UvA-DARE (Digital Academic Repository)

Risk aversion and social networks

Kovářík, J.; van der Leij, M.J.

DOI

[10.1515/rne-2014-0029](https://doi.org/10.1515/rne-2014-0029)

Publication date

2014

Document Version

Submitted manuscript

Published in

The Review of Network Economics

[Link to publication](#)

Citation for published version (APA):

Kovářík, J., & van der Leij, M. J. (2014). Risk aversion and social networks. *The Review of Network Economics*, 13(2), 121-155. <https://doi.org/10.1515/rne-2014-0029>

General rights

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

Risk aversion and social networks*

Jaromír Kovářik^{†‡}

Marco J. van der Leij^{§¶||}

November 7, 2014

Abstract

This paper first investigates empirically the relationship between risk aversion and social network structure in a large group of undergraduate students. We find that risk aversion is strongly correlated to local network clustering, that is, the probability that one has a social tie to friends of friends. We then propose a network formation model that generates this empirical finding, suggesting that locally superior information on benefits makes it more attractive for risk averse individuals to link to friends of friends. Finally, we discuss implications of this model. The model generates a positive correlation between local network clustering and benefits, even if benefits are distributed independently *ex ante*. This provides an alternative explanation of this relationship to the one given by the social capital literature. We also establish a linkage between the uncertainty of the environment and the network structure: risky environments generate more clustered and more unequal networks in terms of connectivity.

Keywords: Network formation, risk aversion, uncertainty/riskiness, clustering coefficient, degree distribution, local/global search

JEL codes: D81, D85

*We thank Michèle Belot, Pablo Brañas, Arthur Campbell, Pascal Courty, María Paz Espinosa, Marcel Fafchamps, James Fowler, Edoardo Gallo, Jayant Ganguli, Sanjeev Goyal, Daan in 't Veld, Matthew Jackson, Friederike Mengel, Paolo Pin, Andrea Prat, Giorgio Topa, Fernando Vega-Redondo, Adrien Vigier, Andrew Zeitlin and audiences at the European University Institute, University of Amsterdam, University of Cambridge, University of Málaga, University of the Basque Country, University of Oxford, University of Warwick, SAE 2009 in Zaragoza, ICFSN 2009 in Paris and EAA-ESEM 2009 in Barcelona. Marco van der Leij would like to thank The Netherlands Organisation for Scientific Research (NWO FinComplexity) and Jaromír Kovářik to the Spanish *Ministerio de Economía y Competitividad* (ECO 2012-31626, ECO 2012-35820), the Government of the Basque Country (IT783-13), and GACR (14-22044S) for financial support. Views expressed are those of the authors and do not necessarily reflect official positions of De Nederlandsche Bank.

[†]Dpto. Fundamentos del Análisis Económico I & Bridge, University of the Basque Country, Av. Lehendakari Aguirre 83, 48015 Bilbao, Spain. Email: jaromir.kovarik@ehu.es

[‡]CERGE-EI, a joint workplace of Charles University in Prague and the Economics Institute of the Academy of Sciences of the Czech Republic, Politických vězňů 7, 111 21 Prague, Czech Republic.

[§]*Corresponding author.* CeNDEF, University of Amsterdam, Valckeniersstraat 65-67, 1018 XE Amsterdam, The Netherlands. E-mail: m.j.vanderleij@uva.nl

[¶]Tinbergen Institute, Gustav Mahlerplein 117, 1082 MS Amsterdam, The Netherlands.

^{||}Research Department, De Nederlandsche Bank, Westeinde 1, 1017 ZN Amsterdam, The Netherlands.

1 Introduction

Social networks play an important role in many socio-economic settings, and particular network architectures influence individual and global economic outcomes.¹ Moreover, there is a large heterogeneity in network positions. Links in a social network are often unequally distributed with few members having many links and central positions, whereas the majority of the nodes lie in the periphery (Goyal et al. 2006). The variation in network positions explains part of the variation in economic outcomes across individuals, groups, and even organizations or countries (Snyder and Kick, 1979; Granovetter, 1985; Burt, 1995; Lin, 1999; Zaheer and Bell, 2005).

This raises the question of how this variety in network positions arises in the first place. One possibility is that ex-ante similar agents endogenously form networks that are ex-post very unequal, a view that has received most of the attention.² An alternative explanation would be that the population was already heterogeneous ex-ante, and that inequality in network positions reflects this heterogeneity. Indeed, Fowler et al. (2009) report that a surprisingly large amount of heterogeneity in friendship networks can be attributed to inheritable traits.³

Such importance of inheritable traits on social network position generates the question of what type of traits drive particular network position. In this paper, we focus on one trait of particular importance to economics, risk aversion, and analyze whether differences in risk attitudes are related to differences in network positions. For this purpose, we consider a unique data set among undergraduate students of the University of Granada, Spain, for which both risk aversion and their social network were elicited. As far as we know, we are the first to analyze this relationship.

We find that risk aversion is not related to the number of friendship nominations a student states or receives. However, risk aversion is strongly related to node's transitivity or clustering coefficient, that is the probability that a friend of her friend is also her friend. In particular, we find that individuals with high aversion to risk tend to move in closely knit network neighborhoods. This relationship cannot be explained by gender differences in network formation and risk aversion.

¹See Goyal (2007) and Jackson (2008) for reviews.

²See e.g. Jackson and Wolinsky (1996) or Goyal and Vega-Redondo (2007) for the models generating star networks, and for example Buskens and Van de Rijt (2008) for equal networks. Goyal and Joshi (2006) specify the conditions on the utility function leading to unequal networks.

³In particular, Fowler et al. (2009) found that “genetic factors account for 46% (...) of the variation in in-degree (how many times a person is named as a friend), but heritability of out-degree (how many friends a person names) is not significant (...). In addition, node transitivity is significantly heritable, with 47% (...) of the variation explained by differences in genes (p.1720).”

We then propose a simple theoretical model that explains the relation between risk aversion and node transitivity and analyze the theoretical implications of this relationship. The model mimics the formation of links in a networking event, in which agents do not know each other initially but learn about one another through social interactions. In the model, agents enter the network one by one and each incoming agent has to decide with whom of the already present agents she wishes to interact. Initially, the newcomer is uninformed about the benefits of interacting with any agent and thus sets up the first link randomly. However, we assume that, after creating this first link, the agent not only learns about the benefits of interaction with this initial friend, but also about the benefits of interacting with the friends of this friend. Thus, the network acts as a local conduit of private information of interaction benefits. More precise information about the friends of friends than about a random agent outside the current social circle leads risk-averse individuals to create connections to friends of friends and the more risk-averse a newcomer is, the stronger this tendency. Since creating a link to a friend of a friend implies the creation of a transitive triple, this process naturally generates a positive relationship between risk aversion and clustering coefficient.

The model has two interesting economic implications. First, an agent's clustering coefficient is *positively* related to agent's income within this model, even though linking benefits are ex-ante distributed independently and even though risk aversion is negatively related to income and positively related to clustering coefficient. The reason for this relationship is what one may call a good neighborhood effect: newcoming agents who learn that friends of their friends deliver high benefits are more likely to link to them, leading to high clustering coefficient. Our model thus provides a new argument for why highly clustered neighborhoods are likely to have high payoffs; some members would have left the cluster if the benefits wouldn't have been so high. This argument is new and radically different from the standard arguments explaining a relationship between clustering coefficient and payoffs based on the idea of social capital benefits on (fixed) networks (Granovetter, 1985; Coleman, 1988). These arguments show that social capital benefits are not necessary to generate such relationship and, conversely, an empirical finding showing a relationship between clustering coefficient and income is not sufficient to prove that social capital benefits exist.

Second, we analyze the impact of shifts in the distribution of benefits from linking on inequality in network positions. We find that, in this model, a mean shift of the benefit distribution does not impact the distribution of network positions, while a mean-preserving spread, that is, more uncertainty, leads to a higher clustering coefficient and more inequality with respect to agents'

in-degree. The reason is that more uncertainty leads agents to create more links locally, to friends of friends, which automatically increases network clustering. Moreover, linking to friends of friends benefits agents who are already well connected, a phenomenon known as preferential attachment, and this phenomenon leads to unequal ‘power-law tail’ in-degree distributions (Barabasi and Albert, 1999; Jackson and Rogers, 2007). Since Jackson and Rogers (2007) show in a similar setup that inequality in network connectivity lowers efficiency, our results indicate that more uncertainty affects negatively the social welfare.

Our paper is the first that relates the variation in risk aversion to variation in social network structure.⁴ Regarding network formation, some literature discusses the role of heterogeneity in network formation theory, e.g. Galeotti et al. (2006). However, no efforts have been made to explain the heterogeneity in network positions. Our network formation model can be considered an extension of the model by Jackson and Rogers (2007). They already show that a higher probability to link to a friend of a friend, instead of a random agent, leads to more inequality in network degree and income. Our model gives a theoretical underpinning for the probability to link a friend of friend, relating this probability to risk aversion and the uncertainty of the environment.⁵ Campbell (2014) provides an alternative, signalling-based justification of why people might link to friends of friends in Jackson and Rogers’ (2007) framework.

The paper is structured as follows. In the next section, we show empirically that risk aversion is positively related to node transitivity. Section 3 proposes a theoretical model that generates this relation. Next, in Sections 4 and 5 we derive the theoretical implications of the model. Section 6 discusses some extensions and Section 7 concludes.

2 Empirical Analysis

In this section, we report an elicited real-life social network, and relate its network characteristics to the level of risk aversion of individuals. The data presented here were collected within a sequence of surveys and experiments in the spring semester of 2005 at the University of Granada,

⁴Our approach differs substantially from the literature on risk-sharing network formation (e.g. Bramoullé and Kranton, 2007, Bloch et al., 2008), but we provide several interesting implications for this literature. In particular, we show that they should consider both the heterogeneity in risk aversion and the prevailing uncertainty of the context, in which the risk-sharing network form.

⁵The model of Jackson and Rogers (2007) has been extended in other directions by many others. See e.g. Bramoullé et al. (2012), Atalay (2013), or Vigier (2014).

Spain.⁶ First-year undergraduate students of Economics were invited to participate in the sessions. In order to stimulate participation, subjects were rewarded with classroom points that served to increase the final grade in the course of Microeconomics I. Subjects were informed that the number of points obtained during the 5 sessions in which they would participate, contributed to their final grade in the course in the following way: the student who obtained the highest number of points would add three extra points (out of ten) to her final grade. Other subjects' grade depended on how close their performance was to the winner's. Subjects were not informed of their own and others' performances in any of the sessions and tasks until all sessions had ended. In this document, we combine the data from two particular sessions: one in which the network was elicited and one in which risk aversion was elicited.

This data set is the only one we know of that contains both data on risk aversion and on network structure and, therefore, gives us a unique opportunity to analyze the relation between risk aversion and positions in a social network.

2.1 Network elicitation

All first-year students (more than 300) were invited to participate in a network elicitation survey and a subsequent game serving to elicit their sharing preferences. The aim of the survey to elicit relationships among subjects was to be able to study experimentally the effects of social distance on sharing behavior, rather than the analysis of the actual underlying network structure. We refer to Kovářík et al. (2012) for details.

The elicitation protocol was very simple. Subjects had to write the names of their friends in class on a piece of paper. There was no restriction on the number of friends people could have named. On the other hand, to incentivize the naming of close friendships a special emphasis was put in the instructions to make clear to the subjects that in the next stage of the experiment they would be given a chance to benefit only one of the friends they named. To this aim, the instructions stated: “the more friends you will name, the lower the probability to be able to reward a particular friend of yours.”

Using this elicitation protocol, students named, on average, 2.03 friends. Friendship nominations were not always going in both ways; in fact, only 65% of all friendship nominations from A to B were reciprocated by a nomination from B to A . Hence, the friendship network, g_D ,

⁶The sequence of sessions carried out by the pool under scrutiny was as follows: a dictator game (March), a GRE-type math test (beginning of April), risk aversion elicitation (end of April), network elicitation jointly with a two-stage variation of the dictator game (May), and a session containing a Traveler's dilemma and self-confidence questionnaire (June).

Figure 1: The network of friendship relationships between 256 undergraduate students at the University of Granada. An arc from node i to j indicates that student i named student j as a friend. The colors of the nodes denote the amount of risk aversion (white: strongly risk loving. red: strongly risk averse. grey: risk aversion omitted).

is directed. As a robustness check and because friendships are usually considered mutual, we also analyze the implied undirected network, g_U , in which a link between A and B is in place if either A names B or B names A . In the undirected network, each student has on average 2.75 links. Figure 1 provides a global view of the elicited social network of the 256 individuals, who participated in the network elicitation part of the experiment.

For the directed network, we focus on the same network statistics as in Fowler et al. (2009): in-degree, out-degree, fraction of transitive triples, and betweenness. In-degree of individual i gives the number of individuals $j \neq i$ that name i as a friend of theirs; out-degree equals the number of individuals i names as her friends. The fraction of transitive triples measures the fraction of triples ijk , for which i has an outgoing link to j and j has an outgoing link to $k \neq i$ (i.e. i names j as a friend, and j names k) and agent i also has an outgoing link to k . Formally, the fraction of transitive triples of agent i is:

$$C_i(g) = \frac{\sum_{j \neq i; k \neq i, j} g_{ij} g_{jk} g_{ik}}{\sum_{j \neq i; k \neq i, j} g_{ij} g_{jk}}. \quad (1)$$

Note that $C_i(g)$ is only well defined if i names at least 1 friend, and this friend names also at least one friend (apart from i). The fraction of transitive triples of network g is:

$$C(g) = \frac{\sum_{i,j \neq i; k \neq i,j} g_{ij} g_{jk} g_{ik}}{\sum_{i,j \neq i; k \neq i,j} g_{ij} g_{jk}}. \quad (2)$$

The betweenness of agent i measures the number of shortest paths between other agents, of which i is part of. Formally,

$$B_i(g) = \sum_{j \neq k; j, k \neq i} \frac{\tau_{j,k}^i(g)}{\tau_{j,k}(g)} \quad (3)$$

where $\tau_{j,k}^i(g)$ is the number of shortest paths from j to k in network g that pass through node i , and $\tau_{j,k}(g)$ is the total number of shortest paths from j to k .⁷

The first two variables measure the connectivity of individuals, while the fraction of transitive triples reflects the density or interconnectivity of agents' neighborhoods. Finally, betweenness measures how central a node is in the whole network architecture. For the undirected network, we use degree, clustering coefficient, and betweenness. Degree is the number of links agent i has, independently of the direction and reciprocity of the link. Clustering coefficient is the corresponding measure for the fraction of transitive triples for an undirected network, defined as:

$$C_i(g) = \frac{\sum_{j \neq i; k \neq i,j} g_{ij} g_{jk} g_{ik}}{\sum_{j \neq i; k \neq i,j} g_{ij} g_{ik}}. \quad (4)$$

This measure is well defined as long as an agent has at least two links. Betweenness is defined as in (3) but ignoring the direction of the link.

Table 1 shows the summary statistics of all the considered network variables. The connectivity of the directed network (in terms of degree) is, by definition, below that of the undirected network. Regarding the transitivity/clustering coefficient, it is .31 and .397, two orders of magnitude higher than that of a random network, in which the expected coefficient would be $1/(n-1) = .004$. Note that transitivity and clustering coefficient is not defined for all agents, leading to a loss of observations.

⁷ A path from j to k is defined as a sequence of distinct nodes, $\{j_0, j_1, \dots, j_{d-1}, j_d\}$ with $j_0 \equiv j$ and $j_d \equiv k$, such that, for each $\ell \in 1, \dots, d$, there is an outgoing link from $j_{\ell-1}$ to j_ℓ . The path length of this path is d . The shortest paths between j and k are then the paths that have the shortest path length.

	Obs.	Mean	St.Dev.	Max
Directed network				
indegree	256	2.031	1.671	8
outdegree	256	2.031	1.444	6
trans.triples	188	0.310	0.326	1
betweenness	256	243.8	513.3	3146.3
Undirected network				
degree	256	2.750	1.887	9
clustering	181	0.397	0.360	1
betweenness	256	425.1	760.5	5250

Table 1: Summary statistics for the friendship network among 256 undergraduate students at the University of Granada. Fraction of transitive triples and clustering are not well defined for some nodes and therefore contain less than 256 observations.

2.2 Risk aversion elicitation

The risk aversion test was performed in a session different from the network elicitation one. There were 187 subjects who coincided in both sessions, allowing us to relate their attitudes toward risk to their position in the class network. To elicit risk attitudes, a variation of the standard Holt and Laury’s (2002) protocol was used. In the first stage, subjects were asked to choose between a two-player game and a fixed amount of money. The presented game represented *de facto* a lottery where a rational player obtains either 4 or 8 experimental points both with the same probability $1/2$.⁸

Each subject faced a series of 10 decisions, in which he had to decide between the game or a fixed amount of money. Decision 1 was to decide between either 1 experimental point for sure or to play the game; Decision 2 between 2 points for sure or the game, and so on, until Decision 10 between 10 points for sure or the game. Our measure of risk aversion is, then, the number of times a student chose to take the fixed amount. This measures ranges between 0 and 10, with 10 being extremely risk averse.

⁸The actual setup was a bit more involved. The two-player game in the second stage was given by the following table.

	A	B
A	4,4	8,3
B	3,8	0,0

Students that chose to play the game were told (i) they would be the row players and (ii) the play of the column player would be simulated by a computer, which would play A with 50% probability or B with the same probability. The aim of this exposition was to somehow check the rationality of subjects, as playing A always leads to a higher payoff. In fact, only one subject, who was removed from the analysis, chose to play B. Moreover, given that the column player was a computer, neither social nor efficiency concerns were at work here. See Brañas-Garza et al. (2004) for more details concerning the experimental protocol.

Figure 2: A histogram of the measure of risk aversion.

Figure 2 shows a histogram of the obtained measure of risk aversion. One direct observation is the vast amount of heterogeneity in the measure. It is the objective in the paper to relate this heterogeneity in risk aversion to network positions.

Unfortunately, the students' choices are so various that it puts into question the rationality of students. Since subjects could earn at least 4 experimental points by playing the game, any decision for a fixed amount equal or smaller than 4 would be irrational. Similarly, since the maximum payoff from the game is 8, any decision against a fixed amount of 8 or more cannot be rational. Translated to our measure of risk aversion, it implies that rational players should have a risk aversion measure between 3 and 6. Unfortunately, Figure 2 shows that quite a number of students, in fact 45 out of 187, did make 'irrational' decisions in the risk elicitation experiments. These students may have been confused by the set up, leading to rather random choices. Since such choices would only add noise and attenuate the power to detect any relation, we analyze the data including and excluding 'irrational' observations.

2.3 Results

Table 2 lists the correlation coefficients of the variable measuring risk aversion with the network measures in the directed and undirected friendship networks. In the former case, the correlations with in-, out-degree, and betweenness are calculated for all students participating in both the network and risk attitude elicitation, but the correlation of transitivity with risk attitudes is calculated only for those subjects, who have at least 1 friend and whose friends have at least 1 friend as well. The table shows that there is no significant relation between risk attitudes of individuals and the connectivity/centrality measures, while risk aversion and transitivity are significantly correlated. Significance is particularly strong when we focus on the students whose choices in the risk elicitation game could be considered rational. A similar picture emerges when we consider the undirected friendship network, in which a link is declared to exist between A and B if either A names B or B names A . Only clustering coefficient is significantly correlated to risk aversion.

These correlations suggest that heterogeneity in risk attitudes is related to heterogeneity in transitivity. To assess the robustness of these findings, we formally control for two important empirical regulations. First, women are typically more averse to risk than men (Croson and Gneezy, 2009) and the detected correlations might thus be due to gender, rather than risk aversion. Second, socially generated networks exhibit important large correlations between different

Whole sample (187 obs.)							
	Directed network				Undirected network		
	indegree	outdegree	tr.triple	between.	degree	clustering	between.
risk aversion							
correlation	0.035	0.002	0.135*	-0.068	-.003	0.226***	-0.052
<i>p</i> -value	0.627	0.978	0.097	0.352	0.964	0.006	0.481
Sample of risk aversion between 3 and 6 only (142 obs.)							
	Directed network				Undirected network		
	indegree	outdegree	tr.triple	between.	degree	clustering	between.
risk aversion							
correlation	0.012	0.014	0.340***	-0.010	-0.006	0.383***	-0.137
<i>p</i> -value	0.882	0.867	0	0.904	0.945	0	0.105

Table 2: Correlation between risk attitudes and network characteristics. Fraction of transitive triples is only defined for agents who have at least one friend who also has at least one friend, reducing the sample size to 152 (upper) and 112 (lower) observations; clustering coefficient is only defined for agents who have at least two friends, reducing the sample size to 146 and 112 observations. *** significant at 1% level; * significant at 10%.

network variables (Jackson and Rogers, 2007). We thus control for these issues using regression analysis, including a female dummy and all the network characteristics into the model.

Tables 4 and 5 report the estimates for the direct and undirected networks, respectively. The results corroborate that the findings from Table 3 are robust to controlling for both gender and potential correlations between the individual network variables. The connectivity and centrality are not significant in any model specification, for any network, and independently of the sample considered. In contrast, the fraction of transitive triples and clustering correlate consistently with risk aversion.

Interestingly, when we consider the students whose choices could be considered rational, we observe a significant gender effect on risk aversion, that is, men are less risk averse than women. However, once we include in the regression the transitivity/clustering coefficient, the gender effect attenuates and becomes insignificant. In contrast, the transitivity/clustering variable is strongly significant. Hence, the relation between network clustering and risk aversion seems much stronger than the relation between gender and risk aversion.

In sum, our empirical analysis uncovers that more risk averse participants are systematically embedded in more clustered network neighborhoods within the friendship network of the group. This finding goes beyond the gender effects in risk attitudes and do not result from different connectivity across subjects.

	Whole sample			Risk aversion betw. 3 to 6		
	(1)	(2)	(3)	(4)	(5)	(6)
in-degree	0.041 (0.102)	—	0.037 (0.102)	-0.039 (0.070)	—	-0.039 (0.070)
out-degree	-0.192 (0.149)	—	-0.182 (0.148)	-0.130 (0.100)	—	-0.123 (0.100)
trans. triples	0.851* (0.493)	—	0.682 (0.505)	1.418*** (0.350)	—	1.315*** (0.358)
betweenness($\times 10^{-4}$)	-1.17 (3.16)	—	1.54 (3.16)	2.58 (2.15)	—	2.16 (2.17)
gender	—	0.489 (0.282)	0.417 (0.291)	—	0.424* (0.207)	0.263 (0.206)
constant	4.89*** (0.39)	4.43*** (0.21)	4.70*** (0.41)	4.21*** (0.27)	3.96*** (0.15)	4.25 (0.28)
# obs.	152	152	152	112	112	112
R^2	0.038	0.020	0.051	0.134	0.037	0.148

Table 3: Estimation results from regressions of risk aversion on the directed network variables and gender. OLS standard errors in parentheses. * significant at 10% level; ***at 1% level.

	Whole sample			Risk aversion betw. 3 to 6		
	(1)	(2)	(3)	(4)	(5)	(6)
degree	-0.107 (0.104)	—	-0.108 (0.104)	0.014 (0.070)	—	0.018 (0.070)
clustering	1.056* (0.433)	—	0.895* (0.467)	1.148*** (0.311)	—	1.021*** (0.326)
betweenness($\times 10^{-4}$)	0.37 (1.91)	—	-0.05 (1.96)	-0.59 (1.37)	—	-1.03 (1.41)
gender	—	0.237 (0.296)	0.283 (0.308)	—	0.406* (0.210)	0.266 (0.211)
constant	4.68*** (0.44)	4.18*** (0.24)	4.62*** (0.44)	4.01*** (0.31)	4.25*** (0.16)	3.92*** (0.21)
# obs.	146	146	146	109	109	109
R^2	0.059	0.017	0.064	0.148	0.034	0.161

Table 4: Estimation results from regressions of risk aversion on the undirected network variables and gender. OLS standard errors in parentheses. * significant at 10% level; ***at 1% level.

3 The Model

In this section, we propose a network formation model that relates risk aversion to transitivity and clustering. The model is a variant of the models in Vázquez (2003) and Jackson and Rogers (2007), but unlike those papers, we propose a mechanism in which the decision to link locally or globally is based on utility maximization. Due to the close relation to Jackson and Rogers (2007) and Vázquez (2003), the model generates typical socially generated architectures (formally shown in Appendix B) and allows us to relate individuals' network characteristics to their risk attitudes.

Let $N(t)$ be the population of agents existing at time t . The directed network among those agents is denoted by $G(t)$; $g_{ij}(t) = 1$ denotes a directed link from i to j at time t . Define $N_i(t) = \{j : g_{ij}(t) = 1\}$ as the out-degree neighborhood of individual i at time t . For notational convenience, dependence on t will be dropped if there is no confusion.

Network formation occurs through the following dynamic process. Each period one new player enters the population. This player is identified by its entrance period i . We assume that individuals have a capacity constraint with respect to out-degree, and agent i is only able to have m links pointing outwards, which she creates when entering the network. On the other hand, we assume that individuals do not have capacity constraints with respect to the in-degree, and therefore no individual j refuses a link ij if offered by i .⁹ Links are only created by the entering node i at time of entrance. Afterwards links cannot be changed.¹⁰

The benefits node i gets from linking with an existing node $j \in N(i) \equiv \{0, 1, \dots, i-1\}$ is denoted as u_{ij} , which is drawn from an i.i.d. distribution F having support on the interval $[a, b]$. This distribution (but not the realizations) is common knowledge and has mean $\bar{u}$. Naturally, the assumption that u_{ij} is independently distributed for each i and j is a very strong assumption. For example, it excludes the possibility that some nodes have intrinsic traits v_j that make them more beneficial for any node i . Such an extension of the model is discussed in Section 6. It also excludes any (indirect) network effects once the network is in place. As we will see, even this simplest case gives rise to a relation between network structure, risk aversion and payoffs, such that it is best to focus on this case first. Section 6 explores some deviations from the independence assumptions.

⁹We are aware that economic agents have limited capacity and/or reject linking proposals in many contexts. However, since the main contribution of this paper comes from the search mechanism, we abstract from this issue here.

¹⁰The network formation process is initialized by letting the first $m+1$ agents create a link with all their predecessors, that is, each agent $k \in \{0, \dots, m\}$ creates $k-1$ links, such that $g_{ki} = 1$ if $i < k$, and $g_{ki} = 0$ if $i > k$.

Let $U_i = U_i(\sum_{N_i} u_{ij})$ be a constant absolute risk aversion (CARA) utility function of agent i , that is

$$U_i \left(\sum_{N_i} u_{ij} \right) = -\frac{1}{\rho_i} e^{-\rho_i \sum_{N_i} u_{ij}}. \quad (5)$$

The risk premium, r_i , increases with ρ_i . This risk premium is constant for a given agent i and benefit distribution F , that is, agent i is indifferent between a sure benefit of $\bar{u} - r_i$ against a random benefit drawn from F . Formally,

$$U_i(x + \bar{u} - r_i) = E[U_i(x + u_{ij})] \text{ for all } x.$$

Let individuals differ in their risk attitudes. In particular, we assume that there are two levels of risk aversion with risk premium, r_H and r_L , with $\bar{u} - b < r_L < r_H < \bar{u} - a$. A new node i has risk premium $r_i = r_H$ with probability θ , and $r_i = r_L$ with $1 - \theta$.¹¹

The decision of node i to link with m nodes $j_1, \dots, j_m$ goes as follows. When entering, individual i initially does not have information on the benefits nor on the number of links of the other individuals. Nonetheless, individual i may obtain information on j by connecting to a friend of j , say k , who is connected to j , $g_{kj} = 1$.¹² The new node i first connects with one randomly drawn existing node j_1 . We assume that by connecting to j_1 , individual i obtains a perfect signal on the benefits of the out-degree neighbors of j_1 .¹³ Individual i then makes a decision on whom to connect next. If

$$\max_{k \in N_{j_1}} u_{ik} > \bar{u} - r_i \quad (6)$$

then i connects to the node $k \in N_{j_1}$ that maximizes u_{ik} , otherwise i connects to a random node outside N_{j_1} .¹⁴ Let this second node to which i links be denoted by j_2 . By connecting to j_2 it again learns about the benefits of the nodes in N_{j_2} . Starting from j_2 the process is repeated, that is, if

$$\max_{k \in N_{j_2}} u_{ik} > \bar{u} - r_i,$$

¹¹The results of this paper do not change qualitatively if we assume more complex distributions of risk aversion.

¹²An alternative approach is to consider a process that creates homophily with respect to quality, such as in Montgomery (1991). In that case, linking to, say, a high type would give us information on the type of the neighbors.

¹³As in Jackson and Rogers (2007) we assume that network search is directed, in particular, channeled through out-degree links. Allowing for network search through in-degree links would complicate the analysis significantly. See Jackson and Rogers (2007) for details.

¹⁴There are only two ways of creating links in the model for simplicity, but all the results go through as long as people have better information about closer individuals than about more distant ones in the network and risk-averse agents tend to connect to nodes about whom they are better informed.

Figure 3: Example of link formation of a new node i , when $m = 3$, $\bar{u} = .5$ and $r_i = .1$. (a) Node i creates a random link to j , and learns about the benefits of linking to j and j 's neighbors. A link to k gives the highest benefit to i , and since $u_{ik} = .8 > \bar{u} - r_i = .4$, a link to k is preferred to a random link. (b) Node i creates a link to k and learns about the benefits of linking to k 's neighbors. (c) Since the benefits of linking to a neighbor of k are all lower than $\bar{u} - r_i$, node i creates a link to a random node in some other part of the network.

then i connects to the node $k \in N_{j_2}$ that maximizes u_{ik} , otherwise i connects to a random node outside N_{j_2} .¹⁵ The linking process of i stops when i has formed m links, after which agent $i + 1$ enters the network, and starts to create links. Figure 3 illustrates one step of the network formation process when $m = 3$.

It is worth noting that with the decision rule in (6), we assume a certain bounded rationality of agents whenever $m \geq 3$. More precisely, suppose that i connects to j in the first linking decision and k , a neighbor of j , afterwards. If j linked up locally to a neighbor of k in one of the previous rounds, then there exists a node l who is an out-degree neighbor of both j and k . Hence, i observed l after the first linking, but has not connected to him (since he linked up to k), and observes l again after linking to k . This means that i observes again the same node, about whom he has full information. Since this occurs with positive (non-negligible) probability, there is a certain (expected) utility loss from linking to a neighbor of a neighbor anytime agents link through the network. Completely rational agents should take this potential utility loss into account while deciding whether to connect locally or through random search. In the main model, we abstract from this possibility. In Appendix A, we illustrate how taking this into account affects the linking decision of individual for $m = 3$. It adds substantial complexity to the model without affecting any of the theoretical results below.

¹⁵We assume that i does not recall (or is unable to contact) the neighbors of previously visited links, that is, at the s -th step, the agent i is unable to recall the benefits from linking to agents in neighborhoods of $N_1, \dots, N_{s-1}$. This is only in order to keep the model tractable. If we would allow for aggregation of information on neighbors, the probability to link a friend would steadily increase during the m linking steps but this would not affect the qualitative nature of our results.

Before we state the results, we introduce some notation. Denote $d_i(t) = \sum_j g_{ji}(t)$ the in-degree of individual i at period t . Let $p(r)$ be the probability that an entering node $t + 1$ of type r , having already linked to agents $j_1, \dots, j_s : 0 < s < m$, decides to link to a friend of j_s instead of linking to someone randomly. The expected probability that a random agent finds it optimal to follow a network-based meeting is then $p_\theta = \theta p(r_H) + (1 - \theta)p(r_L)$. The definition of the clustering coefficient differs depending on the way one keeps track of the direction of the links; below, we focus on the measure defined earlier in (1) and (2).

In Appendix B, we show formally that the proposed network formation process generates empirically observed social networks. In particular, the in-degree distribution of agents for large d has a power-law distribution in the tail - $f(d) \propto d^{-(\frac{m}{(m-1)p_\theta} + 1)}$ for large d - and the clustering coefficient is bounded away from zero (Theorem 6). Moreover, the resulting topologies also exhibit short network distances, assortativity, and the negative clustering/degree correlation (see Appendix B).

The following proposition formalizes the relation between an agent's risk aversion and her position in the network:¹⁶

Proposition 1 *The in-degree of i , $d_i(t)$ is independent of r_i , the degree of risk aversion of i , while r_H types have a higher fraction of transitive triples than r_L types.*

By assumption of the model the out-degree of all individuals is identical, m , and therefore independent of risk aversion. The same holds for in-degree. On average, the in-degree of both r_H and r_L types depends on the distribution of risk aversion in the population, determined by θ , rather than on i 's type. This is due to two facts; first, because link formation is one-sided, that is, only the entering agent $t + 1$ decides on the formation of a link, and second, because the distribution of risk aversion in the population is independent of the distribution of utilities that i conveys for entering agent $t + 1$. As a result, entering node $t + 1$ does not take into account the risk attitude of agents while deciding whether to connect to them or not.

Proposition 1 also shows that in our model more risk averse types with r_H have a higher fraction of transitive triples than r_L types. The intuition is that random links do not contribute to closing triads, whenever the network is large. The fact that r_H types have fewer random links implies that they have larger node transitivity.

¹⁶All proofs can be found in Appendix C.

4 Network Position and Payoffs

One of the main interests in the study of social and economic networks is the relation between individual network position and individual economic outcomes. In particular, the relation between clustering coefficient and payoffs has raised some debate. On the one hand, the theory of network closure (Coleman, 1988) argues that local clustering is beneficial, because it allows for better monitoring, which enforces more cooperation and higher trust levels (see also Granovetter, 1985). On the other hand, the theory of structural holes (Burt, 1995) argues that network positions that bridge different groups allow for better information access and control. These structural hole positions are typically characterized by low local clustering.

Given that our model builds on standard economic assumptions of utility maximization, we are able to give an alternative view on the relation between clustering and payoffs. To this aim, we first show that the expected payoff depends on the type. Define the payoff of i as $\sum_{j \in N_i} u_{ij}$.¹⁷

Proposition 2 *Suppose that $r_H > r_L > 0$. The expected payoff of an individual of type r , $E[\sum_{j \in N_i} u_{ij} | r_i = r]$, is decreasing with the risk premium r .*

Proposition 2 shows that, in our model, individuals with larger risk premium tend to earn less. There is a standard economic interpretation behind this result: risk averse individuals accept sure relatively low payoffs from second-order neighbors in order to avoid risky decisions.

Now we proceed with the analysis of the relation between the individual network position and payoffs. By the construction of the model, there is no relation between in-degree (out-degree and, hence, degree) and payoffs, since only the quality of out-degree neighbors are relevant for payoffs of agents and out-degree is the same for all nodes. Therefore, we focus on clustering coefficient.

So far, we have derived two results regarding risk aversion, Propositions 1 and 2. The former establishes a positive relation between risk aversion of individuals and their clustering coefficient, while the latter proposition proves that the expected payoff is negatively affected by risk aversion. This might suggest that if there is any relation between clustering coefficient and payoffs it should be negative. However, Theorem 3 shows that this is not the case.¹⁸

¹⁷Utilities of individuals $U(\sum_{j \in N_i} u_{ij})$ are non-comparable, and therefore not considered.

¹⁸We were only able to prove the theorem for $m < 5$. The matters are complex for larger m . However, we can easily show that there are upper and lower bounds on the payoff conditional on clustering, both strictly increasing in the level of clustering. Therefore, we conjecture that the theorem holds for any m .

Theorem 3 *For $m < 5$, the expected payoff of individual i , conditional on her clustering coefficient c , $E[\sum_{j \in N_i} u_{ij} | C_i = c]$, is non-decreasing in c .*

At first sight, Theorem 3 seems to contradict Propositions 1 and 2. Nevertheless, a closer look at the forces behind the formation of transitive triples (that determine the level of clustering) reveals a more complex relation between payoffs and clustering.

There are two forces at work. Propositions 1 and 2 captures one direction: Risk averse individuals pay a risk premium for sure payoffs from network-based linking, which leads to larger clustering and this drives payoffs of more clustered individuals down.

However, there is a “neighborhood effect,” which goes into the opposite direction. More precisely, people whose neighborhoods are attractive tend to stay within their neighborhoods, that is, they link to the neighbors of their neighbors, increasing the individual clustering coefficient as well as the average payoff.

Theorem 3 shows that the neighborhood effect always dominates the influences from Propositions 1 and 2.

5 Risk Preferences and Contexts

In this section, we analyze the impact of changes in the socioeconomic environment, measured by the distribution of benefits, on the structure of the social network.

Recent empirical literature provides evidence that who links up with whom is an endogenous process influenced by the socioeconomic environment (de Weerdt, 2004; Krishnan and Sciubba, 2009). Other streams of literature document how existing networks reshape when the environmental conditions change. For instance, Goyal et al. (2006) report how the structure of scientific collaboration has changed over past decades in parallel with the burst of communication technologies, and Eeckhout and Munshi (2010) - while analyzing an informal financial institution that brings agents together in groups - observe that participants rematch immediately following an unexpected exogenous regulatory change. To provide an example outside the domain of economics, it has been documented that the emergence of HIV epidemics has considerably affected the architecture of needle-sharing among drug users (Rothenberg et al., 1998). Hence, networks endogenously reorganize in presence of exogenous shocks. This generate many new questions. Which aspects of the environment trigger the endogenous adaptation of social organization? Why and how do network architectures react to these variables?

The present framework allows us to relate how network properties depend on the economic and social context in which the network formation takes place. If risk attitudes or the distribution of benefits are different in one environment compared to the other, then individual decisions are different, and so is the network formation process and the eventual network structure. Hence, different social and economic contexts lead to different network architectures, and this may have implications on eventual social and economic outcomes as well.

We illustrate formally how the change of the context, characterized by the distribution function of benefits, F , interacts with risk preferences of individuals. To this aim, on top of assuming that agents have constant absolute risk aversion utility function with risk aversion coefficient ρ_i , we also assume that the payoff distribution of linking to individual j , u_{ij} , is normally distributed with mean $\bar{u}$ and variance σ^2 . With these assumptions, the risk premium of an individual will be a function of her risk-aversion coefficient and characteristics of the payoff distribution F :

$$r = r(\rho_i, \bar{u}, \sigma^2)$$

Proposition 4 *Suppose that individual i has utility function (5) with coefficient of absolute risk aversion ρ_i . Let F be a cumulative distribution function of a normal distribution with mean $\bar{u}$ and variance σ^2 . Then,*

- (i) $p[r(\rho_i, \bar{u}, \sigma^2)]$ does not depend on $\bar{u}$, and
- (ii) $p[r(\rho_i, \bar{u}, \sigma^2)]$ increases with σ^2 .

This result has strong implications for the model. It shows that we can observe the same individuals in very different network position (in terms of their random vs. local search), depending on the riskiness of the environment in which a particular network is embedded. More precisely, we show that more risky contexts will drive people to link up to neighbors of their neighbors more often. In contrast, an increase or decrease of the average benefits will not affect the decisions of agents, as long as preferences for absolute riskiness are preserved.

Proposition 4 also has direct implications for the global structure of the model:

Theorem 5 *Let F and F' be two normal distribution functions with means $\bar{u}$ and $\bar{u}'$ and variances σ^2 and σ'^2 respectively. Consider the networks g and g' associated with linking benefit distributions F and F' .*

- (i) *If $\sigma^2 > \sigma'^2$, then the degree distribution of g' second order stochastically dominates the degree distribution of g , and $C(g) > C(g')$.*

(ii) If $\sigma^2 = \sigma'^2$, then the degree distribution of g and g' are identical and $C(g) = C(g')$, independently of $\bar{u}$ and $\bar{u}'$.

This result shows how a change of context affects the network properties. A mean preserving-spread of the payoff distribution has a direct effect on whether the network will be more or less random, since more risky environment enhances local, non-random search. This at the same time affects the probabilities of incumbent nodes to receive a link. More precisely, less connected agents, who mostly rely on global search, become less likely to receive a link, while agents above a certain degree, whose main source of new connections is to be found through the network, are now more likely to receive new incoming links. The global effect, stated in Theorem 5, is a shift of the degree distribution in terms of second-order stochastic dominance; a riskier environment creates more inequality in terms of connectivity.

Concerning the local clustering of the network, riskier environments generate more clustered network architectures. Location shift alone will affect neither the degree distribution nor local clustering of the network.

These findings illustrate how network architectures endogenously adapt to changes of environmental variables. For instance, people might be more careful choosing close friends than mere acquaintanceship, leading to more clustered networks in the former case. Other examples might be that firms' position in technological networks can differ according to the riskiness of the innovation in progress, and that people will search for new sexual partners more locally after the start of the HIV epidemics. Despite that our network model does not allow for relinking, we believe the same logic - which determines which type of connection will establish - will also operate when individuals sever or redirect their links.

6 Extensions and Discussion

Our model is built on strong assumptions in order to keep it tractable. In particular the assumption that benefits are identically and independently distributed is unlikely to be true. These assumptions help us to focus on the role of risk aversion in network formation, the main objective of this paper. In this section, in order to understand what happens if we relax some of these assumptions, we discuss variations on the standard model.

6.1 Common Benefits from Linking

The assumption of idiosyncratic benefits makes the proposed model rather restrictive. There is a large number of applications, where the potential benefits of a particular node are the same for all the members of the population. Examples of such applications can be labor market connections, where some individuals have better access to job opportunities, coauthorship networks, research networks among companies etc. In terms of our model, this would make $u_{ij} = u_j$ for each $i \in N \setminus \{j\}$. The effect of this specification is that network-based linking would become more frequent, because anytime an entering node i links up to j , who formed at least one link through the network (say to a node k), i will create a link to k with probability one if i 's risk premium is equal or larger than j 's one, since, given that $g_{jk} = 1$ and $u_{ik} = u_{jk} = u_k$, $u_k > \bar{u} - r_i$ if $r_i \geq r_j$. Then, network-based search is enhanced under such a specification.

However, the main results of this paper remain qualitatively unchanged. The relation between clustering and risk aversion still holds, since the above argument does apply if $r_i < r_j$. If i is less risk averse than j , he does not necessarily create links to neighbors of j found through the network. As a result, the positive relation between risk aversion and clustering coefficient remains, while in-degree would still be independent of risk attitudes. As a results, the relations between clustering and expected payoffs and between the uncertainty and network structure are also robust to this alternative model specification. Since the network-based search is enhanced, the only effect would be quantitative: larger clustering coefficient and more unequal degree distribution. Therefore, our model also applies to situations where the benefits of a particular node are the same for all the members of the population.

6.2 Public Knowledge

In our model we assumed that an entering agent initially has zero information about the benefits it can obtain by connecting to other agents. In reality, this is not always realistic. For example, in the academic world it is always possible to find information on other scientists by looking up their C.V.

The assumption of no prior information on the benefits of linking is easily relaxed. For example, suppose instead that an individual i has an imperfect signal about the benefits of linking to j , say $\tilde{u}_{ij} = u_{ij} + \epsilon_{ij}$ where ϵ_{ij} is unobserved i.i.d. noise with zero expectation. Initially, the entering agent i links to the agent j about whom it has the best signal, $\max_j \tilde{u}_{ij}$. Next, the agent receives a better, perhaps perfect, signal about the neighbors of j , and again the agent decides to link the best neighbor of j , or to link to the best outside option, that is the node k

with the best signal that is not j or a neighbor of j , $\max_{k \in N \setminus \{j \cup N_j\}} \tilde{u}_{ik}$. Agent i chooses to link to a neighbor of j if and only if

$$\max_{k \in N_{j_1}} u_{ik} > \max_{k \in N \setminus \{j \cup N_j\}} E[u_{ik} | \tilde{u}_{ik}] - \tilde{r}_i$$

Naturally, $\tilde{r}_i$ is smaller than r_i , because given that there is some initial information on non-neighbors, the risk of linking to a non-neighbor is smaller. Nevertheless, given that i still has a better signal about the neighbors of j than about non-neighbors, risk aversion again plays a role; the more risk-averse agent i , the more likely i links to a neighbor of j . This implies the same positive relation between clustering coefficient and risk aversion. Here it is irrelevant that the choice of non-neighbor is not random anymore, that is, agent i would choose the non-neighbor with the best signal, but given that the benefits and the signals are still randomly distributed, for the outside observer the choice of agent i is observably equivalent to random linking if it is not a neighbor of j .

Note that conditional on the signal $\tilde{u}_{ik}$, the expected benefit of linking to k , $E[u_{ik} | \tilde{u}_{ik}]$ will be between the true benefit u_{ik} and the average benefit $\bar{u}$ with the expected value closer to the former when the quality of the signal is better. Therefore,

$$\max_{k \in N \setminus \{j \cup N_j\}} E[u_{ik} | \tilde{u}_{ik}] > \bar{u}.$$

When agents have prior information on non-neighbors, it is therefore more likely that they choose a “random” link than a friend of a friend, compared to the case where agents do not have such prior information. Moreover, the better the signal on the benefits of linking to non-neighbors, the more likely it is that the agent chooses to link to a non-neighbor, which for the outside observer is a “random” link.

An application is the impact of the internet on network formation. The emergence of internet has made it much easier to publish and obtain information on other individuals. For example, it is now standard that scientists put their C.V. on their homepage, which is then publicly available. In our model this implies that individuals have some prior knowledge on the benefit of linking to individuals, and therefore they are much less likely to link to a friend of a friend. That is, random linking should have become much more prominent than local network-based linking. Evidence provided in Fafchamps et al. (2010) indeed suggests that this is the case.

7 Conclusion

This paper shows that inherent characteristics of individuals may play an important role in network formation and in explaining empirical regularities of networks. Ex-ante individual heterogeneity is an issue that has been underexplored; partly in order to keep models tractable, and partly due to the belief the network formation is an endogenous process, and that understanding this endogenous process is what is most important.¹⁹ However, recent work of Fowler et al. (2009) suggests that ex-ante individual differences are very important as well. They do not explore what behavioral heritable aspects lead to this variation, though. Since Cesarini et al. (2009) report that a non-negligible part of risk-taking preferences of people are due to genes, we believe that we uncover a possible heritable aspect explaining social network positions of individuals: the variation in clustering can partly be traced back to variation in risk aversion among individuals. We hope our results will enhance the exploration of the relevance of heterogeneity in social networks.

We propose a simple theory, introducing simple economic reasoning into the models of Vázquez (2003) and Jackson and Rogers (2007), which has the property that risk aversion and clustering are related. Analyzing this theoretical model further, we make some additional contributions. In particular, one contribution of the paper is the new mechanism that relates network position to payoffs. Standard sociological theory and economic theory on network effects (Coleman, 1988; Burt, 1995; Ballester et al., 2006) takes the network as rigid, and proposes mechanisms of social interaction that lead to different payoffs for agents in different social network positions. In contrast, the relation between network position and payoffs in our model is induced by the network formation process and the fact that information on the benefits of linking a friend of a friend is more precise than information of the benefits of linking a stranger.

As a last contribution, we provide one argument for why different network topologies arise in different socioeconomic contexts and why they may be affected by changing environment, such as lower cost of communication and link maintenance, or external interventions or shocks that influence the benefits from linking opportunities. Hence, our model or its variations might provide an interesting tool for the evaluation of policies in networked contexts. Nevertheless, whether networks indeed react this way to external shocks is an empirical question, which we leave for future research.

¹⁹Exceptions are Galeotti et al. (2006) and Jackson and Rogers (2005) in the economics literature, Bianconi and Barabasi (2001) and Kong et al. (2008) in the physics literature, and Burt et al. (1998) in the sociology literature.

References

- [1] Atalay, E. Sources of variation in social networks. *Games and Economic Behavior* 79, 106-131 (2013).
- [2] Ballester, C., Calvó-Armengol, A. and Zenou, Y. Who is Who in Networks. Wanted: The Key Player. *Econometrica* 75(4), 1403-1418 (2006).
- [3] Barabási, A. L. and Albert, R. Emergence of scaling in random networks. *Science*, 286(5439), 509-512 (1999).
- [4] Bianconi, G. and Barabasi, A.L. Bose-Einstein Condensation in Complex Networks, *Physical Review Letters* 86(24), 5632-5635 (2001).
- [5] Bloch, F., Genicot, G., and Ray, D. Informal insurance in social networks. *Journal of Economic Theory* 143, 36-58 (2008).
- [6] Brañas-Garza, P., Jiménez-Jiménez, F., and Morales, A.J. Strategic Uncertainty and Risk Attitudes: The Experimental Connection, *Cuadernos de Economía*, 27, 139-152 (2004).
- [7] Bramoullé, Y., Currarini, S., Jackson, M. O., Pin, P., and Rogers, B. W. Homophily and long-run integration in social networks. *Journal of Economic Theory* 147, 1754-1786 (2012).
- [8] Bramoullé, Y., and Kranton, R. Risk-sharing networks. *Journal of Economic Behavior and Organization* 64, 275-294 (2007).
- [9] Burt, R.S. *Structural Holes: The Social Structure of Competition*, Harvard University Press (1995).
- [10] Burt, R.S., Jannotta, J.E., and Mahoney, J.T., *Personality Correlates of Structural Holes*, *Social Networks* 20, 63-87 (1998)
- [11] Buskens, V. and Van de Rijt, A. Dynamics of Networks if Everyone Strives for Structural Holes. *American Journal of Sociology* 114, 371-407 (2008).
- [12] Campbell, A. Signaling in social networks and social capital formation, *Economic Theory* 57, 303-337 (2014).

- [13] Cesarini, D., Dawes, C., Johannesson, M., Lichtenstein, P. and Wallace, B. Genetic Variation in Preferences for Giving and Risk Taking, *Quarterly Journal of Economics*, 124, 809-842 (2009).
- [14] Coleman, J. Social Capital in the Creation of Human Capital, *American Journal of Sociology*, 94, 95-120 (1988).
- [15] Croson, R., and Gneezy, U. Gender differences in preferences. *Journal of Economic Literature* 47, 448-474 (2009).
- [16] de Weerd J. Risk-sharing and Endogenous Network Formation, in Dercon, S. (ed.): *Insurance Against Poverty*, Oxford University Press 197-216 (2004).
- [17] Eeckhout J. and Munshi K. Matching in Informal Financial Institutions, *Journal of the European Economic Association*, 8, 947-988 (2010).
- [18] Fafchamps, M., Goyal, S., and Van der Leij, M. Matching and Network Effects, *Journal of the European Economic Association*, 8, 203-231 (2010).
- [19] Fowler, J., Dawes, C., and Christakis, N. A Model of Genetic Variation in Human Social Networks, *Proceedings of the National Academy of Sciences*, 106(6), 1720-1724 (2009).
- [20] Galeotti, A., Goyal, S., and Kamphorst, J. Network Formation with Heterogeneous Players, *Games and Economic Behavior*, 54(2), 353-372 (2006).
- [21] Goyal, S. *Connections. An Introduction to the Economics of Networks*, Princeton Univ. Press (2007).
- [22] Goyal, S. and Joshi, S. Unequal connections. *International Journal of Game Theory* 34, 319-349 (2006).
- [23] Goyal, S., van der Leij M.J. and Moraga-Gonzalez J.L. Economics: An Emerging Small World, *Journal of Political Economy* 114(2), 403-412 (2006).
- [24] Goyal, S. and Vega-Redondo, F. Structural holes in social networks. *Journal of Economic Theory* 137, 460-492 (2007).
- [25] Granovetter, M. Economic Action and Social Structure: The Problem of Embeddedness, *American Journal of Sociology* 91: 481-510 (1985).

- [26] Holt, C.A and Laury, S.K. Risk Aversion and Incentive Effects, *American Economic Review* 92, 1644-1655 (2002).
- [27] Jackson, M. *Social and Economic Networks*, Princeton University Press (2008).
- [28] Jackson, M. and Rogers, B. The Economics of Small Worlds, *Journal of the European Economics Association*, 3(2-3), 617-627 (2005).
- [29] Jackson, M. and Rogers, B. Meeting Strangers and Friends of Friends: How Random Are Social Networks, *American Economic Review* 97(3), 890-915 (2007).
- [30] Jackson, M. O. and Wolinsky, A. A strategic model of social and economic networks. *Journal of Economic Theory* 71(1), 44-74 (1996).
- [31] Kong, J., Sarshar, N., and Roychowdhury, V. Experience versus Talent Shapes the Structure of the Web, *Proceedings of the National Academy of Sciences* 105(37), 13724-13729 (2008).
- [32] Kovářík, J., Brañas-Garza, P., Cobo-Reyes, R., Paz Espinosa, M., Jiménez, N., and Ponti, G., Prosocial Norms and Degree Heterogeneity in Social Networks, *Physica A*, 392, 849-853 (2012).
- [33] Krishnan P. and Sciubba E. Links and Architecture in Village Networks, *Economic Journal* 119, 917–949 (2009).
- [34] Lin, N. Building a network theory of social capital. *Connections* 22, 28-51 (1999).
- [35] Montgomery, J. Social Networks and Labor Market Outcomes: Toward an Economic Analysis, *American Economic Review* 81(5), 1408-1418 (1991).
- [36] Rothenberg R.B., Potterat J.J., Woodhouse D.E., Muth S.Q., Darrow W.W. and Klov Dahl A.S. Social network dynamics and HIV transmission, *AIDS* 12(12), 1529-36 (1998).
- [37] Snyder, D., and Kick, E. L. Structural position in the world system and economic growth, 1955-1970: A multiple-network analysis of transnational interactions. *American Journal of Sociology* 84, 1096-1126 (1979).
- [38] Vázquez, A. Growing Network with Local Rules: Preferential Attachment, Clustering Hierarchy, and Degree Correlations, *Physical Review E* 67, 056104 (2003).

- [39] Vigier, A. Meeting friends of friends and homophily: a complementarity. *Economic Theory Bulletin* 2, 45-52 (2014).
- [40] Zaheer, A., and Bell, G. G. Benefiting from network position: firm capabilities, structural holes, and performance. *Strategic Management Journal* 26, 809-825 (2005).

A Perfect rationality for $m = 3$

There is an important issue concerning the rationality of players for $m > 2$. In particular, the clustering of the network can lead to a situation, such that anytime a node i links through network there is a positive probability that one or more of newly observed neighbors of neighbor have already been observed and not chosen in previous linking stages. A completely rational individual should take this into account. Since the entering node has not linked up to such node(s), i has information about them. This affects the mean field analysis, because the expected payoff from linking to such neighbors of neighbors is lower than the expected payoff i gets from observing and linking to someone i has no information about. In this section, we illustrate this argument formally.

Denote the nodes that $t + 1$ connects in each linking stage as j_1 , j_2 and j_3 . First, note that this issue never concerns the first and last linking decision, since the first is always random, while in the last linking decision the entering nodes do not care about who they observe afterwards. Hence, for $m = 3$ the only linking decision, in which he may observe someone, whom he has already observed, is the second one. Suppose that $t + 1$ decides to link to a node $j_2 \in N_{j_1}(t + 1)$ such that $j_2 \in \arg \max_{j \in N_{j_1}(t+1)} u_{ij}$. If so, then there is a positive probability that j_1 is connected to a neighbor of j_2 . If this occurs there exist a node l , an out-degree neighbor of both j_1 and j_2 , who $t + 1$ observes after linking to j_1 and will observe after linking to j_2 . Furthermore, there is an important information in the fact that $t + 1$ observed l , but has not connected to him.

Formally, $t + 1$ links to a $j_2 \in \arg \max_{j \in N_{j_1}(t+1)} u_{ij}$ if

$$\max_{j \in N_{j_1}(t+1)} u_{ij} - \sum_{s=1}^{m-1} \left\{ \frac{C(g)^s(m-s)}{m} \left[\bar{u} - \frac{m-s}{m} \bar{u} - \frac{s}{m} \int_a^{\max_{j \in N_{j_1}(t+1)} u_{ij}} u dF(u) \right] \right\} > \bar{u} - r_{t+1} \quad (7)$$

where $C(g)$ is the fraction of transitive triples in the population and measures the average probability that a triangle exists. $C(g) = \frac{p_\theta(m-1)}{m^2 - \frac{p_\theta}{m}}$ for $m = 3$ and reflects the average probability that $t + 1$ observes $m - 1$ new individuals and 1 individual $t + 1$ has already been observed and have not chosen because the utility he would reported to $t + 1$ was lower than $\max_{j \in N_{j_1}(t+1)} u_{ij}$. The second expression, $\bar{u} - \frac{m-1}{m} \bar{u} - \frac{1}{m} \int_a^{\max_{j \in N_{j_1}(t+1)} u_{ij}} u dF(u)$, reflects the expected utility loss due the fact that $t + 1$ observes only $m - 1$ new individuals (instead of m), taking into account the expected utility from the individual observed and unchosen in the previous linking stage.

After some simplification of (7), we get

$$\max_{j \in N_{j1}(t+1)} u_{ij} - \sum_{s=1}^{m-1} \frac{C(g)^s s^2}{m^s} [\bar{u} - \int_a^{\max_{j \in N_{j1}(t+1)} u_{ij}} u dF(u)] > \bar{u} - r_{t+1} \quad (8)$$

$$\max_{j \in N_{j1}(t+1)} u_{ij} - \left[\sum_{s=1}^{m-1} \frac{C(g)^s s^2}{m^s} \right] \left[\int_{\max_{j \in N_{j1}(t+1)} u_{ij}}^b u dF(u) \right] > \bar{u} - r_{t+1}. \quad (9)$$

As a result, the probability that $t + 1$ links through network search in its second linking decision is

$$p^{2nd}(r_{t+1}) = 1 - F \left[\bar{u} - r_{t+1} + \sum_{s=1}^{m-1} \frac{C(g)^s s^2}{m^s} \int_{\max_{j \in N_{j1}(t+1)} u_{ij}}^b u dF(u)^m \right] > p(r_{t+1}).$$

Then, the expected probability of node $i < t + 1$ to receive a new link in $t + 1$, analogous to expression (15), is

$$\frac{dd_i(t)}{dt} = \frac{1}{t} + \left[\frac{1 - p^{2nd}(r_{t+1})}{t} + p^{2nd}(r_{t+1}) \frac{d_i(t)}{t} \frac{1}{m} \right] + \left[\frac{1 - p(r_{t+1})}{t} + p(r_{t+1}) \frac{d_i(t)}{t} \frac{1}{m} \right]. \quad (10)$$

The only difference between (15) and (10) is the intermediate term. The effect of perfect rationality is to enhance global search. In (10), it increases the probability of receiving a random link and decreases the probability of receiving a link from a neighbor of a neighbor. The overall effect, hence, depends on the current in-degree of each node. In particular, nodes with large in-degrees will be negatively affected by perfect rationality, because a large fraction of nodes they receive is through local linking. Nodes with low connectivity, on the other hand, benefit from the form of rationality we model here, since they receive almost no links through network anyway. Formally, $\frac{\partial \frac{dd_i(t)}{dt}}{\partial p(r_{t+1})} = \frac{d_i(t)}{t} \frac{1}{m} - \frac{1}{t} > 0$ if $d_i(t) > m$. Hence, the effect of the rationality discussed here is the following:

- If $d_i(t) > m$, i receives a link with lower probability than in the original specification,
- If $d_i(t) = m$, i is unaffected by the new specification,
- If $d_i(t) < m$, i receives a link with higher probability in the new specification.

In sum, the effect of the perfect rationality considered here is to enhance random search. This will affect the in-degree of each agent as a function of his connectivity. From the global point of view, the tails of the degree distribution shift down, more frequent global search lowers the clustering coefficient, and the distances would shrink.

B Generated Network Architectures

In this section, we show theoretically and via simulations that the network formation process proposed in Section 3 exhibits typical features of empirical social networks. Given the complexity of the problem (especially due to the dependence of meetings on the network structure), we rely on mean-field analysis of the model. The mean-field approach approximates the complex evolution of a stochastic system by a simpler deterministic system, in which the evolution is determined by the expected change. The first results is theoretical:²⁰

Theorem 6 *Under mean field approximation,*

(i) *if $m > 1$ and $p(r_L) > 0$, the (complementary) cumulative distribution function of in-degrees in period t can be characterized as*

$$1 - F_t(d) = \left(\frac{\frac{m(m+p_\theta - mp_\theta)}{(m-1)p_\theta}}{d + \frac{m(m+p_\theta - mp_\theta)}{(m-1)p_\theta}} \right)^{\frac{m}{(m-1)p_\theta}} \quad (11)$$

(ii) *the average clustering coefficient in the network satisfies*

$$C(g) \geq \frac{p_\theta}{m^2}(m-1). \quad (12)$$

Proof of Theorem 6. Let us first prove part (i) of the theorem. For an entering individual $t+1$ after linking to j_s , $p(r)$ equals the probability that at least one of m friends of j_s is more attractive than the benefits of linking randomly, that is

$$p(r) = 1 - F(\bar{u} - r)^m. \quad (13)$$

This probability naturally depends on the risk aversion of agent $t+1$, such that $p(r_H) > p(r_L)$, and the expected probability that a random agent finds it optimal to follow a network-based meeting is $p_\theta = \theta p(r_H) + (1 - \theta)p(r_L)$.

At entrance, the linking process of $t+1$ is as follows: She first links up randomly. Thus, for a particular agent $i < t+1$ the probability of receiving this link is $\frac{1}{t}$. Once this link has been created, $t+1$ faces $m-1$ decisions between linking locally through the network (by observing

²⁰The present set-up does not allow to solve analytically for the clustering coefficient for general m as - due to the sequential linking process - it is a solution of polynomial of degree $m-2$.

the neighbors of his neighbors), or linking to a randomly chosen agent from the population. In this case, the probability of i to increase its degree in one of these decisions is approximately

$$\frac{1 - p(r_{t+1})}{t} + p(r_{t+1}) \frac{d_i(t)}{t} \frac{1}{m}, \quad (14)$$

where the first part corresponds to the probability that $t + 1$ decides for a random search and links up to i . The second part of the expression is the joint probability of three events: (i) $t + 1$ finds it attractive to connect through the network structure, $p(r_{t+1})$, (ii) she has connected to one of the d_i (in-degree) neighbors j of i in the previous decision, $d_i(t)/t$,²¹ and (iii) i has the largest gain for $t + 1$ out of the (out-degree) neighbors of j , $1/m$.

Given that each link $i < t + 1$ can receive at most one link in each period, that is, multiple links are ruled out, we can write the deterministic change of i 's in-degree in period t as

$$\frac{dd_i(t)}{dt} = \frac{1}{t} + (m - 1) \left[\frac{1 - p_\theta}{t} + p_\theta \frac{d_i(t)}{t} \frac{1}{m} \right]. \quad (15)$$

Note that (15) can be rewritten as $\frac{dd_i(t)}{dt} = a \frac{d_i(t)}{t} + \frac{b}{t} + c$, where $a = \frac{(m-1)p_\theta}{m}$, $b = [1 + (m - 1)(1 - p_\theta)]$, and $c = 0$. Given that $m > 1$ and that $p(r_L) > 0$ ensures $p_\theta > 0$, the first part of Lemma 1 in Jackson and Rogers (2007) applies.

As for the clustering coefficient, consider an agent i . Each agent initially creates 1 random link and afterwards faces $m - 1$ decisions to either link locally or search randomly. The first case occurs with probability $p(r_i)$. Thus the agent has on average $p(r_i) \times (m - 1)$ links that are based on network search. If k is found through network search, then it must be through j to whom i is also linked. So we have $g_{ij} = g_{jk} = g_{ik} = 1$. Each such network-searched link creates *at least* one transitive triple. Given that the amount of triples for which $g_{ij} = g_{jk} = 1$ equals m^2 and $E[p(r_i)] = p_\theta$, we obtain (12). ■

Theorem 6 shows that - as long as there is a positive probability of low-risk individuals to find an attractive agent through network - $f(d) \propto d^{-(\frac{m}{(m-1)p_\theta} + 1)}$ for large d ; that is, the in-degree distribution of agents for large d has a power-law distribution in the tail, and the average clustering coefficient will be strictly positive independently of other characteristics of the model.

To check the precision of the mean-field approximations and to check whether the generated architectures exhibit other stylized facts of empirical social networks, we also run simulations of

²¹This approximation does not take into account that there is positive assortativity in the network. If indegree neighbor j is found through local network search, then the probability that j is found increases in the degree of j , d_j , and given that there is a positive degree correlation, in the degree of i , d_i , as well. Simulation results obtainable from the authors suggest that this ignorance does not have major implications on the results.

Figure 4: (a) The predicted (solid line) and simulated (crosses and circles) degree distributions. (b) The average clustering coefficient. The solid line is the predicted lower bound from Proposition 4.

the model and match them with the approximations. The simulation assumes that u_{ij} is drawn from a standard uniform distribution. Agents have a risk premium of either $r_L = 0$ and $r_H = .25$ with equal probability ($\theta = .5$). We initially set $m = 2$ and we generate a network of 5000 nodes. Figures 4 and 5 contain various plots for four of the five stylized facts of observed social networks that our model predicts. The distances are only discussed at the end of this section.

The simulations confirm the findings from Theorem 6 and Proposition 1. Figure 4a contrasts the predicted in-degree distribution with the simulated one. We can conclude that the mean field approach approximates very well the degree distribution generated by the model. Moreover, we distinguish between the high (crosses) and low (circles) risk premium types and find no systematic difference. Figure 4b plots the average clustering coefficient (in terms of fraction of transitive triples) for several values of m . The figure shows that the clustering coefficient is indeed positive and lies above the lower bound derived in Proposition 6. In fact, the simulated values of the average clustering coefficient are well above and increase over m , suggesting that the more connections the agents of our model form, the more clustered the network becomes.

In addition, the generated networks exhibit short network distances, assortativity, and the negative clustering/degree correlation. The first states that the average network distances and the largest distance between two (reachable) nodes in real-life networks are in general low in relation to the size of the network. The second property, assortativity, is a tendency such that high (low) degree nodes are more likely linked to high (low) degree nodes. Last, negative

Figure 5: (a) Relation between degree and clustering., (b) Assortativity

clustering/degree correlation simply suggests that the larger the degree of the node the lower its clustering coefficient. The reader is referred to Goyal (2007) or Jackson (2008) for formal definitions and evidence.

Figure 5a shows the negative clustering/degree correlation. (Here the clustering coefficient is measured ignoring the direction of the links.) The x- and y-axes plot the degree and clustering, respectively, and there is an obvious negative relation between the two variables in the graph. Moreover, in this plot we also make a distinction between the clustering coefficient of high risk-averse agents and low risk-averse agents. The plot shows that the clustering coefficient is substantially higher for high-risk averse agents, in particular for low degree values, where the majority of the nodes lies. To check for assortativity, in Figure 5b we draw a plot with the degree of a node on the x-axis and the average degree of an out-neighboring node on the y-axis. This plot shows a positive correlation. Nodes with high degree have also high degree neighbors, indicating positive assortativity. We also compute the degree correlation, which is .260, well above zero.

To check for network distances, we compute the average networks distance and the largest distance between two nodes in the resulting simulated network, again ignoring directions. The obtained values are 5.74, and 13, respectively, thus of the order of $\ln(n)$.

C Proofs

Proof of Proposition 1. Using that the initial in-degree of entering agents is 0, solving (15) leads to the in-degree of an agent i at period t :

$$d_i(t) = \left\lceil \frac{m(m + p_\theta - mp_\theta)}{(m-1)p_\theta} \right\rceil \left(\frac{t}{i} \right)^{\frac{(m-1)p_\theta}{m}} - \frac{m(m + p_\theta - mp_\theta)}{(m-1)p_\theta}. \quad (16)$$

Given that it is independent of r_i , the first part of the proposition directly follows. The second follows from that $p(r_H) > p(r_L)$, thus high types are more likely to search through the network. Each time an agent i decides to link through the network at least one transitive triple is created in its neighborhood, whereas the probability that a transitive triple is created after a random linking decision converges to 0 for large t . The proposition directly follows. ■

Proof of Proposition 2. Note that

$$E\left[\sum_{j \in N_i} u_{ij} | r_i = r\right] = \bar{u} + (m-1) \left\{ [1 - p(r)]\bar{u} + p(r)E\left[\max_{j \in N_i} u_{ij} | \max_{j \in N_i} u_{ij} > \bar{u} - r\right] \right\}.$$

Since

$$\begin{aligned} & [1 - p(r_H)]\bar{u} + p(r_H)E\left[\max_{j \in N_i} u_{ij} | \max_{j \in N_i} u_{ij} > \bar{u} - r_H\right] \\ &= \int_a^{\bar{u}-r_H} \bar{u} dF(u)^m + \int_{\bar{u}-r_H}^b u dF(u)^m \\ &= \int_a^{\bar{u}-r_H} \bar{u} dF(u)^m + \int_{\bar{u}-r_H}^{\bar{u}-r_L} u dF(u)^m + \int_{\bar{u}-r_L}^b u dF(u)^m \\ &< \int_a^{\bar{u}-r_H} \bar{u} dF(u)^m + \int_{\bar{u}-r_H}^{\bar{u}-r_L} \bar{u} dF(u)^m + \int_{\bar{u}-r_L}^b u dF(u)^m \\ &= [1 - p(r_L)]\bar{u} + p(r_L)E\left[\max_{j \in N_i} u_{ij} | \max_{j \in N_i} u_{ij} > \bar{u} - r_L\right], \end{aligned}$$

it directly follows that $E[\sum_{j \in N_i} u_{ij} | r_i = r_L] > E[\sum_{j \in N_i} u_{ij} | r_i = r_H]$. ■

Proof of Theorem 3. For any existing link ij , define L_{ij} an indicator that is 1 if i found j through local network search, and 0 if found by random search. For any existing link ij , let

$$\bar{u} \equiv E[u_{ij} | L_{ij} = 0] = \int_a^b u dF(u)$$

and

$$\begin{aligned}
\tilde{u} &\equiv E[u_{ij}|L_{ij} = 1] \\
&= P[r_i = r_H|L_{ij} = 1]E[u_{ij}|L_{ij} = 1, r_i = r_H] + P[r_i = r_L|L_{ij} = 1]E[u_{ij}|L_{ij} = 1, r_i = r_L] \\
&= P[r_i = r_H|L_{ij} = 1] \int_{\bar{u}-r_H}^b u dF(u)^m + P[r_i = r_L|L_{ij} = 1] \int_{\bar{u}-r_L}^b u dF(u)^m
\end{aligned}$$

denote the expected payoff of linking up to a random individual and a neighbor of a neighbor, respectively. Let $L_i = \sum_{j \in N_i} L_{ij}$. Then

$$E\left[\sum_{j \in N_i} u_{ij} | C_i = c\right] = \bar{u}E[m - L_i | C_i = c] + \tilde{u}E[L_i | C_i = c].$$

Naturally, $\tilde{u} > \bar{u}$. Hence, nodes who search more often locally will tend to earn higher payoffs.

To complete the proof, we now show that $E[L_i | C_i = c]$ is weakly increasing in c for $m < 5$. Note that each node i can close at most $\sum_{j=1}^{m-1} j = \frac{m(m-1)}{2}$ triples in period i since links are directed and entering nodes can only link up to older nodes.

Suppose $m = 2$. Then C_i is (approximately) 0 or positive, depending on whether the second link was random or via a friend of friend. Hence, $E[L_i | C_i = 0] = 0$ and $E[L_i | C_i > 0] = 1$.

Next, suppose $m = 3$ and denote the out-degree neighbors of i as j_1, j_2, j_3 . Then L_i is at most 2, and at most 3 triples can be closed. The first link does not close triples, the second link closes one triple if $L_{ij_2} = 1$, and the third link closes one or two triples if $L_{ij_3} = 1$. Hence, $E[L_i | C_i = 0] = 0$, $E[L_i | C_i = 1/9] = 1$, $E[L_i | C_i \geq 2/9] = 2$.

Finally, suppose $m = 4$ and let i have out-degree neighbors $j_1, \dots, j_4$. Then L_i is at most 3, and at most 6 triples can be closed. We have, $E[L_i | C_i = 0] = 0$, $E[L_i | C_i = 1/16] = 1$, $E[L_i | C_i = 2/16] = 2$, $E[L_i | C_i = 3/16] \in (2, 3)$, and $E[L_i | C_i \geq 4/16] = 3$. ■

Proof of Proposition 4. With utility (5), the certainty equivalent y of linking to a random agent solves

$$U_i(y) = E_F[U(x)].$$

Hence, with normal distribution of payoffs, y is given by

$$y = \bar{u} - \frac{\rho_i}{2}\sigma^2$$

The risk premium is defined as $r = \bar{u} - y = \frac{\rho_i}{2}\sigma^2$, leading to

$$p\left(\frac{\rho_i}{2}\sigma^2\right) = 1 - F\left(\bar{u} - \frac{\rho_i}{2}\sigma^2\right)^m.$$

With a normal distribution, $F(\bar{u} - \frac{\rho_i}{2}\sigma^2) = \Phi(-\frac{\rho_i}{2}\sigma)$ where $\Phi(\cdot)$ is the cumulative distribution function of standard normal distribution. Since it is decreasing with σ^2 and independent of $\bar{u}$, the proposition directly follows. ■

Proof of Theorem 5. Since $p_\theta = \theta p(r_H) + (1 - \theta)p(r_L)$, it follows from Proposition 4 that ratio of the global and local search probabilities $\frac{1+(1-p_\theta)(m-1)}{p_\theta(m-1)}$ decreases with σ^2 and is independent of $\bar{u}_F$. It then directly follows from Jackson and Rogers (2007), Theorem 6, that the degree distribution of g' second order stochastically dominates the degree distribution of g whenever $\sigma^2 > \sigma'^2$, independently of $\bar{u}$ and $\bar{u}'$. Moreover, since p_θ increases with σ^2 and remains constant with $\bar{u}$, the results on the clustering coefficient directly follow. ■