Department of Economics
School of Social Sciences

Socioeconomic Networks with Long-Range Interactions

Rui Carvalho [1]
University College London

Giulia Iori [2]
City University

Department of Economics
Discussion Paper Series
No. 07/12

---

[1] Centre for Advanced Spatial Analysis, 1-19 Torrington Place, University College London, WC1E 6BT UK. Email: rui.carvalho@ucl.ac.uk

[2] Department of Economics, City University, Northampton Square, London, EC1V 0HB, UK. Email: g.iori@city.ac.uk

Socioeconomic Networks with Long-Range Interactions

Rui Carvalho*
Centre for Advanced Spatial Analysis, 1-19 Torrington Place, University College London, WC1E 6BT United Kingdom

Giulia Iori†
Department of Economics, School of Social Science City University, Northampton Square, London EC1V 0HB United Kingdom

In well networked communities, information is often shared informally among an individual's direct and indirect acquaintances. Here we study a modified version of a model previously proposed by Jackson and Wolinsky to account for communicating information and allocating goods in socioeconomic networks. The model defines a utility function of node i which is a weighted sum of contributions from all nodes accessible from i. First, we show that scale-free networks are more efficient than Poisson networks for the range of average degree typically found in real world networks. We then study an evolving network mechanism where new nodes attach to existing ones preferentially by utility. We find the presence of three regimes: scale-free (rich-get-richer), fit-get-rich, and Poisson degree distribution. The fit-get-rich regime is characterized by a decrease in average path length.

I. INTRODUCTION

The study of socioeconomic networks is a burgeoning field in the physics and economics literature, with major progress having been attained over the last decade [1, 2, 3, 4, 5, 6]. Individuals and firms interact through networks to share information and resources, exchange goods and credit, make new friendships or partnerships etc. The structure of the network through which interactions take place may thus have an important effect on the success of the individual or the productivity of the firm [1]. Furthermore, the network of interactions among socioeconomic agents plays an important role for the stability and efficiency of socioeconomic systems [7]. Theories about how interaction networks form are thus essential for a deeper understanding of the development and organization of society as a whole.

The economics literature focuses mainly on equilibrium networks and the network formation mechanisms are based on utility maximization and costs minimization. The aim of most economic papers is to identify, among the set of equilibrium networks, the geometry that optimizes efficiency [31] in the sense of social benefit. Likewise, economists are interested in the stability [32] of equilibrium networks under link deletion, addition or rewiring [1, 2]. A shortcoming of these models is that the equilibrium networks are often too simple in their geometry (stars, complete networks, interlinked stars, etc.), typically as a consequence of the symmetries that need to be assumed in the payoff functions in order to make the models analytically tractable [8].

The physics literature, instead, has mainly focused on the characterization of the structure of real networks and proposed dynamic models, mostly based on probabilistic rules, capable of reproducing the observed geometrical structures (Poisson, stretched exponential and scale-free networks) [9, 10, 11].

In this paper we try to combine the physics and economic approaches, by introducing a stochastic network formation mechanism inspired by economists' utility maximization models, which naturally extends the well known physicists' preferential attachment rule [12].

One of the most interesting models of socioeconomic network formation was introduced by Jackson and Wolinsky in 1996 [1]. In their model, the formation and evolution of links is driven by a utility maximization mechanism. The model is based on the assumption that agents may derive benefit not only from the nodes to which they are directly connected (their nearest neighbours), but also from the ones they are connected to indirectly (possibly via long paths). The utility of node i is defined as:

$$u_i = w_{ii} + \sum_{j \neq i} w_{ij} \delta^{d_{ij}} - \sum_{j \in \mathcal{V}(i)} c_{ij} \qquad (1)$$

where the contribution to the utility of i from j may depend on the weight w_{ij} of the edge between i and j (or, alternatively, on the fitness of node j); $0 \leq \delta < 1$ captures the idea that the utility gain from indirect connections decreases with distance; d_{ij} is the number of links in the shortest path between i and j ($d_{ij} = \infty$ if there is no path between i and j); $\mathcal{V}(i)$ is the set of nearest neighbours of i; and c_{ij} are the (node specific) costs to establish a directed connection between i and j [33]. Costs can also be differentiated in costs of initially creating or maintaining an edge [2].

The papers by Jackson and Wolisnky [1], as well as the one by Bala and Goyal [2], are mainly concerned with

*Electronic address: rui.carvalho@ucl.ac.uk
†Electronic address: g.iori@city.ac.uk

stability and efficiency of the network resulting from different dynamic updating rules. In particular, Jackson and Wolisnky study pairwise stability when agents can only update a link at a time (either delete it or create it), while Bala and Goyal allow agents to rearrange all their connections at once. The updating is deterministic in both models, and a new configuration is accepted only if it increases the utility of the agent. These two papers show that the star network is both efficient and stable for a wide range of the parameters when $\delta = 1$. Nonetheless, a multiplicity of network architectures exist in [2] for $0 < \delta < 1$ which could be a strict Nash equilibria, and to which the system may converge depending of the initial conditions. Feri [13] has shown that for sufficiently large networks the star network is stochastically stable for almost all the range of parameters, even for $0 < \delta < 1$.

Here we focus on a simplified version of the Jackson and Wolisnky model, i.e. the case $w_{ij} = 1$, $w_{ii} = 0$ and $c_{ij} = 0$. In this case, the utility can be rewritten as

$$u_i = \sum_{l=1}^{l_{max}^{(i)}} \sum_{\{k|d_{ik}=l\}} \delta^l = \sum_{l=1}^{l_{max}^{(i)}} \delta^l z_l^{(i)} \qquad (2)$$

where the sum in l is over all shortest paths of length l from node i, the sum in k is over all nodes whose shortest path from i is $d_{ik} = l$, $l_{max}^{(i)}$ is the path length of the node the furthest away from node i, and $z_l^{(i)}$ is the number of lth-nearest neighbours of node i. The utility of a node is expressed in (2) as a weighted sum of the number of nodes accessible from i on outward "layers" of increasing distance from i. Thus, we start at node i and multiply δ by the number of nodes that are joined by an edge to i–this being the first layer. We then add δ^2 times the number of nodes that are joined by an edge to a node in the first layer–this is the second layer. We continue in this way until no new nodes are found. Hence, expression (2) incorporates implicitly the well known breath-first search algorithm [14].

We first study how the average utility in a network, when individual utility is defined by (2), depends on the underlying network topology. We derive analytical results in this respect by using the generating function approach [15]. We then focus on network growing mechanisms. We assume that new nodes arrive steadily and create links with existing nodes in a probabilistic way, proportionally to existing nodes' utility. In this way, we build on the preferential attachment growth rule of Barabási and Albert [9, 12] which can be recovered from equation (2) when $l_{max}^{(i)} = 1$.

If each new node attached deterministically to the existing node with maximal utility, the resulting network would be a star. The randomness generated by the probabilistic attaching rule, can be interpreted as costs, barriers, or bounded rationality all of which limit the ability to establish links in an optimal way, thus possibly generating more realistic geometries than the star network. Furthermore, preferential attachment is, arguably, the most extensively studied mechanism of network formation and one that has revealed insights into properties observed in real networks. Therefore, it is important to understand the robustness of the specific rule of linear preferential attachment by node degree, which is one of the aims of this paper.

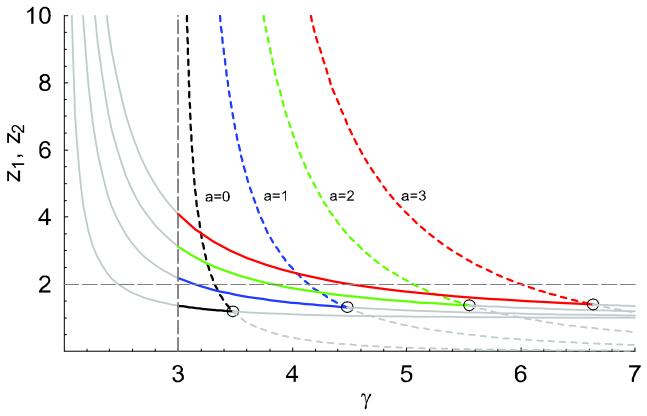

FIG. 1: Average number of first and second-neighbours ($z_1(\gamma, a)$ and $z_2(\gamma, a)$) in networks with degre distribution given by (10). From left to right, we plot $z_1(\gamma, a)$ (full curves) and $z_2(\gamma, a)$ (dashed curves) for $a = 0$ (black), 1 (blue), 2 (green) and 3 (red). The values of $z_1(\gamma \leq 3, a)$ for which $z_2(\gamma, a)$ is not defined are plot in grey, as well as regions of the curves for which $z_1(\gamma, a) > z_2(\gamma, a)$. The circles denote the intersection of the two curves, $z_1(\gamma, a)$ and $z_2(\gamma, a)$, for each value of a.

II. ANALYTICAL RESULTS FOR RANDOM NETWORKS

An interesting question to ask (for example, from the point of view of the social planner) is which network structure maximizes the total, or the average, utility (networks that satisfy this condition are said to be efficient in economics). We show that it is possible to derive analytical results for the average utility in star, Poisson and scale-free networks. By comparing average utility in different network topologies with the same size and the same average degree (which is equivalent to fixing the number of nodes and number of links), we show that scale-free networks are more efficient than Poisson random networks (even though less efficient than the star).

The average utility of a star network is given by:

$$\overline{u}_*(\delta) = \delta z_1 \left(1 + \delta \frac{N-2}{2}\right) \qquad (3)$$

where $z_1 = 2(N-1)/N$. For N large, $z_1 \simeq 2$ and $\overline{u}_*(\delta) \sim N\delta^2$.

To derive an expression for average utility in generic random networks we average both sides of (2):

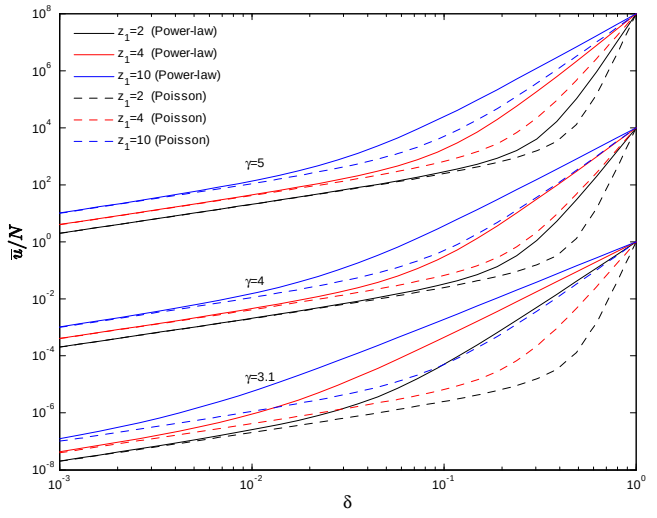

FIG. 2: Analytical results for scaled average utility in networks with power-law (full curves) and Poisson (dashed curves) degree distributions as a function of δ, z_1 and γ for $N = 10^5$. Curves have been shifted vertically for different values of γ for clarity. Values of z_1 increase from bottom to top.

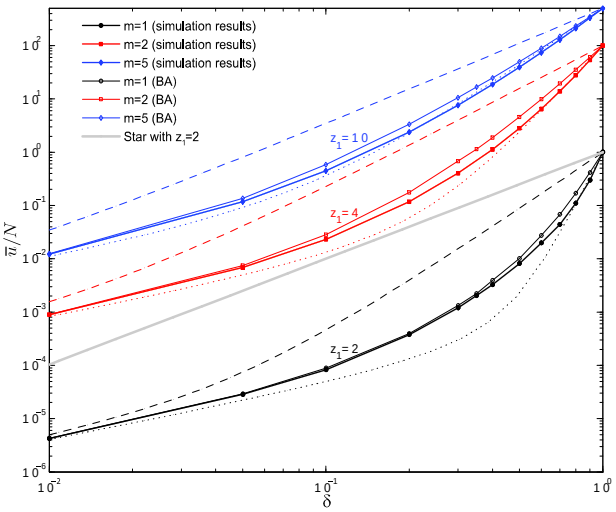

FIG. 3: Average utility for the simulation results (solid curve and symbols) and the BA model (solid curve, open symbols) for $m = 1$ ($z_1 = 2$), 2 ($z_1 = 4$) and 5 ($z_1 = 10$). We also plot the analytical curves for average utility in Poisson (dotted curve) and scale-free (dashed curve) networks for $z_1 = 2, 4$ and 10 and $\gamma = 3.1$ (for scale-free networks). Curves were computed for networks with $N = 5 \times 10^3$ vertices and have been shifted vertically for distinct values of z_1. The curve in gray is the average utility of a star with $N = 5 \times 10^3$ and $z_1 \simeq 2$. Simulation results were averaged over 30 runs.

$$\overline{u}\,(\delta) = \sum_{l=1}^{\bar{l}} \delta^l z_l \tag{4}$$

where z_l is the average number of lth neighbours of a node and $\bar{l}$ is the average path length.

When the number of nodes z_l at a distance l away from a given node is equal to N, then l is roughly equal to the average path length, $\bar{l}$ [17, 18]:

$$1 + \sum_{l=1}^{\bar{l}} z_l = N \tag{5}$$

If the sum in (2) was to be evaluated up to distance $l^{(i)}_{max} = 1$ for every node, expression (4) would simplify to $\overline{u}\,(\delta) = \delta z_1$, *i.e.* average utility would be independent of the specific network topology and all networks with the same number of nodes and links would be equally efficient. Thus we need to introduce long range interactions ($l^{(i)}_{max} > 1$) to be able to rank networks in terms of their efficiency.

Now that we have expressed average utility in terms of the breadth-first search algorithm, we can derive a closed form of expression (4) if we have access to analytical expressions for average path length and the average number of lth neighbours. This can be accomplished by generating functions, which are particularly useful when determining means, standard deviations and moments of distributions [15].

The average number of neighbours (average degree) and the average number of second neighbours of a node can be derived from the probability generating function of node degree, $G_0\,(x) = \sum_{k=0}^{\infty} p_k x^k$, as long as the degree distribution, p_k, is specified. The beauty of the generating function formalism is that one can derive z_l as a function of z_1 and z_2 only [17, 18, 19]:

$$z_l = \left[\frac{z_2}{z_1}\right]^{l-1} z_1 \tag{6}$$

Replacing equation (6) in equation (4) yields

$$\overline{u}\,(\delta) = \delta z_1 \sum_{l=1}^{\bar{l}} (\delta Z)^{l-1} = \frac{\delta z_1((\delta Z)^{\bar{l}} - 1)}{\delta Z - 1} \tag{7}$$

where $Z = z_2/z_1$ and $\bar{l}$ is the average path length. For $Z > 1$ and $N > z_1 + 1$, which are conditions satisfied by most networks, $\bar{l}$ can be calculated as a function of N, z_1 and z_2 from (5) and (6) as [17]:

$$\bar{l} = \frac{\ln[(N-1)\,(Z-1)\,/z_1 + 1]}{\ln\,(Z)} \tag{8}$$

In what follows, we investigate the behaviour of (7) for Poisson and scale-free random networks.

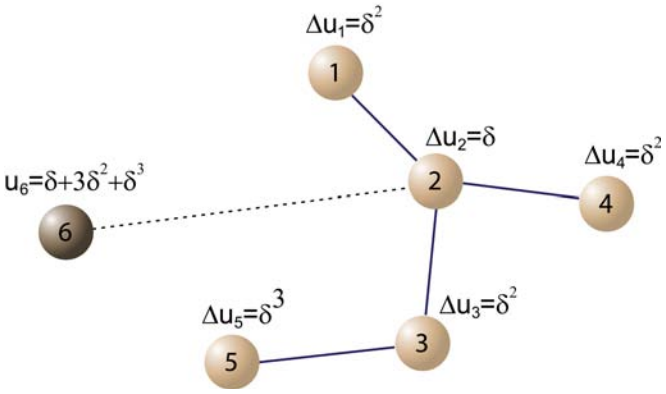

FIG. 4: Schematic layout of network growth when $m = m_0 = 1$. The addition of a new node, 6, implies an increase of the utility of nodes 1 to 5 which is simply δ^d, where d is the path length from node 6. The simplicity of this updating mechanism allowed simulations to be run with $N = 10^5$ when $m = 1$.

Poisson random networks are characterized by $z_1 = pN$ and $z_2 = z_1^2$ [17], thus (8) yields $\overline{l_P} = \ln(\frac{N(z_1-1)+1}{z_1})/\ln(z_1)$. In this case, (7) becomes:

$$\overline{u}_P(N,\delta,z_1) = \frac{z_1\delta\left((\delta z_1)^{\ln(N+\frac{1-N}{z_1})/\ln(z_1)} - 1\right)}{\delta z_1 - 1} \tag{9}$$

for $N > z_1 + 1$, $0 < \delta \leq 1$ and $z_1 > 1$.

Next, we consider scale-free networks with degree distribution of the form:

$$p_k(\gamma,a) = \frac{1}{\zeta(\gamma,1+a)}(a+k)^{-\gamma} \quad ,a \geq 0 \tag{10}$$

where the normalizing factor $\zeta(\gamma,a+1) = \sum_{k=1}^{\infty}(a+k)^{-\gamma}$ is the Hurwitz zeta function ($\gamma > 1$). The generating function for the probability distribution is given by

$$G_0(x,\gamma,a) = \sum_{k=1}^{\infty} p_k x^k = \frac{x\Phi(x,\gamma,a+1)}{\zeta(\gamma,a+1)} \tag{11}$$

where $\Phi(x,\gamma,a) = \sum_{k=0}^{\infty}\frac{x^k}{(a+k)^\gamma}$ is the Lerch transcendent. For our purposes, only the first two derivatives of $\Phi(x,\gamma,a+1)$ with respect to x are relevant as the average number of first and second-neighbours are given, respectively, by $z_1(\gamma,a) = \left.\frac{\partial G_0(x)}{\partial x}\right|_{x=1}$ and $z_2(\gamma,a) = \left.\frac{\partial^2 G_0(x)}{\partial x^2}\right|_{x=1}$. Hence

$$z_1(\gamma,a) = \frac{\Phi(1,\gamma-1,a+1) - a\Phi(1,\gamma,a+1)}{\zeta(\gamma,a+1)} \quad ,\gamma > 2 \wedge a \geq 0 \tag{12}$$

$$z_2(\gamma,a) = \frac{\zeta(\gamma-1,a+1)}{\zeta(\gamma,a+1)}z_1(\gamma-1,a) - (a+1)z_1(\gamma,a) \quad ,\gamma > 3 \wedge a \geq 0 \tag{13}$$

Thus

$$Z(\gamma,a) = \frac{\zeta(\gamma-1,a+1)}{\zeta(\gamma,a+1)}\frac{z_1(\gamma-1,a)}{z_1(\gamma,a)} - a - 1 \quad ,\gamma > 3 \wedge a \geq 0 \tag{14}$$

Substituting (12) and (14) into (8), we find

$$\bar{l}_{SF}(N,\gamma,a) = \frac{\ln\left(-\frac{(a+2)(N-1)}{z_1(\gamma,a)} + \frac{z_1(\gamma-1,a)\zeta(\gamma-1,a+1)(N-1)}{z_1(\gamma,a)^2\zeta(\gamma,a+1)} + 1\right)}{\ln\left(-a + \frac{z_1(\gamma-1,a)\zeta(\gamma-1,a+1)}{z_1(\gamma,a)\zeta(\gamma,a+1)} - 1\right)} \quad ,N > z_1(\gamma,a)+1 \wedge \gamma > 3 \wedge a \geq 0 \wedge Z(\gamma,a) > 1 \tag{15}$$

and thus average utility is given by

$$\overline{u}_{SF}(N,\delta,\gamma,a) = \frac{\delta z_1(\gamma,a)\left((\delta Z(\gamma,a))^{\bar{l}_{SF}(N,\gamma,a)} - 1\right)}{\delta Z(\gamma,a) - 1} \quad ,N > z_1(\gamma,a)+1 \wedge 0 < \delta \leq 1 \wedge \gamma > 3 \wedge a \geq 0 \wedge Z(\gamma,a) > 1 \tag{16}$$

where $z_1(\gamma,a)$, $Z(\gamma,a)$ and $\overline{l_{SF}}(N,\gamma,a)$ are given by (12), (14) and (15), respectively.

When $a = 0$, the distribution of degree, (10), becomes a pure power-law $p_k(\gamma) = \frac{1}{\zeta(\gamma)}k^{-\gamma}$. In this case, we have $\zeta(\gamma,a+1)|_{a=0} = \zeta(\gamma)$ and $\Phi(x,\gamma,a+1)|_{a=0} = \frac{Li_\gamma(x)}{x}$,

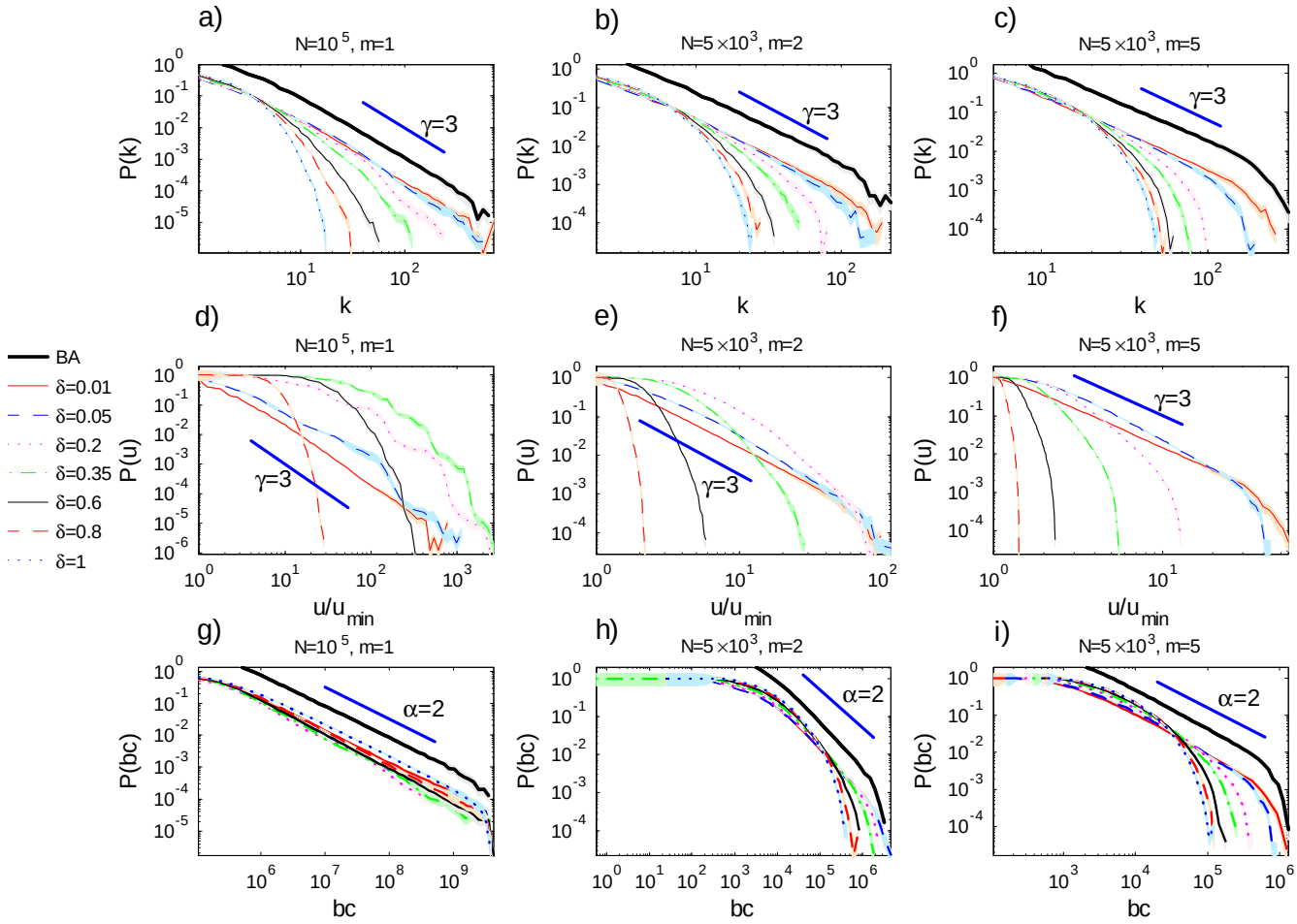

FIG. 5: Cumulative distribution function of degree (panels a), b) and c)), utility (panels d), e) and f)) and betweenness centrality (panels g), h) and i)) for several values of $\delta \in]0,1]$, and $m = 1, 2$ and 5. We also plot the corresponding distribution of degree and betweenness for the BA model (curves were shifted vertically). Simulations were averaged over 30 runs in networks with $N = 10^5$ ($m = 1$) or $N = 5 \times 10^3$ ($m = 2$ and 5). Coloured bands around the curves are 95% confidence intervals.

therefore (11) becomes

$$G_0(x,\gamma) = \frac{\mathrm{Li}_\gamma(x)}{\zeta(\gamma)} \tag{17}$$

This generating function is also obtained for the power-law distribution with exponential cut-off, proposed in [20, 21], $p_k(\gamma,\kappa) = Ck^{-\gamma}e^{-k/\kappa}$, in the limit $\kappa \to \infty$.

Expression (17) implies

$$z_1(\gamma)|_{a=0} = \frac{\zeta(\gamma-1)}{\zeta(\gamma)} \quad ,\gamma > 2 \tag{18}$$

$$z_2(\gamma)|_{a=0} = \frac{\zeta(\gamma-2)-\zeta(\gamma-1)}{\zeta(\gamma)} \quad ,\gamma > 3 \tag{19}$$

Therefore, in pure power-law networks, when $N \to \infty$, the average number of second-neighbours, $z_2(\gamma)$, is finite only for $\gamma > 3$. However, the Riemann zeta function, $\zeta(\gamma)$, is a decreasing function of γ (for $\gamma > 3$) and $z_1(\gamma = 3) = \pi^2/6\zeta(3) \simeq 1.36843$. In other words, the existence of $z_2(\gamma)$ implies $z_1(\gamma) < z_1(\gamma = 3) \simeq 1.36843$, which is a non-realistically low value for average degree in real networks. This explains why we have chosen the modified scale-free distribution (10).

The generating function (11) encapsulates *all* the moments of the degree distribution [17]. Hence, the expressions for $z_1(\gamma,a)$ and $z_2(\gamma,a)$, (12) and (13), are only exact in the limit $N \to \infty$. Further, $\overline{l}_{SF}(N,\gamma,a)$ and $\overline{u}_{SF}(N,\delta,\gamma,a)$, both of which depend on $z_2(\gamma,a)$, are only defined where $z_2(\gamma,a)$ is finite, *i.e.* for $\gamma > 3$. Therefore, it is essential to understand the behaviour of $z_1(\gamma,a)$ and $z_2(\gamma,a)$ in scale-free networks. Figure 1 shows z_1 (full curves) and z_2 (dashed curves) within the range $\gamma > 3 \wedge Z > 1$ (where $\overline{l}_{SF}(N,\gamma,a)$ is defined) for, from left to right, $a = 0, 1, 2$ and 3.

Having deduced closed-form expressions for average utility in Poisson and scale-free networks, we can now compare both networks under the condition that z_1 is the

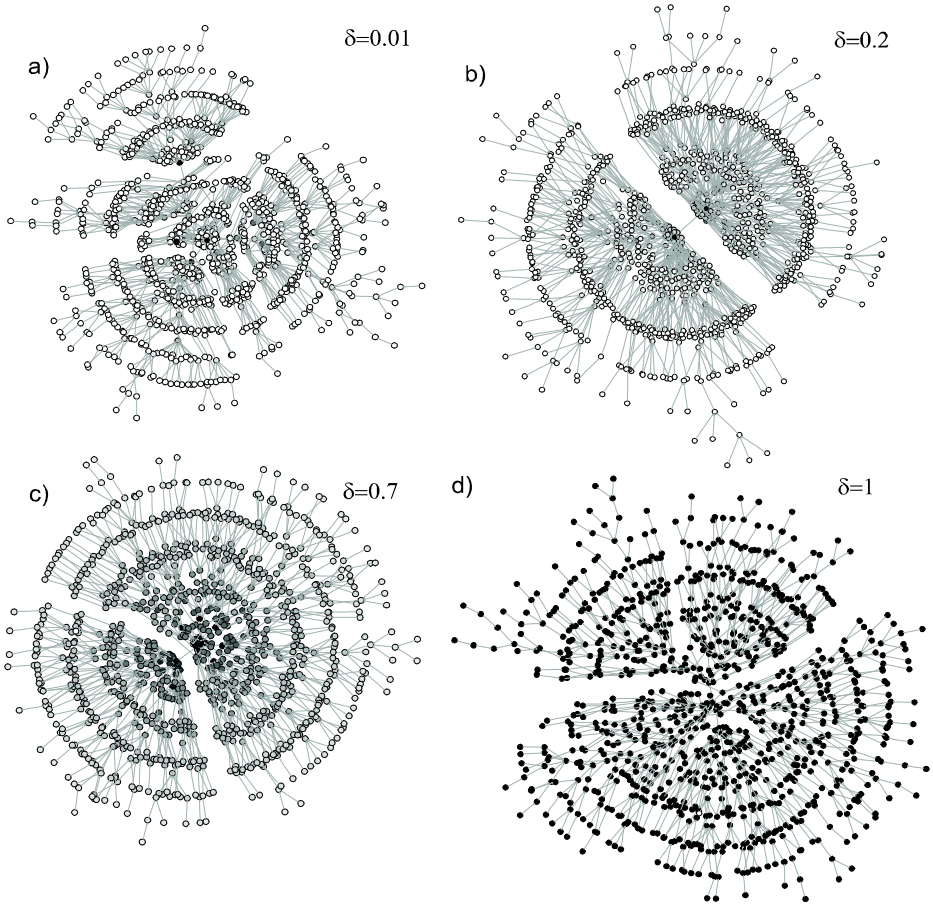

FIG. 6: Kamada-Kawai spring layout [16] for $m = 1$ and $N = 10^3$. Panel a) is a sample layout for $\delta = 0.01$, b) for $\delta = 0.2$, c) for $\delta = 0.7$ and d) for $\delta = 1$. On each panel, nodes are coloured by their utility on a gray scale from minimal (white) to maximal (black) utility.

same. Figure (2) is a plot of scaled average utility versus δ when $z_1 = \{2, 4, 10\}$ and $N = 10^5$ for Poisson and scale-free networks. The average utility of Poisson networks is completely specified by N, δ and z_1, but scale-free networks defined by (10) have one extra degree of freedom in $z_1(\gamma, a)$. In this case, we compute z_1 numerically by solving (12) for $z_1(\gamma, a) = \{2, 4, 10\}$ when $\gamma = \{3.1, 4, 5\}$. Comparisons among the star, Poisson and scale-free networks are only valid when $z_1 \simeq 2$ as this is the average degree of the star network in the limit $N \to \infty$. We have plot the scaled average utility for these networks in Figure 3 when $N = 5 \times 10^3$ and $z_1 = 2$, where the curves for the star, Poisson and scale-free networks appear in gray, black dotted and black dashed, respectively. As is clear in Figure 3, the star network has a much higher average utility than the Poisson or scale-free networks for all values of δ, confirming the well known result that the star network is efficient [1, 2].

Finally, note that both $\overline{u}_P$ and $\overline{u}_{SF}$ grow slower than N. As a consequence, when $N \to \infty$,

$$\widetilde{\overline{u}_{SF}} = \frac{\overline{u}_{SF}}{N} = \widetilde{\overline{u}_P} = \frac{\overline{u}_P}{N} = \begin{cases} 0 & \Leftarrow \ 0 < \delta < 1 \\ 1 & \Leftarrow \quad \delta = 1 \end{cases}, \quad (20)$$

thus comparisons between the curves in Figure (2) are only valid for finite N.

III. EVOLVING NETWORKS

In the classic Barabási and Albert model [12], a network is grown by adding, at every time step, a new node that attaches to m existing nodes with a probability proportional to their degree, $\Pi(k_i) = k_i / \sum_{j=1}^{N} k_j$. At time t, the resulting network has size $N_t = m_0 + t$, where $m_0 \geq m$ is the size of the (fully connected) network at time $t = 0$. Preferential attachment generates a scale-free probability density of incoming links that leads to the stationary result $p(k) = 2m^2/k^\gamma$, with $\gamma = 3$ independently of m. The model is characterized by a clustering coefficient larger than the one found for the Erdös Rényi networks (for $m > 1$) and no clear assortative/disassortative behaviour [9].

The linear preferential attachment hypothesis is very sensitive, as the scale-free nature of the network is destroyed by a non-linear attachment rule $\Pi(k_i) \sim k_i^\alpha$ [9]. Several models have been proposed lately to investi-

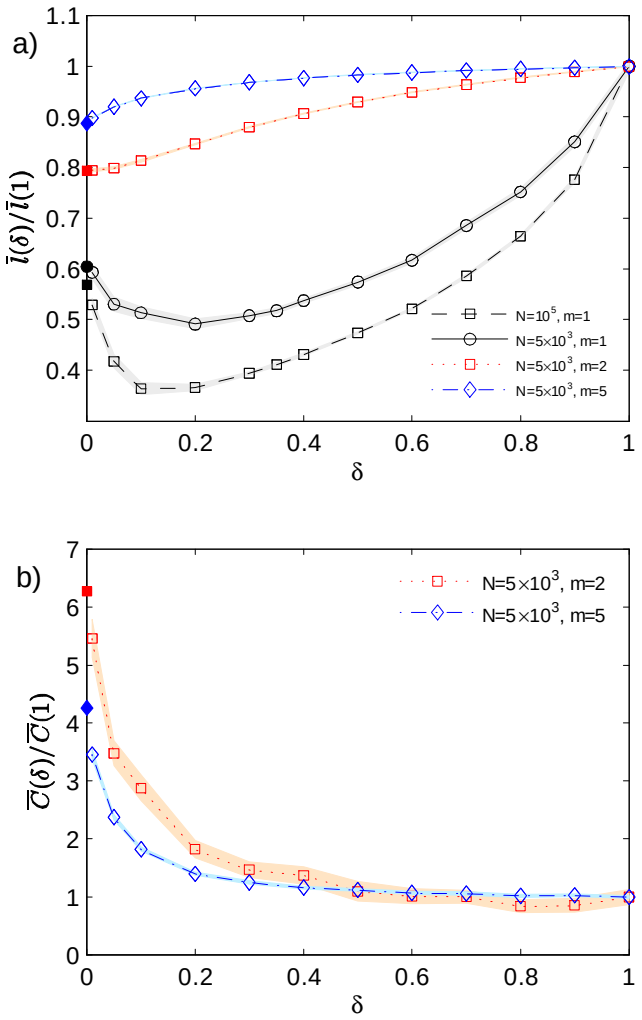

FIG. 7: Average path length, a), and clustering coefficient, b), for the simulation results when $m = 1, 2$ and 5. Curves were scaled by, respectively, average path length and clustering coefficient for $\delta = 1$ (Poisson network). Coloured bands around the curves are 95% confidence intervals.

gate extensions of the preferential attachment mechanism through edge removal and rewiring, inheritance, redirection or copying; node competition, aging and capacity constraints; and accelerated growth of networks to name just a few (see [9, 10, 22] for reviews). Of particular relevance to our approach are fitness models [23, 24, 25, 26], where the probability of attaching to a node is proportional to node fitness

$$\Pi(k_i) \sim \frac{f_i k_i}{\sum_{j=1}^{N} f_j k_j}. \tag{21}$$

Here we extend the preferential attachment rule by introducing a growing mechanism inspired on the work of Jackson and Wolisnky [1]. Our contribution is to propose preferential attachment by node utility. Thus, the probability that a new node j will be connected to an existing node i depends on the utility of i, such that

$$\Pi_i = \frac{u_i}{\sum_{k=1}^{N} u_k} \tag{22}$$

where the utility of node i, u_i, is given by (2). All nodes have the same utility for $\delta = 0$ and $\delta = 1$:

$$\begin{cases} u_i = 0 & \forall i \text{ when } \delta = 0 \\ u_i = N & \forall i \text{ when } \delta = 1 \end{cases} \tag{23}$$

so attachment happens randomly in these cases and we recover a Poisson distribution of node degree. The preferential attachment rule (22) is invariant up to multiplicative factors in (2), so for $\delta \neq 0$ the qualitative behaviour of the model remains unchanged if we define utility as

$$u_i^{'} = \frac{u_i}{\delta} = k_i + \sum_{l=2}^{l_{max}^{(i)}} \sum_{k \in \mathcal{V}_i^k} \delta^l \tag{24}$$

where k_i is the degree of node i. Thus, as $\delta \to 0$ our model converges to the Barabási-Albert model and the network becomes scale-free.

Our model has resemblances with the fitness models discussed above. However, there is a fundamental discrepancy: we regard utility as a time-dependent measure of node fitness, whereas existing models assume that node fitness does not change with time.

At each time step, a new node j joins the network and the utility of existing nodes changes. When $m = 1$, the utility increment to an existing node i at distance l from j is given by $\Delta u_i = \delta^j$ and therefore, at each time step, the computation of utility for the network can be completed in $O(N)$ time. Figure 4 is a diagram of a possible network configuration with $m = m_0 = 1$ after $t = 5$ time steps, showing the change in utility of existing nodes $\Delta u_i = \delta^j$. When $m > 1$, the increment in the utility of node i depends on the existing network geometry and $\Delta u_i > \delta^j$. Therefore, when $m > 1$, we need to re-compute the utility of *all* existing nodes at every time step, and therefore the computation runs in $O(N^2)$ time as it involves running a *breadth-first-search* algorithm from every node. This is the reason why we have ran simulations for $N = 10^5$ when $m = 1$, but only up to $N = 5 \times 10^3$ when $m > 1$.

Existing nodes i at a higher distance than a certain $l_{\max}$ from new node j receive a contribution $\Delta u_i = \delta^{d(j,i)} < 10^{-precison}$ which is less than the number of significant digits that the computer can store (typically $precison = 32$ in double precision), and do not need to have their utility updated in the simulations. This maximal distance $l_{\max}$ is defined as

$$10^{-precision} > \delta^{l_{\max}} \Leftrightarrow l_{\max} > -\frac{precision}{\log_{10} \delta} \tag{25}$$

Our implementation of the algorithm updates the utility of all nodes accessible from the new node j up to distance

$l_{\max} = -32/\log_{10}\delta$. The code was implemented in C++ and ran in a Condor framework (high throughput computing) [27] for several values of δ. Ensemble averages were taken over 30 runs.

Figure 3 shows the simulation results for scaled average utility in our model and the Barabási-Albert model (solid and open symbols on solid curves, respectively) overlaid with the analytical curves for average utility of Poisson and scale-free networks (dotted and dashed lines, respectively) as δ is varied. We observe that average utility in our model is higher than in Poisson networks, but smaller than in the Barabási-Albert model and scale-free networks.

Expressions (23) and (24) predict the existence of two distinct regimes: a scale-free regime as $\delta \to 0$ ($\delta \neq 0$) and a Poisson regime for $\delta = \{0, 1\}$. However, preferential attachment by utility introduces a third, *fit-get-rich* [28], regime between these two when $m = 1$. Next we discuss each of these regimes.

We start with the scale-free regime. Expression (24) implies that, for small enough δ, preferential attachment by degree is indistinguishable from preferential attachment by utility and the probability density of both quantities should decay like $p(x) \sim x^{-\gamma}$ with $\gamma = 3$. Figure 5, panels a) to f), which are plots of the distribution of degree and utility for $m = 1, 2$ and 5 show, for small δ, a scale-free decay of the distribution of both degree and utility with the same exponent as the Barabási-Albert model. However, the distribution would only be scale-free *exactly* at $\delta = 0$, whereas in our model there is a discontinuity at this parameter value and the distribution becomes Poisson as all nodes have the same utility. Therefore, the scale-free regime is only an approximation. In the Barabási-Albert model, the degree distribution decays as a power-law with exponent $\gamma = 3$ independently of m. However, this behaviour, which seems to be a consequence of the peculiar linear preferential attachment mechanism, is destroyed with perturbations. In fact, in our model, a higher m has the effect of homogenizing the utility of the network nodes, with the consequence that, for the same values of δ, the degree distribution is closer to a Poisson with increasing m.

The "fit-get-rich" regime, which appears for $m = 1$, is a consequence of the preferential attachment rule by node utility and is characterized by the emergence of a few utility hubs which do not coincide necessarily with the degree hubs. The effect is that nodes with high utility will receive more links. The relevant property of preferential attachment by utility is that, for δ small, the nearest neighbours of nodes with high utility will also have high utility, and thus attract more links than nodes faraway from the utility hubs. The preferential attachment mechanism by node utility rewards the utility hubs and their neighbours, independently of their degree and newcomers which attach to a utility hub will inherit a proportion of the hubs' utility. These newcomers will then have a high probability of receiving a new link in the next time step, although their degree will be $m = 1$. The consequence is a stratification of utility values and a lowering of the network radius. In this regime ($m = 1$), the distribution of utility shows a step-like behaviour which is indicative of the presence of characteristic values in utility, i.e. most nodes can be classified into a small set of distinct utilities. This phenomenon can be observed in Figures 5d) and 6 for networks of $N = 10^5$ and 10^3 nodes, respectively. In parallel, the *fit-get-rich* regime is characterised by the decrease of average path length, as can be observed in Figure 7a) where we plot scaled $\bar{l}$. As δ increases from 0 to 1, the newcomers will be assigned a higher percentage of the utility of the node they are attaching to, but node utility will also become more homogeneously distributed throughout the network nodes and eventually the *fit-get-rich* regime disappears. This regime is clearly identifiable for $m = 1$, but seems to be destroyed for higher values of m.

Finally, the Poisson regime happens at $\delta = 1$, when all nodes have the same utility, *i.e.* $u = N$. In this case, preferential attachment by node utility is equivalent to random attachment and the degree distribution is Poisson as can be observed in Figure 5a) to c).

Betweenness centrality is plot in Figure 5g), h) and i) as m is varied. Recent results show also that the distribution of loads (or betweenness) scales with a power law [29, 30] $p(g) \sim g^{-\alpha}$ where $\alpha = 2$ for a tree (and hence for $m = 1$). This justifies the collapse of the curves of the distribution of betweenness in Figure 5g). Also the load is fixed given the degree $g \sim k^{\eta}$ and $\eta = \frac{\gamma-1}{\alpha-1}$. As can be observed in Figure 5h) and i), the distribution of betweenness deviates from the power-law behaviour as m is increased. We plot the clustering coefficient in Figure 7b) and observe that it decreases with the increase of δ from the value for the Barabási-Albert model, until, for $\delta = 1$, the value for Poisson random networks.

IV. DISCUSSION

We have studied models of socioeconomic networks with long-range interactions inspired by the work of Jackson and Wolinsky [1] where the connection costs are reduced to zero. The last assumption is justified by the fact that if costs were node independent they would not play any role in the growing model. Similarly, costs do not play a significative role if we restrict the comparison of average utility in section II to networks with the same size and the same degree.

The derivation of analytical expressions for average utility in Poisson and scale-free networks reveals that the latter have higher $\overline{u}$ for the range of parameters that is of significance in real-world networks ($z_1 \geq 2$). This suggests a novel mechanism which may explain the ubiquitous presence of scale-free networks, in particular in situations where collaboration, interaction and information sharing among the nodes are of paramount relevance.

Comparisons with the star network can only be made

for $z_1 = 2$ and, in this case, the star has the highest utility of the networks studied here. Nevertheless, the star network is of little practical relevance as the average degree of large real networks is not necessarily constrained to $z_1 = 2$.

The growth mechanism we have proposed is a natural extension of the Barabási-Albert preferential attachment by degree to preferential attachment by node utility. Our analysis shows that for small values of δ, the utility decay parameter, the network retains a scale-free structure that is nonetheless destroyed when δ increases. We have identified a *fit-get-richer* regime in δ where the network is characterized by a lower average path length than the scale-free network and a step-like distribution of utility. As δ approaches one, the range of different node utilities is reduced and, eventually, at $\delta = 1$, all nodes have the same utility equal to the network size, $u = N$.

[1] M. O. Jackson and A. Wolinsky, Journal of Economic Theory **71**, 44 (1996).

[2] V. Bala and S. Goyal, Econometrica **68**, 1181 (2000).

[3] M. Marsili, F. Vega-Redondo, and F. Slanina, Proceedings of the National Academy of Sciences U.S.A. **101**, 1439 (2004).

[4] D. O. Cajueiro, Physical Review E **72**, 047104 (2005).

[5] G. C. M. A. Ehrhardt and M. Marsili, Physical Review E **74**, 036106 (2006).

[6] A. D. Martino and M. Marsili, Journal of Physics A: Mathematical and General **39**, R465 (2006).

[7] G. Iori, S. Jafarey, and F. G. Padilla, Journal of Economic Behaviour and Organization **61**, 525 (2006).

[8] M. O. Jackson, The economics of social networks, in *Advances in Economics and Econometrics, Theory and Applications: Ninth World Congress of the Econometric Society*, edited by W. N. Richard Blundell and T. Persson, volume I, Cambridge University Press, 2006.

[9] R. Albert and A.-L. Barabási, Reviews of Modern Physics **74**, 47 (2002).

[10] M. Newman, SIAM Review **45**, 167 (2003).

[11] S. N. Dorogovtsev and J. F. F. Mendes, *Evolution of Networks: From Biological Nets to the Internet and WWW*, Oxford University Press, 2003.

[12] A.-L. Barabási and R. Albert, Science **286**, 509 (1999).

[13] F. Feri, Journal of Economic Theory (2007, forthcoming).

[14] J. Kleinberg and E. Tardos, *Algorithm Design*, Addison Wesley, 2006.

[15] H. S. Wilf, *Generatingfunctionology*, Academic Press, 1994.

[16] T. Kamada and S. Kawai, Information Processing Letters **31**, 7 (1989).

[17] M. Newman, S. H. Strogatz, and D. J. Watts, Physical Review E **64**, 026118 (2001).

[18] M. Newman, D. J. Watts, and S. H. Strogatz, Proceedings of the National Academy of Sciences U.S.A. **99**, 2566 (2002).

[19] D. S. Callaway, M. Newman, S. H. Strogatz, and D. J. Watts, Physical Review Letters **25**, 5468 (2000).

[20] M. Aldana, Physica D **185**, 45 (2003).

[21] M. E. J. Newman, Physical Review Letters **95**, 108701 (2005).

[22] S. N. Dorogovtsev and J. F. F. Mendes, Advances in Physics **51**, 1079 (2002).

[23] C. Bedogne and G. J. Rodgers, Physical Review E **74**, 046115 (2006).

[24] G. Caldarelli, A. Capocci, P. D. L. Rios, and M. A. Munõz, Physical Review Letters **89**, 258702 (2002).

[25] G. D. Masi, G. Iori, and G. Caldarelli, **74**, 066112 (2006).

[26] V. D. P. Servedio, G. Caldarelli, and P. Buttà, Physical Review E **70**, 056126 (2004).

[27] D. Thain, T. Tannenbaum, and M. Livny, Concurrency and Computation: Practice and Experience **17**, 323 (2005).

[28] G. Bianconi and A.-L. Barabási, Physical Review Letters **86**, 5632 (2001).

[29] K.-I. Goh, B. Kahng, and D. Kim, Physical Review Letters **87**, 278701 (2001).

[30] M. Barthélemy, Eur. Phys. J. B **38**, 163 (2004).

[31] A network g is efficient with respect to an aggregate utility measure u if $u(g) \geq u\left(g^{'}\right)\ \forall g^{'} \in \mathcal{G}$ [1].

[32] A network is pairwise stable when no node would benefit from severing an existing link, and no two nodes would benefit from forming a new link [1].

[33] In [1] costs are assumed to be equally, or cooperatively, shared by i and j, but extensions to the non cooperative case have also been explored.