

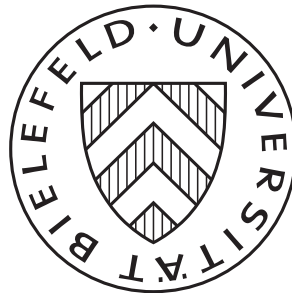
Center for
Mathematical Economics
Working Papers

515

July 2014

Continuous Homophily and Clustering in Random Networks

Florian Gauer and Jakob Landwehr



Center for Mathematical Economics (IMW)
Bielefeld University
Universitätsstraße 25
D-33615 Bielefeld · Germany

e-mail: imw@uni-bielefeld.de
<http://www.imw.uni-bielefeld.de/wp/>
ISSN: 0931-6558

Continuous Homophily and Clustering in Random Networks

Florian Gauer* and Jakob Landwehr†

July 21, 2014

We propose a random network model incorporating heterogeneity of agents and a continuous notion of homophily. Unlike the vast majority of the corresponding economic literature, we capture homophily in terms of similarity rather than equality of agents. We show that if links between similar agents are indeed more likely, our homophilous random network model exhibits clustering. Moreover, simulations indicate that the well-known small-world phenomenon is preserved even at high homophily levels. As a possible application we provide a stylized labor market model, where a firm can hire a worker via the social network.

Keywords: Random Graphs, Homophily, Clustering, Small-World Phenomenon, Network Formation, Labor Market Search

JEL-Classification: D85, J64, Z13.

*Center for Mathematical Economics, Bielefeld University, postbox 100131, D-33501 Bielefeld, Germany. Email: florian.gauer@uni-bielefeld.de, phone: +49 521 106 4918.

†Center for Mathematical Economics, Bielefeld University, postbox 100131, D-33501 Bielefeld, Germany. Email: jakob.landwehr@uni-bielefeld.de, phone: +49 521 106 4918. The authors would like to thank Tim Hellmann, Christoph Kuzmics, Agnieszka Rusinowska, Fernando Vega-Redondo, Dunia López-Pintado, Emily Tanimura and Michael Günther, as well as the participants of various seminars at Bielefeld University and University Paris 1 for valuable comments and suggestions. This research was carried out within the International Research Training Group "Economic Behavior and Interaction Models" (EBIM) financed by the German Research Foundation (DFG) under contract GRK 1134/2.

1 Introduction

Suppose you own a firm and want to fill an open vacancy through the social contacts of one of your current employees. Whom would you ask to recommend someone? Most probably you would address the worker who would himself perform best in the position in question. While this seems to be intuitively reasonable, why do we expect it to be optimal? One important reason is that people tend to connect to similar others. This phenomenon is known as homophily (Lazarsfeld et al., 1954).

In this paper we introduce a continuous notion of homophily based on incorporating heterogeneity of agents into the Bernoulli Random Graph (BRG) model as examined by Erdős and Rényi (1959). To this end we propose a two-stage random process. First, agents are assigned characteristics independently drawn from a continuous interval and second a network realizes, linking probabilities being contingent on a homophily parameter and the pairwise distance between agents' characteristics. This enables us to account for homophily in terms of similarity rather than equality of agents, capturing the original sociological definition instead of the stylized version up to now commonly used in the economic literature.

As a first result we determine the expected linking probabilities between agents (Proposition 1) as well as the expected number of links (Corollary 2). We then calculate the expected probability that an agent has a certain number of links (Proposition 2), showing that the according binomial distribution of the original BRG model is preserved to some degree. In Proposition 3 we establish a threshold theorem for any given agent to be connected. For all these results we demonstrate that the BRG model is recuperated as the limit case of no homophily and we thus provide a generalization thereof.

As a main result, we show that in our model homophily induces clustering (Proposition 4), two stylized facts frequently observed in real-world networks which are not captured by the BRG model. Furthermore, clustering proves to be strictly increasing in homophily. Additionally, two simulations will indicate that even at high homophily levels the well-known small-world phenomenon is preserved. We finally provide an application of the homophilous random network model within a stylized labor market setting to answer the introductory questions.

In the literature the presence of homophily has been established in a wide range of sociological and economic settings. Empirical studies on social networks discovered strong evidence for the similarity of connected individuals with respect to age (e.g. Verbrugge, 1977; Marsden, 1988; Burt, 1991), education (e.g. Marsden, 1987; Kalmijn, 2006), income (e.g. Laumann, 1966, 1973), ethnicity (e.g. Baerveldt et al., 2004; Ibarra, 1995) or geographical distance (e.g. Campbell, 1990; Wellman, 1996).

For an extensive survey see McPherson et al. (2001). In recent years economists have developed an understanding of the relevance of network effects in a range of economic contexts. Thus, bearing in mind the presence of homophily in real-world networks can be of great importance for creating meaningful economic models.

There already exists a strand of economic literature examining homophily effects in different settings (see e.g. Currarini et al., 2009). Most of the models assume a finite type-space and binary homophily in the sense that an agent prefers to connect to others that are of the same type while not distinguishing between other types.¹ Thus, these models rather capture the idea of equality than of similarity. However, in reality people are in many respects neither “equal” nor “different”. We therefore believe that a notion that provides an ordering of the “degree of similarity” with respect to which an agent orders his preference for connections can capture real-world effects more accurately. This gives rise to a continuous notion of homophily in networks.

Besides the presence of homophily, further stylized facts about real-world networks have been identified in empirical studies, such as the well-known small-world phenomenon and high levels of clustering (see e.g. Milgram, 1967; Watts and Strogatz, 1998).² As in many cases these networks are very large and remain unknown for an analysis, typically random networks are used as an approximation. This constitutes a challenge to design the random network formation process in a way to ensure it complies with the observed stylized facts.

Since the seminal work of Erdős and Rényi (1959), who developed and analyzed a random graph model where a fixed number out of all possible bilateral connections is randomly chosen, a lot of different models have been proposed (see e.g. Wasserman and Pattison, 1996; Watts and Strogatz, 1998; Barabási and Albert, 1999). The most commonly used until today is the BRG model, where connections between any two agents are established with the same constant probability. It has been shown that for large networks this model is almost equal to the original model of Erdős and Rényi (1959) (for details see Jackson, 2005; Bollobás, 2001).³ It is well understood that this model reproduces the small-world phenomenon but does not exhibit clustering. Equally, a notion of homophily is not present as the described random process does not rely on individual characteristics.

The latter is also true for the small-world model proposed by Watts and Strogatz

¹For several homophily measures of this kind see Currarini et al. (2009).

²The small-world phenomenon describes the observation that even in large networks on average there exist relatively short paths between two individuals, while clustering means that two individuals with a common neighbor have an increased probability of being connected.

³In fact, the BRG model rather than their original one is nowadays also known as the Erdős-Rényi model.

(1998). Starting from a network built on a low-dimensional regular lattice, they reallocate randomly chosen links and obtain a random network showing a small-world phenomenon. According to their notion this encompasses an increased level of clustering. However, the socio-economic causality of this occurrence remains uncertain. In this regard our model can to some extent serve as a socio-economic foundation of the work of Watts and Strogatz (1998).

In general, not much work has yet been dedicated to the incorporation of homophily into random networks. However, some papers exist that include similar ideas. Jackson (2008) analyzes the impact of increasing homophily on network statistics such as clustering and the average distance of nodes. A finite number of types as well as linking probabilities between them are exogenously given. Though linking probabilities may vary among types, which allows for cases where similar types are preferred, his notion of homophily remains binary. Golub and Jackson (2012) also assume a finite number of types as well as the linking probabilities between them to be exogenously given. Based on this they analyze the implications of homophily in the framework of dynamic belief formation on networks. Bramoullé et al. (2012) combine random link formation and local search in a sequentially growing society of heterogeneous agents and establish a version of binary homophily along with a degree distribution.

Besides the continuous notion of homophily, a major distinction of our approach is the sequential combination of two random processes, where agents' characteristics are considered as random variables that influence the random network formation. We thus account for the fact that in many applications in which the network remains unobserved, it seems unnatural to assume that individual characteristics, which in fact may depict attitudes, beliefs or abilities, are perfectly known.

We conclude the paper by providing an application of our model for the labor market, proposing an analysis of the introductory question: When is it optimal for a firm to search for a new employee via the contacts of a current employee? We assume the characteristic of each worker to be her individual ability to fill the open vacancy and use our homophilous random network model as an approximation of the workers' network. Given an agent and her characteristic, we determine the expected characteristic of a random contact (Proposition 5). This gives rise to a simple decision rule, stating in which constellations firms should hire via the social network. In particular, given sufficiently high levels of homophily and the current employee's ability, it proves to be always optimal to hire via the social network.

Within the job search literature, Horváth (2011) and Zaharieva (2013) incorporate homophily among contacts into job search models. However, these models are again based on a binary concept of homophily and do not include an explicit notion of networks. This research strand traces back to the work of Montgomery (1991), who

was the first to address this issue. Finally, our application to some extent captures an idea proposed by Ioannides and Loury (2004) to combine this class of models with a random network setting à la Erdős-Rényi.

The rest of the paper is organized as follows. In Section 2 we set up the model. Section 3 reveals basic properties of homophilous random networks, while results on clustering can be found in Section 4. In Section 5 we simulate the model focusing on the small-world phenomenon. Section 6 contains the labor market application and Section 7 concludes.

2 The Model

We set up a model of random network formation, where first each agent is randomly assigned a continuous characteristic which then will influence the respective linking probabilities.

Consider a set of agents $N = \{1, 2, \dots, n\}$, who will be connected via a non-directed network. A connection of two agents $i, j \in N$ will be denoted by ij , and we will denote by $g^N = \{ij \mid i, j \in N\}$ the complete network, that is the network where any two agents are connected. Then, we let $\mathbb{G} = \{g \mid g \subseteq g^N\}$ be the set of all possible non-directed networks. Further, we define $N_i(g) = \{j \in N \mid ij \in g\}$ to be the set of neighbors of agent i in network g , and let $\eta_i(g) = |N_i(g)|$ denote the number of her neighbors.

Each agent will be assigned a characteristic p_i , where the vector $p = (p_1, p_2, \dots, p_n)$ will be a realization of the random variable $P = (P_1, P_2, \dots, P_n)$. The underlying distribution of each P_i is assumed to be standard uniform, hence all P_i are identically and independently distributed.

Subsequent to the assignment of characteristics a random network forms. Here, we assume the following variation of the Bernoulli Random Graph (BRG) model as introduced by Erdős and Rényi (1959). The linking probability of two agents $i, j \in N$ is given by

$$q(p_i, p_j) = \lambda a^{|p_i - p_j|}, \quad (2.1)$$

where $\lambda, a \in [0, 1]$ are exogenous parameters independent of agents i and j .

Figure 1 depicts the linking probabilities $q(p_i, p_j)$ for different parameters a , first as a function of the distance of characteristics and second as a function of p_j for given $p_i = 0.25$. As in our model λ simply serves as a scaling parameter corresponding to the linking probability in the BRG model, in Figure 1 it is fixed to 1 for simplicity.

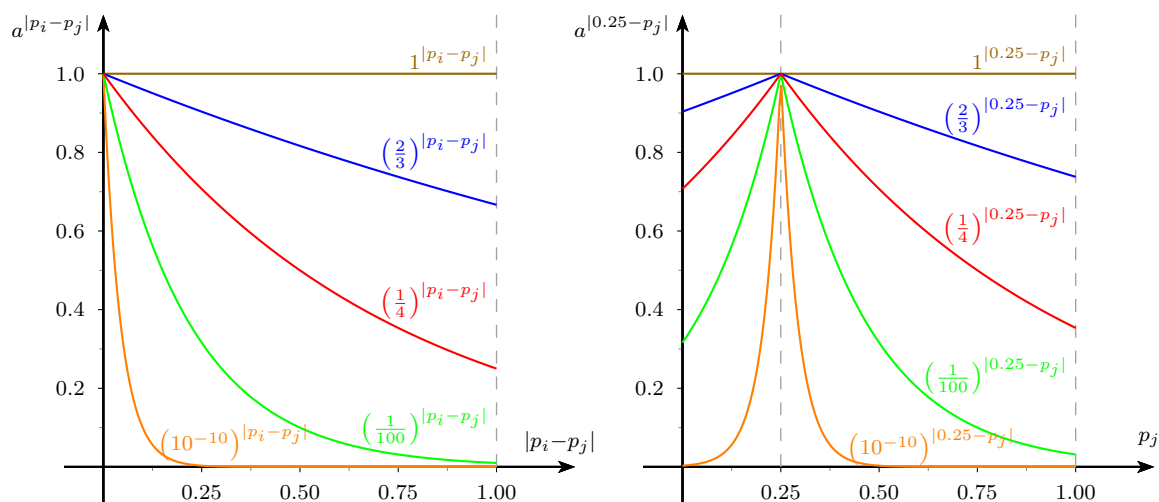


Figure 1: (i) Linking probability for all distances of characteristics, for several homophily parameters a ; (ii) Linking probabilities for an agent with characteristic $p_i = 0.25$, for several homophily parameters a .

In addition, let us shortly elaborate on the role of parameter a . Observe that the linking probability q is decreasing in $|p_i - p_j|$, as a takes values only in $[0, 1]$. In particular, for $a = 1$ the model is equal to the BRG model, as all linking probabilities are equal to λ and hence independent of the agents' characteristics, whereas if $a = 0$ solely agents with identical characteristics $p_i = p_j$ will connect with probability λ , while all other linking probabilities are zero.

Insofar, the parameter a serves as a measure of homophily in the model, where lower values correspond to a higher homophily level in the network. The notion at hand measures homophily in a continuous instead of a binary manner, since the distance function $|\cdot|$ is continuous.

Finally, it is important to understand that in situations where the vector of characteristics is unknown the linking probability $q(P_i, P_j)$ is in fact a conditional probability.

3 Basic Properties of Homophilous Random Networks

This section will constitute a foundation for the upcoming main results. To this end we first need to collect several important properties of the homophilous random

network model, such as the expected linking probabilities and the number of links of agents. We moreover discuss a threshold theorem for an agent to be isolated, what will particularly be of importance for the labor market application provided in Section 6.

Throughout this section we explore on the one hand situations in which the realization of one considered agent $i \in N$ is known while all others are not, and on the other hand situations in which the whole vector of characteristics is unknown. In any case we demonstrate that the BRG model is recuperated as the limit case of no homophily and we thus provide a generalization thereof.

We start by determining the expected linking probabilities for two given agents $i, j \in N$.

Proposition 1.

The expected probability that the link ij forms, given agent i 's realized characteristic is $P_i = p_i$ while all other characteristics p_{-i} are unknown is

$$\mathbb{E}^P [\mathbb{P}^G (ij \in G \mid P) \mid P_i = p_i] = \frac{\lambda}{\ln(a)} (a^{p_i} + a^{1-p_i} - 2) =: \varphi(\lambda, a, p_i). \quad (3.1)$$

The expected probability that the link ij forms if the vector p is unknown is

$$\mathbb{E}^P [\mathbb{P}^G (ij \in G \mid P)] = \frac{2\lambda}{\ln(a)^2} (a - 1 - \ln(a)) =: \Phi(\lambda, a). \quad (3.2)$$

The proof of Proposition 1 as well as all subsequent proofs can be found in the Appendix. It is straightforward to understand that the function φ indeed has to depend on characteristic p_i , as it makes a difference whether p_i tends to the center or to the boundaries of the interval $[0, 1]$. The closer p_i is to 0.5 the smaller is the expected distance to other agents' characteristics, hence the higher is the expected linking probability φ . In particular, it is $\operatorname{argmax}_{p_i} \varphi = 0.5$ and $\operatorname{argmin}_{p_i} \varphi = \{0, 1\}$ for all $a \in (0, 1)$. To this respect it is obvious that $\varphi(\lambda, a, 0) \leq \Phi(\lambda, a) \leq \varphi(\lambda, a, 0.5)$ for all $\lambda, a \in [0, 1]$.

It is also important to notice that the expected linking probability is decreasing in homophily, that is for all $a \in (0, 1]$

$$\frac{\partial}{\partial a} \Phi(\lambda, a) = \frac{\partial}{\partial a} \left[2\lambda \frac{a - 1 - \ln(a)}{\ln(a)^2} \right] = 2\lambda \frac{2(1 - a) + \ln(a)(1 + a)}{a \ln(a)^3} > 0.^4$$

⁴We indeed can include the value $a = 1$ here, as it happens to be a removable discontinuity of the derivative. On the contrary at $a = 0$ the right-handed derivative is infinity as the expected number of links is zero with probability one.

To verify intuition that our model reproduces the BRG model as a limit case and to gain insights on the behavior in boundary cases the following corollary is concerned with the limits of the expected linking probabilities with respect to the homophily parameter a .

Corollary 1.

For maximal homophily, meaning $a \rightarrow 0$ the expected linking probability is

$$\lim_{a \rightarrow 0} \varphi(\lambda, a, p_i) = \lim_{a \rightarrow 0} \Phi(\lambda, a) = 0. \quad (3.3)$$

In case of no homophily, meaning $a \rightarrow 1$ the expected linking probability is

$$\lim_{a \rightarrow 1} \varphi(\lambda, a, p_i) = \lim_{a \rightarrow 1} \Phi(\lambda, a) = \lambda. \quad (3.4)$$

Maximal homophily in this model means that only agents with identical characteristics would have a strictly positive linking probability. However, since the standard uniform distribution has no mass point such two agents do not exist with positive probability. Therefore, both according expected linking probabilities φ and Φ tend to zero.

In case of no homophily, as mentioned before the model indeed reproduces the BRG model, such that all linking probabilities are alike, independent of individual characteristics p .

Based on Proposition 1 we also immediately get the expected number of links of an agent.

Corollary 2.

The expected number of links of an agent with given characteristic $P_i = p_i$ is

$$\mathbb{E}^P [\mathbb{E}^G [\eta_i(G) \mid P] \mid P_i = p_i] = (n - 1)\varphi(\lambda, a, p_i), \quad (3.5)$$

and likewise if p is unknown

$$\mathbb{E}^P [\mathbb{E}^G [\eta_i(G) \mid P]] = (n - 1)\Phi(\lambda, a). \quad (3.6)$$

A proof of Corollary 2 is omitted as it is clear that all expected linking probabilities are independent and the result hence directly follows from the proof of Proposition 1. Observe that from this result we can also directly calculate the expected number of links in a network to be

$$\frac{n(n - 1)}{2} \Phi(\lambda, a).$$

From Corollary 1, we deduce that the expected number of links is zero for maximal homophily while in case of no homophily one gets $\lambda n(n-1)/2$ links in expectation, again as in the BRG model.

In the following we calculate the expected probability for an agent with given characteristic to have a certain number of links and thus show that the model inherits a version of the binomial distribution known from the BRG model.

Proposition 2.

The expected probability that an agent with given characteristic $P_i = p_i$ has exactly $k \in \{0, 1, \dots, n-1\}$ links is given by

$$\mathbb{E}^P [\mathbb{P}^G (\eta_i(G) = k \mid P) \mid P_i = p_i] = \binom{n-1}{k} \cdot \varphi(\lambda, a, p_i)^k \cdot (1 - \varphi(\lambda, a, p_i))^{n-k-1}. \quad (3.7)$$

Observe that this form can be interpreted as a binomial distribution with parameters $\varphi(\lambda, a, p_i)$ and $n-1$. It is also worth noting that the extreme cases meet the expected outcome, as it is

$$\lim_{a \rightarrow 0} \mathbb{E}^P [\mathbb{P}^G (\eta_i(G) = k \mid P) \mid P_i = p_i] \stackrel{(3.3)}{=} \binom{n-1}{k} \cdot 0^k \cdot 1^{n-k-1} = \begin{cases} 1, & \text{if } k = 0 \\ 0, & \text{else} \end{cases},$$

$$\lim_{a \rightarrow 1} \mathbb{E}^P [\mathbb{P}^G (\eta_i(G) = k \mid P) \mid P_i = p_i] \stackrel{(3.4)}{=} \binom{n-1}{k} \cdot \lambda^k \cdot (1 - \lambda)^{n-k-1},$$

where the latter case unsurprisingly is exactly the probability for any agent to have exactly k links in the BRG model, where the independent linking probability is λ . Unfortunately, the calculation of such a form in case of the whole vector of characteristics p being unknown is analytically not tractable.

One major reason why random network models are used frequently is to match qualitative characteristics of real world networks. The Law of Large Numbers in this case yields that large networks indeed meet these characteristics with a high probability (cf. e.g. Jackson, 2010b, Chapter 4). A seminal contribution of Erdős and Rényi (1959) was to give so-called threshold theorems for the case of the BRG model. These results state that if the network size n goes to infinity while the linking probability $\lambda(n)$ goes to zero slower than some threshold $t(n)$, the limit network has a certain property with probability one, while if $\lambda(n)$ goes to zero faster than $t(n)$ then the limit network has the same property only with probability zero.⁵

⁵For a more elaborate characterization of thresholds as well as several results see Bollobás (1998).

It is clear that this kind of results can only be found for monotone properties, that is for those which yield that if any network g has the property then also any network $g' \supseteq g$ has it. One example is the property that a given agent has at least one link. Observe that particularly regarding our application of the labor market (Section 6) this feature will be a prerequisite and therefore of great importance, as determining the expected characteristic of a given agent's contact is meaningful only if this agent is not isolated. Thus, we now establish a threshold function for this particular monotone property.

Proposition 3.

Assume a minimal level of homophily to be guaranteed in the model as the network size becomes large. Then the function $t(n) = 1/(n - 1)$ is a threshold for a given agent to be non-isolated in the following sense:

$$\mathbb{E}^P [\mathbb{P}^G (\eta_i(G) \geq 1 \mid P) \mid P_i = p_i] \rightarrow 1 \quad \forall p_i \in [0, 1] \quad \text{if } \frac{-\lambda(n)/\ln(a(n))}{t(n)} \rightarrow \infty,$$

$$\mathbb{E}^P [\mathbb{P}^G (\eta_i(G) \geq 1 \mid P) \mid P_i = p_i] \rightarrow 0 \quad \forall p_i \in [0, 1] \quad \text{if } \frac{-\lambda(n)/\ln(a(n))}{t(n)} \rightarrow 0.$$

Notice first that in Proposition 3 the right-hand side conditions are equivalent to $\varphi(\lambda(n), a(n), \hat{p})/t(n)$ converging to infinity or 0, respectively, for any arbitrary $\hat{p} \in [0, 1]$. For details refer to the proof in the Appendix.

What is surprising about this as well as about other threshold theorems is the sharp distinction made by the threshold $t(n)$, in the sense that if the growth of the probability φ passes the threshold $t(n)$, the probability of any agent being isolated changes “directly” from 0 to 1. What is more, notice that the threshold $t(n) = 1/(n - 1)$ is actually the same as in the BRG model, however it has to hold for φ rather than just for λ , as in this model both λ and a may vary on the size of the network. Indeed, it does not seem farfetched to assume that homophily increases with the network size, as the assortment of similar agents gets larger. Having understood this one can directly deduce the cases where only one of the two parameters varies with n :

Corollary 3.

If $a \equiv a(n)$ depends on n but λ does not, one gets that if $a(n)$ goes toward zero faster than $\exp(-n)$ then any given agent will be isolated with probability one in the limit, while if $a(n)$ does not go toward zero or at least slower than $\exp(-n)$ then any given agent will have at least one link with probability one in the limit.

If $\lambda \equiv \lambda(n)$ depends on n but a does not, the condition collapses to the threshold

of $t(n)$ for $\lambda(n)$ as in the BRG model, where any given agent has at least one link if $\lambda(n)$ grows faster than $t(n)$ while if $\lambda(n)$ grows slower than $t(n)$ any given agent is isolated with probability 1.

Both parts of Corollary 3 follow directly from Proposition 3, such that a proof can be omitted.

4 Clustering

As mentioned in the Introduction a main criticism of the Bernoulli Random Graph (BRG) model is that the resulting networks do not exhibit clustering, while most examples of real-world networks do (see e.g. Watts and Strogatz, 1998; Newman, 2003, 2006). In this section we will show that our model indeed exhibits clustering and one can use the homophily parameter a to calibrate the model to a broad range of clustering degrees.

The notion of clustering in general captures the extent to which connections in networks are transitive, that is the frequency with which two agents are linked to each other if they have a common neighbor. Watts and Strogatz (1998), who introduced this concept, measure the transitivity of a network by a global clustering coefficient C which denotes the average probability that two neighbors of a given agent are also directly linked. A random graph model is said to exhibit clustering, if the coefficient C is larger than the general, unconditional linking probability of two agents (cf. Newman, 2006). Defining the set of networks that include some link $ij \in g^N$ as $\mathbb{G}_{ij} = \{g \subseteq g^N \mid ij \in g\} \subset \mathbb{G}$, this can be transferred to our model in the following way:

Definition 1 (*Clustering*):

For the model as introduced in Section 2 with $\lambda \in [0, 1]$ and $a \in (0, 1)$, the **clustering coefficient** is defined as

$$C(\lambda, a) := \mathbb{E}^P \left[\mathbb{P}^G (G \in \mathbb{G}_{jk} \mid P) \mid G \in \mathbb{G}_{ij} \cap \mathbb{G}_{ik} \right],$$

where $i, j, k \in N$.

The model is said to exhibit clustering if $C(\lambda, a) > \Phi(\lambda, a)$.

The choice of the agents i, j and k obviously cannot have an influence in this context, since ex-ante all agents are equal. Remember also that Φ gives the probability

of two agents to be connected, characteristics being unknown. The function C as well captures this probability, however conditional on the existence of a common neighbor.

It should be clear that the original BRG model does not exhibit clustering since every link is formed with the same probability independent of the presence of common neighbors. However, as we will discover next, apart from the limit case of no homophily the model at hand possesses this feature and is insofar more realistic.

Proposition 4.

In the homophilous random network model the clustering coefficient is given by

$$C(\lambda, a) = \lambda \frac{3(\ln(a)a^2 + \ln(a) - a^2 + 1)}{2(2\ln(a)a + 4\ln(a) + a^2 - 8a + 7)}.$$

Given a non-extreme homophily parameter the model exhibits clustering, that is for all $\lambda \in (0, 1]$, $a \in (0, 1)$ it is

$$C(\lambda, a) > \Phi(\lambda, a).$$

The intuition for Proposition 4 is the following: If there is homophily to some degree and two agents have a common neighbor, then this fact contains additional information. The expected distance between these two agents is smaller than if the assumption of a common neighbor had not been given. Again due to homophily, it is therefore more likely that a link between these two agents will form. Beyond that, Figure 2 might also contribute to a better understanding of the situation. Notice here that it is $C(\lambda, a)/\lambda = C(1, a)$ and $\Phi(\lambda, a)/\lambda = \Phi(1, a)$. One can perceive that the difference $C(\lambda, a) - \Phi(\lambda, a)$ is strictly decreasing in $a \in (0, 1)$ for all $\lambda \in (0, 1]$, that is clustering is strictly increasing in the degree of homophily. Moreover, it again turns out to be interesting to consider the limit cases of maximal and no homophily:

Corollary 4.

For maximal homophily, meaning $a \rightarrow 0$ it is

$$\lim_{a \rightarrow 0} C(\lambda, a) = \lim_{a \rightarrow 0} [C(\lambda, a) - \Phi(\lambda, a)] = \frac{3\lambda}{8}.$$

In case of no homophily, meaning $a \rightarrow 1$ we get

$$\lim_{a \rightarrow 1} C(\lambda, a) = \lim_{a \rightarrow 1} \Phi(\lambda, a) = \lambda.$$

If there is no homophily, we are again back in the BRG model about which we already knew that it does not exhibit clustering. The second part of Corollary 4 confirms this. However, the more interesting case is the one of maximal homophily. Though in the limit no link forms with positive probability, from this analysis one can deduce properties in case of homophily being high, yet non-maximal, due to continuity of the functional forms.

Let us clarify the intuition why the clustering coefficient for maximal homophily takes a value strictly between zero and λ . Recall first that it is $\lim_{a \rightarrow 0} \Phi(\lambda, a) = 0$, since for maximal homophily only agents with identical characteristics are linked with positive probability and such two agents exist with probability zero. However, the clustering coefficient is a probability already conditioned on the existence of links to a common neighbor. This additional information implies that either characteristics are equal or links have formed despite differing characteristics. Though both events occur only with probability zero, this does not preclude them per se. Having understood this, it should be clear that in the first case the probability of the third link would indeed be λ , while in the second case it would still be zero. Taken together, this yields $\lim_{a \rightarrow 0} C(\lambda, a) \in (0, \lambda)$. However, it remains surprising that the clustering coefficient takes the specific value $\frac{3\lambda}{8}$.

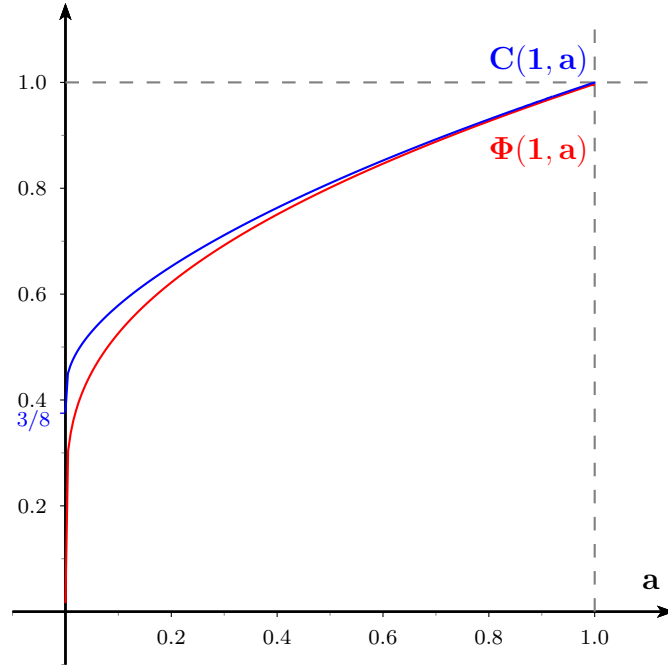


Figure 2: Clustering

5 The Small-World Phenomenon

Besides the presence of homophily and clustering another stylized fact is frequently observed in many real-world networks, which is widely known as the small-world phenomenon. It captures the finding that even in large networks there typically exist remarkably short paths between two individuals. The original BRG model is known to reproduce this characteristic (see e.g. Bollobás, 2001; Chung and Lu, 2002). Thus, in this section we aim to establish the small-world phenomenon to be preserved within our homophilous random network (HRN) model even in case of high homophily. For this purpose we will simulate a variety of homophilous random networks, since this issue seems to be no longer analytically tractable. The simulations will provide a strong indication that also in cases of high homophily the small-world phenomenon remains present. Additionally, we will apply two alternative statistical notions of clustering. It will turn out that their values are not significantly different from the analytical measure given in Definition 1.

Figure 3 may already provide an intuition regarding the differences between cases of high and low homophily. In particular, while the total number of links is almost the same in both simulated 100 agent networks one observes clustering merely in the first case.

The notion of the small-world phenomenon usually grounds on the average shortest path length between all pairs of agents belonging to a network and having a connecting path. With regard to real-world networks the small-world phenomenon is a rather vague concept, since it is typically based on subjective assessments of path lengths rather than on verifiable, definite criteria. However, most people will agree that the values for several real-world networks as for instance compiled by Watts and Strogatz (1998) and Newman (2003) are surprisingly low. Insofar it could be said that most of these networks exhibit the small-world phenomenon. A formal definition of the small-world phenomenon applicable to most random network models is given by Newman (2003):

Definition 2 (*Small-world Phenomenon*):

*A network is said to show the **small-world phenomenon** if the average shortest path length \bar{d} between pairs of agents having a connecting path scales logarithmically or slower with network size n while keeping agents' expected degree constant, that is if $\bar{d}/\ln(n)$ is non-increasing in n .*

As already mentioned it has been established that the original BRG model exhibits the small-world phenomenon according to Definition 2 (see e.g. Bollobás, 2001; Chung and Lu, 2002). However, it is not clear whether this still holds for our

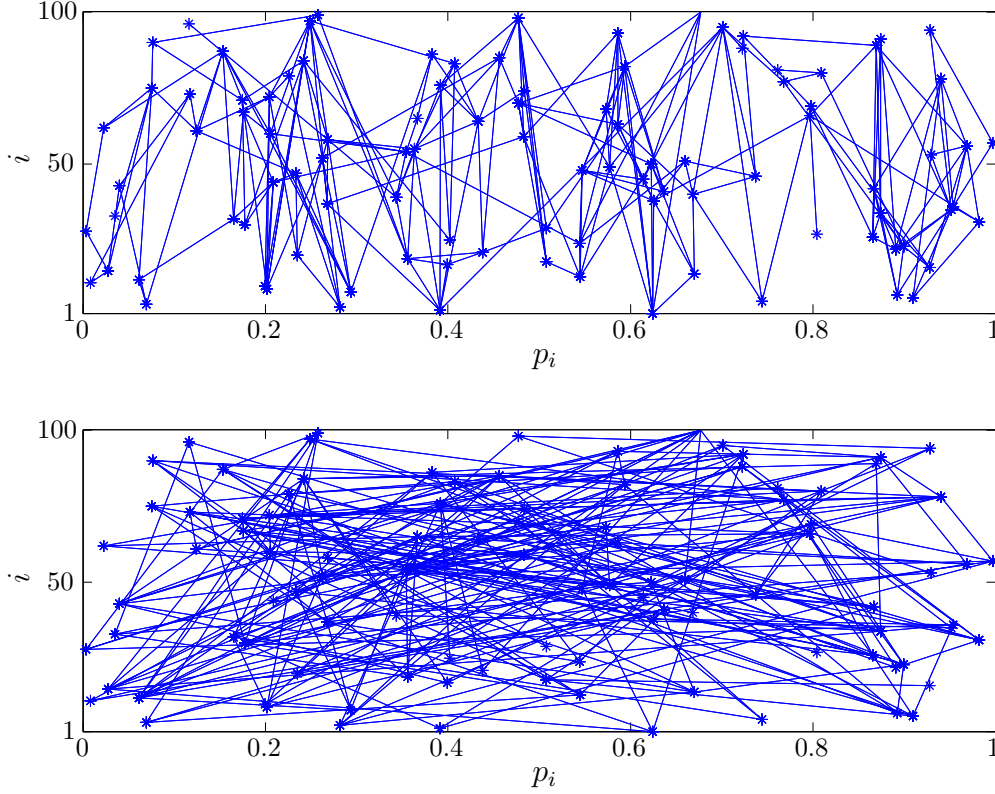


Figure 3: (1) HRN with $\lambda = 0.5, a = 10^{-8}$; $\#links = 484$
(2) BRG with linking probability $\Phi(0.5, 10^{-8}) = 0.0513$; $\#links = 496$.

generalization given a considerably high level of homophily, but the results of the following simulations will give some indication.

Prior to this, let us additionally introduce two statistical notions of clustering, which are frequently used in the literature and closely related to the one given in Definition 1. The simulations will offer the possibility to compare these. Here clustering is associated with an increased number of triangles in the network. More precisely, both alternative clustering measures are defined based on the ratio of the number of triangles and the number of connected triples. A triangle is a subset of three agents all of whom being connected to each other while a connected triple is a subset of three agents such that at least one of them is linked to the other two. Formally, this means the following:

Definition 3 (Statistical Clustering):

For a given network with set of agents $N = \{1, \dots, n\}$ the *clustering coefficients*

$C^{(1)}$ and $C^{(2)}$ are determined by

$$C^{(1)} = \frac{3 \times \text{number of triangles in the network}}{\text{number of connected triples in the network}} \quad \text{and}$$

$$C^{(2)} = \frac{1}{n} \sum_{i \in N} \frac{\text{number of triangles containing agent } i}{\text{number of connected triples centered on agent } i}.$$

The coefficient $C^{(1)}$ counts the overall number of triangles and relates it to the overall number of connected triples in the network. The factor of three accounts for the fact that each triangle contributes to three connected triples and ensures that $C^{(1)} \in [0, 1]$. The second one, $C^{(2)}$, which goes back to Watts and Strogatz (1998), first calculates an individual clustering coefficient for each agent and then averages these. Compared to the first one, $C^{(2)}$ gives more weight to low-degree agents.⁶ Additionally, notice that $C^{(2)}$ is only well-defined if there are no isolated or loose-end agents in the network.

To capture both the heuristic and the formal approach to the small-world phenomenon, we conduct two different simulations. In the first we fix the number of agents $n = 500$ and the ex-ante expected degree of any agent to $\mathbb{E}[\eta_i] = 15$. Furthermore, we select several homophily levels ranging from no homophily, i.e. the limit case of the BRG model, to very high homophily, represented by $a = 10^{-8}$. For each a we then simulate a homophilous random network $R = 1000$ times and assess the averaged network statistics. All parameters and network statistics of the simulation are stated in Table 1.

Fixing the expected degree enables us to compare the results for the different homophily levels, as this leads to identical values for $\Phi(\lambda, a)$ in all cases. Recall that Φ captures the expected probability of two agents to be connected, characteristics being unknown (cf. Proposition 1).

Regarding the results of the simulation, we find that the average path length increases in homophily. This is in line with intuition as agents with distant characteristics are increasingly likely to be distant in the network. However, it increases by less than one step from no to highest homophily and an average distance of less than 3.4 between two agents can still be considered relatively small in a network of 500 agents with about 15 links on average. Thus regarding the heuristic approach it seems reasonable to accept the small-world phenomenon to be present for all homophily levels.⁷

⁶ $C^{(2)}$ calculates the mean of the ratios while $C^{(1)}$ rather constitutes the ratio of the means (see Newman, 2003).

⁷To calculate average shortest paths one commonly restricts to agents having a connecting path if the network is not connected. However, such a network realized extremely rarely within

Parameter / Statistics	$a = 1$	$a = 10^{-2}$	$a = 10^{-4}$	$a = 10^{-6}$	$a = 10^{-8}$
n			500		
R			1000		
Exp. Degree $\mathbb{E}[\eta_i]$			15		
Exp. Linking Prob. Φ			0.0301		
λ	0.0301	0.0882	0.1553	0.2239	0.2928
Avg. Degree $\bar{\eta}$	14.9990 (0.2475)	15.0074 (0.3064)	15.0098 (0.2986)	14.9899 (0.2925)	15.0037 (0.2839)
Avg. Shortest Path \bar{d}	2.5944 (0.0113)	2.6288 (0.0164)	2.8086 (0.0277)	3.0806 (0.0429)	3.3939 (0.0611)
$\bar{d}/\ln(n)$	0.4175 (0.0018)	0.4230 (0.0026)	0.4519 (0.0045)	0.4957 (0.0069)	0.5461 (0.0098)
Clustering Coeff. C	0.0301	0.0411	0.0641	0.0892	0.1147
Clustering Coeff. $C^{(1)}$	0.0301 (0.0013)	0.0411 (0.0016)	0.0642 (0.0023)	0.0891 (0.0029)	0.1147 (0.0035)
Clustering Coeff. $C^{(2)}$	0.0301 (0.0015)	0.0411 (0.0019)	0.0642 (0.0026)	0.0892 (0.0032)	0.1148 (0.0039)

Table 1: Results of Simulation 1 comparing network statistics for different homophily levels ranging from no homophily (BRG) to extreme homophily; Standard errors stated in parentheses.

Furthermore, we observe an increasing level of clustering for the simulated homophilous random networks, what is in line with the findings in Section 4. If homophily is highest, the probability that two agents are linked given they have a common neighbor is about four times as high as in the case of the Bernoulli Random Graphs, where this probability coincides with the unconditional linking probability $\Phi(\lambda, a)$. Another expectable yet important observation is that there are no significant differences between the expected clustering coefficient C (cf. Definition 1) and the values we determined for the statistical coefficients $C^{(1)}$ and $C^{(2)}$ (cf. Definition 3).⁸

All in all, Simulation 1 indicates that the homophilous random network model exhibits the small-world phenomenon and clustering at the same time for $a \in (0, 1)$. In the following we will consider the most interesting case of highest homophily captured by $a = 10^{-8}$ in more detail.

The second simulation focuses on the formal Definition 2 of the small-world phe-

this simulation, namely only in 0.06% of all cases.

⁸Notice that isolated and loose-end agents never appeared in the simulation guaranteeing that $C^{(2)}$ was steadily well-defined.

nomenon. For this purpose we simulate a collection of $R = 100$ networks for each size $n = 150, 200, 250, \dots, 1000$ and compute the respective averages of the relevant network statistics. To this end we use the parameter of highest homophily that is considered in Simulation 1. The precise data is stated in Table 2. Notice that the simulation for each network size is structurally the same as in the first simulation, merely a smaller number of iterations is chosen due to computational restrictions. However, as can be seen in Table 1, all standard errors and especially the one of the ratio $\bar{d}/\ln(n)$ are very low. Hence, 100 iterations should be sufficient to generate a precise estimate.

Figure 4, where the ratio of the average shortest path length and the logarithm of the network size $\bar{d}/\ln(n)$ is plotted for the different network sizes n , reveals that this ratio decreases. We thus deduce that the average path length \bar{d} increases slower in n than $\ln(n)$ does. Therefore, the homophilous random networks exhibit the small-world phenomenon according to Definition 2.

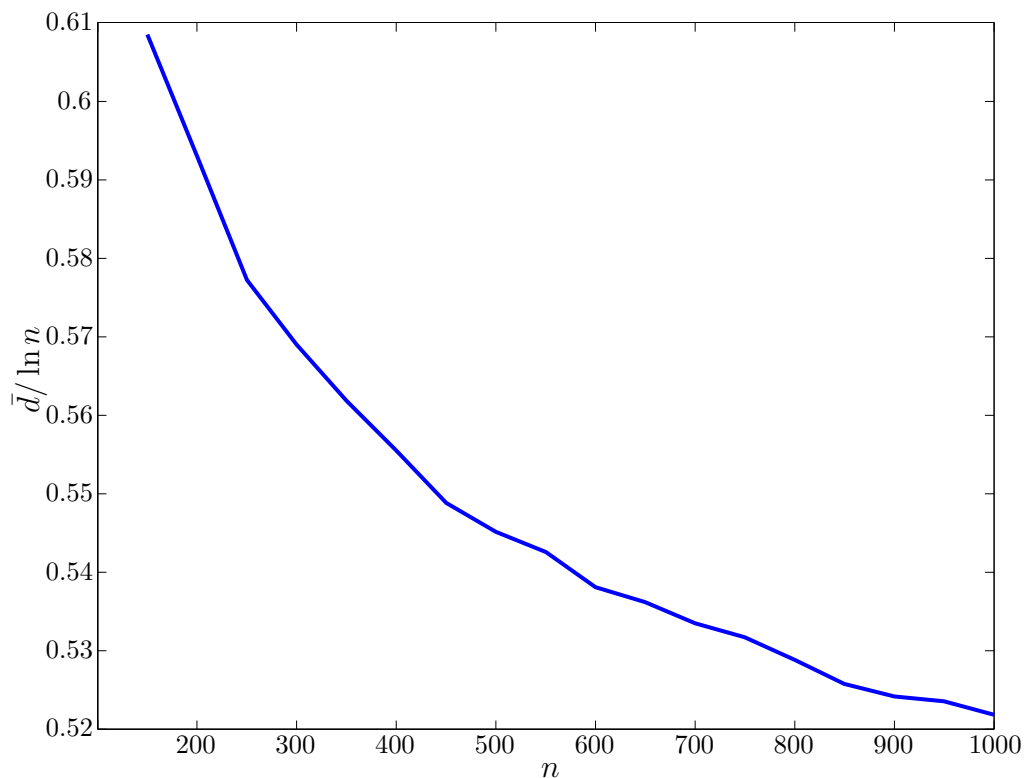


Figure 4: Small World of HRN with n from 150 to 1000 and constant expected degree 15.

Parameter / Statistics	$n = 150$	200	250	300	350	400	450	500	550
R					100				
a					10^{-8}				
Expected Degree					15				
Average Degree $\bar{\eta}$	14,99	15,02	14,98	15,02	14,97	15,00	15,01	15,03	15,02
Average Shortest Path \bar{d}	3,05	3,14	3,19	3,25	3,29	3,33	3,35	3,39	3,42
$\bar{d}/\ln(n)$	0,609	0,593	0,577	0,569	0,562	0,556	0,549	0,545	0,543

Parameter / Statistics	$n = 600$	650	700	750	800	850	900	950	1000
R					100				
a					10^{-8}				
Expected Degree					15				
Average Degree $\bar{\eta}$	15,01	15,00	15,01	14,99	14,98	15,03	15,04	14,97	15,01
Average Shortest Path \bar{d}	3,44	3,47	3,50	3,52	3,54	3,55	3,57	3,59	3,61
$\bar{d}/\ln(n)$	0,538	0,536	0,534	0,532	0,529	0,526	0,524	0,524	0,522

Table 2: Results of Simulation 2 computing average degrees, shortest paths and small world ratios of the HRN model for a growing network size.

6 An Example of the Labour Market

So far we gave a theoretical analysis of the suggested homophilous random network model. In this section however, we want to provide one possible economic application.

In recent years more and more research in the field of labor economics was dedicated to understanding the mechanisms of different hiring channels. One of these channels which is commonly used in reality relies on the contacts of current employees. Starting with the seminal contribution of Montgomery (1991), a lot of researchers decided to model the contacts between workers as a social network (cf. e.g. Calvó-Armengol, 2004; Calvó-Armengol and Jackson, 2007; Dawid and Gemkow, 2013).⁹ As known from the extensive sociological literature (cf. Section 1), within these social networks one should expect to observe homophily with respect to skills or competence, performance, education, level of income, and geographical distance. While there are lots of empirical studies confirming the existence of homophily in worker's social contacts and analyzing the implications (e.g. Mayer and Puller, 2008; Rees, 1966), only few work has yet been dedicated to developing theoretical models capturing this effect.¹⁰

In our application we consider a firm that wants to fill an open vacancy. Two possible hiring channels are available, on the one hand the formal job market and on the other hand the possibility to hire a contact of its current employee.

Based on the model introduced in Section 2, consider a network of n workers and a vector of characteristics p capturing the ability of each worker to do the vacant job. Further, assume that agent 1 is the current employee of the firm while all other agents $2, \dots, n$ are supposed to be available on the job market. While we fix p_1 as a parameter of the model, meaning that the firm knows the ability of its current employee, $p_{-1} = (p_2, \dots, p_n)$ is as before a realization of the $(n - 1)$ -dimensional random variable P_{-1} . Finally, we assume that a homophilous network randomly forms according to individual linking probabilities (2.1), for given parameters $\lambda, a \in (0, 1)$.

Knowing the distribution function of the random variable P_{-1} and the conditional linking probabilities but not the realization, the firm has to decide on one hiring channel. The expected characteristic of a contact of agent 1 can be calculated as

⁹For an extensive survey including both empirical and theoretic literature from sociology and economics see Ioannides and Loury (2004).

¹⁰Exceptions are Horváth (2011), Van der Leij and Buhai (2008) and Zaharieva (2013), all using binary notions of homophily.

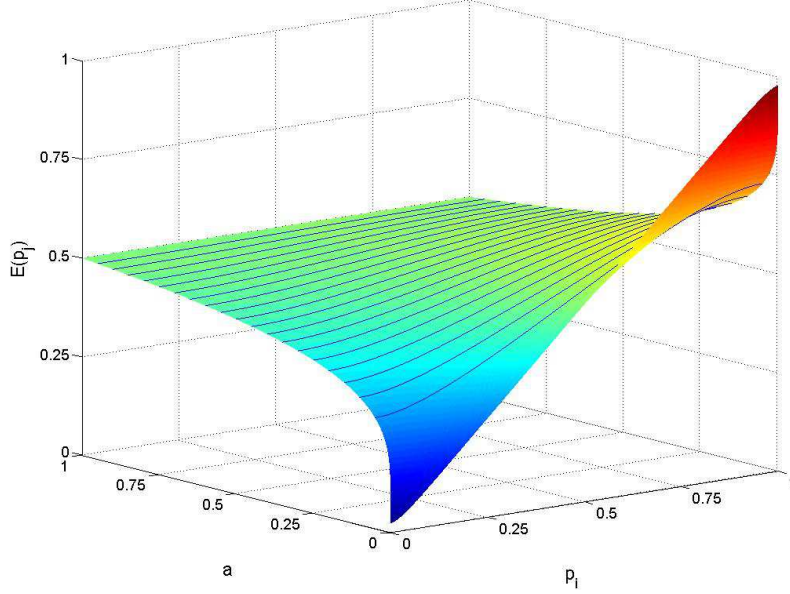


Figure 5: Expected characteristic of a contact.

follows.¹¹

Proposition 5.

Given some homophily parameter $a \in (0, 1)$, the expected characteristic of a neighbor $j \in \{2, \dots, n\}$ of agent 1 with given characteristic $p_1 \in [0, 1]$ is

$$\mathbb{E}^P [P_j \mid G \in \mathbb{G}_{1j}] = \frac{1}{2} + \frac{(a^{p_1} - a^{1-p_1})(\frac{1}{2} - \frac{1}{\ln(a)}) + 2p_1 - 1}{2 - a^{p_1} - a^{1-p_1}}. \quad (6.1)$$

A plot of function (6.1) is given in Figure 5. However, on investigating the expected characteristic (6.1), it turns out that it has some intuitive properties for some special cases which should yield some insight to the appearance of the rather complicated functional form. We collect these in the following Corollary

¹¹Notice that this probability is meaningful only if the given agent 1 has at least one link. For large networks however this is guaranteed whenever the respective condition of the threshold theorem (cf. Proposition 3) is fulfilled.

Corollary 5.

The functional form (6.1) yields:

- $\mathbb{E}^P[P_j \mid G \in \mathbb{G}_{1j}] \Big|_{p_1=\frac{1}{2}} = \frac{1}{2} \quad \forall a \in (0, 1),$
- $\lim_{a \rightarrow 0} \mathbb{E}^P[P_j \mid G \in \mathbb{G}_{1j}] = p_1 \quad \forall p_1 \in [0, 1],$
- $\lim_{a \rightarrow 1} \mathbb{E}^P[P_j \mid G \in \mathbb{G}_{1j}] = \frac{1}{2} \quad \forall p_1 \in [0, 1].$

Again, all of these properties can be identified also in Figure 5.

Assuming for simplicity that the expected ability of a worker hired via the formal job market is some value $\bar{p} \in (0, 1)$ independent of the homophily parameter a and the ability of the current employee p_1 , it becomes clear that the firm faces a simple decision rule when to hire via the social network. Namely, for sufficiently high p_1 and low a , respectively, the expected ability of a contact exceeds any ability level \bar{p} , such that in this case the firm should prefer to hire a randomly chosen contact.

7 Conclusion

In this work we try to set up a novel homophilous random network model incorporating heterogeneity of agents. In a two-stage random process, first a one-dimensional characteristic is assigned to each vertex, throughout the paper denoted as agents. Second, based on the realized characteristics the links of a random network form whilst taking into account a continuous notion of homophily that captures the frequently observed propensity of individuals to connect with similar others. Due to a continuous formalization of homophily our approach allows for a broad range of homophily levels ranging from the extreme case of maximal homophily where only equal agents get linked with positive probability to the case where there is no homophily at all. The latter case corresponds to the Bernoulli Random Graph (BRG) model, often referred to as the Erdős-Rényi model. Insofar, our model can also be regarded as a generalization hereto.

Most importantly, unlike the vast majority of related economic models we indeed capture homophily as it is defined and used in the sociological literature in terms of similarity rather than equality.

In Section 3 we reveal some basic properties and network statistics of the homophilous random network model and establish a threshold theorem. The comparison with the BRG model provides additional insight. In Section 4 we focus on another stylized fact of real-world networks, namely the occurrence of clustering,

a form of transitivity among connections. Though homophily and clustering are frequently observed in reality, both phenomena are not captured by the original BRG model. While we reveal by simulations that the small-world phenomenon is apparently preserved, we are able to show analytically that in our model homophily induces clustering. This gives rise to the conjecture that also in reality there might be a considerable causality between the two. It might be worthwhile for future research to pursue this question.

Finally, we provide an easily accessible application of our model for labor economics (Section 6). Assuming homophily with respect to abilities, we consider a network of workers according to the setting of the introduced homophilous random network model. We determine the expected ability of a given worker's random contact to do a certain job. This yields a simple decision rule for a firm which wants to fill a vacancy and needs to decide whether to hire through a current employee's contacts or the formal job market. It proves to be always optimal to rely on the current employee's contacts if this worker's ability as well as the level of homophily in the network are sufficiently high.

While our simulation results already yield a strong indication, for future work it still remains open to show analytically that even in cases of high homophily the small-world phenomenon is preserved in homophilous random networks.

As a second point it would be a natural yet analytically challenging extension to check the qualitative robustness of the findings for different distributions of characteristics. For many applications a distribution that puts more weight on intermediate characteristics would without doubt capture reality more accurately. Also, an extension of the model to multidimensional characteristics would be valuable, in particular if one would succeed to combine characteristics of both continuous and binary nature.

Finally, a calibration of the model to real-world data is yet to be done. Doing this in a meaningful way is most certainly a challenge, especially as the level of homophily within a network is not clearly observable. However, one way to deal with this would be to calibrate the model to the observable degree of clustering, which we showed to be directly connected to homophily in our model.

8 Appendix

Proof of Proposition 1. Calculate the expected probability:

$$\begin{aligned}
\mathbb{E}^P [\mathbb{P}^G (ij \in G \mid P) \mid P_i = p_i] &= \mathbb{E}^P [\lambda a^{|P_i - P_j|} \mid P_i = p_i] \\
&= \lambda \left(\int_0^1 \underbrace{f_{P_j}(p_j)}_1 a^{|p_i - p_j|} dp_j \right) \\
&= \lambda \left(\int_0^{p_i} a^{p_i - p_j} dp_j + \int_{p_i}^1 a^{p_j - p_i} dp_j \right) \\
&= \lambda \left(a^{p_i} \int_0^{p_i} a^{-p_j} dp_j + a^{-p_i} \int_{p_i}^1 a^{p_j} dp_j \right) \\
&= \lambda \left(a^{p_i} \frac{1 - a^{-p_i}}{\ln(a)} + a^{-p_i} \frac{a - a^{p_i}}{\ln(a)} \right) \\
&= \frac{\lambda}{\ln(a)} (a^{p_i} + a^{1-p_i} - 2). \tag{8.1}
\end{aligned}$$

What is more, by integrating (8.1) with respect to p_i we get the expected probability if p is unknown.

$$\begin{aligned}
\mathbb{E}^P [\mathbb{P}^G [ij \in G \mid P]] &= \mathbb{E}^P [\lambda a^{|P_i - P_j|}] \\
&= \lambda \left(\int_{[0,1]^2} \underbrace{f_{P_i, P_j}(p_i, p_j)}_{=f_{P_i}(p_i)f_{P_j}(p_j)=1} a^{|p_i - p_j|} d(p_i, p_j) \right) \\
&\stackrel{(8.1)}{=} \lambda \left(\int_0^1 \frac{(a^{p_i} + a^{1-p_i} - 2)}{\ln(a)} dp_i \right) \\
&= \frac{\lambda}{\ln(a)} \left[\frac{a^{p_i} - a^{1-p_i} - 2p_i \ln(a)}{\ln(a)} \right] \Big|_0^1 \\
&= \frac{\lambda}{\ln(a)^2} [a - 1 - 2 \ln(a) - 1 + a] \\
&= \frac{2\lambda}{\ln(a)^2} [a - 1 - \ln(a)].
\end{aligned}$$

□

Proof of Corollary 1. Using l'Hôpital's rule, calculate the limit of φ as

$$\begin{aligned}\lim_{a \rightarrow 0} \varphi(\lambda, a, p_i) &= \lim_{a \rightarrow 0} \frac{\lambda(a^{p_i} + a^{1-p_i} - 2)}{\ln(a)} \\ &= \lim_{a \rightarrow 0} \frac{\lambda(p_i a^{p_i-1} + (1-p_i)a^{-p_i})}{1/a} \\ &= \lim_{a \rightarrow 0} \lambda(p_i a^{p_i} + (1-p_i)a^{1-p_i}) = 0,\end{aligned}$$

and likewise

$$\begin{aligned}\lim_{a \rightarrow 1} \varphi(\lambda, a, p_i) &= \lim_{a \rightarrow 1} \frac{\lambda(a^{p_i} + a^{1-p_i} - 2)}{\ln(a)} \\ &= \lim_{a \rightarrow 1} \frac{\lambda(p_i a^{p_i-1} + (1-p_i)a^{-p_i})}{1/a} \\ &= \lim_{a \rightarrow 1} \lambda(p_i a^{p_i} + (1-p_i)a^{1-p_i}) = \lambda.\end{aligned}$$

For the case of Φ , we get by using l'Hôpital's rule twice

$$\lim_{a \rightarrow 0} \Phi(\lambda, a) = \lim_{a \rightarrow 0} 2\lambda \frac{a - 1 - \ln(a)}{\ln(a)^2} = \lim_{a \rightarrow 0} 2\lambda \frac{1 - 1/a}{2 \ln(a)/a} = \lim_{a \rightarrow 0} \lambda \frac{a - 1}{\ln(a)} = 0,$$

as well as

$$\lim_{a \rightarrow 1} \Phi(\lambda, a) = \lim_{a \rightarrow 1} 2\lambda \frac{a - 1 - \ln(a)}{\ln(a)^2} = \lim_{a \rightarrow 1} 2\lambda \frac{a - 1}{2 \ln(a)} = \lim_{a \rightarrow 1} \lambda \frac{1}{1/a} = \lambda.$$

□

Proof of Proposition 2. Calculate

$$\begin{aligned}
& \mathbb{E}^P \left[\mathbb{P}^G (\eta_i(G) = k \mid P) \mid P_i = p_i \right] \\
&= \mathbb{E}^P \left[\sum_{K \subseteq N \setminus \{i\}: |K|=k} \left(\prod_{j \in K} (q(P_i, P_j)) \cdot \prod_{l \in N \setminus K \setminus \{i\}} (1 - q(P_i, P_l)) \right) \mid P_i = p_i \right] \\
&= \sum_{K \subseteq N \setminus \{i\}: |K|=k} \left(\mathbb{E}^P \left[\prod_{j \in K} (q(P_i, P_j)) \cdot \prod_{l \in N \setminus K \setminus \{i\}} (1 - q(P_i, P_l)) \mid P_i = p_i \right] \right) \\
&= \sum_{K \subseteq N \setminus \{i\}: |K|=k} \left(\int_{[0,1]^{n-1}} \left(\underbrace{f_{P_{-i}}(p_{-i})}_{=1} \cdot \prod_{j \in K} (q(p_i, p_j)) \cdot \prod_{l \in N \setminus K \setminus \{i\}} (1 - q(p_i, p_l)) \right) dp_{-i} \right) \\
&= \sum_{K \subseteq N \setminus \{i\}: |K|=k} \left(\prod_{j \in K} \left(\int_0^1 (q(p_i, p_j)) dp_j \right) \cdot \prod_{l \in N \setminus K \setminus \{i\}} \left(\int_0^1 (1 - q(p_i, p_l)) dp_l \right) \right) \\
&\stackrel{(3.1)}{=} \sum_{K \subseteq N \setminus \{i\}: |K|=k} \left(\left(\frac{\lambda}{\ln(a)} (a^{p_i} + a^{1-p_i} - 2) \right)^k \cdot \left(1 - \frac{\lambda}{\ln(a)} (a^{p_i} + a^{1-p_i} - 2) \right)^{n-k-1} \right) \\
&\stackrel{(3.1)}{=} \binom{n-1}{k} \cdot (\varphi(\lambda, a, p_i))^k \cdot (1 - \varphi(\lambda, a, p_i))^{n-k-1}.
\end{aligned}$$

□

Proof of Proposition 3. The probability that an agent with given characteristic p_i is isolated is

$$\mathbb{E}^P \left[\mathbb{P}^G (\eta_i(G) = 0 \mid P) \mid P_i = p_i \right] \stackrel{(3.7)}{=} (1 - \varphi(\lambda(n), a(n), p_i))^{n-1}.$$

If we assume that there will be at least some homophily as the size of the network becomes large, that is formally

$$\exists \tilde{\epsilon} > 0, \bar{n} \in \mathbb{N} : a(n) \leq 1 - \tilde{\epsilon} \quad \forall n \geq \bar{n}$$

then we have that

$$\exists \epsilon > 0 : 2 - a(n)^{\hat{p}} - a(n)^{1-\hat{p}} \in [\epsilon, 2] \quad \forall n \geq \bar{n}.$$

Now it holds that if $\lim_{n \rightarrow \infty} [-\lambda(n)/(\ln(a(n))t(n))] = \infty$ then

$$\begin{aligned}
& \lim_{n \rightarrow \infty} (1 - \varphi(\lambda(n), a(n), p_i))^{n-1} \\
&= \lim_{n \rightarrow \infty} \left(1 - \frac{\varphi(\lambda(n), a(n), p_i)/t(n)}{n-1} \right)^{n-1} \\
&\stackrel{(3.1)}{=} \lim_{n \rightarrow \infty} \left(1 - \frac{\frac{\lambda(n)(n-1)}{\ln(a(n))} (a(n)^{p_i} + a(n)^{1-p_i} - 2)}{n-1} \right)^{n-1} \\
&= \lim_{n \rightarrow \infty} \exp \left(\underbrace{-\frac{\lambda(n)(n-1)}{\ln(a(n))}}_{\rightarrow \infty} \underbrace{(a(n)^{p_i} + a(n)^{1-p_i} - 2)}_{\in [-2, -\epsilon]} \right) \\
&= 0,
\end{aligned}$$

while if $\lim_{n \rightarrow \infty} [-\lambda(n)/(\ln(a(n))t(n))] = 0$ we have

$$\begin{aligned}
& \lim_{n \rightarrow \infty} (1 - \varphi(\lambda(n), a(n), p_i))^{n-1} \\
&= \lim_{n \rightarrow \infty} \exp \left(\underbrace{-\frac{\lambda(n)(n-1)}{\ln(a(n))}}_{\rightarrow 0} \underbrace{(a(n)^{p_i} + a(n)^{1-p_i} - 2)}_{\in [-2, -\epsilon]} \right) \\
&= 1.
\end{aligned}$$

□

Proof of Proposition 4. We calculate

$$\begin{aligned}
C(\lambda, a) &= \mathbb{E}^P [\lambda a^{|P_j - P_k|} \mid G \in \mathbb{G}_{ij} \cap \mathbb{G}_{ik}] \\
&= \lambda \int_{[0,1]^n} a^{|p_j - p_k|} f_P(p \mid G \in \mathbb{G}_{ij} \cap \mathbb{G}_{ik}) dp \\
&= \lambda \int_{[0,1]^n} a^{|p_j - p_k|} \frac{f_{P,G}(p, \mathbb{G}_{ij} \cap \mathbb{G}_{ik})}{f_G(\mathbb{G}_{ij} \cap \mathbb{G}_{ik})} dp \\
&= \frac{\lambda}{f_G(\mathbb{G}_{ij} \cap \mathbb{G}_{ik})} \int_{[0,1]^n} a^{|p_j - p_k|} f_{P,G}(p, \mathbb{G}_{ij} \cap \mathbb{G}_{ik}) dp \\
&= \frac{\lambda}{f_G(\mathbb{G}_{ij} \cap \mathbb{G}_{ik})} \int_{[0,1]^n} a^{|p_j - p_k|} f_G(\mathbb{G}_{ij} \cap \mathbb{G}_{ik} \mid P = p) \overbrace{f_P(p)}^{=1} dp \\
&= \frac{\lambda}{\int_{[0,1]^n} \underbrace{f_P(x)}_{=1} f_G(\mathbb{G}_{ij} \cap \mathbb{G}_{ik} \mid P = x) dx} \int_{[0,1]^n} a^{|p_j - p_k|} f_G(\mathbb{G}_{ij} \cap \mathbb{G}_{ik} \mid P = p) dp \\
&= \frac{\lambda}{\int_{[0,1]^n} \mathbb{P}^G(G \in \mathbb{G}_{ij} \cap \mathbb{G}_{ik} \mid P = x) dx} \int_{[0,1]^n} a^{|p_j - p_k|} \mathbb{P}^G(G \in \mathbb{G}_{ij} \cap \mathbb{G}_{ik} \mid P = p) dp \\
&= \frac{\lambda}{\int_{[0,1]^n} \lambda a^{|x_i - x_j|} \lambda a^{|x_i - x_k|} dx} \int_{[0,1]^n} a^{|p_j - p_k|} \lambda a^{|p_i - p_j|} \lambda a^{|p_i - p_k|} dp \\
&= \lambda \frac{\int_{[0,1]^n} a^{|p_j - p_k| + |p_i - p_j| + |p_i - p_k|} dp}{\int_{[0,1]^n} a^{|x_i - x_j| + |x_i - x_k|} dx} \\
&= \lambda \frac{\int_{[0,1]^3} a^{|p_j - p_k| + |p_i - p_j| + |p_i - p_k|} d(p_i, p_j, p_k)}{\int_{[0,1]^3} a^{|x_i - x_j| + |x_i - x_k|} d(x_i, x_j, x_k)}. \tag{8.2}
\end{aligned}$$

Let us solve the integral in the denominator first. For the sake of readability denote $x = (x_i, x_j, x_k)$.

$$\begin{aligned}
& \int_{[0,1]^3} a^{|x_i-x_j|+|x_i-x_k|} dx \\
&= \int_{\substack{x \in [0,1]^3: \\ x_j, x_k \leq x_i}} a^{2x_i-x_j-x_k} dx + \int_{\substack{x \in [0,1]^3: \\ x_i \leq x_j, x_k}} a^{x_j+x_k-2x_i} dx \\
& \quad + \int_{\substack{x \in [0,1]^3: \\ x_j \leq x_i \leq x_k}} a^{x_k-x_j} dx + \int_{\substack{x \in [0,1]^3: \\ x_k \leq x_i \leq x_j}} a^{x_j-x_k} dx \\
&= \frac{2 \ln(a) - 4a + a^2 + 3}{2(\ln(a))^3} + \frac{2 \ln(a) - 4a + a^2 + 3}{2(\ln(a))^3} \\
& \quad + \frac{2 \ln(a) - 4a + 2a \ln(a) + 4}{2(\ln(a))^3} + \frac{2 \ln(a) - 4a + 2a \ln(a) + 4}{2(\ln(a))^3} \\
&= \frac{1}{2(\ln(a))^3} [8 \ln(a) - 16a + 2a^2 + 4 \ln(a)a + 14].
\end{aligned}$$

Now solve the integral in the nominator of (8.2), substituting x for p in order to use the same notation as above.

$$\begin{aligned}
& \int_{[0,1]^3} a^{|x_j-x_k|+|x_i-x_j|+|x_i-x_k|} dx \\
&= \int_{\substack{x \in [0,1]^3: \\ x_i \leq x_j \leq x_k}} a^{2x_k-2x_i} dx + \int_{\substack{x \in [0,1]^3: \\ x_i \leq x_k \leq x_j}} a^{2x_j-2x_i} dx \\
& \quad + \int_{\substack{x \in [0,1]^3: \\ x_j \leq x_i \leq x_k}} a^{2x_k-2x_j} dx + \int_{\substack{x \in [0,1]^3: \\ x_j \leq x_k \leq x_i}} a^{2x_i-2x_j} dx \\
& \quad + \int_{\substack{x \in [0,1]^3: \\ x_k \leq x_i \leq x_j}} a^{2x_j-2x_k} dx + \int_{\substack{x \in [0,1]^3: \\ x_k \leq x_j \leq x_i}} a^{2x_i-2x_k} dx \\
&= 6 \frac{\ln(a) - a^2 + a^2 \ln(a) + 1}{4(\ln(a))^3} \\
&= \frac{1}{2(\ln(a))^3} [3 \ln(a) - 3a^2 + 3a^2 \ln(a) + 3].
\end{aligned}$$

All in all, we get

$$C(\lambda, a) = \lambda \frac{3 \ln(a) - 3a^2 + 3a^2 \ln(a) + 3}{8 \ln(a) - 16a + 2a^2 + 4 \ln(a)a + 14}.$$

By using this, we can now start with the actual proof. It is

$$\begin{aligned}
& C(\lambda, a) - \Phi(\lambda, a) \\
&= \lambda \left(\frac{3(\ln(a)a^2 + \ln(a) - a^2 + 1)}{2(2\ln(a)a + 4\ln(a) + a^2 - 8a + 7)} + \frac{2(\ln(a) - a + 1)}{\ln(a)^2} \right) \\
&= \lambda \frac{3\ln(a)^3(a^2+1) + \ln(a)^2(-3a^2+8a+19) + \ln(a)(-4a^2-40a+44) + (-4a^3+36a^2-60a+28)}{2\ln(a)^2(2\ln(a)a+4\ln(a)+a^2-8a+7)} \tag{8.3}
\end{aligned}$$

In the following, we will use that for $a \in (0, 1)$

$$\ln(a) = - \sum_{m=0}^{\infty} \frac{(1-a)^{m+1}}{m+1}$$

and therefore $\ln(a) < - \sum_{m=0}^M \frac{(1-a)^{m+1}}{m+1} < 0$ for all $M \in \mathbb{N}$. The first and easier part is to show that the denominator of (8.3) is negative for all $a \in (0, 1)$:

$$\begin{aligned}
& 2\ln(a)a + 4\ln(a) + a^2 - 8a + 7 \\
&= 2(a+2)\ln(a) + a^2 - 8a + 7 \\
&< -2(a+2)\left(1-a + \frac{1}{2}(1-a)^2 + \frac{1}{3}(1-a)^3\right) + a^2 - 8a + 7 \\
&= \frac{1}{3}(a+2)(2a^3 - 9a^2 + 18a - 11) + a^2 - 8a + 7 \\
&= \frac{1}{3}(2a^4 - 5a^3 + 3a^2 + a - 1) \\
&= -\frac{1}{3}(1-a)^3(2a+1) \\
&< 0
\end{aligned}$$

Furthermore, we define

$$g(a) := 3\ln(a)^3(a^2+1) + \ln(a)^2(-3a^2+8a+19) + \ln(a)(-4a^2-40a+44) + (-4a^3+36a^2-60a+28).$$

Then $\lambda g(a)$ is the nominator of (8.3). We calculate

$$\begin{aligned}
\frac{dg}{da}(a) &= \frac{1}{a} \left[6\ln(a)^3a^2 + \ln(a)^2(3a^2 + 8a + 9) + 2\ln(a)(-7a^2 - 12a + 19) + 4(-3a^3 + 17a^2 - 25a + 11) \right], \\
\frac{d^2g}{da^2}(a) &= \frac{1}{a^2} \left[6\ln(a)^3a^2 + 3\ln(a)^2(7a^2 - 3) + 4\ln(a)(-2a^2 + 4a - 5) + 6(-4a^3 + 9a^2 - 4a - 1) \right], \\
\frac{d^3g}{da^3}(a) &= \frac{1}{a^3} \left[18\ln(a)^2(a^2 + 1) + 2\ln(a)(21a^2 - 8a + 11) + 8(-3a^3 - a^2 + 5a - 1) \right], \\
\frac{d^4g}{da^4}(a) &= \frac{1}{a^4} \left[18\ln(a)^2(-a^2 - 3) + 2\ln(a)(-3a^2 + 16a - 15) + 2(25a^2 - 48a + 23) \right], \\
\frac{d^5g}{da^5}(a) &= \frac{1}{a^5} \left[36\ln(a)^2(a^2 + 6) + 12\ln(a)(-2a^2 - 8a + 1) + 2(-53a^2 + 160a - 107) \right], \\
\frac{d^6g}{da^6}(a) &= \frac{1}{a^6} \left[108\ln(a)^2(-a^2 - 10) + 12\ln(a)(12a^2 + 32a + 31) + 2(147a^2 - 688a + 541) \right].
\end{aligned}$$

Notice that

$$g(1) = \frac{dg}{da}(1) = \frac{d^2g}{da^2}(1) = \frac{d^3g}{da^3}(1) = \frac{d^4g}{da^4}(1) = \frac{d^5g}{da^5}(1) = 0$$

and moreover

$$\begin{aligned} \frac{d^6g}{da^6}(a) &= \frac{1}{a^6} \left[108 \ln(a)^2 \underbrace{(-a^2 - 10)}_{<0} + 12 \ln(a) \underbrace{(12a^2 + 32a + 31)}_{>0} + 2(147a^2 - 688a + 541) \right] \\ &< \frac{1}{a^6} \left[108(1-a)^2(-a^2 - 10) - 12(1-a)(12a^2 + 32a + 31) + 2(147a^2 - 688a + 541) \right] \\ &= \frac{2}{a^6} \left[-54a^4 + 180a^3 - 327a^2 + 386a - 185 \right] \\ &= \frac{2}{a^6} (1-a) \left[54\left(a - \frac{7}{9}\right)^3 + 103\left(a - \frac{7}{9}\right) - \frac{2146}{27} \right] \\ &< \frac{2}{a^6} (1-a) \left[54\frac{2^3}{9} + 103\frac{2}{9} - \frac{2146}{27} \right] \\ &= -\frac{112}{a^6} (1-a) \\ &< 0. \end{aligned}$$

Combining this it follows for all $a \in (0, 1)$

$$\frac{d^5g}{da^5}(a) > 0 \Rightarrow \frac{d^4g}{da^4}(a) < 0 \Rightarrow \frac{d^3g}{da^3}(a) > 0 \Rightarrow \frac{d^2g}{da^2}(a) < 0 \Rightarrow \frac{dg}{da}(a) > 0 \Rightarrow g(a) < 0.$$

Taken together we have indeed that

$$C(\lambda, a) - \Phi(\lambda, a) = \lambda \frac{g(a)}{2 \ln(a)^2 (2 \ln(a)a + 4 \ln(a) + a^2 - 8a + 7)} > 0.$$

□

Proof of Corollary 4. By applying l'Hôpital's rule three times we calculate

$$\begin{aligned}
\lim_{a \rightarrow 0} C(\lambda, a) &= \lambda \lim_{a \rightarrow 0} \frac{3 \ln(a) - 3a^2 + 3a^2 \ln(a) + 3}{8 \ln(a) - 16a + 2a^2 + 4 \ln(a)a + 14} \\
&= \lambda \lim_{a \rightarrow 0} \frac{3/a - 6a + 6a \ln(a) + 3a}{8/a - 16 + 4a + 4 \ln(a) + 4} \\
&= \frac{3\lambda}{4} \lim_{a \rightarrow 0} \frac{1 - a^2 + 2a^2 \ln(a)}{2 - 3a + a^2 + a \ln(a)} \\
&= \frac{3\lambda}{4} \frac{\lim_{a \rightarrow 0}[1 - a^2 + 2a^2 \ln(a)]}{\lim_{a \rightarrow 0}[2 - 3a + a^2 + a \ln(a)]} \\
&= \frac{3\lambda}{4} \frac{\lim_{a \rightarrow 0}[1] - \lim_{a \rightarrow 0}[a^2] + \lim_{a \rightarrow 0}[2a^2 \ln(a)]}{\lim_{a \rightarrow 0}[2] - \lim_{a \rightarrow 0}[3a] + \lim_{a \rightarrow 0}[a^2] + \lim_{a \rightarrow 0}[a \ln(a)]} \\
&= \frac{3\lambda}{4} \frac{1 - 0 + \lim_{x \rightarrow \infty}[2 \ln(1/x)/x^2]}{2 - 0 + 0 + \lim_{x \rightarrow \infty}[\ln(1/x)/x]} \\
&= \frac{3\lambda}{4} \frac{1 + \lim_{x \rightarrow \infty}[-2x(1/x^2)/2x]}{2 + \lim_{x \rightarrow \infty}[-x(1/x^2)/1]} \\
&= \frac{3\lambda}{4} \frac{1 + \lim_{x \rightarrow \infty}[-1/x^2]}{2 + \lim_{x \rightarrow \infty}[-1/x]} = \frac{3\lambda}{8}.
\end{aligned}$$

The stated result follows immediately, since we established in Corollary 1 that $\lim_{a \rightarrow 0} \Phi(\lambda, a) = 0$.

On the other side, by again using l'Hôpital's rule three times we get

$$\begin{aligned}
\lim_{a \rightarrow 1} C(\lambda, a) &= \lambda \lim_{a \rightarrow 1} \frac{3 \ln(a) - 3a^2 + 3a^2 \ln(a) + 3}{8 \ln(a) - 16a + 2a^2 + 4 \ln(a)a + 14} \\
&= \lambda \lim_{a \rightarrow 1} \frac{3/a - 6a + 6a \ln(a) + 3a}{8/a - 16 + 4a + 4 \ln(a) + 4} \\
&= \lambda \lim_{a \rightarrow 1} \frac{3 - 3a^2 + 6a^2 \ln(a)}{8 - 12a + 4a^2 + 4a \ln(a)} \\
&= \lambda \lim_{a \rightarrow 1} \frac{-6a + 12a \ln(a) + 6a}{-12 + 8a + 4 \ln(a) + 4} \\
&= \lambda \lim_{a \rightarrow 1} \frac{12 \ln(a) + 12}{8 + 4/a} \\
&= \lambda.
\end{aligned}$$

According to Corollary 1 it is as well $\lim_{a \rightarrow 1} \Phi(\lambda, a) = \lambda$ which concludes the proof. \square

Proof of Proposition 5. We calculate

$$\begin{aligned}
\mathbb{E}^P [P_j \mid G \in \mathbb{G}_{1j}] &= \int_0^1 p_j f_{P_j|G}(p_j, \mathbb{G}_{1j}) dp_j \\
&= \int_0^1 p_j f_{P_j}(p_j \mid G \in \mathbb{G}_{1j}) dp_j \\
&= \int_0^1 p_j \frac{f_{P_j, G}(p_j, \mathbb{G}_{1j})}{f_G(\mathbb{G}_{1j})} dp_j \\
&= \int_0^1 p_j \frac{f_G(\mathbb{G}_{1j} \mid P_j = p_j) \overbrace{f_{P_j}(p_j)}^1}{f_G(\mathbb{G}_{1j})} dp_j \\
&= \int_0^1 p_j \frac{f_G(\mathbb{G}_{1j} \mid P_j = p_j)}{\underbrace{\int_0^1 \underbrace{f_{P_j}(x)}_1 \underbrace{f_G(\mathbb{G}_{1j} \mid P_j = x)}_{\mathbb{P}(G \in \mathbb{G}_{1j} \mid P_j = x)} dx}} dp_j \\
&= \int_0^1 p_j \frac{\overbrace{f_G(\mathbb{G}_{1j} \mid P_j = p_j)}^{\lambda a^{|p_1 - p_j|}}}{\underbrace{\int_0^1 \lambda a^{|p_1 - x|} dx}_{\frac{\lambda}{\ln(a)}(a^{p_1} + a^{1-p_1} - 2)}} dp_j \\
&= \frac{\ln(a)}{a^{p_1} + a^{1-p_1} - 2} \int_0^1 p_j a^{|p_1 - p_j|} dp_j.
\end{aligned}$$

The integral can be calculated as follows:

$$\begin{aligned}
\int_0^1 p_j a^{|p_1 - p_j|} dp_j &= \int_0^{p_1} p_j a^{(p_1 - p_j)} dp_j + \int_{p_1}^1 p_j a^{(p_j - p_1)} dp_j \\
&= \frac{a^{p_1} - p_1 \ln(a) - 1}{\ln(a)^2} + \frac{a^{1-p_1}(\ln(a) - 1) - p_1 \ln(a) + 1}{\ln(a)^2}.
\end{aligned}$$

It follows that

$$\mathbb{E}^P(P_j \mid G \in \mathbb{G}_{1j}) = \frac{a^{p_1} + a^{1-p_1}(\ln(a) - 1) - 2p_1 \ln(a)}{\ln(a)(a^{p_1} + a^{1-p_1} - 2)} \quad (8.4)$$

$$= \frac{1}{2} + \frac{(a^{p_1} - a^{1-p_1})(\frac{1}{2} - \frac{1}{\ln(a)}) + 2p_1 - 1}{2 - a^{p_1} - a^{1-p_1}}. \quad (8.5)$$

\square

Proof of Corollary 5. Consider the functional form (6.1). We calculate the properties in question, where the first and second one turn out to be straightforward. For $a \in (0, 1)$ it is

$$\mathbb{E}^P[P_j \mid G \in \mathbb{G}_{1j}] \Big|_{p_1=\frac{1}{2}} = \frac{1}{2} + \frac{(\sqrt{a} - \sqrt{a})(\frac{1}{2} - \frac{1}{\ln(a)}) + 1 - 1}{2 - \sqrt{a} - \sqrt{a}} = \frac{1}{2}.$$

Furthermore, we get for $p_1 \in (0, 1)$

$$\lim_{a \rightarrow 0} \mathbb{E}^P[P_j \mid G \in \mathbb{G}_{1j}] = \frac{1}{2} + \frac{(0 - 0)(\frac{1}{2} + 0) + 2p_1 - 1}{2 - 0 - 0} = p_1$$

and for the marginals

$$\begin{aligned} \lim_{a \rightarrow 0} \mathbb{E}^P[P_j \mid G \in \mathbb{G}_{1j}] \Big|_{p_1=0} &= \frac{1}{2} + \frac{(1 - 0)(\frac{1}{2} + 0) + 0 - 1}{2 - 1 - 0} = 0, \\ \lim_{a \rightarrow 0} \mathbb{E}^P[P_j \mid G \in \mathbb{G}_{1j}] \Big|_{p_1=1} &= \frac{1}{2} + \frac{(0 - 1)(\frac{1}{2} + 0) + 2 - 1}{2 - 0 - 1} = 1. \end{aligned}$$

And finally, by using l'Hôpital's rule twice, we get for $p_1 \in [0, 1]$

$$\begin{aligned} &\lim_{a \rightarrow 1} \mathbb{E}^P[P_j \mid G \in \mathbb{G}_{1j}] \\ &\stackrel{(8.4)}{=} \lim_{a \rightarrow 1} \frac{a^{p_1} + a^{1-p_1}(\ln(a) - 1) - 2p_1 \ln(a)}{\ln(a)(a^{p_1} + a^{1-p_1} - 2)} \\ &= \lim_{a \rightarrow 1} \frac{p_1 a^{p_1-1} + (1-p_1)a^{-p_1}(\ln(a) - 1) + a^{-p_1} - \frac{2p_1}{a}}{\frac{1}{a}(a^{p_1} + a^{1-p_1} - 2) + \ln(a)(p_1 a^{p_1-1} + (1-p_1)a^{-p_1})} \\ &= \lim_{a \rightarrow 1} \frac{p_1(p_1 - 1)a^{p_1-2} + p_1(p_1 - 1)a^{-p_1-1}(\ln(a) - 1) + (1-p_1)a^{-p_1-1} - p_1 a^{-p_1-1} + \frac{2p_1}{a^2}}{-\frac{1}{a^2}(a^{p_1} + a^{1-p_1} - 2) + \frac{2}{a}(p_1 a^{p_1-1} + (1-p_1)a^{-p_1}) + \ln(a)(p_1(p_1 - 1)a^{p_1-2} + p_1(p_1 - 1)a^{-p_1-1})} \\ &= \frac{p_1(p_1 - 1) + p_1(p_1 - 1)(0 - 1) + (1 - p_1) - p_1 + 2p_1}{-(1 + 1 - 2) + 2(p_1 + (1 - p_1)) + 0} \\ &= \frac{1}{2}. \end{aligned}$$

□

References

- Baerveldt, C., Van Duijn, M. A., Vermeij, L., and Van Hemert, D. A. (2004). Ethnic boundaries and personal choice. assessing the influence of individual inclinations to choose intra-ethnic relationships on pupils networks. *Social Networks*, 26(1):55–74.
- Barabási, A.-L. and Albert, R. (1999). Emergence of scaling in random networks. *science*, 286(5439):509–512.
- Bollobás, B. (1998). *Modern graph theory*, volume 184. Springer.
- Bollobás, B. (2001). *Random Graphs: Second Edition*, volume 73. Cambridge University Press.
- Bramoullé, Y., Currarini, S., Jackson, M. O. ., Pin, P., and Rogers, B. W. (2012). Homophily and long-run integration in social networks. *Journal of Economic Theory*, 147(5):1754–1786.
- Burt, R. S. (1991). Measuring age as a structural concept. *Social Networks*, 13(1):1–34.
- Calvó-Armengol, A. (2004). Job contact networks. *Journal of Economic Theory*, 115(1):191–206.
- Calvó-Armengol, A. and Jackson, M. O. (2007). Networks in labor markets: Wage and employment dynamics and inequality. *Journal of Economic Theory*, 132(1):27–46.
- Campbell, K. E. (1990). Networks past: A 1939 bloomington neighborhood. *Social Forces*, 69(1):139–155.
- Chung, F. and Lu, L. (2002). The average distances in random graphs with given expected degrees. *Proceedings of the National Academy of Sciences*, 99(25):15879–15882.
- Currarini, S., Jackson, M. O., and Pin, P. (2009). An economic model of friendship: Homophily, minorities, and segregation. *Econometrica*, 77(4):1003–1045.
- Dawid, H. and Gemkow, S. (2013). How do social networks contribute to wage inequality? insights from an agent-based analysis. *Working Paper*.
- Erdős, P. and Rényi, A. (1959). On random graphs. *Publicationes Mathematicae Debrecen*, 6:290–297.

- Golub, B. and Jackson, M. O. (2012). How homophily affects the speed of learning and best-response dynamics. *The Quarterly Journal of Economics*, 127(3):1287–1338.
- Horváth, G. (2011). Occupational mismatch and social networks. *Available at SSRN 1794082*.
- Ibarra, H. (1995). Race, opportunity, and diversity of social circles in managerial networks. *Academy of Management Journal*, 38(3):673–703.
- Ioannides, Y. M. and Loury, L. D. (2004). Job information networks, neighborhood effects, and inequality. *Journal of Economic Literature*, 42(4):1056–1093.
- Jackson, M. O. (2005). The economics of social networks. *Social Science Working Paper*, 1237.
- Jackson, M. O. (2008). Average distance, diameter, and clustering in social networks with homophily. In *Internet and Network Economics*, pages 4–11. Springer.
- Jackson, M. O. (2010a). An overview of social networks and economic applications. *The handbook of social economics*, 1:511–85.
- Jackson, M. O. (2010b). *Social and economic networks*. Princeton University Press.
- Kalmijn, M. (2006). Educational inequality and family relationships: Influences on contact and proximity. *European sociological review*, 22(1):1–16.
- Laumann, E. O. (1966). *Prestige and association in an urban community: An analysis of an urban stratification system*. Bobbs-Merrill Indianapolis and New York.
- Laumann, E. O. (1973). *Bonds of pluralism: The form and substance of urban social networks*. J. Wiley.
- Lazarsfeld, P. F., Merton, R. K., et al. (1954). Friendship as a social process: A substantive and methodological analysis. *Freedom and control in modern society*, 18(1):18–66.
- Marsden, P. V. (1987). Core discussion networks of americans. *American sociological review*, pages 122–131.
- Marsden, P. V. (1988). Homogeneity in confiding relations. *Social networks*, 10(1):57–76.
- Mayer, A. and Puller, S. L. (2008). The old boy (and girl) network: Social network formation on university campuses. *Journal of public economics*, 92(1):329–347.

- McPherson, M., Smith-Lovin, L., and Cook, J. M. (2001). Birds of a feather: Homophily in social networks. *Annual Review of Sociology*, pages 415–444.
- Milgram, S. (1967). The small world problem. *Psychology today*, 2(1):60–67.
- Montgomery, J. D. (1991). Social networks and labor-market outcomes: Toward an economic analysis. *The American Economic Review*, 81(5):1408–1418.
- Newman, M. E. (2003). The structure and function of complex networks. *SIAM review*, 45(2):167–256.
- Newman, M. E. (2006). Random graphs as models of networks. *Handbook of graphs and networks: From the genome to the internet*.
- Rees, A. (1966). Information networks in labor markets. *The American Economic Review*, 56(1/2):559–566.
- Van der Leij, M. and Buhai, S. (2008). A social network analysis of occupational segregation.
- Verbrugge, L. M. (1977). The structure of adult friendship choices. *Social Forces*, 56(2):576–597.
- Wasserman, S. and Pattison, P. (1996). Logit models and logistic regressions for social networks: I. an introduction to markov graphs andp. *Psychometrika*, 61(3):401–425.
- Watts, D. J. and Strogatz, S. H. (1998). Collective dynamics of 'small-world' networks. *nature*, 393(6684):440–442.
- Wellman, B. (1996). Are personal communities local? a dumptarian reconsideration. *Social networks*, 18(4):347–354.
- Zaharieva, A. (2013). Social welfare and wage inequality in search equilibrium with personal contacts. *Labour Economics*, 23:107–121.