' costs* $\{c_s\}_{s\in S}$ *and buyers' valuations* $\{v_b\}_{b\in B}$, *some PS network* $G^*$ *always exists.*

**Proof**: See Appendix A.2.

Having allowed the trading network to be determined endogenously, we now want to understand its implications in terms of some of the key "benchmark" features that distinguish the operation of frictionless markets. We focus on two of them. First, we ask whether the induced network will span enough trading possibilities to remove any arbitrage opportunities. Second, we inquire whether the bargaining outcome resulting from the induced network ensures allocative efficiency. These two questions are addressed by the next two results. As advanced, while the first question is responded positively, a stark negative answer is provided to the second one.

PROPOSITION 6 *Given any profile of sellers' costs $\{c_s\}_{s\in S}$ and buyers' valuations $\{v_b\}_{b\in B}$, every Pairwise-Stable trading network $G^*$ is Arbitrage-Free.*

**Proof**: See Appendix A.2.

PROPOSITION 7 *Assume that sellers' costs and buyers' valuations are all distinct*[9] *and more than just one pair of agents in the economy could profitably trade. Then, every Pairwise-Stable network $G^*$ is inefficient.*

**Proof**: See Appendix A.2.

The previous two results indicate that, even though individual linking incentives lead agents to eliminate arbitrage opportunities, this force is not enough to guarantee efficiency – in fact, inefficiency must be expected to obtain almost always, except for "coincidental" arrays of costs and valuations. The intuition for the first result, Proposition 6, is quite clear. If a trading network were to allow for some arbitrage opportunities, agents would anticipate it at the network-formation stage and form mutually profitable links accordingly. On the other hand, to grasp the essential intuition underlying Proposition 7, the key point to note is that pairwise stability rules out the establishment of strategically redundant links as well as too much segmentation of the trading network. Both considerations together lead to a "hybrid situation" where a decentralized matching mechanism cannot operate effectively. In particular, it cannot guarantee an efficient pattern of matchings, thus producing some unavoidable wastefulness in the allocation of resources.

Such inefficiency has multiple sides to it, but heterogeneity does not seem to be a crucial one. Even though the statement of Proposition 7 does presume some heterogeneity, the nature (and need) of it is not at all as in Gale (1987). As we illustrated in Section 4, the main issue in Gale's setup concerns the (typical) discrepancy to be expected between the total surplus earned by the two sides of the market if they were to trade at the (efficiency-supporting) Walrasian price. In our case, even if there were no such discrepancy (say, because the type distributions are uniform, as in the example discussed in that section), the outcome would continue being inefficient as long as the costs and valuations are distinct.[10] This leads us to making two separate remarks.

[9] This can be simply understood as a genericity requirement. Specifically, it would apply except for a set of Lebesgue measure zero in the corresponding space of real values.

[10] In establishing our result, such type diversity is required simply to ensure that, no matter how intricate the *endogenous* network might be, at least one inefficient trade could take place for some matching outcome (see the proof of the result in the Appendix for details).

A first one concerns the aforementioned result in Gale (1987). Our analysis indicates that his conclusion that heterogeneity and complete connectivity typically bring about inefficiency is robust to the endogenization of the trading network. That is, even if the network is not fixed and complete but is endogenously determined, inefficiency also obtains generically. The second remark revolves around the following conjecture: since heterogeneity is not the fundamental basis for inefficiency in our context, it must instead be the result of how agents' incentives shape the trading network structure. Admittedly, this conjecture is too vague and imprecise to be truly useful. So, in order to clarify matters, we conclude this section with a brief study of a setup where all sellers and buyers are alike.

When all buyers' valuations are set to one and all sellers' costs are normalized to zero, we are in the realm of binary trading networks. Also bear in mind that, as throughout, we maintain the assumption that the operating matching mechanism is decentralized in the sense of Definition 1. Under these conditions, we want to compare two situations: one where the trading network is fixed and complete; another one where it is endogenously determined, as prescribed by pairwise stability.

In the first scenario (fixed complete network), we can readily apply Proposition 4 to compute the LBP, which is simply equal to

$$p^* = \frac{\#(B)}{\#(B) + \#(S)}. \tag{9}$$

Under this price, every agent in the short side of the market trades and the allocation is efficient.

Consider now the case where the trading network is endogenous, in the sense considered throughout this section – i.e. assume that it is PS. Then, we know from Proposition 6 that trade will be conducted at the same price $p^*$, given in (9), that prevails in the complete network. The induced outcome will be in general inefficient (in fact, arbitrarily so), as stated in the following result.

PROPOSITION 8 *Suppose that the seller and buyer population sizes, $\#(S)$ and $\#(B)$, are relative primes and satisfy $|\#(B) - \#(S)| < \alpha \min\{\#(B)\,,\, \#(S)\}$ for some $\alpha < 1$.*[11] *Then, there is some $\hat{N}$ such that, if $N \equiv [\#(B) + \#(S)] \geq \hat{N}$, a PS network exists that is inefficient. Furthermore, if $\alpha < 1/3$, for any $\zeta > 0$ there exists some $\tilde{N}(\zeta)$ such that if $N \geq \tilde{N}(\zeta)$ there exists a PS (inefficient) network that induces, with positive probability in any given period, a welfare loss no smaller than $\zeta$.*[12]

[11] The condition that the number of sellers and buyers are relative primes (i.e. their only common factor is 1) rules out, in particular, the knife-edge case where both populations are of the same size. If $\#(B) = \#(S)$, it can be seen that the unique PS network consists of $\#(B)$ components comprising seller-buyer dyads, and this network is efficient.

[12] Note that the maximum aggregate payoff that the economy can generate in any given period is equal to the cardinality of the short size of the market. Therefore, for any matching outcome $\hat{m}$, prevailing in a period, the induced welfare loss is bounded below by $[\min\{\#(B)\,,\, \#(S)\} - \#(\hat{m})]$. In general, one may be interested in the expected value of this magnitude, which of course would depend on the properties of

**Proof**: See Appendix A.2.

To illustrate matters, consider a simple case where, say, $\#(B) = \#(S) + 1$. For binary networks, pairwise stability is equivalent to the requirement that the trading network be *fully connected* and *minimally* so. Hence one possible such network is a "bipartite line" with one buyer at each extreme. Within this trading structure, if the matching is decentralized, a matching outcome that has positive probability involves having almost 1/3 of the sellers unmatched – simply consider the maximal matching where every third buyer is left unmatched once her two partners have already been matched. Such a matching, of course, entails a large efficiency loss. A comparison of this situation with that explained before for a *fixed complete* network, underscores the major consequences on the allocation of network endogeneity. While efficiency is guaranteed when the complete trading network remains exogenously fixed, it fails to materialize if agents endogneously shape their trading network.

Finally, a related important concern is to understand the source of such inefficiency. In line with the conjecture advanced, we find that, in our case, it is the endogeneity of the network that underlies the phenomenon. This sharply contrasts with the aforementioned model of Gale (1987), where the key factor causing inefficiency is agents heterogeneity. (As explained, buyer and seller homogeneity always lead to an efficient allocation if the network is complete and remains fixed.) It can be argued, therefore, that the inefficiency arising in our model is a consequence of the demand that the *full* allocation mechanism be genuinely decentralized. That is, not only trading and matching must be conducted in a decentralized fashion, but this must apply as well to the procedure that determines the network itself (which must be compatible with agents' linking incentives).

# 7 Conclusion

In this paper, we have studied a context where agents, sellers and buyers, bargain along a bipartite trading network to shape the terms of trade with their partners. As a model of trade in a decentralized economy, this approach represents the polar opposite to that of competitive (Walrasian) markets. We have therefore set ourselves to compare the implications of both in terms of two important features:

- their effectiveness in removing arbitrage opportunities, which we characterize in terms of a generalization of the well-known Marriage Theorem for binary bipartite graphs,
- their ability to ensure efficiency, which implies avoiding both delay and imperfect matching outcomes.

the p.d.f. $\varphi_G$. In order to avoid contemplating conditions on this function, we focus on the maximum possible welfare loss. But it is clear from the construction of the proof of this proposition that, if $\varphi_G$ were, say, uniform on the set of maximal matchings, then the expected loss would grow unboundedly with the population size.

As it is well known, Walrasian markets fare well in both respects. The present network approach, however, does comparably well concerning the first one but fails almost always in terms of the second. Since trading in the real world does have features that are best conceived as network-based, this result raises some concern about the operation of decentralized allocation mechanisms whose specific details are grounded in agents' individual incentives.

Future research should address at least two important issues that have been left aside in the present endeavor. First, it would be important to introduce some degree of incomplete information and compare the outcome under this condition with that discussed in this work. For a complete market, this has been undertaken in a similar bargaining setup by Lauermann (2012) and also by Shneyerov and Wong (2010) and Serrano (2002). In particular, Lauermann (2012) finds that asymmetric information - perhaps surprisingly - can improve the allocative efficiency in bilateral markets. A second important question concerns realistic measures in the design of the network formation mechanism or/and the operation of the bargaining process that could eliminate or mitigate inefficiency.

# Appendix

## A.1 K-sided markets with no trading restrictions: equilibrium characterization

This section generalizes bilateral complete markets to K-sided complete markets. First, we note that we can reinterpret a buyer-seller market as a situation, where two sellers of two different inputs can combine them and produce a single unit of the output with the (market) value $v$. To see this, think of buyer $b$ as another seller with her specific input and define her cost $c_b$ as $c_b = v - v_b$. Then, the surplus that each link $bs$ can produce is given by,

$$v^{sb} = v_b - c_s = v - c_b - c_s.$$

The market equilibrium in the bargaining game is, obviously, no affected by this reinterpretation.

Generalizing this setup to $K$-inputs markets, we partition the set of players $N$ into the subsets $S_1, ..., S_K$, where $S_k$ contains all sellers of the input $k = 1, ..., K$. We assume a Leontief-type production function that requires one unit of each of the $K$ inputs to produce one unit of the output. Then, the surplus of the coalition $C$ that contains one player of each type, $\#C = K$ and $\#(C \cap S_k) = 1$, is given as,

$$v^C = v - \textstyle\sum_{s \in C} c_s. \tag{10}$$

where $c_s$ is the cost that $s \in C$ incurs to produce "her" input. Sellers of all types interact in the multi-input market game that is a direct generalization of the game in Section 2: A coalition $C$ is matched with a positive probability $\pi^C$ whenever it contains exactly one seller of each type. When matched, the members of $C$ bargain and the production can only take place if all members of $C$ agree on the terms of trade. The latter are negotiated in an ultimatum bargaining game, in which a player $p \in C$ is chosen as proposer with probability $\alpha_p^C = \alpha_p / \sum_{s \in C} \alpha_s \in (0,1)$, where $\alpha_p > 0$ is the bargaining power of player $p$. Importantly, we assume the same bargaining power across the sellers of the same input. The proposer makes offer sequentially (in any order) and responders either accept or reject the offer. If an offer is rejected, the coalition dissolves immediately, i.e., without any further offers being made, and the same population of players proceeds to the next date. Otherwise, there is an agreement, in which case the value $v^C$ is created, the responders receive the accepted shares and the proposer obtains the residual surplus. The agreeing coalition is removed and the departing agents are replaced by their replicas with the same type profile. Therefore, the type profile remains the same throughout the entire history of the game.

In an SSPE of this game, each replica of player $i \in N$ expects the same payoff $x_i$ at the beginning of any date. We calculate the payoffs $\mathbf{x} = (x_i)_{i \in N}$ from a system of equations that generalizes the Bellman-like conditions in (4),

$$\begin{aligned} x_i &= \textstyle\sum_{C: i \in C} \pi^C \left( \alpha_i^C \max\{v^C - \delta \sum_{i \in C \setminus i} x_i, \delta x_i\} + (1 - \alpha_i^C) \delta x_i \right) + (1 - \sum_{C: i \in C} \pi^C) \delta x_i \\ &= \delta x_i + \textstyle\sum_{C: i \in C} \pi^C \alpha_i^C \max\{v^C - \delta \sum_{i \in C} x_i, 0\}, \quad \forall i \in N, \end{aligned} \tag{11}$$

where $\alpha_i^C = \alpha_i / \sum_{s \in C} \alpha_i$ is the relative bargaining power of agent $i$ in $C$. Polanski and Lazarova (2011) showed that (11) has a unique solution when $\delta < 1$. We can use then the limit $\delta \to 1$ payoff $x_{i_k}^*$ to compute the input price $p_k := x_{i_k}^* + c_{i_k}$ that the seller $i_k$ with the cost $c_{i_k}$ expects for her input $k$. It turns out that the price $p_k$ is the same across all sellers of input $k$ and across all coalitions. The next proposition shows that the input prices can be calculated by solving a system of $K$ equations, where the $K$ prices determine endogenously the sets of active agents and, at the same time, are computed from averages and cardinalities of the latter sets.

Proposition 9 *Assume the bargaining game where sellers in the set $S_k$, $k = 1, ..., K$, have the cost profile $\{c_i\}_{i \in S_k}$, each of them sells the same input $k$ and has the same bargaining power $\alpha_k$. For any coalition $C$ let $\pi^C > 0$ if and only if $\#C = K$ and $\#(C \cap S_k) = 1$. Let $v^C = v - \sum_{i \in C} c_i$, where $v \geq 0$ is a constant that we interpret as the (market) value of the output. Then, the unique LBO $\mathbf{x}^* = \lim_{\delta \uparrow 1} \mathbf{x}(G^c, \delta)$ that solves (11) is such that the payoff to the seller $s$ of input $k$ is $x_s^* = p_k - c_s$, where the unique input prices $\{p_k\}_{k=1}^K$ solve the system,*

$$p_k = \langle [C(S_k)]^{\leq p_k} \rangle + (v - \textstyle\sum_{t=1}^T \langle [C(S_k)]^{\leq p_k} \rangle)(\sum_{t=1}^T \dfrac{\alpha_t \#[C(S_k)]^{\leq p_k}}{\alpha_k \#[C(S_t)]^{\leq p_t}})^{-1}, \quad k = 1, ..., K. \tag{12}$$

**Proof of Proposition 9:**

The general proof that (11) has a unique solution $\mathbf{x}(G^c, \delta)$ can be found in Polanski and Lazarova (2011). The latter result specifies that the limit solution $\mathbf{x}^* = \lim_{\delta \uparrow 1} \mathbf{x}(G^c, \delta)$ satisfies $\sum_{i \in C} x_i^* \geq v^C$ for any coalition $C \subseteq N$ for which $\pi^C > 0$, with equality holding if $C$ trades in the limit SSPE (LSSPE). Moreover, $x_i^*$ does not depend on the coalition, in which player $i$ agrees. Here, we focus on games where all (and only) coalitions with a single seller of each input are matched with positive probability. Then, for a coalition $C : \#C = K$ and $\#(C \cap S_k) = 1$ that trades in the LSSPE,

$$\textstyle\sum_{i\in C} x_i^* = v^C = v - \sum_{i\in C} c_i \Rightarrow v = \sum_{i\in C}(c_i + x_i^*) = \sum_{k\in 1}^{K} p_k(i_k), \tag{13}$$

where the price $p_k(i_k) := c_{i_k} + x_{i_k}^*$ of input $k$ that the seller $i_k$ obtains does not depend on the agreeing coalition $C$ (because $c_{i_k}$ and $x_{i_k}^*$ do not depend on $C$). In order to see that $p_k(i_k) = p_k$ does not depend on the seller $i_k$, we substitute $i_k$ in $C$ by another seller $j_k \in S_k$ of the same input. For the new trading coalition, say $C'$, it holds that

$$\textstyle\sum_{i\in C'} x_i^* = v^{C'} = v - \sum_{i\in C'} c_i.$$

Then, $p_k(i_k) \neq p_k(j_k)$ leads to a contradiction,

$$p_k(i_k) = c_{i_k} + x_{i_k}^* > (<) c_{j_k} + x_{j_k}^* = p_k(j_k) \Rightarrow$$
$$\textstyle v = \sum_{i\in C\setminus i_k} x_i^* + \sum_{i\in C\setminus i_k} c_i + x_{i_k}^* + c_{i_k} > (<) \sum_{i\in C\setminus i_k} x_i^* + \sum_{i\in C\setminus i_k} c_i + x_{j_k}^* + c_{j_k} = v.$$

Therefore, the total payoff $x(S_k, p_k)$ of trading sellers in $S_k$, i.e., sellers in $S_k$ with costs below $p_k$, is computed as,

$$\textstyle x(S_k, p_k) := \sum_{i\in S_k : c_i \le p_k} x_i^* = \#[C(S_k)]^{\le p_k} \cdot p_k - \sum_{i\in S_k : c_i \le p_k} c_i. \tag{14}$$

On the other hand, by Lemma 2 in Polanski and Lazarova (2011), the total limit payoffs of seller types $t$ and $k$ are proportional to their respective bargaining powers,

$$\alpha_t x(S_k, p_k) = \alpha_k x(S_k, p_k), \quad \forall t, k = 1, ..., K. \tag{15}$$

B substituting (14) into (15), summing the resulting $K$ equations over $t = 1, ..., K$, and substituting from (13) $v = \sum_{k\in 1}^{K} p_k$ we obtain the system (12). ■

The solution to (12) exhausts the value $v$, i.e., $\sum_{t=1}^{K} p_t = v$ (cf. equation 13). For $K = 2$ this simplifies to $p_1 = v - p_2$ and we can transform a two-inputs market into a buyer-seller market by defining the value of input 1 for each agent $b \in S_2$ as $v_b \equiv v - c_b$ for a constant $v$. In order to keep each valuation $v_b$ non-negative, $v$ should be sufficiently large, e.g., $v \equiv \max_{b\in S_2} c_b$.[13] It can be readily verified for the set $V(S_2) \equiv \{v_b\}_{b\in S_2}$ that,

$$\langle [C(S_2)]^{\le v-p_1} \rangle = v - \langle [V(S_2)]^{\ge p_1} \rangle, \quad \#[C(S_2)]^{\le v-p_1} = \#[V(S_2)]^{\ge p_1}. \tag{16}$$

We use then (16) to solve (12) for the price $p_1$ of input 1,

$$\begin{aligned} p_1 &= \langle [C(S_1)]^{\le p_1} \rangle + (v - \langle [C(S_1)]^{\le p_1} \rangle - \langle [C(S_2)]^{\le v-p_1} \rangle)/(1 + \frac{\alpha_2 \#[C(S_1)]^{\le p_1}}{\alpha_1 \#[C(S_2)]^{\le v-p_1}}) \\ &= \frac{\alpha_1 \langle [V(S_2)]^{\ge p_1} \rangle \#[V(S_2)]^{\ge p_1} + \alpha_2 \langle [C(S_1)]^{\le p_1} \rangle \#[C(S_1)]^{\le p_1}}{\alpha_1 \#[V(S_2)]^{\ge p_1} + \alpha_2 \#[C(S_1)]^{\le p_1}}. \end{aligned}$$

[13] Conversely, from the valuation profile $\{v_b\}$, we can construct a cost profile $\{c_b = v - v_b \ge 0\}$, where $v \equiv \max_b v_b$.

With equal bargaining powers, $\alpha_1 = \alpha_2$, the latter formula simplifies to the LBP (5),

$$p_1 = \langle [C(S_1)]^{\leq p_1} \cup [V(S_2)]^{\geq p_1} \rangle.$$

## A.2 Proofs of the main results

LEMMA 1 *Let the bipartite network $G = \{B, S, L\}$ be AF with the price $p^* = P(B, S)$ and let the matching mechanism be $\pi_G$. Then, if we change the marginal matching probability $\pi_G^{sb}$ to $\widetilde{\pi}_G^{sb} > 0$ for a linked pair $sb \in L$, the network $G$ remains AF with the same trading price $p^*$.*

Note that this lemma implies that:

a) Adding links to an AF network will not change the payoff of any player.

b) A network is AF independently of the matching probabilities of its links as long as they are positive.

**Proof of Lemma 1**:

Let the network $G = \{B, S, L\}$ be AF with the price $p^* = P(B, S)$ when the matching mechanism is $\pi_G$. Denote by $x(\delta)$ $(y(\delta))$ the SSPE payoffs when $s \in S$ and $b \in B$ are matched with probability $\pi_G^{sb}$ $(\widetilde{\pi}_G^{sb})$ and by $x$ $(y)$ the respective LSSPE payoffs. Hence, $x(\delta)$ solves the system (4), which we write below in a simplified form,

$$x_k(\delta) = \delta x_k(\delta) + \textstyle\sum_{n \in N} \pi_G^{kn} \max\{v^{kn} - \delta x_k(\delta) - \delta x_n(\delta), 0\}/2, \quad \forall k = 1, ..., N. \tag{17}$$

The system for $y(\delta)$ is obtained from (17) by substituting $\pi_G^{sb}$ by $\widetilde{\pi}_G^{sb}$. Our objective is to prove that $y = x$. We assume w.l.o.g.[14] that $\widetilde{\pi}_G^{sb} > \pi_G^{sb}$ and set up the following auxiliary system,

$$x_k^\varepsilon = \delta x_k^\varepsilon + \mathbf{I}_{k \in \{s,b\}} \lambda(\varepsilon, \delta) + \textstyle\sum_{n \in N} \pi_G^{kn} \max\{v^{kn} - \delta x_k^\varepsilon - \delta x_n^\varepsilon, 0\}/2, \quad \forall k = 1, ..., N, \tag{18}$$
$$\textit{where} \quad \lambda(\varepsilon, \delta) = (\widetilde{\pi}_G^{sb} - \pi_G^{sb}) \max\{v^{sb} - \varepsilon - \delta x_s^\varepsilon - \delta x_b^\varepsilon, 0\}/2 \geq 0,$$

where $\mathbf{I}_c = 0$ if $c$ is true and $\mathbf{I}_c = 0$ otherwise. Note that the solution $x^\varepsilon(\delta)$ to (18) is equal to $x(\delta)$ if $\lambda(\varepsilon, \delta) = 0$ and is equal to $y(\delta)$ if $\varepsilon = 0$. As $G$ is AF, LSSPE trades in $G$ occur at the price $p^*$ and the sum of payoffs for $s$ and $b$ under $\pi_G$ satisfies,

$$x_s + x_b = \max\{p^* - c_s, 0\} + \max\{v_b - p^*, 0\} \geq \max\{p^* - c_s + v_b - p^*, 0\} \geq v_b - c_s = v^{sb}.$$

Hence, it holds for the solution $x^\varepsilon(\delta)$ to (18) that for any $\varepsilon > 0$ and sufficiently high $\delta$,

$$x_s^\varepsilon(\delta) + x_b^\varepsilon(\delta) = x_s(\delta) + x_b(\delta) > v^{sb} - \varepsilon,$$

as for such $\delta$ the link $sb$ will be inactive, i.e., $\lambda(\varepsilon, \delta) = 0$. Hence,

$$\lim_{\varepsilon \downarrow 0} \lim_{\delta \uparrow 1} x^\varepsilon(\delta) = \lim_{\varepsilon \downarrow 0} x = x. \tag{19}$$

[14] If $\pi_G^{sb} > \widetilde{\pi}_G^{sb} > 0$, then we start we the network where $s$ and $b$ are matched with probability $\widetilde{\pi}_G^{sb}$ and prove that increasing this probability to $\pi_G^{sb}$ does not change the payoffs.

On the other hand,

$$\lim_{\delta\uparrow 1}\lim_{\varepsilon\downarrow 0} x^{\varepsilon}(\delta) = \lim_{\delta\uparrow 1} x^{0}(\delta) = \lim_{\delta\uparrow 1} y(\delta) = y. \tag{20}$$

as $x^0(\delta) = y(\delta)$ and $x^{\varepsilon}(\delta)$ is continuous in $\varepsilon$. As $x^{\varepsilon}(\delta)$ is also left-continuous in $\delta \in [0,1]$, the order of limits in (20) can be interchanged and $x = y$ follows then from (19) and (20). ∎

LEMMA 2 *Let $G$ be a connected bipartite network with the sets $B$ and $S$ of buyers and sellers, respectively. If $\#B = \#S$ and each maximal matching in $G$ is perfect, then $G$ is complete.*

**Proof of Lemma 2**: For the sake of contradiction assume that there is an incomplete network $G$ where $s$ and $b$ are not connected and where each maximal matching is perfect. Then, there does not exist a perfect matching in the subnetwork $G'$ induced by $G$ on the sets $B\backslash b$ and $S\backslash s$. Otherwise, a perfect matching in $G'$ would be a maximal matching on $G$ that is not perfect as $s$ and $b$ cannot be matched. Hence, as a perfect matching in $G'$ does not exists, Hall's theorem implies that,

$$\exists S' \subseteq S\backslash s, \quad \#N_{G'}(S') < \#S'. \tag{21}$$

Furthermore, there is a $b^* \in N_{G'}(S')$ connected in $G$ to a node $s^* \in S\backslash S'$. Otherwise, the nodes in $S'$ and $N_{G'}(S')$ would be disconnected from other nodes in $G$ contradicting the assumption that $G$ is connected. If we match $b^*$ and $s^*$, it follows from (21) that it is impossible to match all nodes in $S'$ to nodes in $B\backslash b^*$ because the former set cannot be connected in $G$ to a set with at least the same number of nodes,

$$\#(b \cup N_{G'}(S')\backslash b^*) < \#S'.$$

It follows, therefore, that any network with a disconnected pair has a maximal matching that is not perfect. ∎

**Proof of Proposition 3**:

(8)⇒AF: For the sake of contradiction, assume that the condition (8) holds but $G$ is not AF. Then, there is a seller $s$ and a buyer $b$ that trade at the price $p \neq p^* = P(B,S)$. Let $H$ be the trading component that trades at the highest price $p_H > p^*$ (if $p_H \leq p^*$ then the following argument can be easily adapted to the case $p_L < p^*$). Let $H_B$ ($H_S$) be the (non-empty) set of active buyers (sellers) in $H$ and note that an active seller $s \in N_G(H_B)$ with $c_s \leq p_H$ sells only at the highest (available to her) price $p_H$, i.e., $s \in H_S$ iff $s \in N_G(H_B)$ and $c_s \leq p_H$. Therefore,

$$P(H_B, N_G(H_B)) = P(H_B, H_S) = p_H > p^*,$$

which contradicts (8).

AF⇒(8): For the sake of contradiction assume that $G$ is AF but the first condition in (8) does not hold (a similar argument can be constructed if the second condition in (8) is violated). Then, there exists a non-empty set $B' \subseteq B$ such that,

$$P(B', N_G(B')) > p^*. \tag{22}$$

We add all missing links between $B'$ and $N_G(B')$ until these two sets are connected by a complete subnetwork $G_1$ and we do the same for the sets $S' = S \backslash N_G(B')$ and $N_G(S')$ obtaining the complete subnetwork $G_2$ ($G_1$ and $G_2$ are only connected by links between $N_G(S')$ and $N_G(B')$). We denote by $\widetilde{G}$ the entire network that resulted from the link addition to $G$ and note that $\widetilde{G}$ is AF by Lemma 1.

Considering now $G_1$ and $G_2$ separately (i.e., ignoring all links between them), they cover disjoint sets of nodes, whose union is $S \cup B$, and both subnetworks are AF due to their completeness. By (22), the LBP in $G_1$ is,

$$p_1 = P(B', N_{G_1}(B')) = P(B', N_G(B')) > p^*. \tag{23}$$

On the other hand, the following inequality for the LBP in $G_2$,

$$p_2 = P(B \backslash B', N_{G_2}(B \backslash B')) = P(N_G(S'), S') \leq p^*,$$

can be easily shown from the definition (7) when $P(B', N_G(B')) > P(B, S)$ and $B = B' \cup N_G(S')$, $S = S' \cup N_G(B')$. Therefore, $p_1 > p_2$ and for each pair $(s, b) \in N_G(B') \times N_G(S')$, i.e., $s \in G_1$, $b \in G_2$,

$$x_s(G_1) + x_b(G_2) = \max\{p_1 - c_s, 0\} + \max\{v_b - p_2, 0\} \geq \max\{p_1 - c_s + v_b - p_2, 0\} > v_b - c_s,$$

as $G_1$ and $G_2$ are AF. Considering now the entire network $\widetilde{G}$, trading at $p_k$ in its subnetwork $G_k$, $k = 1, 2$, and disagreement for any connected pair $(s, b) \in G_1 \times G_2$ forms an LSSPE. As $p_1 > p_2$, this LSSPE contradicts the fact that $\widetilde{G}$ is AF. ■

**Proof of Proposition 4**:

The LBP $p$ in a trading component, in which all sellers in $S'$ and all buyers in $B'$ trade, is computed by (5),

$$\begin{aligned} p &= \left\langle [C(S')]^{\leqslant p} \cup [V(B')]^{\geqslant p} \right\rangle = \left\langle C(S') \cup V(B') \right\rangle \\ &= \frac{\#S'}{\#S' + \#B'} \left\langle C(S') \right\rangle + \frac{\#B'}{\#S' + \#B'} \left\langle V(B') \right\rangle. \end{aligned}$$

Hence,

$$\begin{aligned} p(\#S' + \#B') = \#S' \left\langle C(S') \right\rangle + \#B' \left\langle V(B') \right\rangle \Rightarrow \\ \#S'(p - \left\langle C(S') \right\rangle) = \#B'(\left\langle V(B') \right\rangle - p) \Rightarrow \frac{\langle v(B') \rangle - p}{p - \langle c(S') \rangle} = \frac{\#S'}{\#B'}. \quad ■ \end{aligned}$$

**Proof of Proposition 5**:

If $\max v_b \leq \min c_s$ then the empty network is PS. Otherwise, let $p^* = P(B, S)$ and define the sets $\widetilde{B} \subseteq B$ ($\widetilde{S} \subseteq S$) of buyers (sellers) with valuations (costs) strictly above (below) $p^*$. Note that the complete network $G^c$ on $\widetilde{S} \cup \widetilde{B}$ is AF and forms a single trading component (TC), where all trades take place at price $p^*$. We delete links successively from $G^c$ until we reach a network $G^*$, where no link can be severed without breaking up a TC into two TCs such that one of these TCs trades at a price different to $p^*$. Note that all links of nodes in $B \backslash \widetilde{B}$ and in $S \backslash \widetilde{S}$ will be deleted in the process. We will prove now that $G^*$ is PS as no player will want to sever one of her links and no buyer-seller pair will want to add a common link.

Assume that a buyer $b \in \widetilde{B}$ can sever her link with a seller $s \in \widetilde{S}$ and buy at a price below $p^*$ in an LSSPE on $G^* \ominus sb$. Let $H$ be the union of all TCs that trade at a price below $p^*$ in this equilibrium and let $H_B$ ($H_S$) be the (non-empty) set of active buyers (sellers) in $H$. As $b \in H_B$ then $b \in N_{G^* \ominus sb}(H_S)$ and, hence, $N_{G^* \ominus sb}(H_S) = N_{G^*}(H_S)$. As $G^*$ is AF, the condition (8) holds for the set $H_S$,

$$P(H_S, N_{G^* \ominus sb}(H_S)) = P(H_S, N_{G^*}(H_S)) \geq p^*.$$

From the definition of $H_S$ and $H_B$ follows that $P(H_S, H_B) < p^*$. Hence, there is at least one buyer $b' \in N_{G^*}(H_S) \backslash H_B$, i.e., a buyer that buys at a price weakly above $p^*$ ($b' \notin H_B$) although he is connected to a seller that sells at a price strictly below $p^*$ ($b' \in N_{G^*}(H_S)$). This is, however, incompatible with an LSSPE. We conclude, therefore, that no buyer can delete one of his links in $G^*$ and trade at a price below $p^*$. A symmetric conclusion holds for each seller.

On the other hand, from Lemma 1 follows that $G^* \oplus sb$ is AF if $G^*$ is AF and that all trades occur at the same price in both networks. Hence, no pair of players has an incentive to add a link to $G^*$. We conclude, therefore, that $G^*$ is PS. ■

**Proof of Proposition 6**:

Suppose that a PS network $G^*$ is not AF. Then, there are two trading components, where trade takes place at different prices $p > q$. Hence, there is a buyer $b$ that pays $p$ and a seller $s$ that sells at $q$. Their respective limit payoffs in $G^*$ are $x_b(G^*) = v_b - p$ and $x_s(G^*) = q - c_s$. The sum of their equilibrium payoffs in $G' = G^* \oplus sb$ must be at least equal to the value of their link,

$$x_b(G') + x_s(G') \geq \max\{v_b - c_s, 0\} > v_b - c_s + q - p = x_b(G^*) + x_s(G^*). \tag{24}$$

Then, $s$ and $b$ must be in the same TC in the LSSPE of the game on $G'$. Otherwise, the link $sb$ would be inactive which, together with (24), would imply multiple solutions to the system (4) for $G'$. In this TC, buyer $b$ pays at most $p$ and seller $s$ gets at least $q$. Otherwise, the disadvantaged agent would disagree in the link $sb$ and obtain her LSSPE payoff in $G^*$. Hence, the link $sb$ benefits both involved agents and at least one of them strictly. But then, $G^*$ is inconsistent with the PS condition (ii) in Definition 3. ∎

**Proof of Proposition 7**:

First, we sort the valuations (costs) in a strict order: $v_1 > v_2 > ... > v_{\#B}$ ($c_1 < c_2 < ... < c_{\#S}$). Given this order, we note that a network $G$ is inefficient if it connects unequal number of buyers and sellers. For example, if there are $\#S'$ sellers and $\#B' < \#S'$ buyers connected in $G$, a maximal matching $m$ that covers the highest cost seller $\#S'$ in $G$ exists and it is inefficient,

$$\sum_{sb\in m}(v_b - c_s) \leq \sum_{k=1}^{\#B'-1} \max\{v_k - c_k, 0\} + \max\{v_{\#B'} - c_{\#S'}, 0\} < \sum_{k=1}^{\#B'} \max\{v_k - c_k, 0\}.$$

Hence, in a PS network $G^*$ both market sides must have the same size $2 \leq H \leq \min\{\#S, \#B\}$ and $G^*$ must connect $H$ buyers with the highest valuations to $H$ sellers with lowest costs such that $v_k > p^* > c_k$ for all $k = 1, ..., H$. We note that a complete network $G^c$ that connects $H$ buyers with $H$ sellers is AF with the LBP $p^* = \sum_{k=1}^{H}(c_k + v_k)/2H$. If we remove the link $l$ between the lowest valuation seller and the highest cost buyer from $G^c$, the network $G^c \ominus l$ is still AF with the same price $p^*$. To see this, it suffices to verify the condition (8) for the subsets $B' = \{H\}$ (buyer with the lowest valuation) and $S' = \{H\}$ (seller with the highest cost). For all other subsets the condition (8) is the same in $G^c \ominus l$ and in $G^c$. For example, for $B'$ the condition (8) holds,

$$P(B', N_{G^c\ominus l}(B')) = (v_H + \textstyle\sum_{k=1}^{H} c_k - c_H)/H =$$
$$(2(v_H - c_H) + 2\textstyle\sum_{k=1}^{H} c_k)/2H \leq \textstyle\sum_{k=1}^{H}(c_k + v_k)/2H = p^*,$$

because

$$2(v_H - c_H) < v_1 - c_1 + v_2 - c_2 \leq \textstyle\sum_{k=1}^{H}(v_k - c_k).$$

Similarly, the condition (8) holds also for $S'$. Hence, the network $G^c \ominus l$ is AF with all agents trading at $p^*$ and obtaining the same payoffs as in the complete network $G^c$. Therefore, the condition (i) in Definition 3 is violated for $G^c$ and this network cannot be PS. On the other hand, Lemma 2 shows that the complete network is the only bipartite network with equal sides, where each maximal matching is perfect. Hence, as a PS network $G^*$ cannot be complete, it must have a maximal matching $m$ that is not perfect, i.e., that does not cover at least one buyer-seller pair. The value of such a matching must be strictly less that the maximal surplus in this market,

$$\sum_{sb\in m}(v_b - c_s) < \sum_{k=1}^{H}(v_k - c_k) = \sum_{k=1}^{\min\{\#S,\#B\}} \max\{v_k - c_k, 0\}. \quad ∎$$

**Proof of Proposition 8**:

We start the proof by a constructive argument that shows that, under the specified conditions, a PS network exists with certain useful properties. For notational simplicity, suppose without loss of generality that $\#(B) < \#(S)$ and then denote $H \equiv \#(B)$ and $\Delta \equiv \#(S) - \#(B)$. We shall also find it convenient to label buyers by indices $b_i$ with $i = 1, 2, ..., \#(B)$ and sellers by indices $s_j$ with $j = 1, 2, ..., \#(S)$.

The trading network to be constructed here, which is denoted by $\hat{G}(H, \Delta)$, displays the following structure. First, we have the first $\#(B)+1$ sellers connected to the set of all buyers along a "line" with each buyer $b_i$ linked to sellers $s_i$ and $s_{i+1}$. Then, the residual sellers, $\Delta - 1$ of them, are connected to the buyers as indicated by binary variables $w_i$ that specify whether any given buyer $b_i$ has *one* of the remaining sellers connected to him ($w_i = 1$) or none at all ($w_i = 0$). First, the following lemma specifies conditions that characterize when any given such network is PS.

LEMMA 3 *Let* $\{w_i\}_{i=1}^H$ *be a set of binary variables specifying a trading network as described above. Denote* $W_i \equiv w_1 + w_2 + \cdots + w_i$ *and let* $W_H = \sum_{i=1}^H w_i = \Delta - 1$. *The induced trading network is PS if, and only if,*

$$i\frac{\Delta}{H} > W_i > i\frac{\Delta}{H} - 1 \quad (i = 1, 2, ..., H). \tag{25}$$

*Proof*: Given any set of binary variables $\{w_i\}_{i=1}^H$ as specified, suppose that buyers are connected along a bipartite line that starts with seller $s_1$ and ends with seller $s_{H+1}$, each buyer $b_i$ also having one additional link to a residual seller in the set $\{s_{H+2}, \ldots, s_{H+\Delta}\}$ if, and only if, $w_i = 1$. Consider any particular buyer $b_i$ and focus on the links $b_i s_{i+1}$ and $b_i s_i$ that embed this buyer in the aforementioned line. For each of those two links, we want identify conditions on the set of binary variables $\{w_i\}_{i=1}^H$ that characterize when they are consistent with being part of a PS network. Recall that any such network is arbitrage-free (Proposition 6) and therefore all buyers and sellers trade at the same price $p^*$ given in (9).

Let us start with the link $b_i s_{i+1}$. Buyer $b_i$ cannot profit from destroying this link if, in the absence of it, he would be left in a component with a ratio of buyers to sellers that is higher than in the whole network (cf. Proposition 4). The required condition, therefore, is that

$$\frac{i}{i + W_i} > \frac{\#(B)}{\#(S)}. \tag{26}$$

Clearly, if the above condition holds for the component in which $b_i$ would lie after deletion of the link $b_i s_{i+1}$, a reciprocal condition must apply for the complementary component including seller $s_{i+1}$. In this component, that is, the relative proportion of buyers should be lower than in the overall population, i.e.

$$\frac{\#(B) - i}{\#(S) - i - W_i} < \frac{\#(B)}{\#(S)},$$

or, equivalently,

$$\frac{\#(S) - i - W_i}{\#(B) - i} > \frac{\#(S)}{\#(B)}. \tag{27}$$

The latter condition implies that, in the absence of the link $b_i s_{i+1}$, seller $s_{i+1}$ would be left in a component with a higher seller-buyer ratio and hence the deletion of that link is detrimental. Thus we conclude that if (26) holds, neither buyer $b_i$ nor seller $s_{i+1}$ will find it profitable to eliminate the link $b_i s_{i+1}$.

An analogue reasoning may be applied to the link $b_i s_i$ for any $i = 1, 2, \ldots, H$. In this case, the counterpart of condition (26) is given by

$$\frac{\#(B) - i + 1}{\#(S) - i - W_{i-1}} > \frac{\#(B)}{\#(S)} \tag{28}$$

which is equivalent to the suitable condition for sellers (i.e. the counterpart of (27)) that can be written as

$$\frac{i + W_{i-1}}{i - 1} > \frac{\#(S)}{\#(B)}. \tag{29}$$

Hence condition (28) alone guarantees that neither buyer $b_i$ nor seller $s_i$ may profit from destroying their link.

Consider any set of binary variables $\{w_i\}_{i=1}^{H}$ characterizing a trading network of the sort studied here. The requirement that (26)-(28) hold for all $i = 1, 2, \ldots, H$ is, in effect, a requirement on the corresponding set of cumulative values $\{W_i\}_{i=1}^{H}$. Noting that any given $W_i$ appears both in Condition (26) for the link $b_i s_{i+1}$ and in Condition (28) for the link $b_i s_i$, it is straightforward to check that the implied combined requirement on such $W_i$ can be compactly expressed as indicated in (25) for all $i = 1, 2, \ldots, H$.

The fact that (25) is a *necessary* condition for the induced trading network to be PS follows directly from the condition itself – any violation would mean that some agent (buyer or seller) has incentives to eliminate one of the links in the underlying line subnetwork. As for the sufficiency, it follows from the following three observations.

(i) Given any buyer $b_i$ (with $w_i = 1$) that has a link $b_i s_j$ to some $j \in \{H+1, \ldots, H+\Delta\}$, this link is always profitable for both $b_i$ and $s_j$. Hence the maintenance of all those links is consistent with pairwise stability.

(ii) In view of (i) and the line of reasoning used in the first part of the proof, any link $b_i s_i$ and $b_i s_{i+1}$ ($i = 1, 2, \ldots, H$) in the underpinning line subnetwork is profitable to both parties involved, provided that all others in the whole induced network are in place.

(iii) The full network induced by $\{w_i\}_{i=1}^{H}$ is connected (i.e. consists of a single component) and all links are active (i.e. support trade). Hence, by virtue of Proposition 4, no additional link can affect the LBP (which, as indicated, equals the value $p^*$ given in (9)) and thus cannot affect payoffs either. This, in sum, implies that the absence of these links is consistent with pairwise stability.

Combining all of the above considerations, the proof of the Lemma is complete. $\Diamond$

Next, we build upon the previous Lemma to establish the existence of a PS network. This we do by first proposing an inductive procedure that yields a set of binary variables $\{w_i\}_{i=1}^{H}$

(or its cumulative counterparts $\{W_i\}_{i=1}^H$) that satisfy a weaker version of (25), formulated in terms of non-strict inequalities. This procedure can be decomposed as follows.

- **Induction Hypothesis** $IH(k)$: There is a set $\{w_i\}_{i=1}^k$ such that
$$i\frac{\Delta}{H} \geq W_i = \sum_{s=1}^{i} w_s \geq i\frac{\Delta}{H} - 1 \quad (i = 1, 2, ..., k). \tag{30}$$
- **Induction Step**: For any $k < H$, $IH(k) \Rightarrow IH(k+1)$.

So it is enough to show that $IH(1)$ is true and then that the Induction Step holds.

First, in order to show that $IH(1)$ is true, since $\Delta/H < 1$, it is obviously possible to satisfy
$$\frac{\Delta}{H} \geq W_1 \geq \frac{\Delta}{H} - 1 \tag{31}$$
by simply choosing $w_1 = W_1 = 0$.

Second, to establish the Inductive Step, suppose that $IH(k)$ holds for any given $k < H$. Then, given the corresponding $\{w_i\}_{i=1}^k$, we need to show that some $w_{k+1}$ can be specified such that $IH(k+1)$ holds. But this is a direct consequence of the fact that $\Delta/H < 1$, which implies that the upper and lower bounds in (30) are shifted by a magnitude that is less than one for the additional $i = k+1$ considered in new induction step.

Now we argue that the inequalities considered in the initial induction step as well as in the Induction Hypothesis can be chosen strict. As for (31), this is obvious since, again, $\Delta/H < 1$. Concerning (30), on the other hand, it is enough to show that $i(\Delta/H)$ cannot be an integer for any step $k$ along the induction procedure. Suppose otherwise, for the sake of contradiction. If $k(\Delta/H)$ is an integer at some particular step $k$, this means that for at least one of the feasible choices of $w_k \in \{0, 1\}$ one of the two inequalities in (30) for $i = k$ can be satisfied with equality. For concreteness, suppose that the choice is such that we have $k(\Delta/H) = W_k$. From the conditions formalized by those inequalities, the preceding equality implies that the deletion of the link $b_i s_{i+1}$ would divide the network in two components with an identical ratio of buyers to sellers. An analogous conclusion would be true if $k(\Delta/H) = W_k$ for the link $b_i s_{i+1}$. But this is incompatible with the assumption that $\#(B)$ and $\#(S)$ are relative primes, which then shows that the induction procedure can be made to operate in terms of strict inequalities in (30)-(31).

Finally, we are now in a position to prove the stated result. Consider any $\alpha < 1$ and suppose that $[\#(S) - \#(B)] = \Delta < \alpha[\#(B)] = \alpha H$. Clearly, one can always choose some $\hat{N}$ such that, if $N \equiv [\#(B) + \#(S)] = 2H + \Delta \geq \hat{N}$, then $H - \Delta > 2$. This means that, in the PS network $\hat{G}(H, \Delta)$ constructed before, there is at least one buyer $b_i$, with $1 < i < H$, whose only two links are to sellers $s_i$ and $s_{i+1}$. There is maximal matching $\hat{m}$ on that network with $\{s_i b_{i-1},\ s_{i+1} b_{i+1}\} \subset \hat{m}$. In such a matching, buyer $b_i$ is left unmatched. This is incompatible with efficiency, which in our case is equivalent to the requirement that all buyers (the short side of the market) be matched and trade. This establishes the first claim of the result.

For the second claim, suppose that, within the trading network $\hat{G}(H,\Delta)$, the following (maximal) matching $\tilde{m}$ materializes:

$$\tilde{m} = \{b_i s_{i+1} : \ i = 3r+1 \leq H, \ r = 0,1,2,\ldots\} \cup \{b_i s_i : \ i = 3r+3 \leq H, \ r = 0,1,2,\ldots\}.$$

Given matching $\tilde{m}$, define

$$S \equiv \{b_i : \ i = 3r+2 < H-2\}$$

and then denote by

$$T \equiv \{b_i : \ w_i = 0\}$$

the set of buyers enjoying no connection to the sellers *not* embedded in the line subnetwork underlying $\hat{G}(H,\Delta)$. The set of buyers unmatched under $\tilde{m}$ is simply $U = S \cap T$. Since, as indicated, efficiency in the present case is equivalent to having all buyers trade, the cardinality of the set $U$ provides a lower bound on the inefficiency (i.e. forgone surplus) induced by the contemplated matching outcome. We turn, therefore, to assessing its magnitude.

Assume, as stated, that for some $\alpha < 1/3$, we have $\Delta < \alpha H$. Then, a lower bound for the cardinality of $U$ can be formulated as follows:

$$\#(U) \geq \#(S) - \#(T) \geq \frac{1}{3}H - 2 - \alpha H = (\frac{1}{3} - \alpha)H - 2.$$

It is straightforward to check, therefore, that for any given $\zeta > 0$, if $\tilde{N}(\zeta)$ is chosen so that

$$\tilde{N}(\zeta) > \frac{3}{2}\frac{\zeta + 2}{1 - 3\alpha},$$

the number $\#(U)$ of unmatched players exceeds $\zeta$ whenever $N \geq \tilde{N}(\zeta)$. This establishes the second claim of the proposition and completes the proof. ■

## References


[1] Chartrand, G. (1985): *Introductory Graph Theory*, New York: Dover.

[2] Corominas-Bosch, M. (2004): "Bargaining in a network of buyers and sellers," *Journal of Economic Theory* **115**, 35-77.

[3] Gale, D. (1987): "Limit theorems for markets with sequential bargaining," *Journal of Economic Theory* **43**, 20-54.

[4] Fraja, G.D. and J. Sákovics (2001): "Walras retrouve: Decentralized trading mechanisms and the competitive price," *Journal of Political Economy* **109**(4), 842–863

[5] Hall, P. (1935): "On representatives of subsets," *Journal of the London Mathematical Society* **10**, 26-30.

[6] Jackson, M. and A. Wolinsky (1996): "A strategic model of social and economic networks," *Journal of Economic Theory* **71**, 44-74.

[7] Lauermann, S. (2012): "Asymmetric information in bilateral trade and in markets: An inversion result," *Journal of Economic Theory* **147**, 1969–1997.

[8] Manea, M. (2011): "Bargaining in stationary networks," *American Economic Review* **101**, 2042-2080 .

[9] Nakkas, A. (2012): "Bargaining in a network with heterogeneous buyers," mimeo, January, Vanderbilt University.

[10] Polanski, A. (2007): "Bilateral bargaining in networks," *Journal of Economic Theory* **134**, 557-565.

[11] Polanski, A. and E. Lazarova (2011): "Dynamic multilateral markets," working paper, University of East Anglia.

[12] Rubinstein, A. and A. Wolinsky (1985): "Equilibrium in a market with sequential bargaining," *Econometrica* **53**, 295-328.

[13] Rubinstein, A. and A. Wolinsky (1990): "Decentralized trading, strategic behaviour and the Walrasian outcome," *Rev. Econ. Stud.* **57**(1), 63–78

[14] Serrano, R. (2002): "Decentralized information and the Walrasian outcome: A pairwise meetings market with private values," *J. Math. Econ.* **38**(1–2), 65–89.

[15] Shneyerov, A. and C.L. Wong (2010): "Bilateral matching and bargaining with private information," *Games Econ. Behav.* **68**(2), 748–762