An Econometric Model of Network Formation with an Application to Board Interlocks between Firms

Cristina Gualdani*

This version: November 18, 2016
Click here for latest version

Abstract

This paper provides a framework for studying identification in a network formation model. Network formation is modelled as a static game with complete information and pure strategy equilibrium. Links have directions. Payoffs depend on some players' characteristics partially observed by the researcher and on an externality, or spillover effect (hereafter SE), that goes beyond direct connections - i.e., player i's payoff from forming a link with player j monotonically depends on the number of additional players creating a link with j. This implies that parameters in players' payoffs are partially identified without further assumptions on equilibrium selection. The set of admissible parameter values (sharp identified set) is derived. Even if restrictions are added, conducting inference on the sharp identified set is prohibitively complex when there are four or more players. To attenuate the computational difficulties, the focus is on a larger set of parameter values (outer set) obtained by bounding the empirical probability of any network section being the unique equilibrium, and the probability of such a network section being a possible equilibrium, in a local game of the network formation game. The suggested outer set shows advantages over other outer sets in the literature (Tamer, 2003; Ciliberto and Tamer, 2009; Sheng, 2014), in terms of computational tractability and width. A 95% confidence region for the characterised outer set is computed using data on board interlocks between Italian firms. Results reveal that SE has a positive sign, i.e., firm i's payoff from forming a board interlock with rival j increases with the number of additional competing firms creating a board interlock with j. In view of the co-optation theory in corporate governance, this seems to support the idea according to which the higher the number of competitors with a director sitting on j's board, the stronger their capacity to influence j's decisions and align them with the group's interests.

*I am very grateful to Adam Rosen and Aureo de Paula for their suggestions, guidance and encouragement. I also thank Federico Bugni, Antonio Cabrales, Luis Enrique Candelaria, Andrew Chesher, Allan Collard-Wexler, Alan Crawford, Martin Cripps, Antonio Guarino, Jia Li, Marta C. Lopes, Lorenzo Magnolfi, Matthew Masten, Konrad Mierendorff, Francesca Molinari, Lars Nesheim, Elena Pastorino, Franco Peracchi, Imran Rasul, Daniel Yi Xu. The views expressed in this paper are those of the author and do not necessarily reflect those of the Bank of Italy. All errors are mine.

*Email: c.gualdani.11@ucl.ac.uk. University College London.

1 Introduction

There is a successful literature studying the impact of networks on outcomes, such as technology adoption, smoking behaviour, labour outcomes, academic achievement, and firms' spending for research (e.g., Bandiera and Rasul, 2006; Nakajima, 2007; Bayer, Ross, and Topa, 2008; Graham, 2008; Calvó-Armengol, Patacchini and Zenou, 2009; Conley and Udry, 2010; Sacerdote, 2010; Helmers, Patnam and Rau, 2015). This literature imposes network exogeneity or uses instruments to control for network endogeneity. Therefore, it does not help understanding why certain networks emerge and how to design policies affecting the structure of networks. An estimable model of network formation can allow these questions to be answered by learning about agents' preferences for links through data. Hence, the present work provides a framework for studying identification in a network formation model and applies the proposed methodology to data on board interlocks between Italian firms.

Network formation is modelled as a static game with complete information and pure strategy equilibrium. Links formed by players have directions. Payoffs depend on some players' characteristics partially observed by the researcher and on an externality, or spillover effect (hereafter SE), that goes beyond direct connections - i.e., player i's payoff from forming a link with player j monotonically depends on the number of additional players creating a link with j. The spillover effect SE appears in models of formation of board interlocks between rival firms as a consequence of firms' monitoring incentives towards competitors, as explained below in detail. It also characterises models of formation of friendship networks, trust networks, and advice networks among individuals, where the number of additional individuals creating a link with individual j may respectively proxy j's social popularity or free time, j's trustworthiness, and j's competence or availability.

Identifying the spillover effect SE and other parameters shaping players' preferences is not easy because the presence of the spillover effect SE causes the network formation game to admit multiple equilibria, which in turn complicate identification unless assumptions are imposed regarding how players select the outcome observed in the data from the equilibrium set (hereafter *selection mechanism*). However, as the *selection mechanism* is unknown by the researcher and economic theory provides no guidance on its form, exploiting any restriction on it without good prior evidence may be problematic. Hence, the present work designs a framework to study identification remaining agnostic as to equilibrium selection.

The identification arguments are as follows. Existence of an equilibrium is shown by decomposing the network formation game into smaller (local) games and by establishing equilibrium existence in each local game using Tarski's fixed point theorem and a result from the entry game in Berry (1992)[1]. Equilibrium uniqueness is not guaranteed; i.e., there is more than one mapping from parameters and observed and unobserved exogenous variables, to endogenous variables. As the *selection mechanism* is unknown by the researcher and economic theory provides no guidance regarding its form, any assumption on it may be inappropriate and could bias estimates. Therefore, the econometric analysis proceeds by leaving the *selection mechanism* totally unrestricted, as seen in the most recent empirical literature on entry games (Tamer, 2003; Ciliberto and Tamer 2009 - hereafter CT; Aradillas-Lopez and Rosen, 2014). This in turn implies that the model is only partially identified; i.e., there may be more than one parameter value able to generate the empirical probability distribution of observables for some data generating process consistent with the model's assumptions. In the language of Beresteanu, Molchanov and Molinari (2011) (hereafter BMM) and Chesher and Rosen (2012), this set of admissible parameter values is denominated the sharp identified set.

Following Berry and Tamer (2006), the sharp identified set for this model can be expressed as the set of parameter values for which one can find a *selection mechanism* that, combined with the model, delivers the empirical joint probability distribution of observables. However, this representation does not facilitate inference because it involves the *selection mechanism* which, as totally unrestricted, is a function representing an infinite dimensional nuisance parameter. BMM (2011) offer a powerful alternative by showing that the sharp identified set in some classes of models can be characterised as the set of parameter values satisfying a finite number of inequalities that do not contain the *selection mechanism.* Hence, after showing that under general assumptions this model belongs to the class of models analysed by BMM (2011), the sharp identified set is derived following their approach under increasingly restrictive assumptions on the distribution of unobservables. As assumptions become more restrictive, a significant number of inequalities are redundant and, hence, can be deleted. Even so, conducting inference on the sharp identified set remains prohibitively complex when there are four or more players because it requires verification of a huge number of inequalities.

These computational difficulties can be attenuated by considering a subset of inequalities. In the language of BMM (2011), a subset of inequalities defines an outer set, i.e., a set containing the sharp identified set. This paper proposes a novel outer set obtained by bounding the empirical

[1]See the part of *Result* at p.894 in Berry (1992) concerning equilibrium existence in an entry game with negative competitive effects.

probability of any network section composed by all nodes and the links pointing to a node between the probability of such a network section being the unique equilibrium, and the probability of such a network section being a possible equilibrium, in a local game of the network formation game.

The considered outer set brings computational advantages. Specifically, inequalities defining the proposed outer set contain complicated multi-dimensional integrals that must be computed during the inference procedure for each value of parameters and exogenous observables. Computation can be done via simulation (McFadden, 1989; Pakes and Pollard, 1989) which, in principle, generates a demanding routine. However, simplifications relying on Tarski's fixed point theorem and a result from the entry game in Berry (1992)[2] can significantly speed up the whole process. Overall, Monte Carlo experiments reveal that conducting inference on the characterised outer set is computationally feasible with relatively limited computational resources when the number of players is equal to or smaller than 20. Moreover, if one is willing to impose exchangeability across unobservables, several inequalities are redundant and, hence, can be deleted, allowing further computational gains.

Lastly, the suggested outer set is compared with other outer sets in the literature (Tamer, 2003; CT, 2009; Sheng, 2014) and it is shown to offer advantages in terms of computational tractability and/or width under different sets of assumptions.

In the second part of the paper the proposed methodology is illustrated in action using data on board interlocks between Italian firms. To provide more detail on the context, most organisations are governed by a board of directors composed of executive and non-executive members. The former lead the decision-making process at the firm, and the latter are involved in the monitoring and advising of executives to make sure that they stay aligned with shareholders' interests. Directors can be members of other firms' boards. The term board interlock (or interlocking directorate) refers to any situation in which two firms share one or more directors (Allen, 1974). Moreover, a board interlock is horizontal when linked firms are competitors (Carrington, 1981).

The existence of horizontal board interlocks is a phenomenon that exists in several countries. Deeply analysed by corporate governance experts, it has also drawn the attention of economists as it raises serious antitrust concerns. In fact, horizontal board interlocks may be formed by firms to monitor rivals' decisions with the aim of enforcing potentially non-competitive agreements (Carrington, 1981; Bianco and Pagnoni, 1997; Leslie, 2004; OECD, 2008; Gabrielsen, Hjelmeng,

[2]See the part of *Result* at p.894 in Berry (1992) claiming that every equilibrium in an entry game with negative competitive effects is characterised by the same number of firms entering the market.

and Sørgard, 2011). At the same time, horizontal board interlocks may be used by firms to expand competencies by exchanging advise and expertise (Mace, 1971; Koenig and Goegel, 1981; Mizruchi and Stearns, 1994; Mizruchi, 1996).

Legislation on horizontal board interlocks is not uniform across countries. For example, in the U.S., horizontal board interlocks that meet certain jurisdictional thresholds are illegal under the Clayton Act of 1914 and subsequent ancillary legislation. By contrast, European countries do not impose any clear and general prohibition on horizontal board interlocks and mainly address them by applying domestic competition rules. As such decentralised patchwork legislation is likely to be chaotic, European institutions are discussing the possibility of harmonising regulations on horizontal board interlocks. Within this context, the policy debate on firms' incentives for participating in horizontal board interlocks is intense due to the crucial need to understand their impact on market structures. The empirical application in the present work contributes to this policy debate.

Specifically, firms are assumed to play the network formation game to create horizontal board interlocks with the purpose of monitoring and advising each other. Payoffs are modelled as functions of some firms' characteristics partially observed by the researcher and of the spillover effect SE - i.e., firm i's payoff from forming a board interlock with competitor j monotonically depends on the number of additional rival firms creating a board interlock with j[3].

The presence of the spillover effect SE is suggested by the co-optation theory in corporate governance (Selznick, 1949; Thompson and McEwen, 1958; Pfeffer and Salancik, 1978; Palmer, 1983; Mizruchi, 1996), according to which board interlocks reflect attempts by organisations to co-opt (monitor, anticipate, restrain) sources of environmental uncertainty stemming from the potentially disruptive unilateral actions of other corporations. For example, horizontal board

[3]For simplicity of exposition, this introduction omits any explanation on why horizontal board interlocks can be represented as directed links. As discussed in Section 5, the empirical application focuses on primary horizontal board interlocks. A primary horizontal board interlock arises when two competing firms share a director holding an executive role at at least one of two firms involved (Stokman, Van Der Knoop and Wasseur, 1988). The choice to concentrate on primary horizontal board interlocks is motivated by the fact that primary horizontal board interlocks, by involving individuals appointed with executive duties, are more likely to represent the long-term economic and institutional relations between firms (Mizruchi and Bunting, 1981; Stokman and Wasseur, 1985; Stokman, Van Der Knoop and Wasseur, 1988). Now consider two rival firms i and j that share a director who has an executive role at i and a non-executive role at j. In such a case, the corporate governance literature (Palmer, 1983; Richardson, 1987; Koenig and Goegel, 1981; Mizruchi and Stearn, 1994) suggests that i can monitor and advise j, but the converse does not hold. This creates asymmetries in i, j's payoffs when playing the network formation game and justifies links' directions.

interlocks may help companies to strengthen possibly collusive behaviours by monitoring each other's decisions. One may then conclude that the higher the number of competitors with a director sitting on j's board, the stronger their capacity to influence j's decisions and align these with the group's interests (positive spillover effect SE). On the other hand, the micro-economic theory suggests that, if horizontal board interlocks are formed by firms to enforce a cartel, then the higher the number of competitors with a director sitting on j's board, the lower the gains when pre-empting deviations by j from the collusive agreement (negative spillover effect SE). Hence, including the spillover effect SE in firms' payoffs (without restricting its sign), allows to understand which view is supported by the empirical evidence, when firms' heterogeneity is also considered.

A 95% confidence region for the proposed outer set is constructed following the method of Andrews and Soares (2010) (hereafter AS) and using a sample of Italian joint stock companies belonging to a cross-section of industries in 2010. Among the results, the confidence interval for the spillover effect SE, obtained via projection, has a positive sign; i.e., firm i's payoff from forming a board interlock with rival j increases with the number of additional competing firms creating a board interlock with j.

The rest of the paper is organised as follows. Section 2 contains a summary of the literature. Section 3 illustrates the model. Sections 4 discusses identification. Section 5 presents the empirical application. Section 6 provides some conclusions and directions for future research. Proofs of results are in Appendix C.

2 Related literature

This section contains a discussion of the relevant literature. Readers familiar with it may safely skip the section.

A summary of the literature on identification and estimation of network formation models is provided first. Detailed reviews are in Graham (2015) and de Paula (2015).

Among network formation models, random graph models are characterised by a focus on the probability distribution of the graph as the direct object of interest. Examples include the Erdós and Rényi model that imposes a uniform probability on the class of graphs with a given number of nodes and edges (Erdós and Rényi, 1959; 1960; Zheng, Salganik and Gelman, 2006; Hong and Xu, 2014), the Poisson random graph model that assumes independent and identical probability of link formation for each pair of nodes (Gilbert, 1959; Erdós and Rényi, 1960), models in which nodes are born sequentially and meet existing vertices according to random meetings and network-based meetings (Price, 1976; Watts and Strogatz, 1998; Barabási and Albert, 1999; Kleinber, et al., 1999; Watts, 1999; Pennock, et al., 2002; Newman, 2003; Vázquez, 2003; Jackson and Rogers, 2007; Clauset, Shalizi and Newman, 2009; Jackson, 2009; Kolaczyk, 2009), static models that include dependencies in link formation into the probability function, such as probabilistic graphical models (Frank and Strauss, 1986; Koller and Friedman, 2009) and exponential random graph models (Bollobás, 2001; Robins, Pattison, Kalish and Lusher, 2007; Jackson, 2009; Kolaczyk, 2009).

Even if random graph models can reproduce relatively well the main characteristics of real world networks, they are usually lacking micro-fundation, essential for counterfactual analysis. Conversely, in strategic models of network formation, agents form links according to specific rules (decisions can be made simultaneously or sequentially, unilaterally or bilaterally; information can be complete or incomplete), an explicit equilibrium concept (e.g., pure strategy Nash equilibrium, Nash stability[4], pairwise stability[5], pure strategy pairwise Nash equilibrium[6], etc.), and a payoff structure depending on some features of the network.

Among strategic models of network formation, iterative network formation models (Currarini, Jackson and Pin, 2009; Christakis, et al., 2010; Badev, 2014; Mele, 2015) are models in which, at each iteration of a meeting protocol, a pair of agents is randomly drawn and determines the formation, maintenance or dissolution of a link according to a payoff structure. Conversely, in static games of network formation, as the one considered in the present work, agents simultaneously choose the desired links. The key challenge of static games of network formation is the possible multiplicity of equilibria for a given value of payoff-relevant variables and parameters. When data are composed of a large number of relatively small networks, it is possible to leave unrestricted the *selection mechanism* to avoid imposing additional assumptions that could bias estimates, as seen in the most recent empirical literature on entry games (Tamer, 2003; CT, 2009; de Paula, 2013; Aradillas-Lopez and Rosen, 2014). Consequently, the model may be only partially identified. After having derived the sharp identified set (BMM, 2011; Galichon and Herny, 2011; Chesher and Rosen, 2012; Molchanov and Molinari, 2014), techniques for the estimation of partially identified models can be implemented (if unconditional moment inequalities: Chernozhukov, Hong, and Tamer, 2007; Beresteanu and Molinari, 2008; Romano and Shaikh, 2008; 2010; Rosen, 2008; Stoye, 2009; AS, 2010; Bugni, 2010; Canay, 2010; Romano, Shaikh, and

[4] Myerson (1991).
[5] Jackson and Wolinski (1996).
[6] Jackson and Wolinski (1996); Calvó-Armengol (2004).

Wolf, 2014; Bugni, Canay and Shi, 2015; Kaido, Molinari and Stoye, 2015; Pakes, et al., 2015; if conditional moment inequalities: Chetverikov, 2012; Andrews and Shi, 2013; Chernozhukov, Lee, and Rosen, 2013; Lee, Song, and Whang, 2013; 2014; Amstrong, 2014).

However, inference with multiple equilibria becomes computationally very intensive as the number of players increases because of the proliferation of possible networks. These computational difficulties can be attenuated by considering an outer set, as proposed by Sheng (2014) and in this paper, or exploring features of the model, e.g., supermodularity as in Miyauchi (2014) and Boucher (2016). In comparison with Sheng (2014), while Sheng focuses on undirected links and uses pairwise stability as solution concept, the present work features directed links and pure strategy equilibria. Sheng characterises an outer set by decomposing the network formation game into local games. A similar strategy is followed in this paper but different local games are considered so that the proposed outer sets are not equivalent. Lastly, Section 4.8 shows that, under some assumptions, the outer set designed here is contained in the outer set obtained by applying the strategy outlined by Sheng to this setting, as the second outer set ignores important interdependencies across players' decisions when links have directions.

Finally, Boucher and Mourifié (2015), Leung (2015a; 2015b), de Paula, Richards-Shubik and Tamer (2015) and Menzel (2016) develop econometric analysis when the researcher has access to only one or few large networks.

In view of the final empirical application, the existing literature on board interlocks is now summarised. Board interlocks arise for several reasons. According to the inter-organizational linkage perspective (Palmer, 1983; Ornstein, 1984; Zajac, 1988), companies are entities that possess interests. In pursuit of these interests, they form links with other firms with the main purpose of exchanging information. Board interlocks are considered relations between firms and directors are seen as agents of these relations. In this scenario, board interlocks may reflect attempts by organisations to co-opt (monitor, anticipate, restrain) sources of environmental uncertainty stemming from the potentially disruptive unilateral actions of other corporations (Selznick, 1949; Thompson and McEwen, 1958; Pfeffer and Salancik, 1978; Palmer, 1983; Mizruchi, 1996). Reduced uncertainty comes at a cost because a company's action must now reflect the co-opted group's influence (Selznick, 1949; Pfeffer and Salancik, 1978; Aldrich, 1979; Stiglitz, 1985; Bearden, 1986; Eisenhardt, 1989). For example, horizontal board interlocks may help companies to strengthen possibly collusive behaviours by monitoring and influencing each other's decisions (Carrington, 1981; Mizruchi, 1996; Bianco and Pagnoni, 1997; Leslie, 2004; Carbonai and Di Bartolomeo, 2006; Gabrielsen, Hjelmeng, and Sørgard, 2011; Di Bartolomeo and Canofari, 2015). At the same time, board interlocks may be a way to expand competencies by exchanging advice and expertise (Mace, 1971; Koenig and Goegel, 1981; Mizruchi and Stearns, 1994; Mizruchi, 1996).

Several authors have attempted to empirically investigate the co-opting role of board interlocks by estimating the correlation between a firm's profitability and size, and the intensity of board interlocks. One of the underlying ideas is that, if board interlocks arise to co-opt sources of environmental uncertainty, then larger and more profitable corporations should have more board interlocks because they represent a major source of uncertainty for other companies. Although evidence for a positive association between a firm's profitability and intensity of board interlocking is mixed (Bunting, 1976; Pennings, 1980; Carrington, 1981; Burt, 1983; Meeusen and Cuyvers, 1985; Kaplan and Reishus, 1990; Khwaja, Mian and Qamar, 2011), many studies confirm a positive relation between a firm's size and intensity of board interlocking (Dooley, 1969; Pfeffer, 1972; Allen, 1974; Mizruchi and Stearns, 1988; Booth and Deli, 1996). In line with this part of the literature, the final empirical application considers firms' size and profitability as exogenous characteristics affecting companies' decisions on horizontal board interlocks.

Conversely, according to the class alliance view (Palmer, 1983; Ornstein, 1984; Zajac, 1988; Mizruchi, 1996), directors are actors who possess interests and companies are agents of these actors. In pursuit of their interests, directors form relationships with other peers and board interlocks represent one way of doing it.

As clarified in Section 5, the empirical application in this paper adopts the inter-organizational linkage perspective. This modelling choice is suggested by the Italian legal framework, stating that a director needs her board's approval to join the board of a competing firm (Article 2390 of the Italian Civil Code). Similar provisions are laid down in most European countries.

3 The network formation game

Useful tools of network theory In order to facilitate the illustration of the network formation game, useful tools of network theory are introduced. A directed network of size N can be graphically represented as a collection of N nodes (or vertices), some of them connected by links (or edges) with directions. Nodes are labelled by the integers in $\mathcal{N}_N := \{1, 2, ..., N\}$. A link from node i to node j is denominated link ij. The collection of links from node i to other nodes constitutes node i's outgoing links. Vice versa, the collection of links from any other node to node i represents node i's incoming links. Alternatively, a directed network of size N can be expressed as an $N \times N$ matrix $\boldsymbol{G}$ with ijth component

$$G_{ij} := \begin{cases} 1 & \text{if link } ij \text{ exists} \\ 0 & \text{otherwise} \end{cases}$$

Hence, $\boldsymbol{G}$ is a zero-diagonal matrix and it can be asymmetric. As an example, Figure 1 reports a directed network of size 3.

Consider a group of $3 \leq N < \infty$ players, labelled with the integers in $\mathcal{N}_N := \{1, 2, ..., N\}$ and assume that each player $i \in \mathcal{N}_N$ controls node i. In the network formation game players, endowed with some preferences for links that are shared knowledge, simultaneously decide which links to form according to certain payoffs and equilibrium concept.

Unilateral versus bilateral network formation games The formation of link ij is unilateral when it requires the consent of player i only, as in the case of advice networks, trust networks, and, sometimes, friendship networks[7]. The formation of link ij is bilateral when it requires the consent of both players i and j, as in the case of board interlocks.

In what follows, for simplicity of exposition, identification results are illustrated for the unilateral case. Section 4.9 discusses how to obtain analogous results for the bilateral case. All proofs in the Appendix are presented, when necessary, for both cases.

Players' preferences for links Players' preferences for links depend on characteristics that are partially observed by the researcher. For any $i, j \in \mathcal{N}_N$ with $i \neq j$, let X_i and ϵ_{ij} denote, respectively, a $K \times 1$ vector of observed (to the researcher) characteristics of player i, and the unobserved (to the researcher) heterogeneity of player i affecting its preference for link ij. Let $\boldsymbol{X}$ be an $N \times K$ matrix collecting X_i $\forall i \in \mathcal{N}_N$, and $\epsilon_{i\cdot}$ be an $(N-1) \times 1$ vector collecting ϵ_{ij} $\forall j \neq i \in \mathcal{N}_N$. Lastly, let ϵ be an $N(N-1) \times 1$ vector collecting $\epsilon_{i\cdot}$ $\forall i \in \mathcal{N}_N$.

Players have complete information on their preferences for links. This assumption is based on the idea that observed networks are realisations of a long-run equilibrium. When players are individuals, complete information is postulated also in Sheng (2014) and Miyauchi (2014). When players are firms, as in the empirical application, such a restriction mimics the complete information assumption often imposed in the empirical literature on entry games (Bresnahan and

[7]Unilateralism of decisions in case of friendship networks is controversial. Some empirical papers consider friendship formation as a bilateral process (Christakis, et al., 2010; Miyauchi, 2014). Others (Badev, 2014; Mele, 2015) interpret friendships as model relationships rather than symmetric human ties and describe friendship formation as a unilateral process.

Reiss, 1991; Berry, 1992; Jia, 2008; CT, 2009; Bajari, Hong and Ryan, 2010; Aradillas-Lopez and Rosen, 2014; Fox and Lazzati, 2016)

Players' choices For any $i \in \mathcal{N}_N$, let $G_{i\cdot}$ be an $(N-1) \times 1$ vector collecting G_{ij} $\forall j \neq i \in \mathcal{N}_N$. A pure strategy vector of player i is $G_{i\cdot} \in \{0,1\}^{N-1}$ $\forall i \in \mathcal{N}_N$ and a pure strategy profile of the game is $\boldsymbol{G} \in \mathcal{G}_N$.

Players' payoffs Let $\boldsymbol{G}_{-\{\cdot i\}}$ denote the matrix $\boldsymbol{G}$ with ith column deleted. Each player $i \in \mathcal{N}_N$ gets a payoff

$$U_i(\boldsymbol{G}_{-\{\cdot i\}}, \boldsymbol{X}, \epsilon_{i\cdot}; \theta_u) := \sum_{j=1}^{N} G_{ij} \times \Big[z(X_i, X_j; \beta) + v(\sum_{k \neq i}^{N} G_{kj}; \delta) + \epsilon_{ij}\Big] \tag{3.1}$$

where $z(\cdot;\beta)$ is any function of X_i and X_j known up to a vector of parameters β, $v(\cdot;\delta)$ is any function monotone in $\sum_{k\neq i}^{N} G_{kj}$ and known up to a vector of parameters δ, $\theta_u := (\beta, \delta) \in \Theta_u \subseteq \mathbb{R}^{d_\beta + d_\delta}$, with d_β and d_δ denoting the dimensions of β and δ.

It should be noticed that the derivative of the function $v(\cdot;\delta)$ with respect to $\sum_{k\neq i}^{N} G_{kj}$ delivers the spillover effect SE illustrated in Section 1. Moreover, $U_i(\cdot;\theta_u)$ is additively separable over player i's outgoing links (hereafter *additive separability*). *Additive separability* is a common assumption in empirical models of formation of social networks (Badev, 2014; Sheng, 2014; Mele, 2014; Miyauchi, 2014; Leung, 2015). When players are firms, as in the empirical application, *additive separability* mimics the additive separability over multi-market entry decisions in the empirical literature on entry games (see references above).

All results up to Section B, included, also hold if one imposes label-specific functions and parameters, provided that the monotone functions $\{v_i(\cdot;\delta_i)\}_{\forall i \in \mathcal{N}_N}$ are restricted to have the same slope's sign across i. However, it seems natural to proceed with functions and parameters independent of players' labels, as the data generating process that will be illustrated in Assumption 1 assumes that players' identities or roles vary across replications of the network formation game, in accordance with the data used for the empirical application.

Equilibrium Agents play pure strategy Nash equilibrium (hereafter PSNE). Let $\boldsymbol{G}_{-\{i\cdot,\cdot i\}}$ be the matrix $\boldsymbol{G}$ with ith row and column deleted and $\boldsymbol{G}_{-\{\cdot i\}} = (G_{i\cdot}, \boldsymbol{G}_{-\{i\cdot,\cdot i\}})$.

Definition 1 (PSNE of the network formation game)**.** $\boldsymbol{G}$ is a PSNE of the network formation game if

$$U_i(G_{i\cdot}, \boldsymbol{G}_{-\{i\cdot,\cdot i\}}, \boldsymbol{X}, \epsilon_{i\cdot}; \theta_u) \geq U_i(\tilde{G}_{i\cdot}, \boldsymbol{G}_{-\{i\cdot,\cdot i\}}, \boldsymbol{X}, \epsilon_{i\cdot}; \theta_u)$$

$\forall \tilde{G}_{i\cdot} \neq G_{i\cdot} \in \{0,1\}^{N-1}$ and $\forall i \in \mathcal{N}_N$.

Lastly, by exploiting the *additive separability* of $U_i(\cdot;\theta_u)$, Lemma 1 maintains that the inequalities in Definition 1 simplify to a system of $N(N-1)$ equations whose solution is a PSNE of the network formation game. This characterisation is used to show existence of a PSNE equilibrium (Lemma 4).

Lemma 1 (Characterisation of a PSNE of the network formation game)**.** $\boldsymbol{G}$ is a PSNE of the network formation game if and only if

$$G_{ij} = \mathbb{1}\{z(X_i, X_j; \beta) + v(\sum_{k\neq i}^{N} G_{kj}; \delta) + \epsilon_{ij} \geq 0\} \qquad \forall i, j \in \mathcal{N}_N, i \neq j$$

4 Identification

Let $\theta_0 \in \Theta$ denote the unknown population value of $\theta_u \in \Theta_u$ and of potential parameters entering the joint distribution of unobservables. This section discusses identification of θ_0. In fact, interdependence of decisions caused by the function $v(\cdot;\delta)$ precludes using results for multivariate discrete choice models.

Developing identification arguments firstly requires investigating whether a PSNE of the network formation game exists and whether the network formation game admits multiple PSNE. This is carried out in Sections 4.1 and 4.2.

4.1 Existence of a PSPNE of the network formation game

This section proves that the network formation game has a PSNE for every value of payoff-relevant variables and parameters.

If the function $v(\cdot;\delta)$ was assumed monotone increasing, then existence of a PSNE would be guaranteed by Tarski's fixed point theorem. As the function $v(\cdot;\delta)$ is imposed generically monotone, existence of a PSNE is not straightforward and has to be shown.

Equilibrium existence is established as follows. Firstly, the network formation game is decomposed in N small (local) games such that the network formation game has an equilibrium if and only if each local game has an equilibrium. Secondly, existence of an equilibrium in each local game is proved by using Tarski's fixed point theorem and the constructive proof in Berry (1992) that shows equilibrium existence in an entry game with negative competitive effects[8]. Hence, a PSNE of the network formation game exists.

[8]See *Result* at p.894 in Berry (1992).

Describing a local game To give more details on the considered local games, some definitions are first introduced. For any $j \in \mathcal{N}_N$, the network portion collecting all nodes and the links pointing to node j is denominated *section j*. Alternatively, the *section j* of a network $\boldsymbol{G}$ can be expressed as an $(N-1) \times 1$ vector $G_{\cdot j}$ collecting G_{ij} $\forall j \neq i \in \mathcal{N}_N$. For example, Figure 2 reports the *section 2* of the network in Figure 1.

For any $j \in \mathcal{N}_N$, the local game considered is denominated the *section j game*. In the *section j game* players other than player j simultaneously announce whether they want to form a link pointing to j. Hence, the *section j game* can be thought of as an entry game. A pure strategy of agent i is $G_{ij} \in \{0, 1\}$ $\forall i \in \mathcal{N}_{\cdot j,N}$ and a pure strategy profile of the game is $G_{\cdot j} \in \{0, 1\}^{N-1}$, where $\mathcal{N}_{\cdot j,N} := \{1, ..., j-1, j+1, ..., N\}$. Each agent $i \in \mathcal{N}_{\cdot j,N}$ gets a payoff $G_{ij} \times \Big[z(X_i, X_j; \beta) + v(\sum_{k \neq i}^{N} G_{kj}; \delta) + \epsilon_{ij}\Big]$. Agents play PSNE.

Definition 2 (PSNE of the *section j game*)**.** $G_{\cdot j}$ is a PSNE of the *section j game* if

$$G_{ij} = \mathbb{1}\{z(X_i, X_j; \beta) + v(\sum_{k \neq i}^{N} G_{kj}; \delta) + \epsilon_{ij} \geq 0\} \quad \forall i \in \mathcal{N}_{\cdot j,N}. \tag{4.1}$$

Existence statements It can be seen that the payoff of any player within the *section j game* depends exclusively on $G_{\cdot j}$, $\forall j \in \mathcal{N}_N$. Combining this with the fact that the collection of sets of pure strategy profiles of the *section j game* $\forall j \in \mathcal{N}_N$ constitutes an N-partition of the set of pure strategy profiles of the network formation game allows to shows that

Lemma 2 (Decomposing the network formation game)**.** $\boldsymbol{G}$ is a PSNE of the network formation game if and only if $G_{\cdot j}$ is a PSNE of the *section j game* $\forall j \in \mathcal{N}_N$.

Moreover, using Tarski's fixed point theorem when $v(\cdot; \delta)$ is monotone increasing and the constructive proof in Berry (1992) that shows equilibrium existence in an entry game with negative competitive effects when $v(\cdot; \delta)$ is monotone decreasing, it can be proved that

Lemma 3 (Existence of a PSNE of the *section j game*)**.** There exists a PSNE of the *section j game* $\forall j \in \mathcal{N}_N$.

Hence, by Lemmas 2 and 3,

Lemma 4 (Existence of a PSNE of the network formation game)**.** There exists a PSNE of the network formation game.

4.2 Multiplicity of PSNE of the network formation game

The network formation game admits multiple equilibria for some values of payoff-relevant variables and parameters. This means that values of observed and unobserved exogenous variables do not uniquely pin down the value of endogenous variables, or, equivalently, there is no unique mapping from parameters and observed and unobserved exogenous variables, to endogenous variables. Moreover, by running simulations, it can be seen that the equilibrium set may contain outcomes with a diametrically opposite economic meaning, as the empty network and the fully connected network.

As the *selection mechanism* is unobserved by the researcher and economic theory provides no guidance regarding its form, any assumption on it may be inappropriate and could bias estimates. Therefore, the econometric analysis proceeds by leaving the *selection mechanism* totally unrestricted, as seen in the most recent empirical literature on entry games (Tamer, 2003; CT, 2009; Aradillas-Lopez and Rosen, 2014).

This implies that $\theta_0 \in \Theta$ is only partially identified, i.e. there may be more than one parameter value able to generate the empirical probability distribution of observables for some data generating process consistent with the model's assumptions. Nevertheless, the model can still be able to deliver useful information provided that set estimates or confidence regions are sufficiently tight (de Paula, 2013; Ho and Rosen, 2015).

In more detail, in principle one may want to achieve point identification of θ_0 by imposing additional direct or indirect assumptions on the *selection mechanism*. For instance, one might restrict the function $v(\cdot;\delta)$ to be monotone decreasing. This would imply that the *section j game* has a unique number of players linking to player j for any $j \in \mathcal{N}_N$, mimicking the result in Berry (1992) claiming that every equilibrium in an entry game with negative competitive effects is characterised by the same number of firms entering the market[9]. Consequently, θ_0 may be point identified by considering that number as the equilibrium outcome of interest. However, such a restriction could bias results unless the researcher has a strong prior supporting it.

In the same spirit, another possibility might be assuming that the outcome observed by the researcher is chosen by players at random from the equilibrium set (Bjorn and Vuong, 1984; Kooreman, 1994). However, such a strategy would hardly be justifiable within this framework and may produce biased empirical results.

Alternatively, one could re-design the network formation game as an iterative network forma-

[9]See *Result* at p.894 in Berry (1992).

tion model, where, at each iteration of a meeting protocol, a pair of players is drawn uniformly at random and determines the formation, maintenance or dissolution of a link (Currarini, Jackson and Pin, 2009; Christakis, et al., 2010; Badev, 2014; Mele, 2015). However, a potentially unattractive feature of this approach is that the realised sequence of meetings is contained in the set of equilibria predicted by the underlying static game, acting as an indirect restriction on the *selection mechanism* that may bias estimates.

A fourth option could be assigning a parametric form to the *selection mechanism*, as in the entry game of Bajari, Hong and Ryan (2010). However, when applied to this framework, such a strategy could bias empirical results because economic theory provides no guidance on the form of the *selection mechanism*.

Given the inappropriateness of those four approaches, remaining agnostic as to equilibrium selection offers an alternative. The present work adopts this last strategy and discusses partial identification of $\theta_0 \in \Theta$ in what follows.

4.3 Overview of partial identification results

In the language of BMM (2011) and Chesher and Rosen (2012), the set of parameter values generating the empirical probability distribution of observables for some data generating process consistent with the model's assumptions is denominated the sharp identified set and indicated by $\Theta^\star$.

Following Berry and Tamer (2006), $\Theta^\star$ can be expressed as the set of parameter values for which one can find a *selection mechanism* that, combined with the model, delivers the empirical joint probability distribution of observables. However, this representation does not facilitate inference because it involves the *selection mechanism* which, as totally unrestricted, is a function representing an infinite dimensional nuisance parameter. BMM (2011) offer a powerful alternative by showing that the sharp identified set in some classes of models can be characterised as the set of parameter values satisfying a finite number of inequalities that do not contain the *selection mechanism*. Hence, after showing that under general assumptions (Assumption 1) this model belongs to the class of models analysed by BMM (2011), $\Theta^\star$ is characterised in Section 4.5 following their approach[10].

Nevertheless, conducting inference on $\Theta^\star$ is prohibitively complex, as it requires checking a

[10] Following BMM (2011), $\Theta^\star$ is characterised in Section 4.5 by using random sets defined in the space of observables. One could also proceed with random sets defined in the space of unobservables (Chesher and Rosen, 2012).

huge number of inequalities. For example, with four players, one would need to verify $2^{4096}-2$ inequalities for each value of parameters and exogenous observables, which is a number greater than the quantity of atoms in the universe. Furthermore, in the data used for the empirical application some industries host up to 15 firms.

In order to reduce the number of inequalities, one can attempt to impose stronger assumptions on the distribution of unobservables, such as exchangeability (Assumption 2) and independence (Assumption 3). As assumptions become more restrictive, it can be shown that a significant number of inequalities are redundant, and, therefore, can be deleted (Lemmas 5 and 6). Even so, conducting inference on the sharp identified set remains prohibitively complex when there are four or more players.

An alternative strategy to attenuate these computational difficulties is considering a subset of inequalities. In the language of BMM (2011), a subset of inequalities defines an outer set, i.e., a set containing the sharp identified set. This paper proposes an outer set, Θ^o, collecting the parameter values such that the empirical probability of each realisation of $G_{\cdot j}$ is between the probability of such a realisation being the unique equilibrium of the *section j game* and the probability of such a realisation being a possible equilibrium of the *section j game*, conditional on $\boldsymbol{X}$, $\forall j \in \mathcal{N}_N$.

The suggested outer set brings computational advantages. Specifically, inequalities defining Θ^o contain complicated multi-dimensional integrals that are the probability that $g_{\cdot j}$ is the unique PSNE of the *section j game* and the probability that $g_{\cdot j}$ is a possible PSNE of the *section j game*, for any realisation $g_{\cdot j}$ of $G_{\cdot j}$. These integrals must be computed during the inference procedure for each value of parameters and exogenous observables. Computation can be done via the simple frequency simulator proposed by McFadden, (1989) and Pakes and Pollard (1989). In principle, one would need to draw several values of unobservables and verify whether each of all possible 2^{N-1} realisations of $G_{\cdot j}$ is a PSNE of the *section j game* for every drawn value, generating a demanding routine when N is not small. However, by applying Tarski's fixed point theorem when a parameter value is such that the *section j game* is supermodular, or mimicking for the *section j game* the result in Berry (1992) claiming that every equilibrium in an entry game with negative competitive effects is characterised by the same number of firms entering the market[11] in the opposite case, the amount of realisations of $G_{\cdot j}$ to check for every drawn value of unobservables may be substantially reduced, speeding up the whole process. Overall, Monte Carlo experiments reveal that conducting inference on Θ^o is computationally feasible with

[11] See *Result* at p.894 in Berry (1992).

relatively limited computational resources when the number of players is equal to or smaller than 20.

Furthermore, by adding exchangeability across unobservables (Assumption 2), it can be shown that focusing on the *section j game* for a $j \in \mathcal{N}_N$ (instead of $\forall j \in \mathcal{N}_N$) delivers an outer set equivalent to Θ^o, and the inequalities involving realisations of $(G_{\cdot j}, \boldsymbol{X})$ identical up to a permutation of labels other than label j are equivalent and, hence, can be deleted, allowing further computational gains.

Θ^o is compared with other outer sets in the literature. Specifically, Tamer (2003) and CT (2009) illustrate a static entry game with complete information and construct an outer set collecting the parameter values such that the empirical probability of each realisation of endogenous variables is between the probability of such a realisation being the unique equilibrium of the entry game and the probability of such a realisation being a possible equilibrium of the entry game, conditional on players' observed characteristics. Thus, one can characterise an outer set, Θ^o_{CT}, collecting the parameter values such that the empirical probability of each realisation of $\boldsymbol{G}$ is between the probability of such a realisation being the unique equilibrium of the network formation game and the probability of such a realisation being a possible equilibrium of the network formation game, conditional on $\boldsymbol{X}$. However, computational gains may be insufficient because conducting inference on Θ^o_{CT} requires checking $2 \times 2^{N(N-1)}$ inequalities[12] for each value of parameters and exogenous observables. For example, with 15 players as in the data used for the empirical application, one would need to verify 3.291×10^{63} inequalities for each value of parameters and exogenous observables. Instead, under general assumptions (Assumption 1), Θ^o brings greater computational advantages by requiring verification of $2N \times 2^{N-1}$ inequalities[13], for each value of parameters and exogenous observables, with savings, with respect to Θ^o_{CT}, of a factor depending on $\frac{1}{N}2^{(N-1)^2}$. In terms of width, it can be shown that $\Theta^o_{CT} \subseteq \Theta^o$ under general assumptions (Assumption 1) and $\Theta^o = \Theta^o_{CT}$ when some independence across unobservables is imposed (Assumption 3).

Alternatively, Sheng (2014) considers a game of formation of an undirected networks[14] and designs an outer set collecting the parameter values such that the empirical probability of each realisation of a subnetwork[15] is between the probability of such a realisation being the unique

[12] In fact, $\boldsymbol{G}$ has $2^{N(N-1)}$ possible realisations and, for each realisation, 2 inequalities should be checked.

[13] In fact, $G_{\cdot j}$ has 2^{N-1} possible realisations and, for each realisation, 2 inequalities should be checked.

[14] A network is undirected when links have no direction.

[15] Consider a network characterised by a set of nodes, $\mathcal{N}_N$, and a set of links, $\mathcal{E}$. A subnetwork is defined by a subset of vertices, $\tilde{\mathcal{N}}_N \subseteq \mathcal{N}_N$, and a subset of links, $\tilde{\mathcal{E}} \subseteq \mathcal{E}$, such that $\tilde{\mathcal{E}}$ contains all links in $\mathcal{E}$ connecting any

equilibrium of the *subnetwork game*[16] and the probability of such a realisation being a possible equilibrium of the *subnetwork game*, conditional on $\boldsymbol{X}$, for every subnetwork of size equal to or smaller than α, with $2 \leq \alpha \leq N$. α is set by the researcher according to the available computational resources. Let Θ_S^o be the outer set obtained by applying Sheng's strategy to the network formation game considered in this paper. It can be noticed that if $\alpha = N$ then one goes back to Θ_{CT}^o because there exists only one subnetwork of size N, that is the whole network. If $\alpha < N$, then it can be shown that $\Theta^o \subseteq \Theta_S^o$ when some independence across unobservables is imposed (Assumption 3). From a computational point of view, when $\alpha << N$, conducting inference on Θ_S^o could be computationally easier than conducting inference on Θ^o but set estimates or confidence regions might not be sufficiently tight, because several interdependencies across players' decisions may be ignored.

4.4 Assumptions

This section illustrates the assumptions used, in different combinations, for claiming next results.

Assumption 1 (Data generating process)**.** The data generating process is as follows:

(i) M groups of agents are randomly selected from an infinite population of groups

(ii) For each group $m \in \{1, ..., M\}$

- An integer N_m is drawn from $\mathbb{N} \setminus \{1, 2\}$.
- N_m agents are randomly selected and labelled $1, 2, ..., N_m$. Each agent $i \in \mathcal{N}_{N_m}$ is endowed with some characteristics $X_{i,m}, \epsilon_{i\cdot,m}$[17].
- Agents play the network formation game. When the model predicts a non-singleton set of PSNE, players choose one element of it. The researcher only knows which element players have selected and, crucially, does not recognise whether an observed PSNE has been picked by players from a singleton or a non-singleton set of PSNE.

(iii) A sample of i.i.d. observations $\{N_m, \boldsymbol{X_{N_m}}, \boldsymbol{G_{N_m}}\}_{m=1}^M$ is collected, where the subscript N_m highlights the dependence of matrix size on N_m. The sampling scheme is designed

two nodes in $\tilde{\mathcal{N}}_N$. The size of such a subnetwork is $|\tilde{\mathcal{N}}_N|$.

[16] Given a subnetwork, the *subnetwork game* is intended as the static game with complete information underlying the formation of the subnetwork.

[17] By construction these characteristics are i.i.d. across i.

such that the probability distribution of $\boldsymbol{G}_{\boldsymbol{N_m}}$ conditional on $N_m, \boldsymbol{X}_{\boldsymbol{N_m}}$ is identified $\forall m \in \{1, ..., M\}$[18].

In order to simplify the exposition and without loss of generality, the entire identification analysis in the remaining of Section 4 is made assuming that N_m is a degenerate random variable equal to n, and it holds $\forall n \in \mathbb{N}\backslash\{1, 2\}$. Moreover, the subscript n is left out from random matrices/vectors whose size depends on n and from their realisations. Also the subscript m is deleted to clean up the notation.

(iv) Conditional on $\boldsymbol{X}$, ϵ is continuously distributed on $\mathbb{R}^{n(n-1)}$ with cdf denoted by $F_{\boldsymbol{X}}(\cdot; \theta_\epsilon)$ and known up to $\theta_\epsilon \in \Theta_\epsilon \subseteq \mathbb{R}^{d_\epsilon}$. This implies that, conditional on $\boldsymbol{X}$, $\epsilon_{\cdot j}$, which is a $(n-1) \times 1$ vector collecting ϵ_{ij} $\forall i \in \mathcal{N}_{\cdot j,n}$, is continuously distributed on $\mathbb{R}^{n-1}$ with cdf denoted by $\tilde{F}_{\cdot j, \boldsymbol{X}}(\cdot; \theta_\epsilon)$ and known up to $\theta_\epsilon \in \Theta_\epsilon \subseteq \mathbb{R}^{d_\epsilon}$ $\forall j \in \mathcal{N}_n$.

(v) All random variables are defined on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$.

(vi) $\Theta := (\Theta_u \cup \Theta_\epsilon) \ni \theta := (\theta_u' \ \theta_\epsilon')'$ is compact. $\theta_0 := (\theta_{u,0}' \ \theta_{\epsilon,0}')' \in \Theta$ is the unknown population value of θ.

Assumption 2 (Exchangeability)**.** *Conditional on* $\boldsymbol{X}$, $\{\epsilon_{ij}\}_{\forall i,j \in \mathcal{N}, i \neq j}$ *are exchangeable across* ij.

Assumption 3 (Independence)**.** *Conditional on* $\boldsymbol{X}$, $\{\epsilon_{\cdot j}\}_{\forall j \in \mathcal{N}_n}$ *are independently distributed across* j.

Discussion on anonymity in Assumption 1 In accordance with the data used for the empirical application, the data generating process assumes that players' identities or roles vary across networks, meaning that labels are assigned arbitrarily. It naturally follows that payoffs do not depend on players' labels, i.e., they are anonymous (hereafter *anonymity*). Consequently, equilibrium sets and *selection mechanisms* do not depend on players' labels.

All results in the remaining of Section 4, with the exclusion of Lemma 6, also hold if the data generating process assigns specific identities or roles to players. When the data generating process assigns specific identities or roles to players, Lemma 6 holds if one additionally restricts the *selection mechanism* not to depend on labels.

[18] Also a time/space stationary sampling scheme is sufficient to identify the probability distribution of $\boldsymbol{G}_{\boldsymbol{N_m}}$ conditional on N_m and $\boldsymbol{X}_{\boldsymbol{N_m}}$ $\forall m \in \{1, ..., M\}$. Epstein, Kaido and Seo (2015) discuss inference without restricting the data generating process.

Discussion on Assumption 1 (iv) Imposing ϵ continuously distributed on $\mathbb{R}^{n(n-1)}$ is a technical assumption allowing to apply the results in BMM (2011) to characterise $\Theta^\star$. Assuming that the conditional probability distribution of ϵ is known up to a vector of parameters makes the model fully parametric with the exception of the *selection mechanism*. Whether and how it is possible to relax the parametric structure is a complicated question left to future research.

Discussion on Assumptions 2 and 3 Assumption 2 holds if $\epsilon_{ij} = \alpha_i + \xi_{ij}$, where $\{\alpha_i\}_{i\in\mathcal{N}_n}$ are i.i.d. across i, $\{\xi_{ij}\}_{\forall i,j\in\mathcal{N}_n, i\neq j}$ are i.i.d. across ij, and α_i is independent of ξ_{kh} $\forall i,k,h \in \mathcal{N}_n$ with $i \neq k \neq h$. Assumption 2 holds if $\epsilon_{ij} = \beta_j + \xi_{ij}$, where $\{\beta_j\}_{j\in\mathcal{N}_n}$ are i.i.d. across j, $\{\xi_{ij}\}_{\forall i,j\in\mathcal{N}_n, i\neq j}$ are i.i.d. across ij, and β_j is independent of ξ_{kh} $\forall i,k,h \in \mathcal{N}_n$ with $i \neq k \neq h$. Assumption 2 holds if $\epsilon_{ij} = \alpha_i + \beta_j + \xi_{ij}$, where $\{\alpha_i\}_{i\in\mathcal{N}_n}$ are i.i.d. across i, $\{\beta_j\}_{j\in\mathcal{N}_n}$ are i.i.d. across j, $\{\xi_{ij}\}_{\forall i,j\in\mathcal{N}_n, i\neq j}$ are i.i.d. across ij, and $\alpha_i, \beta_j, \xi_{kh}$ are independent of each other $\forall i,j,k,h \in \mathcal{N}_n$ with $i \neq j \neq k \neq h$. Lastly, Assumption 2 holds if $\{\epsilon_{ij}\}_{\forall i,j\in\mathcal{N}_n, i\neq j}$ are i.i.d. across ij.

For any $j \in \mathcal{N}_n$, Assumption 3 allows ϵ_{ij} to be correlated (not necessarily linearly) with ϵ_{kj} $\forall k,i \in \mathcal{N}_n$ with $k \neq i \neq j$. However, for any $i \in \mathcal{N}_n$, it does not allow ϵ_{ij} to be correlated with ϵ_{ik} $\forall k \in N_n$ with $k \neq i$. Lastly, Assumption 3 holds if $\{\epsilon_{ij}\}_{\forall i,j\in\mathcal{N}_n, i\neq j}$ are i.i.d. across ij.

Assumptions 2 and 3 hold together if $\epsilon_{ij} = \beta_j + \xi_{ij}$, where $\{\beta_j\}_{j\in\mathcal{N}_n}$ are i.i.d. across j, $\{\xi_{ij}\}_{\forall i,j\in\mathcal{N}_n, i\neq j}$ are i.i.d. across ij, and β_j is independent of ξ_{kh} $\forall i,k,h \in \mathcal{N}_n$, $i,k,h \in \mathcal{N}_n$ with $i \neq k \neq h$.

Additional notation Let $\mathcal{G}_n$ be the support of $\boldsymbol{G}$, with $|\mathcal{G}_n| = 2^{n(n-1)}$. Let $\mathcal{X}_n$ be the support of $\boldsymbol{X}$.

Let $\boldsymbol{g}$ and $\boldsymbol{x}$ denote, respectively, a realisation of $\boldsymbol{G}$ and a realisation of $\boldsymbol{X}$. For any $j \in \mathcal{N}_n$, let $g_{\cdot j}$ and $K_{\cdot j}$ denote, respectively, a realisation of $G_{\cdot j}$ and a collection of realisations of $G_{\cdot j}$.

Let $\mathcal{K}_{\mathcal{R}}$ denote the family of compact subsets of the set $\mathcal{R}$. Let Δ^{d-1} indicate the unit simplex in $\mathbb{R}^d$ and $\mathcal{V}(\Delta^{d-1})$ the collection of its vertices.

Let $\varphi : \mathcal{N}_n \to \mathcal{N}_n$ be a bijective function representing a permutation of labels. For any $j \in \mathcal{N}_n$, $g_{\cdot j} \in \{0,1\}^{n-1}$ and permutation of labels φ, let $g^{\varphi}_{\cdot\varphi(j)}$ be the realisation of $G_{\cdot j}$ obtained by applying φ to $g_{\cdot j}$. For any $j \in \mathcal{N}_n$, $K_{\cdot j} \in \mathcal{K}_{\{0,1\}^{n-1}}$ and permutation of labels φ, let $K^{\varphi}_{\cdot\varphi(j)}$ be the collection of realisations of $G_{\cdot j}$ obtained by applying φ to each element of $K_{\cdot j}$. For any $\boldsymbol{x} \in \mathcal{X}_n$ and permutation of labels φ, let $\boldsymbol{x}^{\varphi}$ be the realisation of $\boldsymbol{X}$ obtained by applying φ to $\boldsymbol{x}$. Examples of $g^{\varphi}_{\cdot\varphi(j)}, \boldsymbol{x}^{\varphi}$ are in Appendix D.1.

4.5 The sharp identified set under Assumption 1

In this section the sharp identified set, $\Theta^\star$, is characterised under Assumption 1, following BMM (2011).

Arguments are articulated in three parts. Firstly, under Assumption 1, $\Theta^\star$ can be represented as the set of parameter values such that the empirical conditional probability distribution of $\boldsymbol{G}$ is included in the collection of conditional probability distributions of $\boldsymbol{G}$ predicted by the model. Indeed, for each parameter value, the model predicts a collection of conditional probability distributions because the *selection mechanism* is left unspecified. Secondly, such a collection can be conveniently expressed as the conditional expectation of an appropriately defined random closed set (conditional Aumann expectation). Thirdly, checking the inclusion condition is equivalent to verifying a finite number of inequalities that involve every compact subset of $\mathcal{G}_n$ and do not contain the selection mechanism (Artstein's inequalities for $\boldsymbol{G}$).

Formal arguments Let $\mathcal{A}_{\mathcal{G}_n} \subset \mathcal{K}_{\mathcal{G}_n}$ be the collection of compact subsets of $\mathcal{G}_n$ obtained by taking the Cartesian product of all possible ordered n-tuples with repetition from $\mathcal{K}_{\{0,1\}^{n-1}}$[19].

Given $\theta \in \Theta$, consider the map $\mathcal{S}_{\theta_u}(\boldsymbol{X}, \epsilon) : \Omega \to \mathcal{A}_{\mathcal{G}_n}$ such that $\mathcal{S}_{\theta_u}(\boldsymbol{X}(\omega), \epsilon(\omega))$ is the set of PSNE of the network formation game, for any $\omega \in \Omega$. Following Proposition 3.1 in BMM (2011), under Assumption 1 $\mathcal{S}_{\theta_u}(\boldsymbol{X}, \epsilon)$ is a random closed set almost surely non-empty. $\mathcal{S}_{\theta_u}(\boldsymbol{X}, \epsilon)$ takes values in $\mathcal{A}_{\mathcal{G}_n}$ by Lemma 2 according to which the set of PSNE of the network formation game is the Cartesian product of the set of PSNE of the *section j game* across $j \in \mathcal{N}_n$.

Consider the random matrix $\boldsymbol{G_{\theta_u}} : \Omega \to \mathcal{G}_n$ such that $\boldsymbol{G_{\theta_u}}(\omega)$ reveals the equilibrium selected by players from $\mathcal{S}_{\theta_u}(\boldsymbol{X}(\omega), \epsilon(\omega))$, for any $\omega \in \Omega$. $\boldsymbol{G_{\theta_u}}$ is such that $\boldsymbol{G_{\theta_u}}(\omega) \in \mathcal{S}_{\theta_u}(\boldsymbol{X}(\omega), \epsilon(\omega))$ $\forall \omega \in \Omega$ a.s., and, for this reason, it is called a selection of $\mathcal{S}_{\theta_u}(\boldsymbol{X}, \epsilon)$. When $\omega \in \Omega$ is such that $\mathcal{S}_{\theta_u}(\boldsymbol{X}(\omega), \epsilon(\omega))$ is non-singleton, different selections are possible.

Information contained in $\mathcal{S}_{\theta_u}(\boldsymbol{X}, \epsilon)$ can be represented in a more convenient way by introducing another random closed set. Given $\boldsymbol{G_{\theta_u}}$, let $Q_{\boldsymbol{G_{\theta_u}}} : \Omega \to \mathcal{V}(\Delta^{2^{n(n-1)}-1})$ be a $2^{n(n-1)} \times 1$ random vector such that the ith element of $Q_{\boldsymbol{G_{\theta_u}}}(\omega)$ is equal to 1 if $\boldsymbol{G_{\theta_u}}(\omega)$ is equal to the ith element of $\mathcal{G}_n$, for any $\omega \in \Omega$.

Let $\mathcal{A}_{\mathcal{V}(\Delta^{2^{n(n-1)}-1})} \subset \mathcal{K}_{\mathcal{V}(\Delta^{2^{n(n-1)}-1})}$ be the collection of compact subsets of $\mathcal{V}(\Delta^{2^{n(n-1)}-1})$ obtained by taking the Cartesian product of all possible ordered n-tuples with repetition from $\mathcal{K}_{\mathcal{V}(\Delta^{2^{n(n-1)}-1})}$.

[19] $\mathcal{A}_{\mathcal{G}_n}$ has cardinality $(2^{2^{n-1}} - 1)^n < |\mathcal{K}_{\mathcal{G}_n}| = 2^{2^{n(n-1)}} - 1$. Appendix D.2 explains how to construct $\mathcal{A}_{\mathcal{G}_n}$.

Consider the map $\mathcal{Q}_{\theta_u}(\boldsymbol{X}, \epsilon) \rightarrow \mathcal{A}_{\mathcal{V}(\Delta^{2^{n(n-1)}-1})}$ such that $\mathcal{Q}_{\theta_u}(\boldsymbol{X}(\omega), \epsilon(\omega))$ collects $Q_{\boldsymbol{G}_{\boldsymbol{\theta_u}}}(\omega)$ for any possible selection $\boldsymbol{G}_{\boldsymbol{\theta_u}}$ of $\mathcal{S}_{\theta_u}(\boldsymbol{X}, \epsilon)$, for any $\omega \in \Omega$. Following BMM (2011), under Assumption 1 $\mathcal{Q}_{\theta_u}(\boldsymbol{X}, \epsilon)$ is a random closed set almost surely non-empty[20]. Moreover, $Q_{\boldsymbol{G}_{\boldsymbol{\theta_u}}}$ is such that $Q_{\boldsymbol{G}_{\boldsymbol{\theta_u}}}(\omega) \in \mathcal{Q}_{\theta_u}(\boldsymbol{X}(\omega), \epsilon(\omega))$ $\forall \omega \in \Omega$ a.s., and, for this reason, it is called a selection of $\mathcal{Q}_{\theta_u}(\boldsymbol{X}, \epsilon)$.

Consider the conditional expectation of $Q_{\boldsymbol{G}_{\boldsymbol{\theta_u}}}$, $E_{\theta_\epsilon}(Q_{\boldsymbol{G}_{\boldsymbol{\theta_u}}}|\boldsymbol{X} = \boldsymbol{x})$. It turns out that this is the conditional probability distribution of $\boldsymbol{G}$ predicted by the model given $\theta \in \Theta$. Thus, the conditional expectation of $\mathcal{Q}_{\theta_u}(\boldsymbol{X}, \epsilon)$ can be thought of as the set of conditional expectations of all its selections and it represents the collection of conditional probability distributions of $\boldsymbol{G}$ predicted by the model given $\theta \in \Theta$.

More formally, based on the fact that $\mathcal{Q}_{\theta_u}(\boldsymbol{X}, \epsilon)$ is integrably bounded[21],

$$\mathbb{E}_{\theta_\epsilon}(\mathcal{Q}_{\theta_u}(\boldsymbol{X}, \epsilon)|\boldsymbol{X} = \boldsymbol{x}) := \Big\{ E_{\theta_\epsilon}(Q_{\boldsymbol{G}_{\boldsymbol{\theta_u}}}|\boldsymbol{X} = \boldsymbol{x}) | Q_{\boldsymbol{G}_{\boldsymbol{\theta_u}}} \in Sel(\mathcal{Q}_{\theta_u}(\boldsymbol{x}, \epsilon)) \Big\} \tag{4.2}$$

is known as the conditional Aumann expectation of $\mathcal{Q}_{\theta_u}(\boldsymbol{X}, \epsilon)$. By Theorem D.1 in BMM (2011), under Assumption 1 the conditional Aumann expectation of $\mathcal{Q}_{\theta_u}(\boldsymbol{X}, \epsilon)$ is a non-empty, closed and convex subset of $\Delta^{2^{n(n-1)}-1}$ with finitely many extreme points, for any $\boldsymbol{x} \in \mathcal{X}_n$ a.s.

If the model is correctly specified, then the empirical conditional probability distribution of $\boldsymbol{G}$ should be contained in the collection of conditional probability distributions of $\boldsymbol{G}$ predicted by the model given $\theta_0 \in \Theta$, i.e.

$$\mathbb{P}(\boldsymbol{G}|\boldsymbol{X} = \boldsymbol{x}) \in \mathbb{E}_{\theta_{\epsilon,0}}(\mathcal{Q}_{\theta_{u,0}}(\boldsymbol{X}, \epsilon)|\boldsymbol{X} = \boldsymbol{x}), \ \forall \boldsymbol{x} \in \mathcal{X}_n \text{ a.s.}$$

where $\mathbb{P}(\boldsymbol{G}|\boldsymbol{X} = \boldsymbol{x})$ is a $2^{n(n-1)} \times 1$ vector denoting the empirical probability distribution of $\boldsymbol{G}$ conditional on $\boldsymbol{x}$, identified under Assumption 1.

Hence, following BMM (2011), $\Theta^\star$ under Assumption 1 can be written as

$$\Theta^\star := \Big\{ \theta \in \Theta | \mathbb{P}(\boldsymbol{G}|\boldsymbol{X} = \boldsymbol{x}) \in \mathbb{E}_{\theta_\epsilon}(\mathcal{Q}_{\theta_u}(\boldsymbol{X}, \epsilon)|\boldsymbol{X} = \boldsymbol{x}), \ \forall \boldsymbol{x} \in \mathcal{X}_n \text{ a.s.} \Big\} \tag{4.3}$$

Furthermore, by Theorem D.2 in BMM (2011), under Assumption 1 (4.3) is equivalent to

$$\Theta^\star = \Big\{ \theta \in \Theta | \mathbb{P}(\boldsymbol{G} \in K|\boldsymbol{X} = \boldsymbol{x}) \leq T_{\mathcal{S}_{\theta_u}(\boldsymbol{X},\epsilon)|\boldsymbol{X}=\boldsymbol{x}}(K) \ \forall K \in \mathcal{K}_{\mathcal{G}_n}, \ \forall \boldsymbol{x} \in \mathcal{X}_n \text{ a.s.} \Big\} \tag{4.4}$$

where $T_{\mathcal{S}_{\theta_u}(\boldsymbol{X},\epsilon)|\boldsymbol{X}=\boldsymbol{x}} : \mathcal{K}_{\mathcal{G}_n} \rightarrow [0, 1]$ is known as the capacity functional of $\mathcal{S}_{\theta_u}(\boldsymbol{X}, \epsilon)$ conditional on $\boldsymbol{x}$, and it is prescribed by $T_{\mathcal{S}_{\theta_u}(\boldsymbol{X},\epsilon)|\boldsymbol{X}=\boldsymbol{x}}(K) := \mathbb{P}(\mathcal{S}_{\theta_u}(\boldsymbol{X}, \epsilon) \cap K \neq \emptyset | \boldsymbol{X} = \boldsymbol{x})$ for any $K \in \mathcal{K}_{\mathcal{G}_n}$.

[20] This is due to $\mathcal{Q}_{\theta_u}(\boldsymbol{X}, \epsilon)$ being given by a continuous map applied to $\mathcal{S}_{\theta_u}(\boldsymbol{X}, \epsilon)$.

[21] In fact, $\mathcal{Q}_{\theta_u}(\boldsymbol{X}, \epsilon)$ is almost surely non-empty and $\sup\{||Q_{\boldsymbol{G}}|| \text{ s.t. } \boldsymbol{G} \in Sel(\mathcal{Q}_{\theta_u}(\boldsymbol{X}, \epsilon))\}$ is integrable.

Each inequality in (4.4) is known as Artstein's inequality. Notice that, despite $\mathcal{S}_{\theta_u}(\boldsymbol{X},\epsilon)$ takes values in $\mathcal{A}_{\mathcal{G}_n} \subset \mathcal{K}_{\mathcal{G}_n}$, Artstein's inequality should be verified $\forall K \in \mathcal{K}_{\mathcal{G}_n}$. Lastly, this characterisation of $\Theta^\star$ contains $2^{2^{n(n-1)}} - 2$ inequalities $\forall \theta \in \Theta$ and $\forall \boldsymbol{x} \in \mathcal{X}_n$.

4.6 The sharp identified set under Assumptions 1, 3

Conducting inference on $\Theta^\star$ as characterised in (4.4) is prohibitively complex, as it requires checking a huge number of inequalities. For example, with four players, one would need to verify $2^{4096} - 2$ inequalities $\forall \theta \in \Theta$ and $\forall \boldsymbol{x} \in \mathcal{X}_n$, which is a number greater than the quantity of atoms in the universe. Furthermore, in the data used for the empirical application some industries host up to 15 firms.

In order to reduce the number of inequalities, one can impose stronger assumptions on the distribution of unobservables. This section investigates whether, by introducing Assumption 3, $\Theta^\star$ can be characterised by fewer inequalities than in (4.4).

Arguments are articulated as follows. By construction of $\mathcal{A}_{\mathcal{G}_n}$, for any $K \in \mathcal{A}_{\mathcal{G}_n}$, there exists n sets, $K_{\cdot 1} \in \mathcal{K}_{\{0,1\}^{n-1}}$, ..., $K_{\cdot n} \in \mathcal{K}_{\{0,1\}^{n-1}}$, such that their Cartesian product delivers K. Hence, given $\theta \in \Theta$, one may think that checking Artstein's inequality for $\boldsymbol{G}$ involving K is equivalent to checking Artstein's inequality for $G_{\cdot j}$ involving $K_{\cdot j}$ $\forall j \in \mathcal{N}_n$. It turns out that this is true under Assumptions 1 and 3. Indeed, by Assumptions 1 and 3 combined with Lemma 2, Artstein's inequality for $\boldsymbol{G}$ involving $K \in \mathcal{A}_{\mathcal{G}_n}$ is equal to the product across $j \in \mathcal{N}_n$ of Artstein's inequality for $\boldsymbol{G}_{\cdot j}$ involving $K_{\cdot j} \in \mathcal{K}_{\{0,1\}^{n-1}}$. As all terms are between 0 and 1, if Artstein's inequality for $\boldsymbol{G}_{\cdot j}$ involving $K_{\cdot j}$ is satisfied $\forall j \in \mathcal{N}_n$, then, by taking the product across j, Artstein's inequality for $\boldsymbol{G}$ involving K is verified too. Thus, $\Theta^\star$ can be characterised by fewer inequalities than in (4.4) as the cardinality of $\mathcal{K}_{\{0,1\}^{n-1}}$ is smaller than the cardinality of $\mathcal{K}_{\mathcal{G}_n}$[22].

Formal arguments Given $\theta \in \Theta$, consider the map $\mathcal{S}_{\theta_u,\cdot j}(\boldsymbol{X},\epsilon_{\cdot j}) : \Omega \to \mathcal{K}_{\{0,1\}^{n-1}}$ such that $\mathcal{S}_{\theta_u,\cdot j}(\boldsymbol{X}(\omega),\epsilon_{\cdot j}(\omega))$ is the set of PSNE of the *section j game*, for any $\omega \in \Omega$. Following Proposition 3.1 in BMM (2011), under Assumption 1 $\mathcal{S}_{\theta_u,\cdot j}(\boldsymbol{X},\epsilon_{\cdot j})$ is a random closed set almost surely non-empty.

Consider the random vector $G_{\theta_u,\cdot j} : \Omega \to \{0,1\}^{n-1}$ such that $G_{\theta_u,\cdot j}(\omega)$ reveals the equilibrium selected by players from $\mathcal{S}_{\theta_u,\cdot j}(\boldsymbol{X}(\omega),\epsilon_{\cdot j}(\omega))$, for any $\omega \in \Omega$. $G_{\theta_u,\cdot j}$ is such that $G_{\theta_u,\cdot j}(\omega) \in \mathcal{S}_{\theta_u,\cdot j}(\boldsymbol{X}(\omega),\epsilon_{\cdot j}(\omega))$ $\forall \omega \in \Omega$ a.s., and, for this reason, it is called a selection of $\mathcal{S}_{\theta_u,\cdot j}(\boldsymbol{X},\epsilon_{\cdot j})$. When $\omega \in \Omega$ is such that $\mathcal{S}_{\theta_u,\cdot j}(\boldsymbol{X}(\omega),\epsilon_{\cdot j}(\omega))$ is non-singleton, different selections are possible.

[22]Specifically, one can save $(2^{2^{n-1}} - 1)^n - n(2^{2^{n-1}} - 2) - 1$ inequalities $\forall \theta \in \Theta$ and $\forall \boldsymbol{x} \in \mathcal{X}_n$.

Let $\mathcal{K}^{\diamond}_{\mathcal{G}_n} := \mathcal{K}_{\mathcal{G}_n} \setminus \mathcal{A}_{\mathcal{G}_n}$, i.e., $\mathcal{K}^{\diamond}_{\mathcal{G}_n}$ is the collection of sets not included in $\mathcal{A}_{\mathcal{G}_n}$ obtained by taking the union of elements of $\mathcal{A}_{\mathcal{G}_n}$. Let $\mathbb{P}(G_{\cdot j}|\boldsymbol{X} = \boldsymbol{x})$ be a $2^{n-1} \times 1$ vector denoting the empirical probability distribution of $G_{\cdot j}$ conditional on $\boldsymbol{x}$, identified under Assumption 1.

Consider the set

$$\Theta^{\star\star} := \Big\{\theta \in \Theta | \mathbb{P}(G_{\cdot j} \in K_{\cdot j}|\boldsymbol{X} = \boldsymbol{x}) \le T_{\mathcal{S}_{\theta_u,\cdot j}(\boldsymbol{X},\epsilon_{\cdot j})|\boldsymbol{X}=\boldsymbol{x}}(K_{\cdot j}) \ \forall K_{\cdot j} \in \mathcal{K}_{\{0,1\}^{n-1}}, \ \forall j \in \mathcal{N}_n \text{ and}$$
$$\mathbb{P}(\boldsymbol{G} \in K|\boldsymbol{X} = \boldsymbol{x}) \le T_{\mathcal{S}_{\theta_u}(\boldsymbol{X},\epsilon)|\boldsymbol{X}=\boldsymbol{x}}(K) \ \forall K \in \mathcal{K}^{\diamond}_{\mathcal{G}_n}, \ \forall \boldsymbol{x} \in \mathcal{X}_n \text{ a.s.}\Big\} \tag{4.5}$$

where $T_{\mathcal{S}_{\theta_u,\cdot j}(\boldsymbol{X},\epsilon_{\cdot j})|\boldsymbol{X}=\boldsymbol{x}} : \mathcal{K}_{\{0,1\}^{n-1}} \to [0,1]$ is known as the capacity functional of $\mathcal{S}_{\theta_u,\cdot j}(\boldsymbol{X},\epsilon_{\cdot j})$ conditioned on $\boldsymbol{x}$, and it is prescribed by $T_{\mathcal{S}_{\theta_u,\cdot j}(\boldsymbol{X},\epsilon_{\cdot j})|\boldsymbol{X}=\boldsymbol{x}}(K_{\cdot j}) := \mathbb{P}(\mathcal{S}_{\theta_u,\cdot j}(\boldsymbol{X},\epsilon_{\cdot j}) \cap K_{\cdot j} \neq \emptyset|\boldsymbol{X} = \boldsymbol{x})$, for any $K_{\cdot j} \in \mathcal{K}_{\{0,1\}^{n-1}}$.

Lemma 5 (Sharp identification set under Assumptions 1, 3)**.** (i) Under Assumption 1, $\Theta^{\star\star} \supseteq \Theta^{\star}$. (ii) Under Assumptions 1 and 3, $\Theta^{\star\star} = \Theta^{\star}$.

4.7 The sharp identified set under Assumptions 1, 2, 3

This section investigates whether, by introducing Assumption 2, $\Theta^{\star}$ can be characterised by fewer inequalities than in (4.5).

Arguments are articulated as follows. Exploiting *anonymity* of the data generating process and Assumption 2 allows to show that Artstein's inequalities for $G_{\cdot j}$ are identical to Artstein's inequalities for $G_{\cdot h}$, by applying all possible permutations of labels φ such that $\varphi(j) = h$, for any $h \neq j \in \mathcal{N}_n$. Thus, if Assumption 3 is additionally imposed, it follows that, when characterising $\Theta^{\star}$, it is sufficient to consider Artstein's inequalities for $G_{\cdot j}$ for a $j \in \mathcal{N}_n$, instead of $\forall j \in \mathcal{N}_n$[23]. Moreover, via the same strategy, it is proved that Artstein's inequalities for $G_{\cdot j}$, involving elements of $\mathcal{K}_{\{0,1\}^{n-1}}$ and realisations of $\boldsymbol{X}$ equivalent up to a permutation of labels other than label j, are identical.

Formal arguments For any $j \in \mathcal{N}_n$, consider the set

$$\Theta^{\star\star}_{\cdot j} := \Big\{\theta \in \Theta | \mathbb{P}(G_{\cdot j} \in K_{\cdot j}|\boldsymbol{X} = \boldsymbol{x}) \le T_{\mathcal{S}_{\theta_u,\cdot j}(\boldsymbol{X},\epsilon_{\cdot j})|\boldsymbol{X}=\boldsymbol{x}}(K_{\cdot j}) \ \forall K_{\cdot j} \in \mathcal{K}_{\{0,1\}^{n-1}} \text{ and}$$
$$\mathbb{P}(\boldsymbol{G} \in K|\boldsymbol{X} = \boldsymbol{x}) \le T_{\mathcal{S}_{\theta_u}(\boldsymbol{X},\epsilon)|\boldsymbol{X}=\boldsymbol{x}}(K) \ \forall K \in \mathcal{K}^{\diamond}_{\mathcal{G}_n}, \ \forall \boldsymbol{x} \in \mathcal{X}_n \text{ a.s.}\Big\} \tag{4.6}$$

By construction, $\Theta^{\star\star} \subseteq \Theta^{\star\star}_{\cdot j} \ \forall j \in \mathcal{N}_n$. Moreover, it is possible to show that

[23] By doing so, one can save additional $(2^{2^{n-1}} - 1)^n - 2^{2^{n-1}} + 1$ inequalities $\forall \theta \in \Theta$ and $\forall \boldsymbol{x} \in \mathcal{X}_n$.

Lemma 6 (Sharp identification set under Assumptions 1, 2, 3)**.** (i) Under Assumptions 1 and 2, $\Theta^{\star\star}_{\cdot 1} = \Theta^{\star\star}_{\cdot 2} = \ ... \ = \Theta^{\star\star}_{\cdot n} = \Theta^{\star\star}$. (ii) Under Assumptions 1, 2 and 3, $\Theta^{\star\star}_{\cdot 1} = \Theta^{\star\star}_{\cdot 2} = \ ... \ = \Theta^{\star\star}_{\cdot n} = \Theta^{\star\star} = \Theta^{\star}$. (iii) Under Assumptions 1 and 2, $\forall$ permutation of labels φ such that $\varphi(j) = j$,

$$\mathbb{P}(G_{\cdot j} \in K_{\cdot j} | \boldsymbol{X} = \boldsymbol{x}) \leq T_{\mathcal{S}_{\theta_u, \cdot j}(\boldsymbol{X}, \epsilon_{\cdot j}) | \boldsymbol{X} = \boldsymbol{x}}(K_{\cdot j})$$

is equivalent to

$$\mathbb{P}(G_{\cdot j} \in K^{\varphi}_{\cdot j} | \boldsymbol{X} = \boldsymbol{x}^{\boldsymbol{\varphi}}) \leq T_{\mathcal{S}_{\theta_u, \cdot j}(\boldsymbol{X}, \epsilon_{\cdot j}) | \boldsymbol{X} = \boldsymbol{x}^{\boldsymbol{\varphi}}}(K^{\varphi}_{\cdot j})$$

$\forall K_{\cdot j} \in \mathcal{K}_{\{0,1\}^{n-1}}$, $\forall \boldsymbol{x} \in \mathcal{X}_n$ a.s. and $\forall j \in \mathcal{N}_n$, allowing to drop some redundant inequalities from (4.6).

4.8 An outer set

Sections 4.6 and 4.7 show that, as assumptions on unobservables become more restrictive, $\Theta^{\star}$ can be characterised by fewer inequalities. Even so, conducting inference on $\Theta^{\star}$ remains prohibitively complex when there are four or more players. For example, with four players and imposing Assumption 3, one would need to verify $2^{4096} - 4,228,249,610$ inequalities $\forall \theta \in \Theta$ and $\forall \boldsymbol{x} \in \mathcal{X}_n$, which is a number still greater than the quantity of atoms in the universe.

An alternative strategy to attenuate these computational difficulties is considering a subset of inequalities. In the language of BMM (2011), a subset of inequalities defines an outer set, i.e., a set containing $\Theta^{\star}$. This section proposes an outer set, Θ^{o}, collecting the parameter values such that the empirical probability of each realisation of $G_{\cdot j}$ is between the probability of such a realisation being the unique equilibrium of the *section j game* and the probability of such a realisation being a possible equilibrium of the *section j game*, conditional on $\boldsymbol{X}$, $\forall j \in \mathcal{N}_N$.

Formal arguments Consider Artstein's inequalities for $G_{\cdot j}$ involving the compact sets $\{g_{\cdot j}\}$ and $\{0,1\}^{n-1} \setminus \{g_{\cdot j}\}$ $\forall g_{\cdot j} \in \{0,1\}^{n-1}$ and $\forall j \in \mathcal{N}_n$. Let

$$\Theta^{o} := \Big\{ \theta \in \Theta | \mathbb{P}(G_{\cdot j} \in \{0,1\}^{n-1} \setminus \{g_{\cdot j}\} | \boldsymbol{X} = \boldsymbol{x}) \leq T_{\mathcal{S}_{\theta_u, \cdot j}(\boldsymbol{X}, \epsilon_{\cdot j}) | \boldsymbol{X} = \boldsymbol{x}}(\{0,1\}^{n-1} \setminus \{g_{\cdot j}\})$$
$$\mathbb{P}(G_{\cdot j} \in \{g_{\cdot j}\} | \boldsymbol{X} = \boldsymbol{x}) \leq T_{\mathcal{S}_{\theta_u, \cdot j}(\boldsymbol{X}, \epsilon_{\cdot j}) | \boldsymbol{X} = \boldsymbol{x}}\{g_{\cdot j}\}) \ \forall g_{\cdot j} \in \{0,1\}^{n-1}, \forall j \in \mathcal{N}_n, \forall \boldsymbol{x} \in \mathcal{X}_n \text{ a.s.} \Big\} \tag{4.7}$$

It can be observed that, for any $g_{\cdot j} \in \{0,1\}^{n-1}$ and for any $j \in \mathcal{N}_n$,

$$\mathbb{P}(G_{\cdot j} \in \{0,1\}^{n-1} \setminus \{g_{\cdot j}\} | \boldsymbol{X} = \boldsymbol{x}) \leq T_{\mathcal{S}_{\theta_u, \cdot j}(\boldsymbol{X}, \epsilon_{\cdot j}) | \boldsymbol{X} = \boldsymbol{x}}(\{0,1\}^{n-1} \setminus \{g_{\cdot j}\})$$

is equivalent to

$$\mathbb{P}(G_{\cdot j} = g_{\cdot j} | \boldsymbol{X} = \boldsymbol{x}) \geq \int_{e_{\cdot j} \in \mathbb{R}^{(n-1)} \text{ s.t. } \mathcal{S}_{\theta_u, \cdot j}(\boldsymbol{x}, e_{\cdot j}) = \{g_{\cdot j}\}} d\tilde{F}_{\cdot j, \boldsymbol{x}}(e_{\cdot j}; \theta_{\epsilon}) \tag{4.8}$$

and

$$\mathbb{P}(G_{\cdot j} \in \{g_{\cdot j}\}|\boldsymbol{X} = \boldsymbol{x}) \leq T_{\mathcal{S}_{\theta_u,\cdot j}(\boldsymbol{X},\epsilon_{\cdot j})|\boldsymbol{X}=\boldsymbol{x}}(\{g_{\cdot j}\})$$

is equivalent to

$$\mathbb{P}(G_{\cdot j} = g_{\cdot j}|\boldsymbol{X} = \boldsymbol{x}) \leq \int_{e_{\cdot j}\in\mathbb{R}^{(n-1)} \text{ s.t. } g_{\cdot j}\in\mathcal{S}_{\theta_u,\cdot j}(\boldsymbol{x},e_{\cdot j})} d\tilde{F}_{\cdot j,\boldsymbol{x}}(e_{\cdot j};\theta_\epsilon) \tag{4.9}$$

where the first multi-dimensional integral is the probability that $g_{\cdot j}$ is the unique PSNE of the *section j game* and the second multi-dimensional integral is the probability that $g_{\cdot j}$ is a PSNE of the *section j game.*

Moreover, under Assumption 1, $\Theta^o \supseteq \Theta^\star$. In fact, $\Theta_o \supseteq \Theta^{\star\star}$ by construction, and $\Theta^{\star\star} \supseteq \Theta^\star$ by Lemma 5.

Computational gains Conducting inference on Θ^o requires verifying $2n \times 2^{n-1}$ inequalities, $\forall\theta \in \Theta$ and $\forall\boldsymbol{x} \in \mathcal{X}_n$. In turn, this involves the computation of the multi-dimensional integrals (4.8) and (4.9) which can be done via the simple frequency simulator proposed by McFadden (1989) and Pakes and Pollard (1989). In principle, one would need to draw several values of unobservables and verify whether each of all possible 2^{n-1} realisations of $G_{\cdot j}$ is a PSNE of the *section j game* for every drawn value and $\forall j \in \mathcal{N}_n$, generating a demanding routine when n is not small. However, exploiting some properties of the set of PSNE of the *section j game* for any $j \in \mathcal{N}_n$ speeds up the whole process.

Specifically, when during the inference procedure a candidate parameter value is such that $v(\cdot;\delta)$ is monotone increasing, Tarski's fixed point theorem guarantees existence of a greatest and lowest fixed points. These two fixed points can be quickly obtained by implementing the algorithm in Jia (2008). It follows that one only has to check whether each realisation of $G_{\cdot j}$ lying between the greatest and lowest fixed points is a PSNE of the *section j game.*

Instead, when during the inference procedure a candidate a parameter value is such that $v(\cdot;\delta)$ is monotone decreasing, the result in Berry (1992) claiming that every equilibrium in an entry game with negative competitive effects is characterised by the same number of firms entering the market[24] can be applied to the *section j game* to show that all PSNE of the *section j game* feature the same number, n^*_j, of players linking to player j[25]. n^*_j can be quickly obtained by implementing the constructive algorithm used to prove existence of a PSNE of the *section j game* when $v(\cdot;\delta)$ is monotone decreasing[26]. Thus, one only has to check whether each realisation

[24]See *Result* at p.894 in Berry (1992).

[25]See the end of the proof of Lemma 3 in Appendix C.

[26]See the proof of Lemma 3 in Appendix C.

of $G_{\cdot j}$ characterised by n_j^* players linking to player j is a PSNE of the *section j game*, for a total of $\frac{(n-1)!}{n_j^*!(n-1-n_j^*)!} < 2^{n-1}$ realisations.

Overall, Monte Carlo experiments reveal that conducting inference on Θ^o is computationally feasible with relatively limited computational resources when the number of players is equal to or smaller than 20.

In addition, by introducing Assumption 2, statements (i) and (iii) of Lemma 6 imply that some inequalities in (4.7) are redundant and, hence, can be deleted, allowing further computational gains. Specifically, for any $j \in \mathcal{N}_n$, let

$$\begin{aligned}\Theta^o_{\cdot j} := \Big\{\theta \in \Theta | \mathbb{P}(G_{\cdot j} \in \{0,1\}^{n-1} \setminus \{g_{\cdot j}\} | \boldsymbol{X} = \boldsymbol{x}) \leq T_{\mathcal{S}_{\theta_u,\cdot j}(\boldsymbol{X},\epsilon_{\cdot j})|\boldsymbol{X}=\boldsymbol{x}}(\{0,1\}^{n-1} \setminus \{g_{\cdot j}\}) \\ \mathbb{P}(G_{\cdot j} \in \{g_{\cdot j}\} | \boldsymbol{X} = \boldsymbol{x}) \leq T_{\mathcal{S}_{\theta_u,\cdot j}(\boldsymbol{X},\epsilon_{\cdot j})|\boldsymbol{X}=\boldsymbol{x}}\{g_{\cdot j}\}) \; \forall g_{\cdot j} \in \{0,1\}^{n-1}, \forall \boldsymbol{x} \in \mathcal{X}_n \text{ a.s.}\Big\}\end{aligned} \tag{4.10}$$

By construction, $\Theta^o \subseteq \Theta^o_{\cdot j}$ $\forall j \in \mathcal{N}_n$. Moreover,

Implications of Lemma 6 Under Assumptions 1 and 2, (i) $\Theta^o_{\cdot 1} = \Theta^o_{\cdot 2} = ... = \Theta^o_{\cdot n} = \Theta^o$. (ii) $\forall$ permutation of labels φ such that $\varphi(j) = j$,

$$\begin{cases}\mathbb{P}(G_{\cdot j} \in \{0,1\}^{n-1} \setminus \{g_{\cdot j}\} | \boldsymbol{X} = \boldsymbol{x}) \leq T_{\mathcal{S}_{\theta_u,\cdot j}(\boldsymbol{X},\epsilon_{\cdot j})|\boldsymbol{X}=\boldsymbol{x}}(\{0,1\}^{n-1} \setminus \{g_{\cdot j}\}) \\ \mathbb{P}(G_{\cdot j} \in \{g_{\cdot j}\} | \boldsymbol{X} = \boldsymbol{x}) \leq T_{\mathcal{S}_{\theta_u,\cdot j}(\boldsymbol{X},\epsilon_{\cdot j})|\boldsymbol{X}=\boldsymbol{x}}\{g_{\cdot j}\})\end{cases}$$

is equivalent to

$$\begin{cases}\mathbb{P}(G_{\cdot j} \in \{0,1\}^{n-1} \setminus \{g^{\varphi}_{\cdot j}\} | \boldsymbol{X} = \boldsymbol{x}^{\boldsymbol{\varphi}}) \leq T_{\mathcal{S}_{\theta_u,\cdot j}(\boldsymbol{X},\epsilon_{\cdot j})|\boldsymbol{X}=\boldsymbol{x}^{\boldsymbol{\varphi}}}(\{0,1\}^{n-1} \setminus \{g^{\varphi}_{\cdot j}\}) \\ \mathbb{P}(G_{\cdot j} \in \{g^{\varphi}_{\cdot j}\} | \boldsymbol{X} = \boldsymbol{x}^{\boldsymbol{\varphi}}) \leq T_{\mathcal{S}_{\theta_u,\cdot j}(\boldsymbol{X},\epsilon_{\cdot j})|\boldsymbol{X}=\boldsymbol{x}^{\boldsymbol{\varphi}}}\{g^{\varphi}_{\cdot j}\})\end{cases}$$

$\forall g_{\cdot j} \in \{0,1\}^{n-1}$, $\forall \boldsymbol{x} \in \mathcal{X}_n$ a.s. and $\forall j \in \mathcal{N}_n$.

Comparison with other outer sets in the literature Tamer (2003) and CT (2009) illustrate a static entry game with complete information and construct an outer set collecting the parameter values such that the empirical probability of each realisation of endogenous variables is between the probability of such a realisation being the unique equilibrium of the entry game and the probability of such a realisation being a possible equilibrium of the entry game, conditional on players' observed characteristics. Thus, one can characterise an outer set, Θ^o_{CT}, collecting the parameter values such that the empirical probability of each realisation of $\boldsymbol{G}$ is between the probability of such a realisation being the unique equilibrium of the network formation game and the probability of such a realisation being a possible equilibrium of the network formation game, conditional on $\boldsymbol{X}$.

More formally, consider Artstein's inequalities for $\boldsymbol{G}$ involving $K := \{\boldsymbol{g}\}$ and $K := \mathcal{G}_n \setminus \{\boldsymbol{g}\}$ $\forall \boldsymbol{g} \in \mathcal{G}_n$. Let

$$\Theta^o_{CT} := \Big\{\theta \in \Theta | \mathbb{P}(\boldsymbol{G} \in \mathcal{G}_n \setminus \{\boldsymbol{g}\} | \boldsymbol{X} = \boldsymbol{x}) \leq T_{\mathcal{S}_{\theta_u}(\boldsymbol{X},\epsilon)|\boldsymbol{X}=\boldsymbol{x}}(\mathcal{G}_n \setminus \{\boldsymbol{g}\}) \\ \mathbb{P}(\boldsymbol{G} \in \{\boldsymbol{g}\} | \boldsymbol{X} = \boldsymbol{x}) \leq T_{\mathcal{S}_{\theta_u}(\boldsymbol{X},\epsilon)|\boldsymbol{X}=\boldsymbol{x}}\{\boldsymbol{g}\}) \ \forall \boldsymbol{g} \in \mathcal{G}_n, \ \forall \boldsymbol{x} \in \mathcal{X}_n \text{ a.s.}\Big\} \tag{4.11}$$

As for Θ^o, the two Artstein's inequalities above can be rewritten by using the probability that $\boldsymbol{g}$ is the unique PSNE of the network formation game and the probability that $\boldsymbol{g}$ is a PSNE of the network formation game. Moreover, by (4.4), $\Theta^o_{CT} \supseteq \Theta^\star$.

Computational gains brought by Θ^o_{CT} may be insufficient because conducting inference on Θ^o_{CT} requires checking $2 \times 2^{n(n-1)}$ inequalities $\forall \theta \in \Theta$ and $\forall \boldsymbol{x} \in \mathcal{X}_n$. For example, with 15 players, as in the data used for the empirical application, one would need to verify 3.291×10^{63} inequalities $\forall \theta \in \Theta$ and $\forall \boldsymbol{x} \in \mathcal{X}_n$. Instead, Θ^o brings greater computational gains by requiring verification of $2n \times 2^{n-1}$ inequalities, $\forall \theta \in \Theta$ and $\forall \boldsymbol{x} \in \mathcal{X}_n$, with savings, with respect to Θ^o_{CT}, of a factor depending on $\frac{1}{n}2^{(n-1)^2}$. In terms of width, under Assumption 1 $\Theta^o \supseteq \Theta^o_{CT}$. This follows from statement (i) of Lemma 5 combined with the fact that $\{\boldsymbol{g}\} \in \mathcal{A}_{\mathcal{G}_n}$ for any $\boldsymbol{g} \in \mathcal{G}_n$. Moreover, under Assumptions 1 and 3, $\Theta^o = \Theta^o_{CT}$. This follows from statement (ii) of Lemma 5 combined with the fact that $\{\boldsymbol{g}\} \in \mathcal{A}_{\mathcal{G}_n}$ for any $\boldsymbol{g} \in \mathcal{G}_n$.

Alternatively, Sheng (2014) designs an outer set collecting the parameter values such that the empirical probability of each realisation of a subnetwork is between the probability of such a realisation being the unique equilibrium of the *subnetwork game* and the probability of such a realisation being a possible equilibrium of the *subnetwork game*, conditional on $\boldsymbol{X}$, for every subnetwork of size equal to or smaller than α, with $2 \leq \alpha \leq n$. α is set by the researcher according to the available computational resources. Let Θ^o_S be the outer set obtained by applying Sheng's strategy to this setting. It can be noticed that if $\alpha = n$ then one goes back to Θ^o_{CT} because there exists only one subnetwork of size n, that is the whole network. If $\alpha < n$, then $\Theta^o \subseteq \Theta^o_S$ under Assumptions 1 and 3. Indeed, $\Theta^o_S \supseteq \Theta^o_{CT}$ under Assumption 1[27] and $\Theta^o = \Theta^o_{CT}$ under Assumptions 1 and 3. From a computational point of view, when $\alpha << n$, conducting inference on Θ^o_S could be computationally easier than conducting inference on Θ^o but set estimates or confidence regions might not be sufficiently tight, because important interdependencies across players' decisions may be ignored.

[27] This follows from the fact that if a network is a PSNE, then each of its subnetworks of size equal to or smaller than α is a PSNE, but the converse is not necessarily true.

4.9 Extensions to the bilateral case

The formation of link ij is unilateral when it requires the consent of player i only, as in the case of advice networks, trust networks, and, sometimes, friendship networks. The formation of link ij is bilateral when it requires the consent of both players i and j, as in the case of board interlocks. For simplicity of exposition, econometric results have been illustrated for the unilateral case. This section discusses how to obtain analogous results for the bilateral case.

4.9.1 The bilateral network formation game

In the bilateral network formation game, players, endowed with some preferences for links that are shared knowledge, simultaneously announce desired incoming and outgoing links according to certain payoffs and equilibrium concept, and only mutually announced links are formed.

Players' preferences for links Players' preferences for links depend on characteristics that are partially observed by the researcher. For any $i, j \in \mathcal{N}_N$ with $i \neq j$, let X_i and ϵ^i_{ij} denote, respectively, a $K \times 1$ vector of observed (to the researcher) characteristics of player i, and the unobserved (to the researcher) heterogeneity of player i affecting its preference for link ij. Let $\boldsymbol{X}$ be an $N \times K$ matrix collecting X_i $\forall i \in \mathcal{N}_N$, and ϵ^i be a $2(N-1) \times 1$ vector collecting $(\epsilon^i_{ij}, \epsilon^i_{ji})$ $\forall j \neq i \in \mathcal{N}_N$. Let $\epsilon_{\cdot j}$ be a $2(N-1) \times 1$ vector collecting ϵ_{ij} $\forall i \neq j$. Lastly, let ϵ be a $2N(N-1) \times 1$ vector collecting ϵ^i $\forall i \in \mathcal{N}_N$.

Players' choices For any $i \in \mathcal{N}_N$, let s^i_{ij} be a scalar equal to 1 if player i is willing to form link ij and 0 otherwise. Let s^i be a $2(N-1) \times 1$ vector collecting s^i_{ij} and s^i_{ji} $\forall j \neq i \in \mathcal{N}_N$. Lastly, let s be a $2N(N-1) \times 1$ vector collecting s^i $\forall i \in \mathcal{N}_N$.

In the bilateral network formation game, a pure strategy vector of player i is $s^i \in \{0,1\}^{2(N-1)}$ $\forall i \in \mathcal{N}_N$ and a pure strategy profile of the game is $s \in \{0,1\}^{2N(N-1)}$. Mutual consent is needed to form links, i.e., $G_{ij} = s^i_{ij} s^j_{ij}$ $\forall i, j \in \mathcal{N}_N, i \neq j$ but severing a link can be done unilaterally.

Players' payoffs Each player $i \in \mathcal{N}_N$ gets a payoff

$$\begin{aligned} U_i(\boldsymbol{G}, \boldsymbol{X}, \epsilon^i; \theta_u) := &\sum_{j=1}^{N} G_{ij} \times \Big[z(X_i, X_j; \beta) + v(\sum_{k \neq i}^{N} G_{kj}; \delta) + \epsilon^i_{ij} \Big] \\ &+ \sum_{j=1}^{N} G_{ji} \times \Big[b(X_i, X_j; \gamma) + \epsilon^i_{ji} \Big] \end{aligned} \tag{4.12}$$

where $z(\cdot;\beta)$ and $b(\cdot;\gamma)$ are any functions of X_i and X_j known up to vectors of parameters β and γ, $v(\cdot;\delta)$ is any function monotone in $\sum_{k\neq i}^{N} G_{kj}$ and known up to a vector of parameters δ, $\theta_u := (\beta, \gamma, \delta) \in \Theta_u \subseteq \mathbb{R}^{d_\beta+d_\gamma+d_\delta}$, with d_β, d_γ and d_δ denoting the dimensions of β, γ and δ.

Differently from the unilateral case, the utility function (4.12) includes also the payoff that player i gets from her incoming connections. Similarly to the unilateral case, the utility function (4.12) is additively separable over player i's incoming and outgoing links.

Equilibrium Agents play pure strategy pairwise Nash equilibrium (hereafter PSPNE). The resulting network is a pure strategy pairwise Nash stable (hereafter PSPNS) network (Jackson and Wolinsky, 1996; Calvó-Armengol, 2004; Bloch and Jackson, 2006; Goyal and Joshi, 2006; Calvó-Armengol and Ilkiliç, 2009).

Definitions are now given. Let $\boldsymbol{G}$'s dependence on s be denoted by $\boldsymbol{G}(s)$. When $G_{ij}(s) = 0$, let $\boldsymbol{G}(s) + ij$ denote the matrix $\boldsymbol{G}(s)$ with $G_{ij} = 1$ added. Moreover, let s^{-i} be the vector s without s^i and $s = (s^i, s^{-i})$.

Definition 3 (PSPNS network)**.** A pure strategy profile s is a PSPNE of the bilateral network formation game if

1. It is robust to unilateral multi-link deletion, i.e.,

$$U_i(\boldsymbol{G}(s), \boldsymbol{X}, \epsilon^i; \theta_u) \geq U_i(\boldsymbol{G}(s^i_\diamond, s^{-i}), \boldsymbol{X}, \epsilon^i; \theta_u)$$

   $\forall s^i_\diamond \neq s^i \in \{0,1\}^{2(N-1)}$ and $\forall i \in \mathcal{N}_N$.

2. It is robust to bilateral one-link addition, i.e., there does not exist a pair of players $(i, j) \in \mathcal{N}_N$ such that

$$U_i(\boldsymbol{G}(s) + ij, \boldsymbol{X}, \epsilon^i; \theta_u) \geq U_i(\boldsymbol{G}(s), \boldsymbol{X}, \epsilon^i; \theta_u)$$

   and

$$U_j(\boldsymbol{G}(s) + ij, \boldsymbol{X}, \epsilon^j; \theta_u) \geq U_j(\boldsymbol{G}(s), \boldsymbol{X}, \epsilon^j; \theta_u)$$

   with strict inequality for at least one of the two players.

$\boldsymbol{G}$ is a PSPNS network if there exists a PSPNE s of the bilateral network formation game such that $\boldsymbol{G} = \boldsymbol{G}(s)$.

Alternative equilibrium concepts adopted in bilateral network formation games are pairwise stability (Jackson and Wolinsky, 1996) and Nash stability (Myerson, 1991). Pairwise stable networks are robust to unilateral one-link deletion and bilateral one-link formation. Pairwise stability is an equilibrium notion independent of any network formation procedure and has nice computational properties. However, it only considers very simple deviations and, hence, it may be too tolerant in classifying a network as stable, especially when there are few players. On the other hand, Nash stable networks are constructed by letting players announce desired outgoing and incoming links, according to the pure strategy Nash equilibrium, and, then, forming mutually beneficial links. Using the pure strategy Nash equilibrium in a bilateral game induces coordination problems because link creation requires the consent of the two involved parties. This causes the game displaying a multiplicity of Nash stable networks, always including the empty network, as playing zero is weakly optimal even when forming a link would be profitable to both players. In order to solve this issue, PSPNE allows players to coordinate their decisions and, by not leaving aside any reciprocally beneficial link, it refines the set of stable networks. In particular, the set of PSPNS networks is the intersection of the set of Nash stable networks and the set of pairwise stable networks. Lastly, within this model, the set of PSPNS networks and the set of pairwise stable networks coincide, by the additively separability of the utility function (4.12) over player i's incoming and outgoing links (Gilles and Sarangi, 2005).

Finally, by exploiting the additively separability of the utility function (4.12) over player i's incoming and outgoing links, Lemma 7 maintains that the inequalities in Definition 3 simplify to a system of $N(N-1)$ equations whose solution is a PSPNS network.

Lemma 7 (Characterisation of a PSPNS network)**.** $\boldsymbol{G}$ is a PSPNS network if and only if

$$G_{ij} = \mathbb{1}\{z(X_i, X_j; \beta) + v(\sum_{k \neq i}^{N} G_{kj}; \delta) + \epsilon_{ij}^{i} \geq 0\}\mathbb{1}\{b(X_i, X_j; \gamma) + \epsilon_{ij}^{j} \geq 0\} \quad \forall i, j \in \mathcal{N}_N, i \neq j$$

4.9.2 Identification results

As for the unilateral case, the bilateral network formation game has at least one PSPNS network. Equilibrium existence can be show following the steps illustrated for the unilateral case in Section 4, after having adapted the constructive proof in Berry (1992) to a bilateral game setting. Details are in Appendix A.

As for the unilateral case, the bilateral network formation game admits multiple equilibria for some values of payoff-relevant variables and parameters. Moreover, by running simulations, it can be seen that the equilibrium set may contain outcomes with a diametrically opposite economic meaning, as the empty network and the fully connected network. Identification results when the *selection mechanism* is left unrestricted are identical to those illustrated for the unilateral case in Section 4 after having appropriately replaced equilibrium concepts.

5 Empirical application

5.1 Preliminaries

In this section the proposed methodology is illustrated in action by conducting inference on the outer set Θ^o using data on horizontal board interlocks between Italian firms. Specifically, it is assumed that firms play the bilateral network formation game illustrated in Section 4.9 to form horizontal board interlocks with the purpose of monitoring and advising each other. As mentioned in Section 2, the assumption that firms, and not directors, are the board interlock decision-makers is based on the inter-organisational linkage perspective in the corporate governance literature (Palmer, 1983; Ornstein, 1984; Zajac, 1988). This modelling choice is legitimated by the Italian legal framework, which states that a director needs her board' approval to join the board of a competing firm (Article 2390 of the Italian Civil Code). Similar provisions are laid down in most European countries.

The empirical application focuses on primary horizontal board interlocks. A primary horizontal board interlock arises when two firms share a director who holds an executive position at at least one of two companies involved (Stokman, Van Der Knoop and Wasseur, 1988). Indeed, primary horizontal board interlocks, by involving individuals appointed with executive duties, are more likely to represent the long-term economic and institutional relations between firms (Mizruchi and Bunting, 1981; Stokman and Wasseur, 1985; Stokman, Van Der Knoop and Wasseur, 1988).

Firms' choices on primary horizontal board interlocks are represented as a directed network of size N where link ij exists if firm j appoints as director at least one of firm i's executive board members, for any $i, j \in \mathcal{N}_N$ with $i \neq j$. As an example, Figure 3 reports a directed network of size 4, which represents the primary horizontal board interlocks in an industry composed by four firms. For each node i, firm i's board composition is indicated by two sets of letters. Each letter represents an individual. The first set of letters is the set of executive directors. The second set of letters is the set of non-executive members.

Firms' payoffs are modelled according to (4.12). Following the corporate governance literature on board interlocks, when $G_{ij} = 1$ firm i can monitor and advise firm j. Hence, $\left[z(X_i, X_j; \beta) + v(\sum_{k\neq i}^{N} G_{kj}; \delta) + \epsilon_{ij}^{i}\right]$ is the payoff that firm i gets from monitoring and advising firm j. Conversely, when $G_{ji} = 1$ firm i is monitored and advised by firm j. Thus, $\left[b(X_i, X_j; \gamma) + \epsilon_{ji}^{j}\right]$ is the payoff that firm i gets from being monitored and advised by firm j. Payoffs from incoming and outgoing links are asymmetric: when $G_{ij} = 1$ and $G_{ji} = 0$, it is assumed that, while firm i can monitor firm j's decision-making process, the converse is not true, because although firm i's executive director may fulfil a number of functions that are useful to firm j, his primary employment and loyalty is normally with and to firm i (Palmer, 1983; Richardson, 1987). Similarly, advice flows exclusively from firm i to firm j (Koenig and Goegel, 1981; Mizruchi and Stearn, 1994). These asymmetries justify links' directions.

As explained in Section 1, the presence of the spillover effect SE is suggested by the co-optation theory in corporate governance (Selznick, 1949; Thompson and McEwen, 1958; Pfeffer and Salancik, 1978; Palmer, 1983; Mizruchi, 1996), according to which board interlocks reflect attempts by organisations to co-opt (monitor, anticipate, restrain) sources of environmental uncertainty stemming from the potentially disruptive unilateral actions of other corporations. For example, horizontal board interlocks may help companies to strengthen possibly collusive behaviours by monitoring each other's decisions. One may then conclude that the higher the number of competitors with an executive director sitting on j's board, the stronger their capacity to influence j's decisions and align these with the group's interests (positive spillover effect SE). On the other hand, the micro-economic theory suggests that, if horizontal board interlocks are formed by firms to enforce a cartel, then the higher the number of competitors with an executive director sitting on j's board, the lower the gains when pre-empting deviations by j from the collusive agreement (negative spillover effect SE). Hence, including the spillover effect SE in firms' payoffs (without restricting its sign), allows to understand which view is supported by the empirical evidence, when firms' heterogeneity is also considered.

The additively separability over firm i's incoming and outgoing links mimicks the additive separability over multi-market entry decisions in the empirical literature on entry games (Bresnahan and Reiss, 1991; Berry, 1992; Jia, 2008; CT, 2009; Bajari, Hong and Ryan, 2010; Aradillas-Lopez and Rosen, 2014; Fox and Lazzati, 2016).

Lastly, principal-agent issues are not considered, i.e., it is assumed that executive directors' actions are always in line with their company's will.

In the following, a 95% confidence region for the outer set Θ^o is constructed by applying the method of AS (2010) outlined in Appendix B[28] and using using data on primary horizontal board interlocks between Italian joint stock companies belonging to a cross-section of industries in 2010.

[28]Given that the number of moment inequalities is larger than the number of industries in the data, inference under alternative methods, such as Bugni, Caner, Kock, and Lahiri (2016) and Chernozhukov, Chetverikov, and Kato (2016), is currently in progress.

5.2 The Italian context

Board interlocks in Italy Italy represents an ideal environment to conduct the empirical application. Indeed, board interlocks have been an important feature of Italian capitalism since the end of the nineteenth century (Ciocca, 2007). For example, Luzzatto Fegiz (1928), using data for 1923, highlights the importance of the exchange of executives as a way to create relationships between companies to reduce competition. Rinaldi and Vasta (2005) illustrate how board interlocks playes a crucial role in guaranteeing the stability of the system after the second World War, when, in order to reach a balanced coexistence between state-owned enterprises and private firms, they are used to consolidate and defend controlling positions in the main Italian private companies, together with pyramidal groups, cross participations, non-voting shares and statutory regulations aimed at discouraging takeovers. Bianco and Pagnoni (1997), Barbi (2000) and Bianchi, et al. (2005) explicitly identify the widespread recurse to board interlocks with possible consequences in terms of collusion and restriction to competition as a characteristic of the Italian industrial system since the 1950s. In a comparative study, Santella, et al (2009) highlight that Italy, similarly to France and Germany and differently from the UK, is characterised by a high number of companies linked to each other through shared directors. Finally, many takeover attempts on the Italian Stock Exchange cannot be fully understood if the presence of board interlocks among listed firms is neglected (Bertoni and Randone, 2006).

The evolution in the structure of board interlocks between Italian firms is analysed by several authors. For example, Vasta and Baccini (1997) find that the structure of board interlocks between Italian firms across 1911, 1927 and 1936 is overall stable. Rinaldi and Vasta (2005; 2012) highlight that, while around 1960 Italian corporate networks are highly cohesive, between 1972 and 1983 a sharp decrease in their cohesion degree is registered. Focusing on more recent years, Santella, Drago and Polo (2009) observe that the structure of board interlocks among Italian companies is remarkably stable between 1998 and 2006 with very few exits or entries. Similar conclusions are reached by Bellenzier and Grassi (2014) for the period 1998-2011. Stability is also revealed by the small variation observed in some network measures (averaged over industries) during 2005-2010 for the analysed data (Table 3). Overall this legitimates the complete information assumption.

The Italian regulatory framework on horizontal board interlocks The Italian regulatory framework on horizontal board interlocks fits with the model's assumptions. In fact, while horizontal board interlocks that meet certain jurisdictional thresholds are illegal in U.S. under the Clayton Act of 1914 and subsequent ancillary legislation, in Italy there is no clear and general prohibition on horizontal board interlocks, i.e., a breaking of competition law or of regulation on conflict of interests has to be proved and cannot be presumed. The only exception is Law No. 214 of 2011, which forbids board interlocks with companies or groups operating in the banking, insurance and financial services sectors. This law does not affect the present analysis as it entered into force in 2011.

The population of joint stock companies in Italy The empirical application considers a sample of Italian joint stock companies (*Societá per Azioni*, *Societá in Accomandita per Azioni*) because these represent the largest Italian firms on which the anti trust authority's attention is focused. For clarification, joint stock companies are business entities where shareholders' liability is limited to the nominal value of held shares. A joint stock company is not necessarily listed. Figure 4 reports the number of Italian joint stock companies and the number of Italian limited companies (also including joint stock companies) during 2007-2012. Figure 5 shows the percentage of Italian join-stock companies per macro-sector according to the Istat[29] census of 2011.

5.3 Data

Sources of data The sources of data are the Registro Imprese and the Cerved databases whose access has been provided by the Bank of Italy.

The Registro Imprese is a database in which all Italian companies are required to enrol through the Chamber of Commerce in their province and is the primary source of certification of their constituent data. It offers detailed and updated information on individual firms (e.g., legal status, year of registration, composition of corporate bodies, geographical location, principle line of activity) and on important changes related to their existence (e.g., termination, liquidation, bankruptcy, mergers and acquisitions).

The Cerved database contains information useful for measuring the credit risk of Italian limited companies and, among other information, provides their balance sheet details.

Sample of Italian joint stock companies Italian legislation offers several ways to structure the governance of a joint stock company. The analysed sample is composed of the Italian joint stock companies in 2010 governed under the Articles 2380/2409-septies of the Italian Civil Code,

[29]Istat is the Italian National Institute for Statistics.

which outline a legal framework in line with the model's assumptions. More details are in Appendix D.8.

Including in the analysed sample Italian joint stock companies with governance structures different from those outlined above is a delicate task, currently in progress, that requires careful investigation of the power relations among different corporate bodies.

Information collected for each firm in the sample The composition of firms' boards together with the role of each director (simple director or *Amministratore Delegato*) is extracted from the Registro Imprese database.

Industries are constructing considering firms' principal lines of activity provided by the ATECO 2002 code and extracted from the Registro Imprese database. The ATECO 2002 code is similar to the SIC code in the UK and U.S.[30]. It is an alpha-numeric code with varying degrees of detail - the letters indicate the macroeconomic sector while the numbers represent sub-sectors. It is developed in five levels: sections (letter), subsections (two letters, optional), divisions (2 digits), groups (3 digits), classes (4 digits) and categories (5 digits)[31,32].

In line with the empirical literature on board interlocks (among others Dooley, 1969; Pfeffer, 1972; Allen, 1974; Bunting, 1976; Pennings, 1980; Carrington, 1981), firms' dimensions and profitabilities are set as exogenous variables influencing firms' decisions to form primary horizontal board interlocks.

[30]Other empirical studies on board interlocks identifying industries using the SIC code include Burt (1978), Burt, Christman and Kilburn (1980), Pennings (1980), Carrington (1981), Zajac (1988) and Mizruchi and Stearns (1994).

[31]For example:

- A: Agriculture, hunting and fishing
- 01: Agriculture, hunting and related service activities
- 01.1: Crops
- 01.11: Growing of cereals and other arable crops
- 01.11.1: Growing of cereals (rice included)
- 01.11.2: Growing of oil seeds
- ...

[32]In 2008 the ATECO 2002 code has been replaced by the ATECO 2007 code. However, this paper exploits the information on firms' principal line of activity provided by the ATECO 2002 code as data quality is remarkably higher for the year 2010 on which the present analysis is focused. Lastly, the structure of the ATECO 2007 maintains the same general characteristics of the structure of the ATECO 2002 code.

As per Dooley (1969), Allen (1974) and Mizruchi and Stearns (1988; 1994), a firm's dimension is measured using total assets (hereafter TA)[33], extracted from the Cerved database. As per Baysinger and Butler (1985) and Fligstein and Brantley (1992), a firm's profitability is measured using return on equity (hereafter ROE)[34], extracted from the Cerved database. Lastly, in order to apply the inference method proposed by AS (2010) as discussed in Appendix B, TA and ROE are discretised into ten separate bins, according to their 10, 20, ..., 90th quantiles. Consequently, TA and ROE take values in $\{1, 2, ..., 10\}$. Additional data cleaning steps are covered in Appendix D.9.

Descriptive statistics for the sample Some descriptive statistics for industry size, TA and ROE are in Table 1. The total number of firms is 2599, the total number of industries is 386.

Some descriptive statistics for the most relevant network measures are in Table 2. The definitions of the considered network measures are in Appendix D.10. Overall, the constructed networks look disconnected with several isolated nodes.

5.4 A linear model specification

Identification results discussed in Section 4 hold for any form of the functions $z(\cdot;\beta)$, $b(\cdot;\gamma)$ and $v(\cdot;\delta)$ entering firms' payoffs provided that the function $v(\cdot;\delta)$ is monotone, and for any form of the probability distribution of unobservables. However, constructing a confidence region for the outer set Θ^o requires their specification.

As a first step, linear forms are assigned to to the functions $z(\cdot;\beta)$, $b(\cdot;\gamma)$ and $v(\cdot;\delta)$ and independence across unobservables is imposed. Specifically,

$$\begin{aligned} G_{ij} = \mathbb{1}\{\beta_0 + \beta_1(TA_j - TA_i) + &\beta_2(ROE_j - ROE_i) + \delta \sum_{k \neq i} G_{kj} + \epsilon^i_{ij} \geq 0\} \times \\ &\mathbb{1}\{\gamma_0 + \gamma_1(TA_j - TA_i) + \gamma_2(ROE_j - ROE_i) + \epsilon^j_{ij} \geq 0\} \end{aligned} \tag{5.1}$$

[33] Pfeffer (1972) measures a firm's size using total sales. Booth and Deli (1996) propose the natural log of the sum of the market value of the firm equity plus the book value of preferred stock. An interesting discussion on how to measure firm's size is in Shalit and Sankar (1977).

[34] Alternative measures of a firm profitability include: price-cost margins (Collins and Preston, 1969; Carrington, 1981); market value, price-earnings ratio and debt-equity ratio (Fligstein and Brantley, 1992); return on sales (Mizruchi and Stearns, 1988; Fligstein and Brantley, 1992); return on assets (Richardson, 1987; Mizruchi and Stearns, 1988; Fligstein and Brantley, 1992); return on shareholders' investment and net interest on assets (Bernstein, 1978; Pennings, 1980; Richardson, 1987); dividend cuts (Kaplan and Reishus, 1990); return on invested capital (Bunting, 1976); average Tobin's q (Hermalin and Weisbach, 1991).

$\forall i, j \in \mathcal{N}, i \neq j$. $\{\epsilon_{ij}\}_{i,j\in\mathcal{N}, i\neq j}$ are assumed i.i.d. across ij with ϵ_{ij} distributed as $N(0_2, I_2)$, where 0_2 is a 2×1 zero vector and I_2 is the 2×2 identity matrix. According to this specification, the spillover effect SE defined in Section 1 corresponds to the parameter δ.

Results Table 4 reports the hypercube that contains the 95% confidence region for each parameter value in Θ^o. Consider first the sign of various effects as measured by projections of this hypercube. The projection for the parameter δ, corresponding to the spillover effect SE defined in Section 1, is [2.129, 20.909]. Results on SE have a positive sign, i.e., all else equal, firm i's payoff from having an executive director sitting on rival j's board increases with the number of additional competing firms having an executive director sitting on j's board. In view of the co-optation theory in corporate governance, this outcome seems to support the idea according to which the higher the number of competitors with an executive director sitting on j's board, the stronger their capacity to influence j's decisions and align them with the group's interests.

The projections for the parameters β_1 and β_2 are, respectively, $[0.022,\ 8.381]$ and $[0.012,\ 7.486]$ and indicate that, all else equal, firm i prefers its executives sitting on firm j's board when j is larger and more profitable than i. Again, in view of the co-optation theory, it may be that firms prefer their executives sitting on the board of larger and more profitable competitors because these represent a major source of uncertainty which should be monitored. Moreover, this result is in line with other empirical studies in the literature reporting a positive relation between a firm's size and profitability, and the intensity of board interlocking (e.g., Dooley, 1969; Allen, 1974; Bunting; 1976).

Conversely, the projections for the parameters γ_1 and γ_2 are, respectively, $[-4.327,\ -0.004]$ and $[-9.655,\ -0.016]$ and indicate that, all else equal, firm j prefers appointing as directors executives of firm i when firm j is smaller and less profitable than i. Indeed, it may be that smaller and less profitable firms are not considered capable of offering valuable advice.

Firms' payoff functions do not have unit of measure and, therefore, there is no direct interpretation for coefficients' magnitudes. One idea to discuss magnitudes is considering the ratio between the change induced by a given unit increase in one variable relative to the change induced by a one unit increase in a reference variable. Let the reference variable be $(TA_j - TA_i)$. Results reveal that, within each industry, all else equal, the positive effect on firm i's payoff from link ij (u^i_{ij} for simplicity) of a one-bin increase in $(ROE_j - ROE_i)$ is between 0.007 and 22.127 times the positive effect on u^i_{ij} of a one-bin increase in $(TA_j - TA_i)$. Moreover, within each industry, all else equal, the positive effect on u^i_{ij} of a one unit increase in $\sum_{k\neq i} G_{kj}$ is between 0.255 and roughly 480 times the positive effect of u_{ij}^i a one-bin increase in $(TA_j - TA_i)$. Lastly, within each industry, all else equal, the negative effect on firm j's payoff from link ij (u_{ij}^j for simplicity) of a one-bin increase in $(ROE_j - ROE_i)$ is between 0.041 and 323.851 times the negative effect on u_{ij}^j of a one-bin increase in $(TA_j - TA_i)$.

5.5 An alternative model specification

Firms' preferences for heterogeneity in size and profitability may be non-linear. In order to capture some non-linearities, one can consider an alternative model specification assigning non-linear forms to the functions $z(\cdot;\beta)$ and $b(\cdot;\gamma)$ and see whether results are still informative and in line with those obtained under (5.1). In particular, a model specification where the terms (TA_j-TA_i) and (ROE_j-ROE_i) in (5.1) are replaced by several indicator functions is considered

$$\begin{aligned}G_{ij} =&\mathbb{1}\{\beta_0 + \beta_1\mathbb{1}\{-9 \le TA_j - TA_i \le -5\} + \beta_2\mathbb{1}\{-4 \le TA_j - TA_i \le 0\} + \beta_3\mathbb{1}\{1 \le TA_j - TA_i \le 5\}+\\ &+ \beta_4\mathbb{1}\{-9 \le ROE_j - ROE_i \le -5\} + \beta_5\mathbb{1}\{-4 \le ROE_j - ROE_i \le 0\} + \beta_6\mathbb{1}\{1 \le ROE_j - ROE_i \le 5\}+\\ &+ \delta\sum_{k\neq i} G_{kj} + \epsilon_{ij}^i \ge 0\}\times\\ &\mathbb{1}\{\gamma_0 + \gamma_1\mathbb{1}\{-9 \le TA_j - TA_i \le -5\} + \gamma_2\mathbb{1}\{-4 \le TA_j - TA_i \le 0\} + \gamma_3\mathbb{1}\{1 \le TA_j - TA_i \le 5\}+\\ &+ \gamma_4\mathbb{1}\{-9 \le ROE_j - ROE_i \le -5\} + \gamma_5\mathbb{1}\{-4 \le ROE_j - ROE_i \le 0\} + \gamma_6\mathbb{1}\{1 \le ROE_j - ROE_i \le 5\}+\\ &+ \epsilon_{ij}^j \ge 0\}\end{aligned} \tag{5.2}$$

$\forall i, j \in \mathcal{N}, i \neq j$. As before, $\{\epsilon_{ij}\}_{i,j\in\mathcal{N},i\neq j}$ are assumed i.i.d. across ij with ϵ_{ij} distributed as $N(0_2, I_2)$ as above.

Results Table 5 reports the hypercube that contains the 95% confidence region for each $\theta \in \Theta^o$. Consider first the sign of various effects as measured by projections of this hypercube. The projection for the parameter δ is [31.523, 35.902] and has a positive sign as for specification (5.1). Regarding indicator functions, the base group is $TA_j - TA_i$ and $ROE_j - ROE_i$ both between 6 and 9. It represents the case in which firm j is significantly larger and more profitable than firm i. Table 6 reports confidence intervals for sum of pairs of parameters via projections relative to other combinations of realisations of $TA_j - TA_i$ and $ROE_j - ROE_i$. Overall, the base group is always favoured by firm i, i.e., firm i prefers its executives sitting on firm j's board when firm j is significantly larger and more profitable than firm i. An exception is represented by the projection for $\beta_2 + \beta_6$ that includes both positive and negative values. This means that the corresponding indicator functions may have a positive or a negative effect on payoffs. Conversely, the base group is never favoured by firm j, i.e., firm j prefers its executives sitting on firm j's board when firm j is not significantly larger and more profitable than firm i. An exception is when $TA_j - TA_i$ is between -9 and -5 and $ROE_j - ROE_i$ is between -4 and 0, which is less favoured by firm j than the base group, possibly because firm j sees itself excessively vulnerable and exposed in front of firm i. Moreover, the projections for $\gamma_1 + \gamma_4$ and $\gamma_2 + \gamma_5$ include both positive and negative values. Hence, it can be concluded that results on signs under (5.2) confirm and enrich those obtained under (5.1) revealing important non-linearities in firms preferences for heterogeneity in size and profitability.

Regarding magnitudes, ratios between the change induced by a given unit increase in one variable relative to the change induced by a one unit increase in a reference variable may be less meaningful than for specification (5.1) because regressors are binary. Moreover, reporting marginal effects on the probability distribution of $\boldsymbol{G}$ of changes in exogenous variables does not help as $\boldsymbol{G}$ can take too many values. An alternative option is finding how bounds on density, average degree, percentage of isolated nodes and number of links vary as a consequence of changes in exogenous variables. Various experiments are possible. As an example, Table 7 reports the outcome of the following procedure: for each industry and value of parameters in the 95% confidence region, the discretised amount of total assets of the smallest firms is equalised to the discretised amount of total assets of the biggest firms, hence reducing size heterogeneity within industries; several realisations of unobserved exogenous variables are drawn; for each drawn realisation, PSPNS networks are determined; the density, the average degree, the percentage of isolated nodes and the total number of links in each PSPNS network are computed, and their minimum and maximum value across PSPNS networks are recorded; bounds are averaged across drawn realisations and industriess; finally, the smallest lower bound and largest upper bound across values of parameters are reported in the second and third columns of Table 7. The same experiment is repeated keeping the observed values of total assets within each industry and results are reported in the fourth and sixth columns of Table 7. Lastly, observed empirical values are in the fifth column of Table 7. As a consequence of the simulated shift, the upper bound on density, average degree and number of links increases. The lower bound on the percentage of isolated nodes decreases. Hence, by reducing heterogeneity in firms' size, networks can have on average less isolated nodes but can become more disconnected.

Lastly, checking the robustness of the positiveness of the spillover effect SE by introducing correlations across unobservables is left to future research.

6 Conclusions

This paper provides a framework for studying identification in a network formation model. Network formation is modelled as a static game with complete information and pure strategy equilibrium. Links have directions. Payoffs depend on some players' characteristics partially observed by the researcher and on the spillover effect SE. The proposed methodology relies on partial identification arguments because the network formation game admits multiple equilibria, and equilibrium selection is left unrestricted to avoid inappropriate assumptions that could bias estimates. The designed methodology is illustrated in action using data on primary horizontal board interlocks between Italian firms.

There are some avenues of future research. Specifically, there could be other relevant spillover effects affecting players' payoffs. For example, player i's payoff from forming a link with player j may also depend on the number of common connections shared by i and j. In this spirit, it may be worth enriching players' payoffs by introducing additional externalities and investigating whether the identification results proposed here can be extended to such more complicated settings. Another direction could be to examine how the identification analysis changes if we allow for the possibility of missing data within each network. For example, one may wonder how much larger bounds on parameters would be if data were obtained by randomly drawing some players from the original set of players and considering only the links between them. Indeed, at the expense of reducing the informativeness of the final results, such a strategy would attenuate further the computational burden of inference by reducing the size of networks. In other cases, part of the information on links or players could be missing, but not at random, and it would be interesting to understand how to adjust bounds on parameters accordingly. Lastly, with respect to the board interlock application, it would be interesting to refine the model by including directors' compensation and moral hazard issues - for example, following Gayle, Golan and Miller (2015) - as these aspects may be important for a better understanding of firms' incentives behind the formation of board interlocks.

Table 1: Descriptive statistics for the firms' exogenous variables.

	Mean	St.dev	Min	Max	[0.25; 0.50; 0.75]	Skewness	Kurtosis	Number of firms	Number of industries
N	[6.733]	[3.483]	3	15	[4; 6; 9]	0.812	2.645	2599	386
TA ($\times 10^6$ €)	117.281	1,567.453	0.067	73,916.239	[6.998; 15.653; 39.984]	41.471	1,903.552	✓	✓
ROE (%)	1.266	24.589	−128.410	69.820	[−2.382; 2.360; 11.402]	−1.600	9.071	✓	✓

Table 2: Descriptive statistics for some network measures. Definitions are in Appendix D.10.

	Mean	St.dev	Min	Max	[0.25; 0.50; 0.75] quantiles	Skewness	Kurtosis
Density	0.005	0.026	0	0.333	[0; 0; 0]	8.462	88.322
Average degree	[0.023]	[0.096]	0	1	[0; 0; 0]	5.905	45.181
% Isolated nodes	97.666	8.758	33.333	100	[100; 100; 100]	−4.299	22.587
Number of links	[0.163]	[0.617]	0	6	[0; 0; 0]	4.859	32.750

Table 3: Some network measures averaged over industries across the years 2005-2010. Definitions are in Appendix D.10.

Year	Density	Average degree	% Isolated nodes	Number of links
2005	0.006	[0.025]	97.604	[0.167]
2006	0.007	[0.023]	97.64	[0.135]
2007	0.009	[0.028]	96.974	[0.159]
2008	0.006	[0.028]	97.027	[0.208]
2009	0.005	[0.027]	97.666	[0.208]
2010	0.005	[0.023]	97.666	[0.163]

Table 4: Projections of the 95% confidence region for each $\theta \in \Theta^o$ according to specification (5.1).

β_0	$[-15.399, \ -0.783]$
β_1	$[0.022, \ 8.381]$
β_2	$[0.012, \ 7.486]$
δ	$[2.129, \ 20.909]$
γ_0	$[-0.469, \ 37.490]$
γ_1	$[-4.327, \ -0.004]$
γ_2	$[-9.655, \ -0.016]$

Table 5: Projections of the 95% confidence region for each $\theta \in \Theta^o$ according to specification (5.2).

β_0	$[-7.120, \ -3.431]$
β_1	$[-2.006 \times 10^3, \ -1.998 \times 10^3]$
β_2	$[6.977, \ 12.562]$
β_3	$[0.202, \ 2.629]$
β_4	$[-23.251, \ -15.473]$
β_5	$[-25.678, \ -22.059]$
β_6	$[-14.954, \ -9.652]$
δ	$[31.523, \ 35.902]$
γ_0	$[0.845, \ 2.679]$
γ_1	$[-12.762, \ -7.030]$
γ_2	$[-7.958, \ -4.965]$
γ_3	$[-0.723, \ 1.785]$
γ_4	$[7.290, \ 11.360]$
γ_5	$[4.584, \ 7.291]$
γ_6	$[13.156, \ 16.360]$

Table 6: Projections of sums for interpreting signs according to specification (5.2).

$\beta_1 + \beta_4$	$[-2.029 \times 10^3, \ -2.013 \times 10^3]$
$\beta_1 + \beta_5$	$[-2.029 \times 10^3, \ -2.022 \times 10^3]$
$\beta_1 + \beta_6$	$[-2.017 \times 10^3, \ -2.010 \times 10^3]$
$\beta_2 + \beta_4$	$[-14.348, \ -5.527]$
$\beta_2 + \beta_5$	$[-17.541, \ -10.957]$
$\beta_2 + \beta_6$	$[-6.401, \ 1.372]$
$\beta_3 + \beta_4$	$[-22.973, \ -14.692]$
$\beta_3 + \beta_5$	$[-25.097, \ -20.983]$
$\beta_3 + \beta_6$	$[-14.112, \ -7.810]$
$\gamma_1 + \gamma_4$	$[-3.776, \ 1.016]$
$\gamma_1 + \gamma_5$	$[-6.954, \ -2.445]$
$\gamma_1 + \gamma_6$	$[2.224, \ 8.544]$
$\gamma_2 + \gamma_4$	$[1.266, \ 5.295]$
$\gamma_2 + \gamma_5$	$[-1.449, \ 0.111]$
$\gamma_2 + \gamma_6$	$[7.057, \ 10.608]$
$\gamma_3 + \gamma_4$	$[7.096, \ 13.145]$
$\gamma_3 + \gamma_5$	$[4.403, \ 7.670]$
$\gamma_3 + \gamma_6$	$[13.162, \ 17.150]$

Table 7: Bounds of some network measures according to specification (5.2) when the following experiment is run: within each industry, the amount of total assets of the smallest firms is equalised to the amount of total assets of the biggest firms.

	New lower bound	New upper bound	Old lower bound	Empirical average	Old upper bound
Density	0	0.800	0	0.005	0.886
Average Degree	0	[4.640]	0	[0.022]	[5.153]
% Isolated nodes	0.063	100	0.139	97.666	100
Number of links	0	[41.096]	0	[0.163]	[45.808]

Figure 1: Example of a simple and directed network of size 3, with $\boldsymbol{G} = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$.

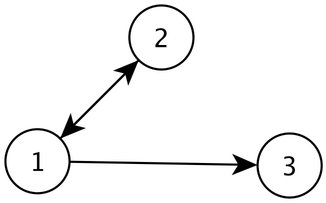


Figure 2: The *section 2* of the network in Figure 1, with $G_{\cdot 2} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$.

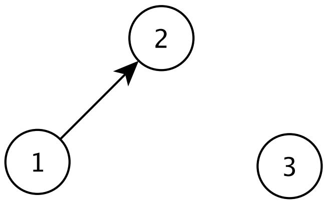
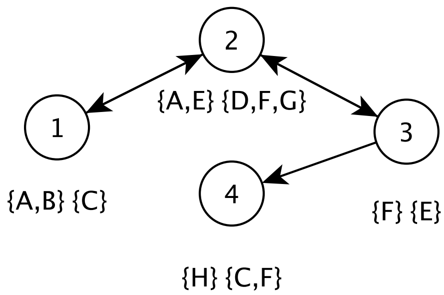


Figure 3: Example of a directed network of size 4 representing the primary horizontal board interlocks in an industry composed by four firms. For each node $i \in \{1, 2, 3, 4\}$, firm i's board composition is indicated by two set of letters. Each letter represents an individual. The first set of letters is the set of executive directors. The second set of letters is the set of non-executive members.



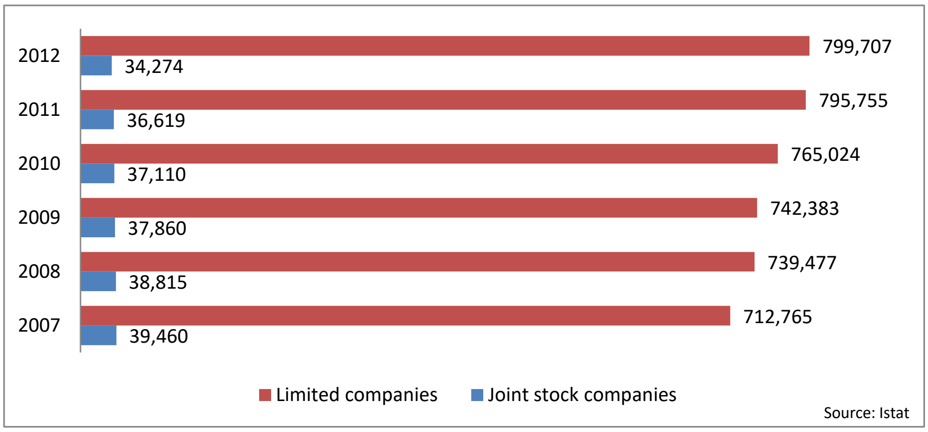


Figure 4: Number of Italian joint stock companies and number of Italian limited companies (also including joint stock companies) during 2007-2012.

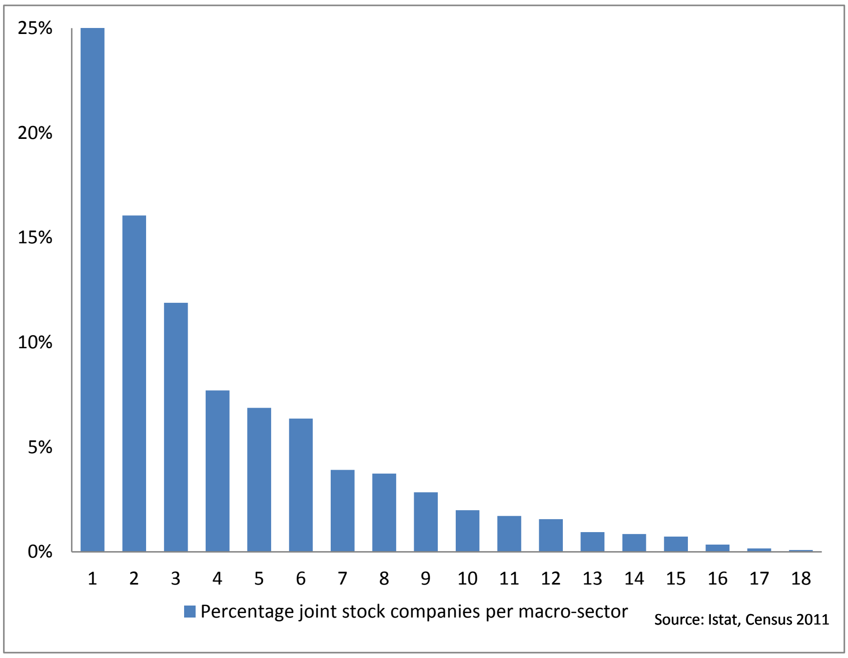


Figure 5: Percentage of Italian join-stock companies per macro-sector according to the Istat census of 2011, where: 1 denotes Manufacturing (in particular of Machinery, Metals and Textiles); 2 denotes Wholesale and Retail trade, and Repair of Motor Vehicles and Motorcycles (in particular Wholesale of Households Goods and Sale of Motor Vehicles); 3 denotes Real Estate Activities; 4 denotes Financial and Insurance Activities; 5 denotes Construction; 6 denotes Professional, Scientific and Technical Activities; 7 denotes Transportation and Storage; 8 denotes Information and Communication; 9 denotes Administrative and Support Service Activities; 10 denotes Water Supply: Sewerage, Waste Management and Remediation Activities; 11 denotes Accommodation and Food Services Activities; 12 denotes Electricity, Gas, Steam and Air Conditioning Supply; 13 denotes Human Health and Social Work Activities; 14 denotes Arts, Entertainment and Recreation; 15 denotes Other Service Activities; 16 denotes Mining, Quarrying; 17 denotes Education; 18 denotes Agricolture, Forestry, Fishing.

References

Aldrich, H.E. (1979): *Organizations and Environments.* Englewood Cliffs, NJ: Prentice-Hall.

Allen, M.P. (1974): "The Structure of Interorganizational Elite Cooptation: Interlocking Corporate Directorates", *American Sociological Review*, 39(3), 393–406.

Armstrong, T.B. (2014): "Weighted KS Statistics for Inference on Conditional Moment Inequalities", *Journal of Econometrics*, 181(2), 92–116.

Andrews, D.W.K., and G. Soares (2010): "Inference for Parameters Defined by Moment Inequalities Using Generalized Moment Selection", *Econometrica*, 78(1), 119–157.

Aradillas-Lopez, A., and A.M. Rosen (2014): "Inference in Ordered Response Games with Complete Information", CeMMAP Working Paper 36/14.

Andrews, D.W.K., and X. Shi (2013): "Inference Based on Conditional Moment Inequalities", *Econometrica*, 81(2), 609–666.

Badev, A. (2014): "Discrete Games with Endogenous Networks: Theory and Policy", University of Pennsylvania Working Paper.

Bajari, P., H. Hong, and S.P. Ryan (2010): "Identification and Estimation of a Discrete Game of Complete Information", *Econometrica*, 78(5), 1529–1568.

Bandiera, O., and I. Rasul (2006): "Social Networks and Technology Adoption in Northern Mozambique", 116(514), *The Economic Journal*, pp.869–902.

Banerjee, A., A.G. Chandrasekhar, and M.O. Jackson (2013): "The Diffusion of Microfinance", *Science*, 341(6144), 1236498 1–7.

Barabási, A., and R.Albert (1999): "Emergence of Scaling in Random Networks", *Science*, 286(5439), 509–512.

Di Bartolomeo, G, and P. Canofari (2015): "Interlocking Directorates and Concentration in the Italian Insurance Market", *Journal of Industry, Competition and Trade*, 15(4), 351–362.

Bayer, P., S. Ross, and G. Topa (2008): "Informal Hiring Networks and Labor Market Outcomes", *Journal of Political Economy*, 116(6), 1150–1196.

Baysinger, B.D., and H.N. Butler (1985): "Corporate Governance and the Board of Directors: Performance Effects of Changes in Board Composition", *Journal of Law, Economics, & Organization*, 1(1), 101–124.

Bearden J. (1986): "Social Capital and Corporate Control: The Singer Company", *Journal of Political and Military Sociology*, 14, 127–148.

Bellenzier, L., and R. Grassi (2014): "Interlocking Directorates in Italy: Persistent Links in Network Dynamics", *Journal of Economic Interaction and Coordination*, 9(2), 183–202.

Beresteanu, A., I. Molchanov, and F. Molinari (2011): "Sharp Identification in Models with Convex Moment Predictions", *Econometrica*, 79(6), 1785–1821.

Beresteanu, A., and F. Molinari (2008): "Asymptotic Properties for a Class of Partially Identified Models", *Econometrica*, 76(4), 763–814.

Bernstein, L.A. (1978): *Financial Statement Analysis.* Georgetown, Ontario: Irwin-Dorsey.

Berry, S. (1992): “Estimation of a Model of Entry in the Airline Industry”, *Econometrica*, 60(4), 889–917.

Berry, S., and E. Tamer (2006): “Identification in Models of Oligopoly Entry”, in *Advances in Economics and Econometrics: Theory and Application: Ninth World Congress of the Econometric Society*, ed. by R. Blundell, W.K. Newey, and T. Persson, vol.2, 46–58, Cambridge: Cambridge University Press.

Bertoni, F., and P. Randone (2006): “The Small-World of Italian Finance: Ownership Interconnections and Board Interlocks amongst Italian Listed Companies”, SSRN Working Paper.

Bianchi, M., M. Bianco, S. Giacomelli, A. M. Pacces, and S. Trento (2005): *Proprietá e Controllo delle Imprese in Italia.* Bologna, Il Mulino.

Bianco, M., and E. Pagnoni (1997): “Board interlocks Across Listed Companies in Italy: The Case of Banks”, *BNL Quarterly Review*, Special Issue on Property, Control and Corporate Governance of Banks.

Bjorn P., and Vuong Q. (1984): “Simultaneous Equations Models for Dummy Endogenous Variables: a Game Theoretic Formulation with an Application to Labor Force Participation. Caltech Work. Pap. 537, Calif. Inst. Technol, Pasadena

Bloch, F., and M. Jackson (2006): “Definitions of Equilibrium in Network Formation Games", *International Journal of Game Theory*, 34(3), 305–318.

Bollobás, B. (2001): *Random Graphs.* Cambridge University Press, Cambridge

Booth, J.R., and D.N. Deli (1996): “Factors affecting the number of outside directorships held by CEOs”, *Journal of Financial Economics*, 40(1), 81–104.

Boucher, V. (2016): “The Estimation of Network Formation Games with Positive Spillovers", Working Paper, Laval University.

Boucher, V., and I. Mourifié (2015): “My Friend Far Far Away: Asymptotic Properties of Pairwise Stable Networks", SSRN Working Paper.

Bresnahan, T.F., and P.C. Reiss (1991): “Entry and Competition in Concentrated Markets”, *Journal of Political Economy*, 99(5), 977–1009.

Bugni, F. (2010): “Bootstrap Inference for Partially Identified Models Defined by Moment Inequalities: Coverage of the Identified Set", *Econometrica*, 78(2), 735–753.

Bugni, F., M. Caner, A.B. Kock, and S. Lahiri (2016): “Inference in Partially Identified Models with Many Moment Inequalities Using Lasso”, Working Paper.

Bugni, F., I. Canay, and X. Shi (2015): “Inference for Functions of Partially Identified Parameters in Moment Inequality Models", CeMMAP Working Paper 54/15.

Bunting, D. (1976): “Corporate interlocking. Part III. Interlocks and return on investment”, *Directors & Boards*, 1(3), 4–11.

Burt, R.S. (1978): “A Structural Theory of Interlocking Corporate Directorates”, *Social Networks*, 1(4),415–435.

Burt, R.S. (1983): *Corporate Profits and Cooptation.* New York: Academic.

Burt, R.S., Christmas K.P., and H.C. Kilburn (1980): “Testing a Structural Theory of Corporate Cooptation: Interorganizational Directorate Ties as a Strategy for Avoiding Market Constraints on Profits”, *American Sociological Review*, 45(5), 821–841.

Caiazza, R., and M. Simoni (2013): "board interlocks' Effects on Economic System's Competitiveness, *Business Strategy Series*,14(1), 30–35.

Calvó-Armengol, A. (2004): "Job Contact Networks", *Journal of Economic Theory*, 115(1), 191–206.

Calvó-Armengol, A., and R. Ilkiliç (2009): "Pairwise-Stability and Nash Equilibria in Network Formation", *International Journal of Game Theory*, 38(1), 51–79.

Calvó-Armengol, A., E. Patacchini, and Y. Zenou (2009): "Peer Effects and Social Networks in Education", *Review of Economic Studies*, 76(4), 1239–1267.

Canay, I. (2010): "EL Inference for Partially Identified Models: Large Deviations Optimality and Bootstrap Validity", *Journal of Econometrics*, 156(2), 408–425.

Carbonai, D., and G. Di Bartolomeo (2006): "board interlocks as a Trust Substitute: The Case of the Italian Non-Life Insurance Industry", University Library of Munich MPRA Paper 4420.

Carrington, P.J. (1981): *Horizontal Co-optation Through Corporate Interlocks.* PhD thesis, Dept. Sociol. Univ. Toronto.

Chernozhukov, V., H. Hong, and E. Tamer (2007): "Estimation and Confidence Regions for Parameter Sets in Econometric Models", *Econometrica*, 75(5), 1243–1284.

Chernozhukov, V., S. Lee, and A. M. Rosen (2013): "Intersection Bounds: Estimation and Inference", *Econometrica*, 81(2), 667–737.

Chernozhukov, V., D. Chetverikov, and K. Kato (2016): "Testing Many Moment Inequalities", CeMMAP Working Paper 42/16.

Chesher, A., and A.M. Rosen (2012): "Simultaneous Equations Models for Discrete Outcomes, Coherence, Completeness, and Identification", CeMMAP Working Paper 21/12.

Chetverikov, D. (2012): "Adaptive Test of Conditional Moment Inequalities", CeMMAP Working Paper 36/12.

Christakis, N., J. Fowler, G. Imbens, and K. Kalianaraman (2010): "An Empirical Model of Strategic Network Formation", Harvard University Working Paper.

Ciliberto, F., and E. Tamer (2009): "Market Structure and Multiple Equilibria in Airline Markets", *Econometrica*, 77(6), 1791–1828.

Ciocca, P. (2007): *Ricchi per Sempre? Una Storia Economica d'Italia (1796-2005).* Torino, Bollati-Boringhieri.

Clauset, A., C.R. Shalizi, and M. Newman (2009): "Power-Law Distributions in Empirical Data", 51(4), 661–703.

Collins, N.R., and L.E. Preston (1969): "Price-Cost Margins and Industry Structure", *Review of Economics and Statistics*, 51(3), 271–286.

Conley, T.G., and C.R. Udry (2010): "Learning about a New Technology: Pineapple in Ghana", *American Economic Review*, 100(1), 35–69.

Currarini, S., M. Jackson, and P. Pin (2009): "An Economic Model of Friendship: Homophily, Minorities and Segregation," Econometrica, 77(4), 1003–1045.

de Paula, A. (2013): "Econometric Analysis of Games with Multiple Equilibria", *Annual Review of Economics*, 5(1), 107–131.

de Paula, A. (2015): "Econometrics of Network Models", CeMMAP Working Paper 52/15.

de Paula, A., S. Richards-Shubik, and E. Tamer (2015): "Identification of Preferences in Network Formation Games", CeMMAP Working Paper 29/15.

Dooley, P.C. (1969): "The Interlocking Directorate", *American Economic Review*, 59(3), 314–323.

Eisenhardt, K.M (1989): "Agency Theory: An assesssment and Review", *The Academy of Management Review*, 14(1), 57–74.

Epstein, L.G., H. Kaido, and K. Seo (2015): "Robust confidence regions for incomplete models", CeMMAP Working Paper 20/15.

Erdös, P., and A. Rényi (1959): "On Random Graphs I", *Publicationes Mathematicae*, 6, 290–297.

Erdös, P., and A. Rényi (1960): "On the Evolution of Random Graphs', *Publications of the Mathematical Institute of the Hungarian Academy of Sciences*, 5, 17–61.

Fligstein, N., and P. Brantley (1992): "Bank Control, Owner Control, or Organizational Dynamics: Who Controls the Large Modern Corporation?", *American Journal of Sociology*, 98(2), 280–307.

Fox, J.T., and N. Lazzati (2016): "Identification of Discrete Choice Models for Bundles and Binary Games", Working Paper.

Frank, O., and D. Strauss (1986): "Markov Graphs", *Journal of the American Statistical Association*, 81(395), 832–842.

Gabrielsen, T.S., E. Hjelmeng, and L. Sørgard (2011): "Rethinking Minority Share Ownership and Interlocking Directorships: The Scope for Competition Law Intervention", *European Law Review*, 6, 837–860.

Gayle, G.-L., L. Golan, and R.A. Miller (2015): "Interlocked Executives and Insider Board Members: An Empirical Analysis", Federal Reserve Bank of St. Louis Working Paper 2015-040A.

Galichon, A., and M. Henry (2011): "Set Identification in Models with Multiple Equilibria", *Review of Economic Studies*, 78(4), 1264–1298.

Gilbert, E.N (1959): "Random Graphs", *Annals of Mathematical Statistics*, 30(4), 1141–1144.

Gilles, R.P., and S. Sarangi (2005): "Stable Networks and Convex Payoffs", Louisiana State University Working Paper.

Goyal, S., and S. Joshi (2006): "Unequal links", *International Journal of Game Theory*, 34(3), 319–349.

Graham, B.S. (2008): "Identifying Social Interactions through Conditional Variance Restrictions", *Econometrica*, 76(3), 643–666.

Graham, B. (2015): "Methods of Identification in Social Networks", *Annual Review of Economics*, 7(1), 465–485.

Helmers, C., M. Patnam, and P.R. Rau (2015): "Do Board Interlocks Increase Innovation? Evidence from a Corporate Governance Reform in India", SSRN Working Paper.

Hermalin, B.E., and M.S. Weisbach (1991): "The Effects of Board Composition and Direct Incentives on Firm Performance", *Financial Management*, 20(4), 101–112.

Ho, K., and A. M. Rosen (2015): "Partial Identification in Applied Research: Benefits and Challenges", CeMMAP working paper CWP64/15.

Hong, H., and J. Xu (2014): "Count Models of Social Networks in Finance", Princeton University Working Paper.

Jackson, M. (2009): *Social and Economic Networks.* Princeton University Press.

Jackson, M., and B. Rogers (2007): “Meeting Strangers and Friends of Friends: How Random Are Social Networks?" *American Economic Review*, 97(3), 890–915.

Jackson, M.O. and J. Wolinsky (1996): “A Strategic Model of Social and Economic Networks”, *Journal of Economic Theory*, 71(1), 44–74.

Jia, P. (2008): “What Happens When Wal-Mart Comes to Town: an Empirical Analysis of the Discount Retail Industry”, *Econometrica*, 76(6), 1263–1316.

Kaido, H., F. Molinari, and J. Stoye (2015): “Inference on Projections of Identified Sets”, Working Paper.

Kaplan, S.N., and D. Reishus (1990): “Outside directorships and corporate performance”, 27(2), 389–410.

Kleinber, J.M., S.R. Kumar, P. Raghavan, S. Rajagopalan, and A. Tomkins (1999): “The Web as a Graph: Measurements, Models and Methods”, in *Proceedings of the International Conference on Combinatorics and Computing (Lecture Notes in Computer Science)*, ed. by T. Asano, H. Imai, D. T. Lee, S. Nakano, and T. Tokuyama, Heidelberg: Springer Berlin.

Koenig, T., and R. Gogel (1981): “Interlocking Corporate Directorships as a Social Network”, *American Journal of Economics and Sociology*, 40(1), 37–50.

Kolaczyk, E. (2009): *Statistical Analysis of Network Data.* Springer-Verlag.

Koller, D., and N. Friedman (2009): *Probabilistic Graphical Models: Principles and Techniques.* MIT Press.

Kooreman, P. (1994): “Estimation of Econometric Models of Some Discrete Games”, *Journal of Applied Econometrics*, 9(3): 255–268.

Khwaja, A.I., A. Mian, and A. Qamar (2011): “Bank Credit and Business Networks”, Working Paper Series rwp11-017, Harvard University, John F. Kennedy School of Government.

Lee, S., K. Song, and Y. Whang (2013): “Testing Functional Inequalities", *Journal of Econometrics*, 172(1), 14–32.

Lee, S., K. Song, and Y. Whang (2014): “Testing for a General Class of Functional Inequalities", CeMMAP Working Paper 09/14.

Leslie, C. (2004): “Trust, Distrust, and Antitrust”, *Texas Law Review*, 83(3), 515–680.

Leung, M. (2015a): “A Random-Field Approach to Inference in Large Network Models", Stanford University Working Paper.

Leung, M. (2015b): “Two-Step Estimation of Network Formation Models with Incomplete Information", *Journal of Econometrics*, 188(1), 182–195.

Luzzatto Fegiz, P. (1928): “Il Consiglio d’Amministrazione e l’Interdipendenza delle Imprese”, *Giornale degli Economisti e Rivista di Statistica*, 68(3), 197–231.

Mace, M.L. (1971): *Directors: Myth and Reality.* Boston: Harvard Bus. Sch.

McFadden, D. (1989): “A Method of Simulated Moments for Estimation of Discrete Response Models without Numerical Integration”, *Econometrica*, 57(2), 995–1026.

Meeusen, W., and L. Cuyvers (1985): “The Interaction between Interlocking Directorships and the Economic Behaviour of Companies”, in *Networks of Corporate Power*, ed. by N. Stokman, R. Ziegler, and J. Scott, Cambridge, England: Polity.

Mele, A. (2015): “A Structural Model of Segregation in Social Networks", Johns Hopkins Working Paper.

Menzel, K. (2016): “Strategic Network Formation with Many Agents”, New York University Working Paper.

Miyauchi, Y. (2014): “Structural Estimation of a Pairwise Stable Network with Nonnegative Externality", MIT Working Paper.

Mizruchi, M.S. (1996): “What Do Interlocks Do? An Analysis, Critique, and assesssment of Research on board interlocks”, *Annual Review of Sociology*, 22, 271–298.

Mizruchi, M.S, and D. Bunting (1981): “Influence in Corporate Networks: An Examination of Four Measures”, *Administrative Science Quarterly*, 26(3), 475–489.

Mizruchi, M.S., and L.B. Stearns (1988): “A Longitudinal Study of the Formation of board interlocks”, *Administrative Science Quarterly*, 33(2), 194–210.

Mizruchi, M.S., and L.B. Stearns (1994): “A Longitudinal Study of Borrowing by Large American Corporations”, *Administrative Science Quarterly*, 39(1), 118–140.

Molchanov, I., and F. Molinari (2014): “Applications of Random Set Theory in Econometrics”, *Annual Review of Economics*, 61(1), 229–225.

Myerson, R. (1991): *Game Theory: Analysis of Conflict*. Harvard University Press.

Nakajima, R. (2007): “Measuring Peer Effects on Youth Smoking Behaviour”, *Review of Economic Studies*, 74(3), 897–935.

Newman, M. (2003): “The Structure and Function of Complex Networks”, *SIAM Review*, 45(2), 167–256.

OECD (2008): *Antitrust Issues Involving Minority Shareholding and Interlocking Directorates*, available at `http://www.oecd.org/competition/mergers/41774055.pdf`

Ornstein, M. (1984): “Board interlocks in Canada: Intercorporate or Class Alliance?", *Administrative Science Quarterly*, 29(2), 210–231.

Pakes, A., and D. Pollard (1989): “Simulation and the Asymptotics of Optimization Estimators”, *Econometrica*, 57(5), 1027–1057.

Pakes, A., J. Porter, K. Ho, and J. Ishii (2015): “Moment Inequalities and Their Application", *Econometrica*, 83(1), 315–334.

Palmer, D. (1983): “Broken Ties: board interlocks and Intercorporate Coordination”, *Administrative Science Quarterly*, 28(1), 40–55.

Patnam, M. (2013): “Corporate Networks and Peer Effects in Firm Policies”, Working Paper.

Pennings, J.M. (1980): *board interlocks*. San Francisco, Calif.: Jossey-Bass.

Pennock, D.M., G.W. Flake, S. Lawrence, E.J. Glover, and C.L. Giles (2002): “Winners Don?t Take All: Characterizing the Competition for Links on the Web”, *PNAS*, 99(8), 5207–52011.

Pfeffer J. (1972): “Size and Composition of Corporate Boards of Directors: The Organization and its Environment”, *Administrative Science Quarterly*, 17(2), 218–228.

Pfeffer J., and G.R. Salancik (1978): *The External Control of Organisations: A Resource Dependence Perspective*. New York: Harper & Row.

Price, D. deSolla (1976): “A General Theory of Bibliometric and Other Cumulative Advantage Processes”, *Journal of the American Society for Information Science*, 27(5), 292–306.

Richardson, R.J. (1987): “Directorship Interlocks and Corporate Profitability”, *Administrative Science Quarterly*, 32(3), 367–386.

Rinaldi, A., and M. Vasta (2005): “The Structure of Italian Capitalism, 1952-1972: New Evidence Using the Interlocking Directorates Technique”, *Financial History Review*, 12(2), 173–198.

Rinaldi, A., and M. Vasta (2012): “The Italian Corporate Network After the Golden Age (1972-1983): From Centrality to Marginalization of State-Owned Enterprises”, *Enterprise and Society*, 13(2), 378–413.

Robins, G., T. Snijders, P. Wang, M. Handcock, and P. Pattison (2007): “Recent Developments in Exponential Random Graph (p^*) Models for Social Networks", *Social Networks*, 29(2), 192–215.

Romano, J.P., and A. Shaikh (2008): “Inference for Identifiable Parameters in Partially Identified Econometric Models", *Journal of Planning and Statistical Inference*, 138(9), 2786–2807.

Romano, J.P., and A. M. Shaikh (2010): “Inference for the Identified Set in Partially Identified Econometric Models," *Econometrica*, 78(1), 169–211.

Romano, J., A. Shaikh, and M. Wolf (2014): “A practical two-step method for testing moment inequalities with an application to inference in partially identified models", *Econometrica*, 82(5), 1979–2002.

Rosen, A. M. (2008): “Confidence Sets for Partially Identified Parameters that Satisfy a Finite Number of Moment Inequalities", *Journal of Econometrics*, 146, 107–117.

Sacerdote, B. (2001): “Peer Effects with Random Assignment: Results for Dartmouth Roommates”, Quarterly Journal of Economics, 116(2), 681–704.

Santella, P., C. Drago, A. Polo, and E. Gagliardi (2009): “A Comparison Among the Director Networks in the Main Listed Companies in France, Germany, Italy, and the United Kingdom", MPRA Paper 16397, University Library of Munich, Germany.

Santella, P., C. Drago, and A. Polo (2009): “The Italian Chamber of Lords Sits on Listed Company Boards: An Empirical Analysis of Italian Listed Company Boards from 1998 to 2006", SSRN Working Paper.

Selznick, P. (1949): *TVA and the Grass Roots*. New York: Harper & Row.

Shalit, S.S, and U. Sankar (1977): “The Measurement of Firm Size”, *Review of Economics and Statistics*, 59(3), 290–298.

Sheng, S. (2014): “A Structural Econometric Analysis of Network Formation Games", UCLA Working Paper.

Stiglitz, J.E. (1985): “Credit Markets and the Control of Capital”, *Journal of Money, Credit and Banking*, 17(2), 133–152.

Stokman, F.N, J. Van Der Knoop, and F.W. Wasseur (1988): “Interlocks in the Netherlands: Stability and careers in the period 1960?1980”, *Social Networks*, 10(2), 183–208.

Stokman, F.N, F.W. Wasseur, and D. Elsas (1985): “The Dutch Network. Types of Interlocks and Network Structure”, in *Networks of Corporate Power. A comparative Analysis of Ten Countries*, ed. by R. Ziegler, and J. Scott, Oxford: Polity Press.

Stoye, J. (2009): “More on Confidence Regions for Partially Identified Parameters", *Econometrica*, 77(4), 1299–1315.

Tamer, E. (2003): “Incomplete Simultaneous Discrete Response Model with Multiple Equilibria”, *Review of Economic Studies*, 70(1), 147–167.

Thompson, J.D., and W. McEwen (1958): “Organizational Goals and Environment: Goal-Setting as an Interaction Process”, *American Sociological Review*, 23(1), 23–31.

Vasta, M., and A. Baccini (1997): “Banks and industry in Italy, 1911?36: New Evidence Using the Interlocking Directorates Technique”, *Financial History Review*, 4(2), 139–159.

Vázquez, A. (2003): “Growing Network with Local Rules: Preferential Attachment, Clustering Hierarchy, and Degree Correlations”, *Physical Review E*, 67(5), 056104.

Watts, D.J. (1999): *Small Worlds*. Princeton: Princeton University Press.

Watts, D.J., and S.H. Strogatz (1988): “Collective Dynamics of ?Small World? Networks”, *Nature*, 393:6684, 440–442.

Zajac, E.J. (1988): “board interlocks as an Interorganizational Strategy: A Test of Critical Assumptions", *The Academy of Management Journal*, 31(2), 428–438.

Zheng, T., M.J. Salganik, and A. Gelman (2006): “How Many People Do You Know in Prison?: Using Overdispersion in Count Data to Estimate Social Structure in Networks”, *Journal of the American Statistical Association*, 101(474), 409–423.

A Existence of a PSPNS network in the bilateral case

Existence of a PSPNS network can be shown following the strategy adopted for the unilateral case. Specifically, for any $j \in \mathcal{N}_N$, the local game considered is denominated the *bilateral section j game*. In the *bilateral section j game* players other than player j simultaneously announce whether they want to form a link pointing to j, j replies and only mutually beneficial links are created.

More formally, let $s^j_{\cdot j}$ be an $(N-1)\times 1$ vector collecting s^j_{ij} $\forall i \in \mathcal{N}_{\cdot j,N}$ and let $s_{\cdot j} \in \{0,1\}^{2(N-1)}$ be a $2(N-1)\times 1$ vector collecting s^i_{ij} $\forall i \in \mathcal{N}_{\cdot j,N}$ and $s^j_{\cdot j}$. A pure strategy of player i is $s^i_{ij} \in \{0,1\}$ $\forall i \in \mathcal{N}_{\cdot j,N}$, a pure strategy vector of player j is $s^j_{\cdot j} \in \{0,1\}^{N-1}$, and a pure strategy profile of the game is $s_{\cdot j} \in \{0,1\}^{2(N-1)}$. Mutual consent is needed to form links, i.e., $G_{ij} = s^i_{ij} s^j_{ij}$ $\forall i \in \mathcal{N}_{\cdot j,N}$. Each player $i \in \mathcal{N}_{\cdot j,N}$ gets a payoff $G_{ij} \times \left[z(X_i, X_j; \beta) + v(\sum_{k\neq i}^N G_{kj}; \delta) + \epsilon^i_{ij}\right]$. Player j gets a payoff $\sum_{i=1}^N G_{ij} \times \left[b(X_i, X_j; \gamma) + \epsilon^j_{ij}\right]$. Agents play PSPNE. The resulting *section j* is a PSPNS *section j*. The definitions of a PSPNE of the *bilateral section j game* and a PSPNS *section j* are now given. Let the dependence of $G_{\cdot j}$ on $s_{\cdot j}$ be denoted by $G_{\cdot j}(s_{\cdot j})$.

Definition A.1 (PSPNS *section j*)**.** A pure strategy profile $s_{\cdot j}$ is a PSPNE of the *section j game* if link ij is beneficial to player i, i.e.,

$$s^i_{ij} = \mathbb{1}\{z(X_i, X_j; \beta) + v(\sum_{k\neq i}^N G_{kj}(s^{(j)}); \delta) + \epsilon^i_{ij} \geq 0\}$$

and link ij is beneficial to player j, i.e.,

$$s^j_{ij} = \mathbb{1}\{b(X_i, X_j; \gamma) + \epsilon^j_{ij} \geq 0\}$$

$\forall i \in \mathcal{N}_{\cdot j,N}$. $G_{\cdot j}$ is a PSPNS *section j* if there exists a PSPNE $s_{\cdot j}$ of the *section j game* such that $G_{\cdot j} = G_{\cdot j}(s_{\cdot j})$, i.e.,

$$G_{ij} = \mathbb{1}\{z(X_i, X_j; \beta) + v(\sum_{k\neq i}^N G_{kj}; \delta) + \epsilon^i_{ij} \geq 0\}\mathbb{1}\{b(X_i, X_j; \gamma) + \epsilon^j_{ij} \geq 0\} \quad \forall i \in \mathcal{N}_{\cdot j,N}$$

Existence statements As for the unilateral case, it can be seen that the payoff of any player within the *bilateral section j game* depends exclusively $G_{\cdot j}$, for any $j \in \mathcal{N}_N$. Combining this with the fact that the collection of sets of pure strategy profiles of the *bilateral section j game* $\forall j \in \mathcal{N}_N$ constitutes an N-partition of the set of pure strategy profiles of the bilateral network formation game allows to show that

Lemma A.1 (Decomposing the bilateral network formation game)**.** $\boldsymbol{G}$ is a PSPNS network if and only if $G_{\cdot j}$ is a PSPNS *section* j $\forall j \in \mathcal{N}_N$.

Moreover, using Tarski's fixed point theorem when $v(\cdot;\delta)$ is monotone increasing and a bilateral game reinterpretation of the constructive proof in Berry (1992) that shows equilibrium existence in an entry game with negative competitive effects[35,36] when $v(\cdot;\delta)$ is monotone decreasing, it can be proved that

Lemma A.2 (Existence of a PSPNS *section* j)**.** There exists a PSPNS *section* j $\forall j \in \mathcal{N}_N$.

Hence, by Lemmas A.1 and A.2,

Lemma A.3 (Existence of a PSPNS network)**.** There exists a PSPNS network.

Proofs of Lemmas A.1, A.2, and A.3 are in Appendix C.

B Inference on Θ^o

This section discusses inference on the outer set Θ^o following the method of AS (2010).

In order to obtain unconditional moment inequalities as required by AS (2010), the observed characteristics of players are supposed to be discrete. Hence, Θ^o can be rewritten as

$$\Theta^o = \Big\{\theta \in \Theta | H^l_{g_{\cdot j},\boldsymbol{x}}(\theta) \leq \mathbb{P}(G_{\cdot j} = g_{\cdot j}, \boldsymbol{X} = \boldsymbol{x}) \leq H^u_{g_{\cdot j},\boldsymbol{x}}(\theta) \quad \forall g_{\cdot j} \in \{0,1\}^{n-1}, \forall j \in \mathcal{N}_n, \forall \boldsymbol{x} \in \mathcal{X}_n, \forall n \in \mathbb{N} \setminus \{1,2\}\Big\} \tag{B.1}$$

where

$$H^l_{g_{\cdot j},\boldsymbol{x}}(\theta) := \int_{e_{\cdot j} \in \mathbb{R}^{(n-1)} \text{ s.t. } \mathcal{S}_{\cdot j,\theta_u}(\boldsymbol{x}, e_{\cdot j}) = \{g_{\cdot j}\}} d\tilde{F}_{\cdot j,\boldsymbol{x}}(e_{\cdot j};\theta_\epsilon)\mathbb{P}(\boldsymbol{X} = \boldsymbol{x}) \tag{B.2}$$

and

$$H^u_{g_{\cdot j},\boldsymbol{x}}(\theta) := \int_{e_{\cdot j} \in \mathbb{R}^{(n-1)} \text{ s.t. } g_{\cdot j} \in \mathcal{S}_{\cdot j,\theta_u}(\boldsymbol{x}, e_{\cdot j})} d\tilde{F}_{\cdot j,\boldsymbol{x}}(e_{\cdot j};\theta_\epsilon)\mathbb{P}(\boldsymbol{X} = \boldsymbol{x}) \tag{B.3}$$

A preliminary step needed to conduct inference on Θ^o is estimation of $\mathbb{P}(G_{\cdot j} = g_{\cdot j}, \boldsymbol{X} = \boldsymbol{x})$ $\forall g_{\cdot j} \in \{0,1\}^{n-1}$, $\forall j \in \mathcal{N}_n$, $\forall \boldsymbol{x} \in \mathcal{X}_n$ and $\forall n \in \mathbb{N} \setminus \{1,2\}$. Moreover, the inference algorithm requires computation of $H^l_{g_{\cdot j},\boldsymbol{x}}(\theta)$ and $H^u_{g_{\cdot j},\boldsymbol{x}}(\theta)$ $\forall g_{\cdot j} \in \{0,1\}^{n-1}$, $\forall j \in \mathcal{N}_n$, $\forall \boldsymbol{x} \in \mathcal{X}_n$, $\forall n \in \mathbb{N} \setminus \{1,2\}$ and $\forall \theta \in \Theta$.

[35] See *Result* at p.894 in Berry (1992).

[36] The proof in the Appendix explains how to adapt the constructive algorithm in Berry (1992) to the bilateral network formation game.

Estimating $\mathbb{P}(G_{\cdot j} = g_{\cdot j}, \boldsymbol{X} = \boldsymbol{x})$, for example via a frequency estimator, is complicated by the fact that players' identities or roles vary across networks and players' observed characteristics have discrete support. Indeed, within each network, there may be observationally identical players which could be labelled arbitrarily by the researcher, hence producing different estimates of $\mathbb{P}(G_{\cdot j} = g_{\cdot j}, \boldsymbol{X} = \boldsymbol{x})$. A similar problem arises when computing $H^l_{g_{\cdot j},\boldsymbol{x}}(\theta)$ and $H^u_{g_{\cdot j},\boldsymbol{x}}(\theta)$. To solve this issue the present work adopts the strategy proposed by Sheng (2014) that relies on Assumptions 1 and 2.

Before illustrating the strategy, it is noticed that under Assumptions 1 and 2, $\Theta^o = \Theta^o_{\cdot j}$ (implication (i) of Lemma 6) and some inequalities defining $\Theta^o_{\cdot j}$ are redundant and, hence, can be deleted (implication (ii) of Lemma 6), for any $j \in \mathcal{N}_n$ and for any $n \in \mathbb{N} \setminus \{1, 2\}$. Thus, under Assumptions 1 and 2, conducting inference on Θ^o is equivalent to conducting inference on

$$\Theta^o_{\cdot 3} = \Big\{\theta \in \Theta | H^l_{g_{\cdot 3},\boldsymbol{x}}(\theta) \leq \mathbb{P}(G_{\cdot 3} = g_{\cdot 3}, \boldsymbol{X} = \boldsymbol{x}) \leq H^u_{g_{\cdot 3},\boldsymbol{x}}(\theta)\ \forall (g_{\cdot 3}, \boldsymbol{x}) \in \mathcal{W}_n,\ \forall n \in \mathbb{N} \setminus \{1,2\}\Big\} \tag{B.4}$$

where the subscript j is fixed to 3 without loss of generality to guarantee that j is contained in $\mathcal{N}_n$ $\forall n \in \mathbb{N} \setminus \{1, 2\}$, and $\mathcal{W}_n \subseteq \{0, 1\}^{n-1} \bigtimes \mathcal{X}_n$ denotes the collection of realisations of $(G_{\cdot 3}, \boldsymbol{X})$ left over after having deleted those generating redundant inequalities according to implication (ii) of Lemma 6[37].

Let $C_{g_{\cdot 3},\boldsymbol{x}} \subset \{0, 1\}^{n-1} \bigtimes \mathcal{X}_n$ be the collection of realisations of $(G_{\cdot 3}, \boldsymbol{X})$ giving rise to inequalities identical to the inequalities involving $(g_{\cdot 3}, \boldsymbol{x})$ according to implication (ii) of Lemma 6[38]. Sheng (2014) observes that, under Assumptions 1 and 2, (B.4) can be rewritten as

$$\Theta^o_{\cdot 3} = \Big\{\theta \in \Theta | H^l_{C_{g_{\cdot 3},\boldsymbol{x}}}(\theta) \leq \mathbb{P}((G_{\cdot 3}, \boldsymbol{X}) \in C_{g_{\cdot 3},\boldsymbol{x}}) \leq H^u_{C_{g_{\cdot 3},\boldsymbol{x}}}(\theta)\ \forall (g_{\cdot 3}, \boldsymbol{x}) \in \mathcal{W}_n,\ \forall n \in \mathbb{N} \setminus \{1,2\}\Big\} \tag{B.5}$$

where $H^l_{C_{g_{\cdot 3},\boldsymbol{x}}}(\theta)$ is the probability that every PSNE of the *section* 3 *game* combined with $\boldsymbol{X}$ falls in $C_{g_{\cdot 3},\boldsymbol{x}}$ and $H^u_{C_{g_{\cdot 3},\boldsymbol{x}}}(\theta)$ is the probability that at least one PSNE of the *section* 3 *game* combined with $\boldsymbol{X}$ falls in $C_{g_{\cdot 3},\boldsymbol{x}}$, given $\theta \in \Theta$. A proof of the equivalence between (B.4) and (B.5) is in Appendix C.

It can be noticed that (B.5) is a convenient way of rewriting $\Theta^o_{\cdot 3}$ as estimates of $\mathbb{P}((G_{\cdot 3}, \boldsymbol{X}) \in C_{g_{\cdot 3},\boldsymbol{x}})$ do not depend on how players are labelled by the researcher. Similarly, the computation of $H^l_{C_{g_{\cdot 3},\boldsymbol{x}}}(\theta)$ and $H^u_{C_{g_{\cdot 3},\boldsymbol{x}}}(\theta)$ is not affected by assigned labels.

[37] Appendix D.3 illustrates an algorithm to construct $\mathcal{W}_n$ when $\mathcal{X}_n$ is finite. It should be noticed that the set $\mathcal{W}_n$ is not unique because one is free to keep any of the realisations of $(G_{\cdot 3}, \boldsymbol{X})$ producing identical inequalities.

[38] In network theory, all realisations of $(G_{\cdot 3}, \boldsymbol{X})$ in $C_{g_{\cdot 3},\boldsymbol{x}}$ are called *isomorphic* and $C_{g_{\cdot 3},\boldsymbol{x}}$ is an *equivalence class* for $(G_{\cdot 3}, \boldsymbol{X})$.

Let $\hat{\mathbb{P}}_{C_{g\cdot 3,\boldsymbol{x}}}$ denote an unbiased estimator of $\mathbb{P}((G_{\cdot 3},\boldsymbol{X}) \in C_{g\cdot 3,\boldsymbol{x}})$[39]. By unbiasedness of $\hat{\mathbb{P}}_{C_{g\cdot 3,\boldsymbol{x}}}$,

$$\Theta^o_{\cdot 3} = \Big\{\theta \in \Theta | E(\hat{\mathbb{P}}_{C_{g\cdot 3,\boldsymbol{x}}} - H^l_{C_{g\cdot 3,\boldsymbol{x}}}(\theta)) \geq 0,\ E(H^u_{C_{g\cdot 3,\boldsymbol{x}}}(\theta) - \hat{\mathbb{P}}_{C_{g\cdot 3,\boldsymbol{x}}}) \geq 0\ \forall (g_{\cdot 3},\boldsymbol{x}) \in \mathcal{W}_n,\ \forall n \in \mathbb{N} \setminus \{1,2\}\Big\} \tag{B.6}$$

Reintroducing the subscript m and collecting the lhs of all inequalities in $E(b_m(\theta))$,

$$\Theta^o_{\cdot 3} = \Big\{\theta \in \Theta | E(b_m(\theta)) \geq 0\Big\} \tag{B.7}$$

Let $\bar{b}_M(\theta) := \frac{1}{M}\sum_{m=1}^M b_m(\theta)$ and $\bar{b}_{k,M}(\theta)$ denote its kth element. Let

$$S_M(\theta) := \sum_k \Big(\min\Big\{\frac{\sqrt{M}\bar{b}_{k,M}(\theta)}{\tilde{\sigma}_{k,M}}, 0\Big\}\Big)^2$$

where $\tilde{\sigma}_{k,M}$ is a consistent estimator of the asymptotic standard deviation of $\sqrt{M}\bar{b}_{k,M}(\theta)$. A $1-\alpha$ confidence region for each $\theta \in \Theta^o_{\cdot 3}$ is

$$CS_M := \Big\{\theta \in \Theta \text{ such that } S_M(\theta) \leq \hat{c}_{M,1-\alpha}(\theta)\Big\} \tag{B.8}$$

where $\hat{c}_{M,1-\alpha}(\theta)$ is an estimate of the $1-\alpha$ quantile of the asymptotic probability distribution of $S_M(\theta)$, obtainable following the bootstrap procedure with hard threshold of AS (2010). More details on the construction of $S_M(\theta)$ and $\hat{c}_{M,1-\alpha}(\theta)$ are in Appendix D.6.

C Proofs

Proof of Lemma 1 (Lemma 7). Let $\boldsymbol{G_{-\{ij,\cdot i\}}}$ be the matrix $\boldsymbol{G}$ with ith row and ijth element deleted, and $\boldsymbol{G_{-\{\cdot i\}}} = (G_{ij}, \boldsymbol{G_{-\{ij,\cdot i\}}})$.

Notice that $G_{ij}\times\Big[z(X_i,X_j;\beta)+v(\sum_{k\neq i}^N G_{kj};\delta)+\epsilon_{ij}\Big] \geq \tilde{G}_{ij}\times\Big[z(X_i,X_j;\beta)+v(\sum_{k\neq i}^N G_{kj};\delta)+\epsilon_{ij}\Big]$ is equivalent to $U(G_{ij},\boldsymbol{G_{-\{ij,\cdot i\}}},X,\epsilon_{i\cdot};\theta_u) \geq U(\tilde{G}_{ij},\boldsymbol{G_{-\{ij,\cdot i\}}},X,\epsilon_{i\cdot};\theta_u)$ for $\tilde{G}_{ij} \neq G_{ij} \in \{0,1\}$, $\forall i,j \in \mathcal{N}_N$, $i \neq j$.

It is firstly proved that if $\boldsymbol{G}$ is a PSNE of the network formation game, then $U(G_{ij},\boldsymbol{G_{-\{ij,\cdot i\}}},X,\epsilon_{i\cdot};\theta_u) \geq U(\tilde{G}_{ij},\boldsymbol{G_{-\{ij,\cdot i\}}},X,\epsilon_{i\cdot};\theta_u)$ for $\tilde{G}_{ij} \neq G_{ij} \in \{0,1\}$, $\forall i,j \in \mathcal{N}_N$, $i \neq j$. Consider players i and j. Let $G_{i\cdot-\{ij\}}$ be the vector $G_{i\cdot}$ with ijth element removed. With some abuse of notation, let $G_{i\cdot} = (G_{ij}, G_{i\cdot-\{ij\}})$. By setting $\tilde{G}_{i\cdot} = (\tilde{G}_{ij}, G_{i\cdot-\{ij\}})$ with $\tilde{G}_{ij} \neq G_{ij}$ in $U(G_{i\cdot},\boldsymbol{G_{-\{i\cdot,\cdot i\}}},\boldsymbol{X},\epsilon_{i\cdot};\theta_u) \geq U(\tilde{G}_{i\cdot},\boldsymbol{G_{-\{i\cdot,\cdot i\}}},\boldsymbol{X},\epsilon_{i\cdot};\theta_u)$, it follows that $U(G_{ij},\boldsymbol{G_{-\{ij,\cdot i\}}},X,\epsilon_{i\cdot};\theta_u) \geq U(\tilde{G}_{ij},\boldsymbol{G_{-\{ij,\cdot i\}}},X,\epsilon_{i\cdot};\theta_u)$ and this is verified $\forall i,j \in \mathcal{N}_N$, $i \neq j$.

[39] Appendix D.4 describes a procedure to estimate $\mathbb{P}((G_{\cdot 3},\boldsymbol{X}) \in C_{g\cdot 3,\boldsymbol{x}})$. Appendix D.5 describes a procedure to compute $H^l_{C_{g\cdot 3,\boldsymbol{x}}}(\theta)$ and $H^u_{C_{g\cdot 3,\boldsymbol{x}}}(\theta)$.

Conversely, it is proved that if $U(G_{ij}, \boldsymbol{G_{-\{ij,\cdot i\}}}, X, \epsilon_{i\cdot}; \theta_u) \geq U(\tilde{G}_{ij}, \boldsymbol{G_{-\{ij,\cdot i\}}}, X, \epsilon_{i\cdot}; \theta_u)$ for $\tilde{G}_{ij} \neq G_{ij} \in \{0, 1\}$, $\forall i, j \in \mathcal{N}_N$, $i \neq j$, then $\boldsymbol{G}$ is a PSNE of the network formation game. Consider player i. From (3.1), if $U(G_{ij}, \boldsymbol{G_{-\{ij,\cdot i\}}}, X, \epsilon_{i\cdot}; \theta_u) \geq U(\tilde{G}_{ij}, \boldsymbol{G_{-\{ij,\cdot i\}}}, X, \epsilon_{i\cdot}; \theta_u)$, then, by *additive separability* of $U(\cdot; \theta_u)$, $U(G_{i\cdot}, \boldsymbol{G_{-\{i\cdot,\cdot i\}}}, \boldsymbol{X}, \epsilon_{i\cdot}; \theta_u) \geq U(\tilde{G}_{i\cdot}, \boldsymbol{G_{-\{i\cdot,\cdot i\}}}, \boldsymbol{X}, \epsilon_{i\cdot}; \theta_u)$ $\forall \tilde{G}_{i\cdot} \neq G_{i\cdot} \in \{0, 1\}^{N-1}$ and this is verified $\forall i \in \mathcal{N}_N$.

Lemma 7 can be shown analogously after having replaced PSNE with PSPNE. $\square$

Theorem C.1 (Tarski's fixed point theorem)**.** *Let $F(x)$ be a monotone increasing function from a non-empty complete lattice X into X. Then,*

(i) the set of fixed points of $F(x)$ in X is non-empty, where $\sup_x(\{x \in X, x \leq F(x)\})$ *and* $\inf_x(\{x \in X, x \geq F(x)\})$ *denote, respectively, the greatest and the least fixed points;*

(ii) the set of fixed points of $F(x)$ in X is a non-empty complete lattice.

Let $h_{ij}(\boldsymbol{G}) := \mathbb{1}\{z(X_i, X_j; \beta) + v(\sum_{k \neq i}^N G_{kj}; \delta) + \epsilon_{ij} \geq 0\}$ and

$$h(\boldsymbol{G}) := \begin{pmatrix} 0 & h_{12}(\boldsymbol{G}) & h_{13}(\boldsymbol{G}) & \dots & h_{1N}(\boldsymbol{G}) \\ h_{21}(\boldsymbol{G}) & 0 & h_{23}(\boldsymbol{G}) & \dots & h_{2N}(\boldsymbol{G}) \\ \vdots & \vdots & \vdots & \vdots\vdots\vdots & \vdots \\ h_{N1}(\boldsymbol{G}) & h_{N2}(\boldsymbol{G}) & h_{23}(\boldsymbol{G}) & \dots & 0 \end{pmatrix}$$

Hence, $h : \mathcal{G}_N \to \mathcal{G}_N$. It is possible to show that the function h satisfies the sufficient conditions of Theorem C.1 when $v(\sum_{k \neq i}^N G_{kj}; \delta)$ is monotone increasing in $\sum_{k \neq i}^N G_{kj}$, meaning that the network formation game has a PSNE when $v(\sum_{k \neq i}^N G_{kj}; \delta)$ is monotone increasing in $\sum_{k \neq i}^N G_{kj}$.

Proof. Let the comparison between matrices be coordinate-wise, i.e. for any $\boldsymbol{G}, \boldsymbol{G'} \in \mathcal{G}_N$

$$\boldsymbol{G} \geq \boldsymbol{G'} \Leftrightarrow G_{ij} \geq G'_{ij} \ \forall i, j \in \mathcal{N}_N,\ i \neq j$$

Thus, $\boldsymbol{G} = \boldsymbol{G'}$ if and only if $\boldsymbol{G} \geq \boldsymbol{G'}$ and $\boldsymbol{G} \leq \boldsymbol{G'}$. Moreover, $\boldsymbol{G}$ and $\boldsymbol{G'}$ are unordered if and only if neither $\boldsymbol{G} \geq \boldsymbol{G'}$ nor $\boldsymbol{G} \leq \boldsymbol{G'}$. Therefore, $\mathcal{G}_N$ is a lattice, i.e. a set with a partial order. As $\mathcal{G}_N$ is a finite lattice, it is complete. Furthermore, if $v(\sum_{k \neq i}^N G_{kj}; \delta)$ is monotone increasing in $\sum_{k \neq i}^N G_{kj}$, then h is a monotone increasing function. In fact, consider two matrices $\boldsymbol{G} \geq \boldsymbol{G'}$. Since $G_{ij} \geq G'_{ij}$ $i, j \in \mathcal{N}_N$, $i \neq j$, then $z(X_i, X_j; \beta) + v(\sum_{k \neq i}^N G_{kj}; \delta) + \epsilon_{ij} \geq z(X_i, X_j; \beta) + v(\sum_{k \neq i}^N G_{kj}; \delta) + \epsilon_{ij}$ $i, j \in \mathcal{N}_N$, $i \neq j$. Hence, $h(\boldsymbol{G}) \geq h(\boldsymbol{G'})$ and the sufficient conditions of the theorem are met, meaning that the network formation game has a PSNE when $v(\sum_{k \neq i}^N G_{kj}; \delta)$ is monotone increasing in $\sum_{k \neq i}^N G_{kj}$. $\square$

Proof of Lemma 2 (Lemma A.1). It is firstly proved that if $\boldsymbol{G}$ is a PSNE of the network formation game, then $G_{\cdot j}$ is a PSNE of the *section* j *game* $\forall j \in \mathcal{N}_N$. By Lemma 1, if $\boldsymbol{G}$ is a PSNE of the network formation game, then $G_{ij} = \mathbb{1}\{z(X_i, X_j; \beta) + v(\sum_{k \neq i}^{N} G_{kj}; \delta) + \epsilon_{ij} \geq 0\}$ $\forall i, j \in \mathcal{N}_N$, $i \neq j$. This set of conditions also includes the conditions defining $G_{\cdot j}$ as a PSNE of the *section* j *game*. Therefore, $G_{\cdot j}$ is a PSNE of the *section* j *game* $\forall j \in \mathcal{N}_N$.

Conversely, it is proved that if $G_{\cdot j}$ is a PSNE of the *section* j *game* $\forall j \in \mathcal{N}_N$, then $\boldsymbol{G}$ is a PSNE of the network formation game. If $G_{\cdot j}$ is a PSNE of the *section* j *game*, then, by Definition 2, $G_{ij} = \mathbb{1}\{z(X_i, X_j; \beta) + v(\sum_{k \neq i}^{N} G_{kj}; \delta) + \epsilon_{ij} \geq 0\}$ $\forall i \in \mathcal{N}_{\cdot j,N}$. Based on the fact that this is true $\forall j \in \mathcal{N}_N$, conditions of Lemma 1 are satisfied and $\boldsymbol{G}$ is a PSNE of the network formation game.

Lemma A.1 can be shown analogously after having replaced PSNE with PSPNE. □

Proof of Lemma 3 (Lemma A.2). It is firstly considered the case in which $v(\sum_{k \neq i}^{N} G_{kj}; \delta)$ is monotone increasing in $\sum_{k \neq i}^{N} G_{kj}$. By Theorem C.1, the network formation game has a PSNE. By Lemma 2, if $\boldsymbol{G}$ is a PSNE of the network formation game, then $G_{\cdot j}$ is a PSNE of the *section* j *game* $\forall j \in \mathcal{N}_N$. Therefore, the *section* j *game* has a PSNE.

Now, the case in which $v(\sum_{k \neq i}^{N} G_{kj}; \delta)$ is monotone decreasing in $\sum_{k \neq i}^{N} G_{kj}$ is considered. In what follows it is shown that a PSNE of the section j game can be constructed following proof of *Result* in Berry (1992). Let $Y_{ij} := z(X_i, X_j; \beta) + \epsilon_{ij}$. The elements $\{Y_{ij}\}_{\forall i \in \mathcal{N}_{\cdot j,N}}$ are ordered from largest to smallest. Let $k \in \{1, ..., N-1\}$ denote the position of Y_{ij} in the ordered list and let π be a function such that $\pi(ij) = k$, $\forall i \in \mathcal{N}_{\cdot j,N}$. By replacing the subscript ij with k, the ordered sequence is

$$Y_1 \geq Y_2 \geq \ \dots \ \geq Y_k \geq \ \dots \ \geq Y_{N-1}$$

Let $Y_0 := \max\left\{Y_1, -v(-1; \delta)\right\}$.

n^* is defined as the largest element of the set of integers $\{0, 1, \ \dots \ , k, \ \dots \ , N-1\}$ satisfying $Y_{n^*} + v(n^* - 1; \delta) \geq 0$, i.e.

$$n^* := \max\{k \in \{0, ..., N-1\} | Y_k + v(k-1; \delta) \geq 0\}$$

$G_{\cdot j}$ is constructed by imposing $G_{ij} = 1$ if $\pi(ij) \leq n^*$ and $G_{ij} = 0$ otherwise. One can see that $G_{\cdot j}$ is a PSNE of the *section* j *game*. In fact choosing n^* according to the previous criterion means that

$$Y_0 + v(-1;\delta) \geq 0 \qquad \text{(a)}$$

$$Y_1 + v(0;\delta) \geq 0 \qquad \text{(b)}$$

$$\ldots$$

$$Y_{n^*} + v(n^* - 1;\delta) \geq 0 \qquad \text{(c)}$$

$$Y_{n^*+1} + v(n^*;\delta) < 0 \qquad \text{(d)}$$

$$Y_{n^*+2} + v(n^* + 1;\delta) < 0 \qquad \text{(e)}$$

$$\ldots$$

$$Y_{N-1} + v(N - 2;\delta) < 0 \qquad \text{(f)}$$

For $G_{\cdot j}$ being a PSNE of the *section j game*, the following inequalities should be satisfied

$$Y_1 + v(n^* - 1;\delta) \geq 0 \qquad \text{(g)}$$

$$Y_2 + v(n^* - 1;\delta) \geq 0 \qquad \text{(h)}$$

$$\ldots$$

$$Y_{n^*} + v(n^* - 1;\delta) \geq 0 \qquad \text{(i)}$$

$$Y_{n^*+1} + v(n^*;\delta) < 0 \qquad \text{(l)}$$

$$Y_{n^*+2} + v(n^*;\delta) < 0 \qquad \text{(m)}$$

$$\ldots$$

$$Y_{N-1} + v(n^*;\delta) < 0 \qquad \text{(n)}$$

By observing that inequalities (g), (h), ... ,(i) are implied by inequality (c) and all other inequalities follow from inequality (d), it can be concluded that $G_{\cdot j}$ is a PSNE of the *section j game*.

Following the proof of *Result* in Berry (1992), one can also show by contradiction that all PSNE of the *section j game* are characterised by n^* players linking to firm j. In fact, suppose there is some equilibrium with $k^* > n^*$ edges. Given the definition of equilibrium with k^* links, it should be $Y_{k^*} + v(k^* - 1;\delta) \geq 0$ which contradicts the definition of n^*. Similarly, for $k^* < n^*$.

Lemma A.2 can be shown analogously after having replaced PSNE with PSPNE and imposed

$$Y_{ij} := \begin{cases} z(X_i, X_j;\beta) + \eta^i_{ij} \text{ if } b(X_i, X_j;\gamma) + \eta^j_{ij} \geq 0 \\ -\infty \text{ otherwise} \end{cases}$$

□

Proof of Lemma 4 (Lemma A.3). By Lemma 3, there exists a PSNE of the *section j game* $\forall j \in \mathcal{N}_N$. By Lemma 2, if $G_{\cdot j}$ is a PSNE of the *section j game* $\forall j \in \mathcal{N}_N$, then $\boldsymbol{G}$ is a PSNE of the network formation game. Thus, the network formation game has a PSNE.

Lemma A.3 can be shown analogously after having replaced PSNE with PSPNE.

□

Proof of Lemma 5. It is firstly shown that, under Assumption 1, $\Theta^{\star\star} \supseteq \Theta^{\star}$. Specifically, it is proved that, $\forall \theta \in \Theta$, if

$$\mathbb{P}(\boldsymbol{G} \in K|\boldsymbol{X} = \boldsymbol{x}) \leq T_{\mathcal{S}_{\theta_u}(\boldsymbol{X},\epsilon)|\boldsymbol{X}=\boldsymbol{x}}(K) \ \forall K \in \mathcal{K}_{\mathcal{G}_n}, \ \forall \boldsymbol{x} \in \mathcal{X}_n \text{ a.s.} \tag{C.1}$$

then

$$\mathbb{P}(G_{\cdot j} \in K_{\cdot j}|\boldsymbol{X} = \boldsymbol{x}) \leq T_{\mathcal{S}_{\theta_u,\cdot j}(\boldsymbol{X},\epsilon_{\cdot j})|\boldsymbol{X}=\boldsymbol{x}}(K_{\cdot j}) \ \forall K_{\cdot j} \in \mathcal{K}_{\{0,1\}^{n-1}}, \ \forall j \in \mathcal{N}_n \text{ and}$$
$$\mathbb{P}(\boldsymbol{G} \in K|\boldsymbol{X} = \boldsymbol{x}) \leq T_{\mathcal{S}_{\theta_u}(\boldsymbol{X},\epsilon)|\boldsymbol{X}=\boldsymbol{x}}(K) \ \forall K \in \mathcal{K}^{\diamond}_{\mathcal{G}_n}, \ \forall \boldsymbol{x} \in \mathcal{X}_n \text{ a.s.} \tag{C.2}$$

This is equivalent to show that, $\forall \theta \in \Theta$, if

$$\mathbb{P}(\boldsymbol{G} \in K|\boldsymbol{X} = \boldsymbol{x}) \leq T_{\mathcal{S}_{\theta_u}(\boldsymbol{X},\epsilon)|\boldsymbol{X}=\boldsymbol{x}}(K) \ \forall K \in \mathcal{A}_{\mathcal{G}_n}, \ \forall \boldsymbol{x} \in \mathcal{X}_n \text{ a.s.} \tag{C.3}$$

then

$$\mathbb{P}(G_{\cdot j} \in K_{\cdot j}|\boldsymbol{X} = \boldsymbol{x}) \leq T_{\mathcal{S}_{\theta_u,\cdot j}(\boldsymbol{X},\epsilon_{\cdot j})|\boldsymbol{X}=\boldsymbol{x}}(K_{\cdot j}) \ \forall K_{\cdot j} \in \mathcal{K}_{\{0,1\}^{n-1}}, \ \forall j \in \mathcal{N}_n, \ \forall \boldsymbol{x} \in \mathcal{X}_n \text{ a.s.} \tag{C.4}$$

Consider any $\theta \in \Theta$, $j \in \mathcal{N}_n$ and $K_{\cdot j} \in \mathcal{K}_{\{0,1\}^{n-1}}$. Take $K \in \mathcal{A}_{\mathcal{G}_n}$ corresponding to

$$\underbrace{\{0,1\}^{n-1} \times ... \times \{0,1\}^{n-1}}_{j-1 \text{ times}} \times K_{\cdot j} \times \underbrace{\{0,1\}^{n-1} \times ... \times \{0,1\}^{n-1}}_{n-j \text{ times}}$$

By (C.3),

$$\mathbb{P}(\boldsymbol{G} \in K|\boldsymbol{X} = \boldsymbol{x}) \leq T_{\mathcal{S}_{\theta_u}(\boldsymbol{X},\epsilon)|\boldsymbol{X}=\boldsymbol{x}}(K)$$

which is equivalent, by Lemma 2, to

$$\begin{aligned}
&\mathbb{P}(G_{\cdot 1} \in \{0,1\}^{n-1}, ..., G_{\cdot j-1} \in \{0,1\}^{n-1}, G_{\cdot j} \in K_{\cdot j}, G_{\cdot j+1} \in \{0,1\}^{n-1}, ..., G_{\cdot n} \in \{0,1\}^{n-1}|\boldsymbol{X} = \boldsymbol{x}) \\
&\leq \mathbb{P}(\mathcal{S}_{\theta_u,\cdot 1}(\boldsymbol{X},\epsilon_{\cdot 1}) \cap \{0,1\}^{n-1} \neq \emptyset, ..., \mathcal{S}_{\theta_u,\cdot j-1}(\boldsymbol{X},\epsilon_{\cdot j-1}) \cap \{0,1\}^{n-1} \neq \emptyset, \mathcal{S}_{\theta_u,\cdot j}(\boldsymbol{X},\epsilon_{\cdot j}) \cap K_{\cdot j} \neq \emptyset, \\
&\qquad \mathcal{S}_{\theta_u,\cdot j+1}(\boldsymbol{X},\epsilon_{\cdot j+1}) \cap \{0,1\}^{n-1} \neq \emptyset, ..., \mathcal{S}_{\theta_u,\cdot n}(\boldsymbol{X},\epsilon_{\cdot n}) \cap \{0,1\}^{n-1} \neq \emptyset|\boldsymbol{X} = \boldsymbol{x})
\end{aligned}$$

which is equivalent to

$$\mathbb{P}(G_{\cdot j} \in K_{\cdot j}|\boldsymbol{X} = \boldsymbol{x}) \leq \mathbb{P}(\mathcal{S}_{\theta_u,\cdot j}(\boldsymbol{X},\epsilon_{\cdot j}) \cap K_{\cdot j} \neq \emptyset|\boldsymbol{X} = \boldsymbol{x})$$

$\forall \boldsymbol{x} \in \mathcal{X}_n$ a.s. This holds $\forall K_{\cdot j} \in \mathcal{K}_{\{0,1\}^{n-1}}$, $\forall j \in \mathcal{N}_n$ and $\forall \theta \in \Theta$. By collecting all inequalities, (C.4) is obtained.

Now it is shown that, under Assumptions 1 and 3, $\Theta^{\star\star} = \Theta^{\star}$. As discussed above, if $\theta \in \Theta^{\star}$, then $\theta \in \Theta^{\star\star}$. Hence, in what follows it is proved that if $\theta \in \Theta^{\star\star}$, then $\theta \in \Theta^{\star}$. This is equivalent to show that, $\forall \theta \in \Theta$, (C.4) implies (C.3).

Consider any $\theta \in \Theta$ and $K \in \mathcal{A}_{\mathcal{G}_n}$ corresponding to $K_{\cdot 1} \times ... \times K_{\cdot n}$ with $K_{\cdot j} \in \mathcal{K}_{\{0,1\}^{n-1}}$ $\forall j \in \mathcal{N}_n$. By (C.4)

$$\begin{pmatrix} \mathbb{P}(G_{\cdot 1} \in K_{\cdot 1} | \boldsymbol{X} = \boldsymbol{x}) \leq \mathbb{P}(\mathcal{S}_{\theta_u, \cdot 1}(\boldsymbol{X}, \epsilon_{\cdot 1}) \cap K_{\cdot 1} \neq \emptyset | \boldsymbol{X} = \boldsymbol{x}) \\ \vdots \\ \mathbb{P}(G_{\cdot n} \in K_{\cdot n} | \boldsymbol{X} = \boldsymbol{x}) \leq \mathbb{P}(\mathcal{S}_{\theta_u, \cdot n}(\boldsymbol{X}, \epsilon_{\cdot n}) \cap K_{\cdot n} \neq \emptyset | \boldsymbol{X} = \boldsymbol{x}) \end{pmatrix}$$

$\forall \boldsymbol{x} \in \mathcal{X}_n$ a.s. By taking the product

$$\prod_{j=1}^{n} \mathbb{P}(G_{\cdot j} \in K_{\cdot j} | \boldsymbol{X} = \boldsymbol{x}) \leq \prod_{j=1}^{n} \mathbb{P}(\mathcal{S}_{\theta_u, \cdot j}(\boldsymbol{X}, \epsilon_{\cdot j}) \cap K_{\cdot j} \neq \emptyset | \boldsymbol{X} = \boldsymbol{x}) \tag{C.5}$$

On the lhs

$$\prod_{j=1}^{n} \mathbb{P}(G_{\cdot j} \in K_{\cdot j} | \boldsymbol{X} = \boldsymbol{x}) \underbrace{=}_{\text{Ass. 3}} \mathbb{P}(G_{\cdot 1} \in K_{\cdot 1}, ..., G_{\cdot j} \in K_{\cdot j} | \boldsymbol{X} = \boldsymbol{x}) \underbrace{=}_{\text{Lemma 2}} \mathbb{P}(\boldsymbol{G} \in K | \boldsymbol{X} = \boldsymbol{x}) \tag{C.6}$$

On the rhs

$$\begin{aligned} \prod_{j=1}^{n} \mathbb{P}(\mathcal{S}_{\theta_u, \cdot j}(\boldsymbol{X}, \epsilon_{\cdot j}) \cap K_{\cdot j} \neq \emptyset | \boldsymbol{X} = \boldsymbol{x}) &\underbrace{=}_{\text{Ass. 3}} \mathbb{P}(\mathcal{S}_{\theta_u, \cdot 1}(\boldsymbol{X}, \epsilon_{\cdot 1}) \cap K_{\cdot 1} \neq \emptyset, ..., \mathcal{S}_{\theta_u, \cdot n}(\boldsymbol{X}, \epsilon_{\cdot n}) \cap K_{\cdot n} \neq \emptyset | \boldsymbol{X} = \boldsymbol{x}) \\ &\underbrace{=}_{\text{Lemma 2}} \mathbb{P}(\mathcal{S}_{\theta_u}(\boldsymbol{X}, \epsilon) \cap K \neq \emptyset | \boldsymbol{X} = \boldsymbol{x}) \end{aligned} \tag{C.7}$$

Therefore, replacing (C.6) and (C.7) in (C.5),

$$\mathbb{P}(\boldsymbol{G} \in K | \boldsymbol{X} = \boldsymbol{x}) \leq \mathbb{P}(\mathcal{S}_{\theta_u}(\boldsymbol{X}, \epsilon) \cap K \neq \emptyset | \boldsymbol{X} = \boldsymbol{x})$$

$\forall \boldsymbol{x} \in \mathcal{X}_n$ a.s. This holds $\forall K \in \mathcal{A}_{\mathcal{G}_n}$ and $\forall \theta \in \Theta$. By collecting all inequalities, (C.3) is obtained. □

Proof of Lemma 6. It is shown that, under Assumptions 1 and 2, $\Theta^{\star\star} = \Theta^{\star\star}_{\cdot j}$ $\forall j \in \mathcal{N}_n$. It can be seen that if $\theta \in \Theta^{\star\star}$, then $\theta \in \Theta^{\star\star}_{\cdot j}$ $\forall j \in \mathcal{N}_n$. Hence, in what follows it is proved that, $\forall j \in \mathcal{N}_n$, if $\theta \in \Theta^{\star\star}_{\cdot j}$, then $\theta \in \Theta^{\star\star}$. This is equivalent to show that, $\forall \theta \in \Theta$ and $\forall j \in \mathcal{N}_n$, if

$$\mathbb{P}(G_{\cdot j} \in K_{\cdot j} | \boldsymbol{X} = \boldsymbol{x}) \leq T_{\mathcal{S}_{\theta_u, \cdot j}(\boldsymbol{X}, \epsilon_{\cdot j}) | \boldsymbol{X} = \boldsymbol{x}}(K_{\cdot j}) \ \forall K_{\cdot j} \in \mathcal{K}_{\{0,1\}^{n-1}}, \forall \boldsymbol{x} \in \mathcal{X}_n \text{ a.s.} \tag{C.8}$$

then

$$\mathbb{P}(G_{\cdot h} \in K_{\cdot h}|\boldsymbol{X}=\boldsymbol{x}) \leq T_{\mathcal{S}_{\theta_u,\cdot h}(\boldsymbol{X},\epsilon_{\cdot h})|\boldsymbol{X}=\boldsymbol{x}}(K_{\cdot h}) \ \forall K_{\cdot h} \in \mathcal{K}_{\{0,1\}^{n-1}},\ \forall h \in \mathcal{N}_{\cdot j,n},\ \forall \boldsymbol{x} \in \mathcal{X}_n \text{ a.s.} \tag{C.9}$$

Consider any $\theta \in \Theta$, $j \in \mathcal{N}_n$ and $K_{\cdot j} \in \mathcal{K}_{\{0,1\}^{n-1}}$. Let φ be a permutation of labels such that $\varphi(j) \neq j$. By Assumption 1, $\{X_i\}_{i\in\mathcal{N}_n}$ are i.i.d. across i and, hence, exchangeable across i. Assumption 2 implies that any finite subsequence of ϵ is exchangeable. Thus,

$$\mathbb{P}(G_{\cdot j} \in K_{\cdot j}|\boldsymbol{X}=\boldsymbol{x}) \underbrace{=}_{\text{Ass. 1, 2}} \mathbb{P}(G_{\cdot\varphi(j)} \in K^{\varphi}_{\cdot\varphi(j)}|\boldsymbol{X}=\boldsymbol{x}^{\boldsymbol{\varphi}}) \tag{C.10}$$

and

$$\mathbb{P}(\mathcal{S}_{\theta_u,\cdot j}(\boldsymbol{X},\epsilon_{\cdot j}) \cap K_{\cdot j} \neq \emptyset|\boldsymbol{X}=\boldsymbol{x}) \underbrace{=}_{\text{Ass. 1, 2}} \mathbb{P}(\mathcal{S}_{\theta_u,\cdot\varphi(j)}(\boldsymbol{X},\epsilon_{\cdot\varphi(j)}) \cap K^{\varphi}_{\cdot\varphi(j)} \neq \emptyset|\boldsymbol{X}=\boldsymbol{x}^{\boldsymbol{\varphi}}) \tag{C.11}$$

$\forall \boldsymbol{x} \in \mathcal{X}_n$.

Therefore, combining (C.10) and (C.11), if

$$\mathbb{P}(G_{\cdot j} \in K_{\cdot j}|\boldsymbol{X}=\boldsymbol{x}) \leq T_{\mathcal{S}_{\theta_u,\cdot j}(\boldsymbol{X},\epsilon_{\cdot j})|\boldsymbol{X}=\boldsymbol{x}}(K_{\cdot j})$$

then

$$\mathbb{P}(G_{\cdot\varphi(j)} \in K^{\varphi}_{\cdot\varphi(j)}|\boldsymbol{X}=\boldsymbol{x}^{\boldsymbol{\varphi}}) \leq T_{\mathcal{S}_{\theta_u,\cdot\varphi(j)}(\boldsymbol{X},\epsilon_{\cdot\varphi(j)})|\boldsymbol{X}=\boldsymbol{x}^{\boldsymbol{\varphi}}}(K^{\varphi}_{\cdot\varphi(j)})$$

$\forall \boldsymbol{x} \in \mathcal{X}_n$.

By repeating the procedure $\forall K_{\cdot j} \in \mathcal{K}_{\{0,1\}^{n-1}}$ and $\forall$ permutation of labels φ such that $\varphi(j) \neq j$, all inequalities in (C.9) are obtained.

In addition, from (C.10) and (C.11), for any permutation of labels φ such that $\varphi(j) = j$,

$$\mathbb{P}(G_{\cdot j} \in K_{\cdot j}|\boldsymbol{X}=\boldsymbol{x}) \leq T_{\mathcal{S}_{\theta_u,\cdot j}(\boldsymbol{X},\epsilon_{\cdot j})|\boldsymbol{X}=\boldsymbol{x}}(K_{\cdot j})$$

is equivalent to

$$\mathbb{P}(G_{\cdot j} \in K^{\varphi}_{\cdot j}|\boldsymbol{X}=\boldsymbol{x}^{\boldsymbol{\varphi}}) \leq T_{\mathcal{S}_{\theta_u,\cdot j}(\boldsymbol{X},\epsilon_{\cdot j})|\boldsymbol{X}=\boldsymbol{x}^{\boldsymbol{\varphi}}}(K^{\varphi}_{\cdot j})$$

$\forall \boldsymbol{x} \in \mathcal{X}_n$ a.s. Doing this $\forall K_{\cdot j} \in \mathcal{K}_{\{0,1\}^{n-1}}$ and $\forall \varphi$ such that $\varphi(j) = j$ allows to drop some redundant inequalities from (C.8).

These arguments hold $\forall j \in \mathcal{N}_n$ and $\forall \theta \in \Theta$.

Finally, adding Assumption 3, by statement (ii) of Lemma 5 $\Theta^{\star\star} = \Theta^{\star}$ and, hence, under Assumptions 1, 2 and 3, $\Theta^{\star\star}_{\cdot 1} = \Theta^{\star\star}_{\cdot 2} = \ \dots\ = \Theta^{\star\star}_{\cdot n} = \Theta^{\star\star} = \Theta^{\star}$. □

Equivalence between (B.4) and (B.5) Under Assumptions 1 and 2,

$$\mathbb{P}(G_{\cdot i} = g_{\cdot i}, \boldsymbol{X} = \boldsymbol{x}) = \mathbb{P}(G_{\cdot \varphi(i)} = g^{\varphi}_{\cdot \varphi(i)}, \boldsymbol{X} = \boldsymbol{x}^{\varphi}) \tag{C.12}$$

$\forall$ permutation of labels φ, $\forall (g_{\cdot i}, \boldsymbol{x}) \in \{0,1\}^{n-1} \times \mathcal{X}_n$, $\forall i \in \mathcal{N}$ and $\forall n \in \mathbb{N} \setminus \{1,2\}$.

By (C.12) applied $\forall \varphi$ such that $\varphi(3) = 3$,

$$\mathbb{P}((G_{\cdot 3}, \boldsymbol{X}) \in C_{g_{\cdot 3}, \boldsymbol{x}}) = |C_{g_{\cdot 3}, \boldsymbol{x}}| \times \mathbb{P}(G_{\cdot 3} = g_{\cdot 3}, \boldsymbol{X} = \boldsymbol{x}) \tag{C.13}$$

In a similar way, $H^l_{C_{g_{\cdot 3}, \boldsymbol{x}}}(\theta)$ and $H^u_{C_{g_{\cdot 3}, \boldsymbol{x}}}(\theta)$ can be shown being, respectively, equivalent to $|C_{g_{\cdot 3}, \boldsymbol{x}}| \times H^l_{g_{\cdot 3}, \boldsymbol{x}}$ and $|C_{g_{\cdot 3}, \boldsymbol{x}}| \times H^u_{g_{\cdot 3}, \boldsymbol{x}}$.

Hence,

$$\begin{aligned}&\Big\{\theta \in \Theta | H^l_{C_{g_{\cdot 3}, \boldsymbol{x}}}(\theta) \leq \mathbb{P}((G_{\cdot 3}, \boldsymbol{X}) \in C_{g_{\cdot 3}, \boldsymbol{x}}) \leq H^u_{C_{g_{\cdot 3}, \boldsymbol{x}}}(\theta) \ \forall (g_{\cdot 3}, \boldsymbol{x}) \in \mathcal{W}_n, \ \forall n \in \mathbb{N} \setminus \{1,2\}\Big\} \\ &\underbrace{=}_{\text{(C.13)}} \Big\{\theta \in \Theta || C_{g_{\cdot 3}, \boldsymbol{x}}| \times H^l_{g_{\cdot 3}, \boldsymbol{x}}(\theta) \leq |C_{g_{\cdot 3}, \boldsymbol{x}}| \times \mathbb{P}(G_{\cdot 3} = g_{\cdot 3}, \boldsymbol{X} = \boldsymbol{x}) \leq |C_{g_{\cdot 3}, \boldsymbol{x}}| \times H^u_{g_{\cdot 3}, \boldsymbol{x}}(\theta) \\ &\qquad \forall (g_{\cdot 3}, \boldsymbol{x}) \in \mathcal{W}_n, \ \forall n \in \mathbb{N} \setminus \{1,2\}\Big\} = \Theta^o_{\cdot 3}\end{aligned}$$

D Additional discussion

D.1 Notation examples

Example 1 ($g^{\varphi}_{\cdot \varphi(j)}$)**.** *Let* $n := 4$, $j = 2$, $g_{\cdot 2} := \begin{pmatrix} g_{12} \\ g_{32} \\ g_{42} \end{pmatrix} := \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ *and* $\varphi(1) = 2$, $\varphi(2) = 4$, $\varphi(3) = 1$, $\varphi(4) = 3$. *Hence,* $g^{\varphi}_{\cdot \varphi(2)} = g^{\varphi}_{\cdot 4} := \begin{pmatrix} g^{\varphi}_{14} \\ g^{\varphi}_{24} \\ g^{\varphi}_{34} \end{pmatrix} := \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$

Example 2 ($\boldsymbol{x}^{\boldsymbol{\varphi}}$)**.** *Let* $n := 4$, $\varphi(1) = 2$, $\varphi(2) = 4$, $\varphi(3) = 1$, $\varphi(4) = 3$ *and* $\boldsymbol{x} := \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} := \begin{pmatrix} 4 & 5 \\ 6 & 7 \\ 8 & 9 \\ 10 & 11 \end{pmatrix}$. *Hence,* $\boldsymbol{x}^{\boldsymbol{\varphi}} = \begin{pmatrix} x^{\varphi}_1 \\ x^{\varphi}_2 \\ x^{\varphi}_3 \\ x^{\varphi}_4 \end{pmatrix} := \begin{pmatrix} 8 & 9 \\ 4 & 5 \\ 10 & 11 \\ 6 & 7 \end{pmatrix}$

D.2 Construction of $\mathcal{A}_{\mathcal{G}_\mathbf{n}}$

Consider $K_{\cdot 1} \in \mathcal{K}_{\{0,1\}^{n-1}}, ..., K_{\cdot n} \in \mathcal{K}_{\{0,1\}^{n-1}}$. Construct $\mathcal{B}_{K_{\cdot 1},...,K_{\cdot n}} := \bigtimes_{j=1}^{n} K_{\cdot j}$, where $\bigtimes$ denotes the Cartesian products of sets. Hence, $\mathcal{B}_{K_{\cdot 1},...,K_{\cdot n}}$ is a collection of $L := \Pi_{j=1}^{n}|K_{\cdot j}|$ sets and it can be written as $\{B_l\}_{l=1}^{L}$. Any $B_l \in \mathcal{B}_{K_{\cdot 1},...,K_{\cdot n}}$ is composed by n vectors of dimension $(n-1)\times 1$. Hence, $B_l := \{b_{l,1}, ..., b_{l,n}\}$ with $b_{l,h} := (b_{l,h}^{1} \ ... \ b_{l,h}^{n-1})'$ $\forall h \in \{1, ..., n\}$, $\forall l \in \{1, ..., L\}$. Create the $n \times n$ matrix

$$C_l := \begin{pmatrix} 0 & b_{l,2}^{1} & b_{l,3}^{1} & ... & b_{l,n}^{1} \\ b_{l,1}^{1} & 0 & b_{l,3}^{2} & ... & b_{l,n}^{2} \\ b_{l,1}^{2} & b_{l,2}^{2} & 0 & ... & b_{l,n}^{3} \\ \vdots & \vdots & \vdots & \vdots\vdots\vdots & \vdots \\ b_{l,1}^{n-2} & b_{l,2}^{n-2} & b_{l,3}^{n-2} & ... & b_{l,n}^{n-1} \\ b_{l,1}^{n-1} & b_{l,2}^{n-1} & b_{l,3}^{n-1} & ... & 0 \end{pmatrix}$$

$\forall l \in \{1, ..., L\}$. Let $A := \{C_1, ..., C_L\}$. Repeat the procedure for all possible ordered n-tuples with repetition of $\mathcal{K}_{\{0,1\}^{n-1}}$ and denominate the family of sets A's as $\mathcal{A}_{\mathcal{G}_n}$. Notice that $|\mathcal{A}_{\mathcal{G}_n}| = (2^{2^{n-1}}-1)^n < |\mathcal{K}_{\mathcal{G}_n}| = 2^{2^{n(n-1)}} - 1$.

For example, suppose $n := 3$. Hence,

$$\{0,1\}^2 := \{\begin{pmatrix}1\\1\end{pmatrix}, \begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}, \begin{pmatrix}0\\0\end{pmatrix}\}$$

with $|\{0,1\}^2| = 4$,

$$\begin{aligned}\mathcal{K}_{\{0,1\}^2} := \Big\{ & \{\begin{pmatrix}1\\1\end{pmatrix}\}, \{\begin{pmatrix}1\\0\end{pmatrix}\}, \{\begin{pmatrix}0\\1\end{pmatrix}\}, \{\begin{pmatrix}0\\0\end{pmatrix}\}, \\ & \{\begin{pmatrix}1\\1\end{pmatrix}, \begin{pmatrix}1\\0\end{pmatrix}\}, \{\begin{pmatrix}1\\1\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}\}, \{\begin{pmatrix}1\\1\end{pmatrix}, \begin{pmatrix}0\\0\end{pmatrix}\}, \{\begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}\}, \{\begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}0\\0\end{pmatrix}\}, \{\begin{pmatrix}0\\1\end{pmatrix}, \begin{pmatrix}0\\0\end{pmatrix}\}, \\ & \{\begin{pmatrix}1\\1\end{pmatrix}, \begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}\}, \{\begin{pmatrix}1\\1\end{pmatrix}, \begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}0\\0\end{pmatrix}\}, \{\begin{pmatrix}1\\1\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}, \begin{pmatrix}0\\0\end{pmatrix}\}, \{\begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}, \begin{pmatrix}0\\0\end{pmatrix}\}, \\ & \{\begin{pmatrix}1\\1\end{pmatrix}, \begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}, \begin{pmatrix}0\\0\end{pmatrix}\}\Big\}\end{aligned}$$

with $|\mathcal{K}_{\{0,1\}^2}| = 15$ and

$$\mathcal{K}_{\mathcal{G}_n} := \Big\{\{\begin{pmatrix}0&1&1\\1&0&1\\1&1&0\end{pmatrix}\}, ...\Big\}$$

with $|\mathcal{K}_{\mathcal{G}_n}| = 2^{64} - 1$.

Let $K_{\cdot 1} := \{\begin{pmatrix}1\\1\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}\}$, $K_{\cdot 2} := \{0,1\}^2$, $K_{\cdot 3} := \{0,1\}^2$. Hence,

$$\mathcal{B}_{K_{\cdot 1},K_{\cdot 2},K_{\cdot 3}} := \{\begin{pmatrix}1\\1\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}\} \bigtimes \{0,1\}^2 \bigtimes \{0,1\}^2$$

$$= \{\begin{pmatrix}1\\1\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}\} \bigtimes \{\begin{pmatrix}1\\1\end{pmatrix}, \begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}, \begin{pmatrix}0\\0\end{pmatrix}\} \bigtimes \{\begin{pmatrix}1\\1\end{pmatrix}, \begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}, \begin{pmatrix}0\\0\end{pmatrix}\}$$

$$= \Big\{\{\begin{pmatrix}1\\1\end{pmatrix}, \begin{pmatrix}1\\1\end{pmatrix}, \begin{pmatrix}1\\1\end{pmatrix}\}, \{\begin{pmatrix}1\\1\end{pmatrix}, \begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}1\\1\end{pmatrix}\}, \{\begin{pmatrix}1\\1\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}, \begin{pmatrix}1\\1\end{pmatrix}\}, \{\begin{pmatrix}1\\1\end{pmatrix}, \begin{pmatrix}0\\0\end{pmatrix}, \begin{pmatrix}1\\1\end{pmatrix}\},$$

$$\{\begin{pmatrix}1\\1\end{pmatrix}, \begin{pmatrix}1\\1\end{pmatrix}, \begin{pmatrix}1\\0\end{pmatrix}\}, \{\begin{pmatrix}1\\1\end{pmatrix}, \begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}1\\0\end{pmatrix}\}, \{\begin{pmatrix}1\\1\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}, \begin{pmatrix}1\\0\end{pmatrix}\}, \{\begin{pmatrix}1\\1\end{pmatrix}, \begin{pmatrix}0\\0\end{pmatrix}, \begin{pmatrix}1\\0\end{pmatrix}\},$$

$$\{\begin{pmatrix}1\\1\end{pmatrix}, \begin{pmatrix}1\\1\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}\}, \{\begin{pmatrix}1\\1\end{pmatrix}, \begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}\}, \{\begin{pmatrix}1\\1\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}\}, \{\begin{pmatrix}1\\1\end{pmatrix}, \begin{pmatrix}0\\0\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}\},$$

$$\{\begin{pmatrix}1\\1\end{pmatrix}, \begin{pmatrix}1\\1\end{pmatrix}, \begin{pmatrix}0\\0\end{pmatrix}\}, \{\begin{pmatrix}1\\1\end{pmatrix}, \begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}0\\0\end{pmatrix}\}, \{\begin{pmatrix}1\\1\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}, \begin{pmatrix}0\\0\end{pmatrix}\}, \{\begin{pmatrix}1\\1\end{pmatrix}, \begin{pmatrix}0\\0\end{pmatrix}, \begin{pmatrix}0\\0\end{pmatrix}\},$$

$$\{\begin{pmatrix}0\\1\end{pmatrix}, \begin{pmatrix}1\\1\end{pmatrix}, \begin{pmatrix}1\\1\end{pmatrix}\}, \{\begin{pmatrix}0\\1\end{pmatrix}, \begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}1\\1\end{pmatrix}\}, \{\begin{pmatrix}0\\1\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}, \begin{pmatrix}1\\1\end{pmatrix}\}, \{\begin{pmatrix}0\\1\end{pmatrix}, \begin{pmatrix}0\\0\end{pmatrix}, \begin{pmatrix}1\\1\end{pmatrix}\},$$

$$\{\begin{pmatrix}0\\1\end{pmatrix}, \begin{pmatrix}1\\1\end{pmatrix}, \begin{pmatrix}1\\0\end{pmatrix}\}, \{\begin{pmatrix}0\\1\end{pmatrix}, \begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}1\\0\end{pmatrix}\}, \{\begin{pmatrix}0\\1\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}, \begin{pmatrix}1\\0\end{pmatrix}\}, \{\begin{pmatrix}0\\1\end{pmatrix}, \begin{pmatrix}0\\0\end{pmatrix}, \begin{pmatrix}1\\0\end{pmatrix}\},$$

$$\{\begin{pmatrix}0\\1\end{pmatrix}, \begin{pmatrix}1\\1\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}\}, \{\begin{pmatrix}0\\1\end{pmatrix}, \begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}\}, \{\begin{pmatrix}0\\1\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}\}, \{\begin{pmatrix}0\\1\end{pmatrix}, \begin{pmatrix}0\\0\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}\},$$

$$\{\begin{pmatrix}0\\1\end{pmatrix}, \begin{pmatrix}1\\1\end{pmatrix}, \begin{pmatrix}0\\0\end{pmatrix}\}, \{\begin{pmatrix}0\\1\end{pmatrix}, \begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}0\\0\end{pmatrix}\}, \{\begin{pmatrix}0\\1\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}, \begin{pmatrix}0\\0\end{pmatrix}\}, \{\begin{pmatrix}0\\1\end{pmatrix}, \begin{pmatrix}0\\0\end{pmatrix}, \begin{pmatrix}0\\0\end{pmatrix}\}\Big\}$$

with cardinality $L = 32$. Therefore,

$$
\begin{aligned}
A := \{ & \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}, \\
& \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \\
& \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}, \\
& \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 1 & 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \\
& \begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}, \\
& \begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & 0 \\ 1 & 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \\
& \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}, \\
& \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 1 & 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \}
\end{aligned}
$$

with $|A| = L$.

D.3 Construction of $\mathcal{W}_\mathbf{n}$

This section illustrates a way to construct the set $\mathcal{W}_n$ defined in Section B of the paper when $\mathcal{X}_n$ is finite.

1. Rewrite each realisation of $(g_{\cdot 3}, \boldsymbol{x}) \in \{0,1\}^{n-1} \times \mathcal{X}_n$ by listing

(i) x_3;

(ii) g_{i3} $\forall i \in \mathcal{N}_{\cdot 3,n}$ such that $g_{i3} := 1$, disposing them with respect to x_i in ascending order; if there are $i, k \in \mathcal{N}_n, i \neq k$ such that $g_{i3} = g_{k3} := 1$ and $x_i = x_k$, any order is allowed;

(iii) g_{i3} $\forall i \in \mathcal{N}_{\cdot 3,n}$ such that $g_{i3} := 0$, disposing them with respect to x_i in ascending order; if there are $i, k \in \mathcal{N}_n, i \neq k$ such that $g_{i3} = g_{k3} := 0$ and $x_i = x_k$, any order is allowed;

(iv) x_i $\forall i \in \mathcal{N}_{\cdot 3,n}$ according to the disposition of players adopted in the previous steps.

2. For each row that is repeated once or more, delete all duplications from the second.

3. Collect the saved rows and rearrange each of them in its original order.

As an example, assume $n := 3$, $j = 3$ and $\mathcal{X}_3 := \{1, 0\}$. The set $\{0, 1\}^2 \times \mathcal{X}_3$ is reported in Table D.1. The realisations of $(G_{\cdot 3}, \boldsymbol{X})$ giving rise to the same inequalities, according to implication (ii) of Lemma 6 highlighted in Section 4.8 of the paper, have a symbol of the same colour together with the appropriate permutation of labels φ. Table D.2 reports in blue the rows of Table D.1 reordered.

Table D.1: Representation of $\{0,1\}^2 \times \mathcal{X}_3$.

G_{13}	G_{23}	X_1	X_2	X_3	$\varphi(1)$	$\varphi(2)$	$\varphi(3)$
1	1	1	1	1			
1	1	1	0	1			
1	1	1	1	0			
1	1	1	0	0			
1	1	0	1	1	2	1	3
1	1	0	0	1			
1	1	0	1	0			
1	1	0	0	0			
1	0	1	1	1			
1	0	1	0	1			
1	0	1	1	0			
0	0	1	0	0			
1	0	0	1	1			
1	0	0	0	1			
1	0	0	1	0			
1	0	0	0	0			
0	1	1	1	1	2	1	3
0	1	1	0	1	2	1	3
0	1	1	1	0	2	1	3
0	1	1	0	0	2	1	3
0	1	0	1	1	2	1	3
0	1	0	0	1	2	1	3
1	0	0	0	1	2	1	3
0	1	0	0	0	2	1	3
0	0	1	1	1			
0	0	1	0	1			
0	0	1	1	0			
0	0	1	0	0			
0	0	0	1	1	2	1	3
0	0	0	0	1			
0	0	0	1	0			
0	0	0	0	0			

Table D.2: Reordering the rows of Table D.1.

G_{12}	G_{32}	X_1	X_2	X_3					
1	1	1	1	1	1	1	1	1	1
1	1	1	1	0	1	1	1	1	0
1	1	1	0	1	0	1	1	1	1
1	1	1	0	0	0	1	1	1	0
1	1	0	1	1	1	1	1	1	0
1	1	0	1	0	1	1	1	0	0
1	1	0	0	1	0	1	1	1	0
1	1	0	0	0	0	1	1	0	0
1	0	1	1	1	1	1	0	1	1
1	0	1	1	0	1	1	0	1	0
1	0	1	0	1	0	1	0	1	1
1	0	1	0	0	0	1	0	1	0
1	0	0	1	1	1	1	0	0	1
1	0	0	1	0	1	1	0	0	0
1	0	0	0	1	0	1	0	0	1
1	0	0	0	0	0	1	0	0	0
0	1	1	1	1	1	1	0	1	1
0	1	1	1	0	1	1	0	0	1
0	1	1	0	1	0	1	0	1	1
0	1	1	0	0	0	1	0	0	1
0	1	0	1	1	1	1	0	1	0
0	1	0	1	0	1	1	0	0	0
0	1	0	0	1	0	1	0	1	0
0	1	0	0	0	0	1	0	0	0
0	0	1	1	1	1	0	0	1	1
0	0	1	1	0	1	0	0	1	0
0	0	1	0	1	0	0	0	1	1
0	0	1	0	0	0	0	0	1	0
0	0	0	1	1	1	0	0	1	0
0	0	0	1	0	1	0	0	0	0
0	0	0	0	1	0	0	0	1	0
0	0	0	0	0	0	0	0	0	0

Table D.3: Representation of $\mathcal{W}_3$.

G_{12}	G_{32}	X_1	X_2	X_3
1	1	1	1	1
1	1	1	1	0
0	1	1	1	1
0	1	1	1	0
1	1	1	0	0
0	1	1	0	0
1	1	0	1	1
1	1	0	1	0
0	1	0	1	1
0	1	0	1	0
1	1	0	0	1
1	1	0	0	0
0	1	0	0	1
0	1	0	0	0
1	0	0	1	1
1	0	0	1	0
0	0	0	1	1
0	0	0	1	0
1	0	0	0	0
0	0	0	0	0

D.4 Computation of $\hat{\mathbb{P}}_{\boldsymbol{C_{g_{\cdot 3},x}}}$

Consider any $i \in \mathcal{N}_n$ and $(\tilde{g}_{\cdot i}, \tilde{\boldsymbol{x}}) \in \{0,1\}^{n-1} \bigtimes \mathcal{X}_n$ such that $\exists$ a permutation φ with $\varphi(i) = 3$ generating $(\tilde{g}^{\varphi}_{\cdot\varphi(i)}, \tilde{\boldsymbol{x}}^{\varphi}) = (g_{\cdot 3}, \boldsymbol{x})$. By (C.12),

$$\mathbb{P}(G_{\cdot i} = \tilde{g}_{\cdot i}, \boldsymbol{X} = \tilde{\boldsymbol{x}}) = \mathbb{P}(G_{\cdot 3} = g_{\cdot 3}, \boldsymbol{X} = \boldsymbol{x}) \tag{D.1}$$

Consider $C_{\tilde{g}_{\cdot i}, \tilde{\boldsymbol{x}}, n} \subseteq \{0,1\}^{n-1} \bigtimes \mathcal{X}_n$. By (C.12) applied $\forall \varphi$ such that $\varphi(i) = i$,

$$\mathbb{P}((G_{\cdot i}, \boldsymbol{X}) \in C_{\tilde{g}_{\cdot i}, \tilde{\boldsymbol{x}}, n}) = |C_{\tilde{g}_{\cdot i}, \tilde{\boldsymbol{x}}, n}| \times \mathbb{P}(G_{\cdot i} = \tilde{g}_{\cdot i}, \boldsymbol{X} = \tilde{\boldsymbol{x}}) \tag{D.2}$$

Hence,

$$\begin{aligned}\mathbb{P}((G_{\cdot i}, \boldsymbol{X}) \in C_{\tilde{g}_{\cdot i}, \tilde{\boldsymbol{x}}, n}) &\underbrace{=}_{\text{(D.2)}} |C_{\tilde{g}_{\cdot i}, \tilde{\boldsymbol{x}}, n}| \times \mathbb{P}(G_{\cdot i} = \tilde{g}_{\cdot i}, \boldsymbol{X} = \tilde{\boldsymbol{x}}) \underbrace{=}_{\text{(D.1)}} |C_{\tilde{g}_{\cdot i}, \tilde{\boldsymbol{x}}, n}| \times \mathbb{P}(G_{\cdot 3} = g_{\cdot 3}, \boldsymbol{X} = \boldsymbol{x}) \\ &\underbrace{=}_{|C_{\tilde{g}_{\cdot i}, \tilde{\boldsymbol{x}}, n}| = |C_{g_{\cdot 3}, \boldsymbol{x}}|} |C_{g_{\cdot 3}, \boldsymbol{x}}| \times \mathbb{P}(G_{\cdot 3} = g_{\cdot 3}, \boldsymbol{X} = \boldsymbol{x}) \underbrace{=}_{\text{(C.13)}} \mathbb{P}((G_{\cdot 3}, \boldsymbol{X}) \in C_{g_{\cdot 3}, \boldsymbol{x}})\end{aligned} \tag{D.3}$$

Let

$$\hat{\mathbb{P}}_{C_{g_{\cdot 3}, \boldsymbol{x}}} := \frac{1}{n} \sum_{i=1}^{n} \mathbb{1}((G_{\cdot i}, \boldsymbol{X}) \in C_{\tilde{g}_{\cdot i}, \tilde{\boldsymbol{x}}, n}) \tag{D.4}$$

From (D.3), $\hat{\mathbb{P}}_{C_{g\cdot 3},\boldsymbol{x}}$ is an unbiased estimator for $|C_{g\cdot 3},\boldsymbol{x}| \times \mathbb{P}(G_{\cdot 3} = g_{\cdot 3}, \boldsymbol{X} = \boldsymbol{x}) = \mathbb{P}((G_{\cdot 3}, \boldsymbol{X}) \in C_{g\cdot 3},\boldsymbol{x})$ and does not depend on assigned labels.

An algorithm to compute $\hat{\mathbb{P}}_{C_{g\cdot 3},\boldsymbol{x}}$ is the following:

1. Rewrite $(g_{\cdot 3}, \boldsymbol{x})$ by listing
   (i) x_3;
   (ii) g_{h3} $\forall h \in \mathcal{N}_{\cdot 3,n}$ such that $g_{h3} := 1$, disposing them with respect to x_h in ascending order; if there are $h, k \in \mathcal{N}_n, h \neq k$ such that $g_{h3} = g_{k3} := 1$ and $x_h = x_k$, any order is allowed;
   (iii) g_{h3} $\forall h \in \mathcal{N}_{\cdot 3,n}$ such that $g_{h3} := 0$, disposing them with respect to x_h in ascending order; if there are $h, k \in \mathcal{N}_n, h \neq k$ such that $g_{h3} = g_{k3} := 0$ and $x_h = x_k$, any order is allowed;
   (iv) x_h $\forall h \in \mathcal{N}_{\cdot 3,n}$ according to the disposition adopted in the previous steps.
2. Call A_3 the obtained row of values.
3. $\forall i \in \mathcal{N}_n$ in the dataset, list
   (i) x_i;
   (ii) g_{hi} $\forall h \in \mathcal{N}_{\cdot i,n}$ such that $g_{hi} =: 1$, disposing them with respect to x_h in ascending order; if there are $h, k \in \mathcal{N}_n, h \neq k$ such that $g_{hi} = g_{ki} := 1$ and $x_h = x_k$, any order is allowed;
   (iii) g_{hi} $\forall h \in \mathcal{N}_{\cdot i,n}$ such that $g_{hi} = 0$, disposing them with respect to x_h in ascending order; if there are $h, k \in \mathcal{N}_n, h \neq k$ such that $g_{hi} = g_{ki} := 0$ and $x_h = x_k$, any order is allowed;
   (iv) x_h $\forall h \in \mathcal{N}_{\cdot i,n}$ according to the disposition adopted in the previous steps.
4. Call A_i the obtained row of values $\forall i \in \mathcal{N}_n$.

Hence,

$$\hat{\mathbb{P}}_{C_{g\cdot 3},\boldsymbol{x}} := \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}\{A_i = A_3\}$$

D.5 Computation of $H^l_{C_{g\cdot 3,\boldsymbol{x}}}(\theta)$ and $H^u_{C_{g\cdot 3,\boldsymbol{x}}}(\theta)$

The computation of $H^l_{C_{g\cdot 3,\boldsymbol{x}}}(\theta)$ and $H^u_{C_{g\cdot 3,\boldsymbol{x}}}(\theta)$ can be done via the simple frequency simulator proposed by McFadden (1989) and Pakes and Pollard (1989). Specifically, $\forall i \in \mathcal{N}_n$, R_M[40] realisations of $\epsilon_{\cdot i}$ are randomly drawn from its distribution. Let $\epsilon_{\cdot i,r}$ denote the random vector for the rth draw $\forall i \in \mathcal{N}_n$. Hence,

$$\hat{H}^l_{C_{g\cdot 3,\boldsymbol{x}}}(\theta) := \frac{1}{R_M \times n}\sum_{r=1}^{R_M}\sum_{i=1}^{n} \mathbb{1}(\text{all outcomes of the } \textit{section } i \textit{ game} \text{ fall in } C_{\tilde{g}\cdot i,\tilde{\boldsymbol{x}},n}) \tag{D.5}$$

and

$$\hat{H}^l_{C_{g\cdot 3,\boldsymbol{x}}}(\theta) := \frac{1}{R_M \times n}\sum_{r=1}^{R_T}\sum_{i=1}^{n} \mathbb{1}(\text{at least one outcome of the } \textit{section } i \textit{ game} \text{ falls in } C_{\tilde{g}\cdot i,\tilde{\boldsymbol{x}},n}) \tag{D.6}$$

In the empirical application, $R_T = 100$. In order to establish the value of the indicators function the algorithm illustrated in Appendix D.4 can be employed.

D.6 Inference procedure

This section illustrates how to obtain the test statistic $S_M(\theta)$ and the critical value $\hat{c}_{M,1-\alpha}(\theta)$ when constructing a $1-\alpha$ confidence region for each $\theta \in \Theta^o_{\cdot 3}$ following AS (2010). After having designed a grid of candidate parameter values[41], for each θ in the grid:

(i) Compute $b_m(\theta)$.

(ii) Compute $\bar{b}_M(\theta) := \frac{1}{M}\sum_{m=1}^{M} b_m(\theta)$. Let $\bar{b}_{k,M}(\theta)$ denote its kth element and $\tilde{\sigma}_{k,M}$ a consistent estimator of the asymptotic standard deviation of $\sqrt{M}\bar{b}_{k,M}(\theta)$.

The computation of $\tilde{\sigma}_{k,M}$ is explained in what follows in more details under the i.i.d. sampling scheme. Let

$$b^l_{C_{g\cdot 3,\boldsymbol{x}},m}(\theta) := \hat{\mathbb{P}}_{C_{g\cdot 3,\boldsymbol{x}},m} - H^l_{C_{g\cdot 3,\boldsymbol{x}},m}(\theta)$$

and

$$b^u_{C_{g\cdot 3,\boldsymbol{x}},m}(\theta) := H^u_{C_{g\cdot 3,\boldsymbol{x}},m}(\theta) - \hat{\mathbb{P}}_{C_{g\cdot 3,\boldsymbol{x}},m}$$

[40] The subscript M reminds that R should increase to infinity with sample size to avoid not vanishing simulations errors (CT, 2009).

[41] See Appendix D.7.

It can be noticed that the asymptotic variances of $\sqrt{M}\sum_{m=1}^{M} b^u_{C_{g.3,\boldsymbol{x}},m}(\theta)$ and $\sqrt{M}\sum_{m=1}^{M} b^l_{C_{g.3,\boldsymbol{x}},m}(\theta)$ coincide with the asymptotic variance of $\sqrt{M}\sum_{m=1}^{M} \hat{\mathbb{P}}_{C_{g.3,\boldsymbol{x}},m}$, which, under the i.i.d. sampling scheme, is equivalent to the variance of $\hat{\mathbb{P}}_{C_{g.3,\boldsymbol{x}},m}$. Let

$$\hat{\sigma}_{C_{g.3,\boldsymbol{x}},M} := \sqrt{\frac{1}{M}\sum_{m=1}^{M}(\hat{\mathbb{P}}_{C_{g.3,\boldsymbol{x}},m} - \frac{1}{M}\sum_{m=1}^{M}\hat{\mathbb{P}}_{C_{g.3,\boldsymbol{x}},m})^2}$$

Let

$$\tilde{\sigma}_{C_{g.3,\boldsymbol{x}},M} := \max\{\tau_M, \hat{\sigma}_{C_{g.3,\boldsymbol{x}},M}\}$$

where τ_M is a sequence of positive constants converging to zero as M goes to infinity at an appropriate rate.

$\tilde{\sigma}_{C_{g.3,\boldsymbol{x}},M}$ is used as a consistent estimator of the asymptotic standard deviations of $\sqrt{M}\sum_{m=1}^{M} b^u_{C_{g.3,\boldsymbol{x}},m}(\theta)$ and $\sqrt{M}\sum_{m=1}^{M} b^l_{C_{g.3,\boldsymbol{x}},m}(\theta)$ because $\hat{\sigma}_{C_{g.3,\boldsymbol{x}},M}$ can be zero. This is undesirable as the test statistic requires the division of $\sqrt{M}\frac{1}{M}\sum_{t=1}^{T} b^u_{C_{g.3,\boldsymbol{x}},m}(\theta)$ and $\sqrt{M}\frac{1}{M}\sum_{t=1}^{T} b^l_{C_{g.3,\boldsymbol{x}},m}(\theta)$ by a consistent estimate of their asymptotic standard deviations.

The choice of the rate of convergence to zero of τ_M might affect the asymptotic power properties of the GMS test. Further investigations are in progress. For the purposes of this work, in order to set τ_M, several Monte Carlo simulations were performed comparing the behaviour of the test statistic $S_M(\theta_0)$ as $M \to \infty$ when θ_0 is the true parameter vector, $\tau_M = \frac{1}{M^3}, \frac{1}{M^2}, \frac{1}{M}, \frac{\log M}{M}, \frac{1}{\sqrt{M}}$. $\tau_M = \frac{\log(M)}{M}$ is proposed for use because it is the only value among those examined that prevents $S_M(\theta_0)$ from blowing up as $M \to \infty$ and correctly allows its probability distribution function to shrink around zero as $M \to \infty$.

(iii) Compute the test statistic $S_M(\theta) := \sum_k \left(\min\left\{\frac{\sqrt{M}\bar{b}_{k,M}(\theta)}{\tilde{\sigma}_{k,M}}, 0\right\}\right)^2$.

(iv) For each k, compute $\xi_{k,M}(\theta) := \frac{1}{\sqrt{\log(M)}}\sqrt{M}\frac{\bar{b}_{k,t}(\theta)}{\tilde{\sigma}_{k,M}}$.

(v) For each k, choose the hard threshold $\zeta_{k,M}(\theta) := \begin{cases} 0 & \text{if } \xi_{k,M}(\theta) \le 1 \\ \infty & \text{otherwise} \end{cases}$

(vi) Draw with replacement R bootstrap samples i.i.d. over r. In the empirical application, $R = 120$.

(vii) For $r = 1, \ldots, R$

(a) Repeat steps (i) and (ii) and obtain $\bar{b}^{\star}_{k,M,r}(\theta)$ and $\tilde{\sigma}^{\star}_{k,M,r}$ for each k.

(b) Compute $L_{M,r}(\theta) := \sum_k \left(\min\left\{ \frac{\sqrt{M}(\bar{b}^{\star}_{k,M,r}(\theta) - \bar{b}_{k,M}(\theta))}{\hat{\sigma}^{\star}_{k,M,r}} + \zeta_{k,M}(\theta), 0 \right\} \right)^2$.

(viii) Take the GMS critical value, $\hat{c}_{M,1-\alpha}(\theta)$, as the $(1-\alpha)$ sample quantile of $\{L_{M,r}(\theta)\}_{r=1}^{R}$.

(ix) Reject if $S_M(\theta) > \hat{c}_{M,1-\alpha}(\theta)$.

Hence, the $1-\alpha$ confidence region for each $\theta \in \Theta^{o}_{:3}$ is

$$CS_M = \{\theta \in \Theta \text{ such that } S_M(\theta) \leq \hat{c}_{M,1-\alpha}(\theta)\}$$

D.7 Construction of the initial grid of parameters

One difficulty with conducting inference on sets is scanning over a multi-dimensional parameter space. In practice, what the researcher can do is exploring the parameter space around the global minimum of $S_M(\theta)$ in some rational way. For the empirical application, the slice sampling method of Neal (2003) used by Kline and Tamer (2016) was employed. The procedure is as follows:

(i) List many starting values for θ, one of which has all entries equal to zero, others are constructed using the results of simple probits.

(ii) From each starting value, minimise $S_M(\theta)$ running a global optimisation algorithm in Matlab; specifically, a pattern search algorithm (*psearch*) with different polling strategies and a genetic algorithm (*ga*) were used.

(iii) Let s be the global minimum of $S_M(\theta)$; s is not exactly zero because of the complex numerical problem involved in the computation of $S_M(\theta)$.

(iv) Save one vector of parameters solving $S_M(\theta) = s$ and call it θ_s.

(v) Run the pre-implemented slice sampling routine in Matlab (*slicesample*) setting $\mathbb{1}\{S_M(\theta) = s\}$ as the un-normalized density and θ_s as the starting value; save the results of each iteration in the course of the algorithm.

(vi) Look at the values that $S_M(\theta)$ has taken over all the parameters used in the course of the algorithm and draw a random sample of 500 points. This sample is the initial grid of parameters.

To guarantee a better exploration of all relevant regions of the parameter space, steps (iv), (v) and (vi) were repeated for each vector of parameters found in step (ii) and solving $S_M(\theta) = s$, and the grids obtained from step (vi) were merged. Moreover, robustness checks on the number of random draws from the un-normalised density were conducted.

Alternative procedures are the simulated annealing method proposed by CT (2009) and the differential evolution algorithm described by BMM (2011).

D.8 Legal framework for Italian joint-stock companies under the Articles 2380/2409-septies of the Italian Civil Code

As to what concerns the present work: (i) management is the responsibility of a board of directors (*Consiglio di Amministrazione*), whose size is freely chosen by shareholders; this legitimates the lack of restrictions on boards' sizes; (ii) directors are elected by the shareholders' meeting (*Assemblea Ordinaria*); hence, as the company's will is identified with shareholders' interests, it is legitimate to suppose that the board's will coincides with the firm's will; (iii) the board of directors can delegate its executive duties to one or more of its members; a delegatee is called *Amministratore Delegato*[42]; the purpose of the mandate is to simplify the decision process[43]; if the mandate is conferred, then delegators have monitoring and advising duties regarding delegatees' conduct on the basis of the information received during board meetings; to that extent, delegatees have to report to the board with a frequency determined by the company's statute, and, in any case, at least every six months; moreover, delegators can ask delegatees to provide the board with any information related to the management of the company; this legitimates to distinguish between executive and non-executive board members, hence conferring directionality to links and justifying firms' payoffs specified as (4.12).

D.9 Data construction and cleaning

In order to correctly extract and merge the information from the Registro Imprese, each firm was uniquely identified by combining its (i) Chamber of Commerce's province, (ii) R.E.A[44] code and (iii) tax code. The R.E.A. code is a number assigned to each company when enrolling at the Registro Imprese. The tax code is a numeric code of 16 digits.

Each board member was uniquely identified by her tax code, which is an alphanumeric code

[42]An *Amministratore Delegato* is to an Italian joint stock company what a CEO is to a U.S. company.

[43]Various powers cannot be delegated, e.g., the issue of convertible bonds, the draft of the balance sheet, the increase of share capital or its decrease because of losses, the draft of projects for merging or demerging. These exceptions are not considered in the present analysis.

[44]R.E.A. stands for *Repertorio Economico Amministrativo.*

of 16 characters, similar to the Social Security Number in the United States or the National Insurance Number in the United Kingdom.

In order to merge the information from the Registro Imprese with that from the Cerved database, firms' tax codes were used.

Moreover, sectors composed of 1 or 2 firms and the sector *Holdings* (ATECO 2002 code: 74.15.0) were dropped.

D.10 Definitions of some network measures

For the purpose of measuring the degree of cohesion, the density of a network $\boldsymbol{G}$ is the fraction between the total number of links in the network and the total number of possible links

$$\frac{\sum_{i,j\in\mathcal{N}_N, i\neq j} G_{ij}}{N(N-1)}$$

with $D \in [0,1]$. A higher density denotes tighter relations between firms. It can be observed that in Table 2 the density of constructed networks varies between 0 and 0.333 and has an average value across industries of 0.005. Another important network measure is the average degree of a node which tells how many links a node has on average with other nodes

$$\frac{1}{N} \sum_{i,j\in\mathcal{N}_N} G_{ij}$$

It can be observed that in Table 2 the average degree of constructed networks varies between 0 and 1 and has an average value (approximated to the nearest integer) across industries of 0. Such a low average value is in line with the low average density commented above. In the same spirit, the percentage of isolated nodes of constructed networks, computed as

$$100 \times \frac{1}{N} \sum_{i\in\mathcal{N}_N} \mathbb{1}\{G_{ij} = G_{ji} = 0 \ \forall j \neq i \in \mathcal{N}_N\}$$

varies between 33.333% and 100% with an average value across industries of 97.665%. Lastly, the total number of links of constructed networks, computed as

$$\sum_{i,j\in\mathcal{N}_N, i\neq j} G_{ij}$$

varies between 0 and 6 with an average value (approximated to the nearest integer) across industries of 0.