Working papers series

WP ECON 21.06

Are close-knit networks good for employment?

María Paz Espinosa
Universidad del País Vasco

Jaromír Kovářík
Universidad del País Vasco

Sofía Ruíz-Palazuelos
Universidad Pablo de Olavide

Keywords: (stochastic) affiliation, clustering, inequality, information transmission, labor markets, network cycles, networks, unemployment, wages.

JEL Classification: A14, D85, J60, J30.

Department of Economics

Are close-knit networks good for employment?*

María Paz Espinosa† Jaromír Kovářík†‡§ Sofía Ruiz-Palazuelos†

February 26, 2021

Abstract

How do short network cycles–the building component of close-knit network neighborhoods–affect diffusion? We study this issue in labor markets by explicitly modeling the flow of information about vacancies in social networks. We show that short network cycles induce stochastic affiliation in the transmission of job information, generating diffusion inefficiencies with important short- and long-run micro- and macroeconomic consequences. Short network cycles affect employment and inequality patterns within and across networked societies. In particular, they organize employment probabilities in the sense of the first-order stochastic dominance. People in close-knit neighborhoods and dense networks exhibit worse labor-market outcomes. Since dense, overlapping neighborhoods is one aspect of strong relationships, this uncovers an alternative mechanism behind the *strength of weak ties* (Granovetter, 1973). Moreover, short network cycles reinforce spatial and temporal correlations in employment status, shaping labor-market transition rates and aggregate employment fluctuations.

Keywords: (stochastic) affiliation, clustering, inequality, information transmission, labor markets, network cycles, networks, unemployment, wages.
JEL: A14, D85, J60, J30.

1 Introduction

The key role of social networks in shaping socio-economic phenomena is well documented across a variety of contexts.[1] The importance of social contacts and the architecture of relationships has been particularly recognized in labor economics as a key source of employment information (Ionamides and Datcher, 2005; Beaman, 2016). Granovetter (2018) concludes that people rely primarily on contacts when finding a job, independently of the occupation, skill, location, or socio-economic background.[2] Cappellari and Tatsiramos (2015) estimate that an additional employed friend increases the probability of finding a job by 3.7%.[3] Topa (2001) and Conley and Topa (2002) document geographic, ethnic, and race correlations in employment, suggesting a network-based flow of information about jobs. Other stream of literature analyzes referral systems in hiring, reporting a better

*We thank comments from Coralio Ballester, Willemien Kets, Federico Valenciano, Fernando Vega-Redondo, and many seminar and conference participants. Financial support from *Ministerio de Economía y Competividad and Fondo Europeo de Desarrollo Regional* (ECO2015-64467-R, ECO2015-66027-P, PID2019-106146GB-I00, PID2019-108718GB-I00 MINECO-FEDER) and the Basque Government (IT-1336-19) is gratefully acknowledged.

†Department of Economics, University of the Basque Country, Bilbao, Spain. *E-mails*: `mariapaz.espinosa@ehu.eus`; `jaromir.kovarik@ehu.eus`; `sofia.ruiz@ehu.eus`.

‡Faculty of Arts & Faculty of Economics, University of West Bohemia, Pilsen, Czech Republic

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

[1] See Goyal (2007), Vega-Redondo (2007), Jackson (2010), and Jackson et al. (2017).

[2] The weight given to the network-driven information differs across these dimensions though. See also Myers and Schultz, 1951; Corcoran, Datcher and Duncan, 1980; Marsden and Campbell, 1990; Montgomery, 1991; Marmaros and Sacerdote, 2002; Munshi, 2003; Franzen and Hangarther, 2006; Pellizari, 2010; and Bentolila et al., 2010, among many others.

[3] Some studies directly explore the variation of employment status of some network members on other network integrants (see e.g. Topa, 2001; Bayer et al., 2008; Beaman, 2011; Beaman and Magruder, 2012).

performance of referred employees (Burks et al., 2015), a higher probability of being hired (Brown et al., 2016; Pallais and Sands, 2016), and a longer tenure in the firm (Dustmann et al., 2016). Thus, employers may benefit from the employees network to increase the search efficiency in the hiring process (Barr et al., 2019).[4]

Most of the above mentioned literature has focused on the effects of network size, whereas finer details of network architecture have received less attention. Granovetter (1973) documents that, while 16.7% of interviewed individuals report having found their jobs through a *strong* tie (i.e. someone they saw regularly), 83.4% found their job through a *weak* tie (a contact they meet "occasionally").[5] He argues that the strength of the tie between two individuals is intimately related to the "*overlap in their friendship circles*", in such a way that individuals linked through strong ties are expected to have a higher proportion of common contacts, integrating densely-connected groups deprived of information from other parts of the social system. In the terminology of this paper, they are more likely to belong to short network cycles. Weak ties, on the contrary, may act as bridges between such densely connected neighborhoods, enjoying an advantage for getting novel information with respect to strong ties. Many studies have tested Granovetter's argument (see e.g. Killworth and Bernard, 1974; Friedkin, 1980; Montgomery, 1992) or applied his idea to different frameworks, ranging from the diffusion of knowledge among organizations (e.g. Hansen, 1999) to innovation or creativity (e.g. Ruef, 2002 and Baer, 2010). Recently, Gee et al. (2017) have analyzed the role of weak and strong ties in job search. Using data on six million Facebook users and creating a proxy for job help by identifying people who eventually worked for the same employer as a pre-existing friend, they report that people are more likely to find a job through a weak tie because of their relatively large number, but strong ties are individually more beneficial at the margin. Note that they measure tie strength between two agents as the number of friends they have in common, or–alternatively–the number of triangles, in which both participate.[6] This evidence notwithstanding, little is known about what drives this *strength of weak ties* phenomenon. Is this effect driven solely by the differences in the content of the information, or is it partly due to the different network embeddedness of weak vs. strong ties?

This paper analyzes theoretically whether close-knittedness influences employment prospects and wages of individuals and groups. Close-knit neighborhoods, one of the most prevalent features of real-world social networks (Jackson and Rogers, 2007; Jackson, 2010), are characterized by high transitivity of relationships: friends of friends are typically friends; and so are friends of friends of friends, and so on. In network terminology, highly clustered networks contain a large number of triangles, squares, and other short network cycles.

To study the role of short cycles in the flow of job market information, we build on a model developed by Calvó-Armengol (2004) and Calvó-Armengol and Jackson (2004, 2007).[7] People are distributed on a network and employed individuals who hear about vacancies pass the information on to their unemployed network neighbors.[8] Calvó-Armengol (2004) shows that higher degree and lower second-order degree increase the individual probability of employment and detects a non-monotonic effect of shifts of the degree distribution on employment in cycle-free regular networks. Calvó-Armengol and Jackson (2004) study patterns of correlations from a dynamic perspective. They show that such a network model generates unemployment correlations across time and path-connected individuals, exhibits duration dependence and persistence, and allows to understand inequalities and drop-out decisions.[9] Calvó-Armengol and Jackson (2004) provide several examples suggesting that other network features might play a role, but they do not isolate their effects formally. Last, Calvó-Armengol and Jackson (2007) analyze wages in a similar model. Our contribution to this literature is to model explicitly the role of local network density of links, measured by the presence of short cycles, isolating their effect from that of the first- and second-order degree.

We report two sets of results. The first result has implications that go beyond the labor-market application

[4]Beaman et al. (2018) identify a potential cause of the different labor outcomes of males and females by documenting a significant tendency of men to refer few women, compared to men.

[5]Granovetter (1973) defines the strength of a tie as the "*amount of time, the emotional intensity, the intimacy (mutual confiding), and the reciprocal services which characterize the tie.*"

[6]Bian et al. (2015) show empirically that weak ties are better able to provide job information than strong ones, but strong ties are better at mobilizing diverse forms of favoritism that is particularly relevant for high-earning positions.

[7]Calvó-Armengol (2004) and Calvó-Armengol and Jackson (2004, 2007) have in turn been partially inspired by the early contributions of Boorman (1975) and Diamond (1981).

[8]See also Montgomery (1991). Galeotti and Merlino (2014) model endogenous formation of networks.

[9]Unemployed agents obtain more (less) information when their contacts are employed (unemployed), giving rise to robust forms of correlations in wages and employment status of path connected agents.

of this study: we show that short network cycles generate *stochastic affiliation* in diffusion processes.[10] More precisely, if two friends of individual i are friends or have another common friend, the information flows from these two friends to i are schochastically affiliated even in a one-preriod model, whereas these flows are independent if they are connected neither directly nor indirectly in any other way.[11] This implies that the diffusion does not only depend on the number of first- and second-order connections but also on the geometry created by these links. Thus, two neighborhoods or two networks where the players have the same number of first- and second-order neighbors (i.e. the same *joint* distribution of degree and second-order degree) but differing clustering patterns induce different diffusion dynamics and thus different employment and wage distributions. Since this result is fairly general and does not rely on many of the model assumptions, our work provides a microfoundation for *why* short networks cycles, namely triangles and squares, shape diffusion on networks.

Our second set of results concerns the micro- and macroeconomic implications of short network cycles. We prove that, *ceteris paribus*, people involved in three- or four-cycles have worse employment prospects. The intuition stems from information affiliation: stochastic affiliation increases the probability of mismatch between vacancies and job candidates, leading to diffusion inefficiencies. That is, network cycles are a source of labor-market frictions when employers rely on referrals. Our numerical simulations show that this result persists in the steady state of a dynamic Markov chain. In the long run, network cycles organize the employment probabilities in the sense of the first-order stochastic dominance. People in densely-knit neighborhoods have lower employment prospects than individuals in comparable positions and close-knit networks exhibit higher unemployment rates than more loosely-knit societies with the same joint distributions of first- and second-order degrees. Hence, short network cycles have important short- and long-run consequences for inequality within and across networks. We further show that economically relevant and statistically strong effects are exclusively limited to short cycles in our model. Importantly, none of these results relies on spatial segregation of low- and high-clustering people. Although spatial correlations in the employment of path-connected people, characterized in Calvó-Armengol and Jackson (2004, 2007), weaken the negative impact of short cycles in integrated societies, we show that their detrimental role can persistent in the steady state. Hence, policies aiming at the integration of dense and loosely-knit agents may have a limited effect.

These findings extend for wages. Leaving constant other network features, three- and four- cycles reduce the expected wage of agents involved in these cycles. However, this negative effect is driven by the unemployment probability. Conditional on having a job, the expected wages of individuals involved is short cycles are actually higher. The affiliation of information flows increases the probability of multiple job offers in dense neighborhoods but such multiplicity is not longer redundant if we analyze wages: multiple job offers allow agents to select among offers with different wages if unemployed or accept better-paid jobs if employed. Hence, the direction of the effect of short network cycles may vary in function of one's socio-economic environment.

Short network cycles further affect other features of the employment dynamics. Most importantly, they increase serial correlations of employment. This has two implications. Firstly, since the variability of the steady-state employment is virtually unaffected by network cycles but serial correlation increases considerably, the fluctuations of employment within denser neighborhoods or in closely-knit societies exhibit larger amplitude. This effect is generated by a combination of factors. First, network diffusion generates employment time correlations, as well as correlations between connected nodes. Employed friends maintain their contacts employed due to the diffusion channel, while unemployed agents make their contacts more vulnerable (Calvó-Armengol and Jackson (2004). However, the affiliation in job-market information flows caused by network short cycles reinforces these effects and extends them to network neighborhoods. On the one hand, it slows down the transition between different employment states: it maintains employment within cycled groups of employed individuals while it also perpetuates unemployment in unemployed cycled neighborhoods. On the other hand, the spatial correlation makes the transition from employment to unemployment in a neighborhood more difficult. However, if several members of an employed neighborhood lose their job, they make their circles more vulnerable and drag the whole neighborhood towards unemployment; in contrast, if a positive shock hits an unemployed neighborhood the network externality spills over more easily in cycled neighborhoods. The combination of these factors results

[10]Stochastic affiliation is a strong form of correlation; see Section 2 for a formal definition.

[11]Some results in Calvó-Armengol (2004) and Calvó-Armengol and Jackson (2004, 2007) rely on the assumption that the inflow of information from two neighbors of i to i is independent. Our contribution is to show that this assumption fails even in the static model if the two neighbors are connected or if they share a friend $j \neq i$. We further explore the consequences.

in longer, more persistent employment fluctuations, with more pronounced booms and troughs in closely-knit neighborhoods and networks. Secondly, the combination of higher time persistence of employment and labor market transitions with lower unemployment causes that, even though the probability of maintaining a job is unaffected, the likelihood of remaining unemployed across periods is enhanced by close-knittedness.

The findings of this paper have important implications for two fields, labor economics and network theory. As for the former, we deepen our understanding of the micro- and macroeconomic effects of social networks in labor markets and the economy as a whole. First, we uncover one possible mechanism behind the strength of weak ties (Granovetter, 1973): since weak connections are less likely to be embedded in short cycles,[12] the strength of weak ties lies not only on the informational content but also on the lack of correlation in the information they provide.[13] Granovetter's (1973) argues that the benefits of weak ties derive from their ability to transmit information to larger audiences and to provide novel information. We show that individuals immersed in loose-knit communities may enjoy information advantages even when they are in contact with the same number of people and when the kind of information they receive is of the same nature as in highly cohesive neighborhoods. Thus, weak ties might be relevant not solely because of their bridging role, but due to their capacity to provide independent information.

Second, taking into account that the empirical evidence consistently corroborates that real-life labor markets heavily rely on employee referrals, we show that short network cycles–and consequently, the clustering coefficient, social cohesion, and network close-knittedness–are important determinants of individual and aggregate labor-market performance. Short cycles affect inequality and they may be a relevant source of labor-market frictions. They reinforce the propagation of idiosyncratic shocks, shaping labor fluctuations over time and space in our model. The detected impact is particularly in line with the literature that shows that search frictions lead to sluggish aggregate employment dynamics and labor market churning (e.g. Pissarides, 1985; Bentolilla and Bertola, 1999; Mortensen and Pissarides, 1999; Burgess et al., 2001). There is a literature on how unemployment entry and exit rates are affected by personal characteristics, such as education (Maarten and Wolbers, 2000); Calvó-Armengol and Jackson (2004) position networks among the determinants. We identify another source of sluggishness: the geometry of one's social environment. This channel deserves a deeper empirical investigation. More generally, since employment fluctuations and employment transitions are strongly correlated with business cycles and aggregate macroeconomic volatility cannot be explained without consideration of labor-market fluctuations (Kydland, 1995; Mortensen and Pissardies, 1999), our results may contribute to the understanding of key stylized facts of macroeconomic dynamics.

Last, we add a new potentially negative aspect to the list of economic consequences of recruitment process that rely on employee referrals. Calvó-Armengol (2004) shows that popular friends and too much connectivity in a network may be detrimental and Calvó-Armengol and Jackson (2004, 2007) implies that, under employee referral programs, initially disadvantaged groups are persistently worse off. Munshi and Rosenzweig (2006) report that networks may prevent people from exploiting new opportunities. We show that, apart from connectivity, network cohesion contributes to inequality and lead to labor-market search inefficiencies.

As for network theory, we provide three novel insights. First, we show that network close-knittedness goes beyond the clustering coefficient as both the number and the organization of cycles of different lengths determine to what extent one benefits from the social capital embedded in social structures. Triadic closure and clustering (i.e. cycles of length three) have received a great deal of attention in the literature across fields. Our focus on cycles links naturally two concepts based on triangles, the clustering coefficient and the concept of network support proposed recently by Jackson et al. (2012). However, both concepts ignore the role of longer cycles in shaping socio-economic phenomena. We show formally that cycles of length larger than three also play a role and why.

Second, numerous studies across fields suggest that local clustering may affect diffusion on networks (e.g. Calvó-Armengol and Jackson, 2004; Centola, 2010; Acemoglu et al., 2011; Campbell, 2013). However, it is

[12]Granovetter (1973) argues that *transitivity, (...), is claimed to be a function of the strength of ties, rather than a general feature of social structure."* According to Marsden and Campbell (1984), the overlap of social circles is a particularly good measure of the strength of a tie. Louch (2002) indeed observes that when two contacts of an individual speak to each other frequently or know each other for a long time, i.e. they can be thought of as *strong ties* according to Granovetter (1973) definition, the probability that they are mutually linked increases by 178% and 278%, respectively.

[13]Conversely, this mechanism provides a rationale behind the *structural holes* argument of Burt (2006): one of the advantages of bridging holes is the diversification of information flows from otherwise disconnected parts of the network.

difficult to disentangle the effect of clustering from that of other network features, which is a stumbling block for any causal claims regarding the effect of short cycles and for a microfoundation of their role in network processes. Since the transmission of information in this paper resembles the network diffusion of many other phenomena, our results may have implications beyond the labor market literature. We show formally that clustering matters keeping constant the first- and second-order degree distributions–the two features typically considered in the diffusion literature, and that short network cycles induce correlation in diffusion. Since cycles already impact diffusion in our static, one-period model, they will likely play a relevant role in the dynamics of other network phenomena and abstracting from them may thus provide an incomplete picture of the mechanisms behind network diffusion.

Last, we contribute to the long-standing debate in economic sociology regarding whether dense neighborhoods are beneficial or detrimental (Burt, 2001, 2009; Jackson, Rogers and Zenou, 2017).[14] We report that, even in the same context, the answer depends on the particular question: close-knittedness may be both beneficial and detrimental and both directions coexist even in the same network and application. More precisely, network cohesion may be beneficial in well-off neighborhoods that are successful maintaining high employment and in times of high employment whereas it may hurt the same people in times of economic unease.

The paper is organized as follows. Section 2 presents the model. Section 3 analyzes the static version of the model while the dynamic analysis can be found in Section 4. Section 5 introduces wages into the model. Finally, Section 6 concludes.

2 Model

To study the role of short cycles in labor markets, we build on the model of Calvó-Armengol (2004) and Calvó-Armengol and Jackson (2004, 2007). First, we present and analyze the static, one-period version of the model where we introduce and consider the impact of short cycles. See Section 4 for the analysis of the dynamics and long-run consequences.

2.1 The network

People are distributed on an undirected fixed network that is used to disseminate information about job openings. The network $g = (N, E)$ is characterized by a set of nodes $N = \{1, .., n\}$ and a set of edges or links E between them. We write $g_{ij} = 1$ if individuals i and j are directly linked in g and $g_{ij} = 0$ otherwise. Let $A = (g_{ij})_{i,j \in N}$ be the $n \times n$ symmetric adjacency matrix of the network, with $g_{ii} = 0$. To simplify notation, we sometimes denote a link between i and j by ij. G is the set of all feasible networks.

The set of i's direct contacts in g is defined as $N_i(g) = \{j \in N : g_{ij} = 1\}$; let $n_i(g) = |N_i(g)|$ be the (first-order) *degree* of individual i. Analogously, denote the set of i's second-order or indirect neighbors (neighbors of i's neighbors) as $N_i^2(g) = \{k \in N : g_{ij}g_{jk} = 1 \text{ for some } j \in N,\ i \neq k\}$ and $n_i^2(g) = \left|N_i^2(g)\right|$. Observe from the previous definitions that $N_i(g)$ and $N_i^2(g)$ may have a non-empty intersection; that is, some contacts of i may simultaneously be i's indirect contacts. For a pair of nodes i and j, define $N_{ij}(g) = \{k \in N : k \in N_i(g) \cap N_j(g)\}$ as the set of common contacts of both i and j, with $n_{ij}(g) = |N_{ij}(g)|$. The set of contacts of i that are not shared with j is $N_{i-j}(g) = N_i(g) \backslash N_{ij}(g)$, with $n_{i-j}(g) = |N_{i-j}(g)|$. Distance between nodes i and j in network g is the length of the shortest path between them, denoted as $d_{ij}(g)$. Naturally, $d_{ij}(g) = 1$ if $ij \in E$, $d_{ij}(g) = 2$ if there is $k \in N$ such that $ik, kj \in E$ but $ij \notin E$, and so on.

This paper focuses on the effect of cycles on the probability of receiving information about job vacancies through network contacts.

Definition 1 (Cycle) *A $K-$cycle $Z_K(g)$ is a sequence of distinct nodes $i_1, i_2, \ldots, i_{K-1}, i_K$, such that $K>2$, $g_{i_k i_{k+1}} = 1$ for each $k \in (1, \ldots, K-1)$, and $g_{i_1 i_K} = 1$.*

[14] Relationships in close-knit networks are typically stronger, enabling trust and cooperation (Granovetter, 1973; Burt, 1992; Coleman, 1988; Bloch et al., 2008; Lippert and Spagnolo, 2011; Jackson et al., 2012). However, close-knit networks inhibit the flow of novel information and individuals in tight neighborhoods may receive redundant information (Granovetter, 1973, 2005; Blau, 1986; Burt, 2001).

In words, a $K-$cycle in a network g is a sequence of K linked nodes starting and ending in the same node. A cycle may be equivalently defined as the set of edges. For example, a *three-cycle* or a *triangle* is a path passing through three edges: ij, jk, ki. Equivalently, we may refer to a triangle as the sequence $\{i, j, k\}$.[15] Analogously, a *four-cycle* or a *square* is a path through four edges ij, jk, kl, li or a sequence $\{i, j, k, l\}$. We denote $S_S(g)$ the set of all three- and four-cycles and $S_L(g)$ the set of all K-cycles for $K > 4$ in g; $S(g) = S_S(g) \cup S_L(g)$ thus corresponds to the set of all K-cycles in g. Last, $S_S^i(g)$, $S_L^i(g)$, and $S^i(g)$ are the sets of the corresponding cycles, in which i is involved.

The clustering coefficient of individual i is the fraction of i's direct contacts who are neighbors themselves. In the above terminology, the coefficient measures the number of triangles in i's neighborhood divided by the number of all possible triangles among all i's contacts. Formally,[16]

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

Clustering coefficient reflects the density or closeknittedness within a node's neighborhood; when the coefficient is high, the neighborhood is densely interconnected because most of i's contacts are linked. The average clustering coefficient of network g is simply $C(g) = \frac{1}{n}\sum_{i=1}^{n} C_i(g)$ and measures the overall level of local density within the network.

Note that the clustering coefficient counts the triangles but abstracts from longer cycles. Lind et al. (2005) propose a coefficient that keeps track of the fraction of four-cycles as follows:

$$C_i^4(g) = \frac{\sum g_{ij}g_{jm}g_{mk}g_{ik}}{\sum g_{ij}g_{ik}g_{jm}},$$

where i, j, k and m are nodes of the network. For node i, the number of squares is given by the number of common neighbors among i's contacts (i.e. that is sequences $\{i, j, k, m\}$ such that $g_{ij}g_{jm}g_{mk}g_{ik} = 1$). Again, $C^4(g) = \frac{1}{n}\sum_{i=1}^{n} C_i^4(g)$.

A distinct but related concept is *support*, proposed by Jackson et al. (2012): A link jk is supported if $\exists\ i \in N : i \in N_{jk}(g)$; the link is unsupported if $N_{jk}(g) = \oslash$. A link jk is supported by i if $i \in N_{jk}(g)$; the link is unsupported by i if $i \notin N_{jk}(g)$. The more links are supported by i, the higher is i's embeddedness. Hence, network support is another measure of close-knittedness, but it differs from the clustering coefficient (see Jackson et al., 2012).

Certain network architectures play a particular role in our analysis. First, a *tree* is a network with no cycles of any length; that is, $S(g) = \emptyset$ and $C_i(g) = C_i^4(g) = 0$ for each $i \in N$. A *cycle network* is the only graph of n nodes and n links containing one n-cycle. In Figure 2, networks $g_b - g_d$ depict a triangle, square, and pentagon, respectively; network g_g represents a hexagon. A *regular network* is a network in which $|n_i(g)|$ is equal for all $i \in N$; a regular network is symmetric in the number of direct and indirect contacts of each individual but not necessarily in other network features.[17] A special case of regular network is a *vertex-transitive network*, where no node can be distinguished from any other based on its neighborhood since they all have a structurally identical neighborhood, second-order neighborhood, and so on. There are several formal definitions of these networks (see e.g. Weisstein, 2016), but an important property of these graphs is that each node occupies a structurally equivalent position in the network.[18] Hence, every vertex-transitive graph is regular, and each node has the same degree, second-order degree, and the clustering coefficient. However, the converse is not true; not all regular networks are vertex-transitive, and not all networks in which all nodes have the same clustering coefficient are vertex-transitive. For instance, every node in a network may have the same number of links but they can still

[15] In graph theory, the term n-cycle is sometimes used as a description of a circle network of n nodes. In Figure 2, g_b and g_c would be examples of three- and four-cycles, respectively, under that terminology. In this paper, cycles may also represent a part of more complex architectures, rather than the whole network.

[16] The coefficient is not defined for $n_i(g) < 2$. In such a case, some authors consider clustering to be equal to zero (e.g. Vega-Redondo, 2007), while others leave it undefined.

[17] Some authors call regular networks symmetric (e.g. Calvó-Armengol, 2004). However, the term symmetric network exists in graph theory and it is more restrictive than regularity.

[18] In fact, vertex-transitive graphs are also called node-symmetric (Chiang and Chen, 1995), a name that reflects better the main feature of these networks.

differ in other characteristics, such as the clustering coefficient or global centrality. Examples of vertex-transitive networks include empty and complete networks, circles, cubes, many lattice networks, and Caley graphs.

Figure 1 displays examples of vertex-transitive networks; all nodes in the five networks occupy an identical position. In case of the first three networks, all vertices have four connections and twelve second-order neighbors, but the number of triangles involving a given node increases from zero in case of the leftmost network to two in case of the middle network. The fourth and fifth networks are also vertex-transitive with $n_i(g) = 4$ for all i, but each node is involved in several three- and four-cycles. To see the difference between clustering and support, note that the clustering coefficient of each node increases from zero to one as we move from the left to the right. Nevertheless, the support is already maximal in the third and fourth networks where every single edge is supported, while the clustering is lower than one in these networks because each node has contacts who are not connected.

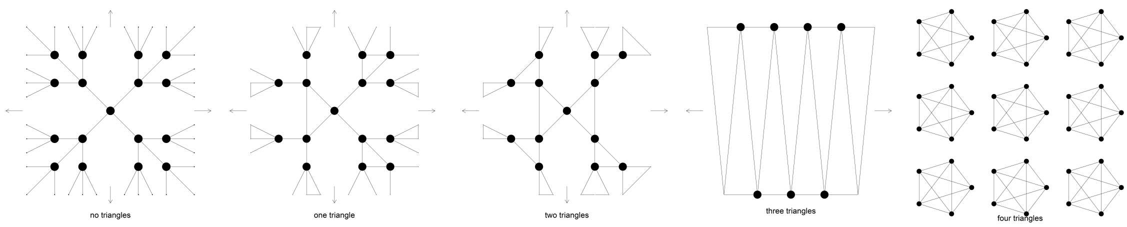


Figure 1: Vertex-transitive networks with degree four but increasing clustering from the left to the right.

2.2 Information flows

In our model, initially each worker is employed.[19] Then, every agent looses her job with probability $b \in (0, 1)$. Afterwards, each node may hear about a vacancy with probability $a \in (0, 1)$, regardless of her employment status. We assume that all jobs are identical.[20] Loosing the job and hearing about a vacancy are independently distributed and independent across individuals.

At this point, each worker can be in one out of four possible situations or status:

Status 1. Each person is employed and has heard about a new job with probability $\alpha = a(1 - b)$.

Status 2. She is unemployed and has not heard about any offer with probability $\beta = b(1 - a)$.

Status 3. She is unemployed but has heard about an offer with probability $\delta = ab$.

Status 4. She is employed with no offer to pass to contacts with probability $\gamma = (1 - a)(1 - b)$.

We label these different status as 1 to 4 respectively, and define the random variables $X_1^i(g)$, $X_2^i(g)$, $X_3^i(g)$, $X_4^i(g)$ as the number of i's contacts who are in each status. Particular realizations of these variables are denoted as $x_1^i(g)$, $x_2^i(g)$, $x_3^i(g)$ and $x_4^i(g)$, with $x_1^i(g) + x_2^i(g) + x_3^i(g) + x_4^i(g) = n_i(g)$. Similarly, the variables $X_m^{jk}(g)$ and $X_m^{j-k}(g)$ measure respectively the number of agents in $N_{jk}(g)$ and $N_{j-k}(g)$ who are in status $m \in \{1, 2, 3, 4\}$; $x_m^{jk}(g)$ and $x_m^{j-k}(g)$ denote again their realizations. Define $Y_i^s(g)$ a random variable, such that $y_i^s(g) = 1$ if i is in state $s = \{1, 2, 3, 4\}$; $y_k^s(g) = 0$ otherwise. Unemployed individuals who hear about a job (those in status 3) immediately accept the offer. Workers who are employed and have heard about a vacancy (in status 1) pass the offer to one of their unemployed neighbors (in status 2) uniformly at random. As a result, status 1 and 2 play a key role in our analysis.

The model assumes that all agents at any point in time have perfect information about the labor status of their direct contacts.[21] Observe that only individuals in state 1 can pass an offer to one of their direct contacts

[19]This assumption is inconsequential. All the results hold if all people start with the same probability of being employed.

[20]In Section 5, we relax this assumption and allow for wages to differ across jobs.

[21]In Calvó-Armengol (2004), employed individuals with a job offer pass the vacancy to one of their contacts who lost their job. Unemployed individuals may have later received an offer but this is not observed by other agents. Though this is not explicitly

in status 2 (i.e. those who lost their job but have not heard about any offer) who immediately accepts the offer. Therefore, an individual's contacts who are in status 1 will be called *potential providers* and her second-order contacts (the contacts of her contacts) who happen to be in status 2 will be named *competitors*. By construction, it is possible that an individual in status 2 simultaneously receives several offers from her different contacts. In such a case, she accepts one of them at random and the others remain unfilled. These redundant offers may generate search frictions in the labor market.

Assume that node i is unemployed. We represent the information flow from $j \in N_i(g)$ to i through a random variable $I_j^i(g)$ taking value 1 when i receives information from j and 0 otherwise. $I_j^i(g)$ is a function of the state of j and j's network position. First, it depends on whether j is in state 1 (a provider), an event with probability α. If so, she passes an offer to one of her unemployed neighbors with equal probability. Thus, i receives an offer from j with probability $\frac{1}{X_2^{j\setminus i}(g)+1}$, where $X_2^{j\setminus i}(g)$ denotes the number of j's contacts other than i who are in state 2. Then,

$$I_j^i(g) = \begin{cases} 1 & \text{with probability} \quad \left(\frac{\alpha}{X_2^{j\setminus i}(g)+1}\right) \\ 0 & \text{with probability} \quad \left(1 - \frac{\alpha}{X_2^{j\setminus i}(g)+1}\right) \end{cases} \tag{1}$$

2.3 The probability of receiving an offer through the network and the unemployment rate

Denote $P^i(g)$ the probability that node i receives a job offer from at least one neighbor in network g when she is is status 2 (i.e. unemployed). With this notation, we can write the employment probability of node i as $E_i(g) = (1-b)+ba+b(1-a)P^i(g)$, which can be interpreted as the individual employment prospect in network g. The employment rate of the network is $E(g) = \frac{1}{n}\sum_{i\in N} E_i(g)$; the unemployment rate is thus $U(g) = 1 - E(g)$.

Since a and b are exogenously given and the same for all individuals, the only difference in (un)employment rates across nodes and networks arises from $P^i(g)$ so we analyze how this probability depends on network cycles.

$P^i(g)$ depends on the employment status of i's direct and indirect (second-order) contacts. The status of i's direct neighbors determines the number of i's potential providers, $X_1^i(g)$, while the status of the contacts of the potential providers determines the number of competitors $(X_2^{j\setminus i}(g))$ and thus the probability with which any potential provider j passes an offer to i (see (1)). As we show below, both three- and four-cycles affect the probability with which i receives at least one offer through her contacts.

Assume that $y_i^2(g) = 1$ and $y_j^1(g) = 1$, $j \in N_i(g)$. Since each node $k \in N_j(g) \setminus i$ may be unemployed (that is, in status 2) with probability β, the probability that i does not receive any offer from j is[22]

$$q_j(n_j(g)) = \sum_{h=0}^{n_j(g)-1} \binom{n_j(g)-1}{h} \beta^h (1-\beta)^{n_j(g)-1-h} \frac{h}{h+1}. \tag{2}$$

In a similar vein, the probability that i does not receive any offer from j, conditionall on knowing the status of $k \in N_j(g) \setminus i$, is

$$q_j(n_j(g) \mid y_k^2(g) = x) = \sum_{h=0}^{n_j(g)-2} \binom{n_j(g)-2}{h} \beta^h (1-\beta)^{n_j(g)-2-h} \; \frac{h+x}{h+x+1}. \tag{3}$$

Expressions (2) and (3) directly lead to the following claim:

stated, it can be seen in the proof of his Proposition 1, where the decision to pass information depends on b instead of β. We change slightly this assumption and assume that the decision depends on β. If we maintained the assumption of Calvó-Armengol (2004), our results would in fact be reinforced.

[22]Since $q_j(n_j(g))$ is identical for any $k \in N_j(g)$ including i, we do not index this probability by i throughout the paper to simplify the notation.

Claim 1 (Individual probability) *Assume that* $y_i^2(g) = 1$ *(unemployed) and let* $y_j^1(g) = 1$ *for a node* $j \in N_i(g)$ *(potential provider of* i*). Then,* $q_j(n_j(g))$ *is increasing in* $n_j(g)$ *and independent of network cycles, and* $q_j(n_j(g) \mid y_k^2(g) = 1) > q_j(n_j(g) \mid y_k^2(g) = 0) = q_j(n_j(g) - 1)$.

Claim 1 illustrates first that the probability with which i does not receive any offer from a particular provider j (or, conversely, the probability that she does, i.e. $1 - q_j(n_j(g))$) only depends on the number of (potential) competitors in j's neighborhood and this probability is not affected by the presence of cycles within i's or j's neighborhoods. Hence, network cycles do not affect the *individual* decision to pass information about jobs. However, Example 1 below illustrates that cycles do affect the probability of getting an offer from *at least one* contact, altering one's employment prospects. Second, Claim 1 shows how knowing the status of a neighbor of a neighbor affects this probability and that knowing that one neighbor does not need a job is equivalent to having one competitor less. These latter results play a key role if a neighbor of a neighbor can simultaneously be a direct neighbor.

Example 1 (Effect of short cycles) *Consider networks* g_b, g_c, g_e, *and* g_f *in Figure 2 and assume that* $y_i^2(g) = 1$ *in each network. The probability that node* 1 *receives a job offer from at least one neighbor in network* g, $P^1(g)$, *satisfies the following:*[23]

$$P^1(g_e) > P^1(g_c) > P^1(g_b). \tag{4}$$

Note that node 1*'s degree is the same in the three networks, while her second-order degree is the same in networks* g_b *and* g_e *but lower in* g_c*. However, individual* 1 *is more likely to be employed in the tree network* g_e*, followed by the square network* g_c*, as compared to the triangle network* g_b *even though the number of competitors is the same in the first two networks and even though node* 1 *has one competitor less in* g_c *than in* g_e*. Similarly,* $P^1(g_f)$ *may be larger or lower than* $P^1(g_b)$*, depending on the parameters. If, say,* $\alpha = 0.9$ *and* $\beta = 0.01$*,* $P^1(g_f) = 0.9886 > P^1(g_b) = 0.9810$*, despite the fact that the number of* 1*'s potential competitors is greater in* g_f *than in* g_b*.* □

Claim 1 and Example 1 jointly deliver several messages. First, the geometry of network neighborhoods affects the probability of receiving a job offer, beyond the number of first- and second-order neighbors. We particularly point to the role of short cycles. Since Claim 1 shows that their presence does not affect the probability of receiving an offer from one particular individual, their role in Example 1 stems from the lack of independence of information flows coming from different contacts j and k if they belong to a cycle with i (operationalized by a multivariate random variable $(I_j^i(g), I_k^i(g))$). The example particularly suggests that network cycles affect individuals negatively and their length may matter. Last but not least, the impact of cycles is economically relevant: their effect in Example 1 rivals with that of two-link-away connections. We formalize these observations in the next section.

3 Results

3.1 Cycles and affiliation of information flows

In cycle-free networks (e.g. trees), $N_i(g) \cap N_i^2(g) = \emptyset$ and $N_{jk}(g) = \{i\}$ if $j, k \in N_i(g)$, implying that information flows from neighbors j and k to i are independent. The probability that neighbor j passes an offer to i depends on j's status and on $X_2^{j \setminus i}$, but it depends neither on the status of any other $k \in N_i(g)$ nor on the status of any second-order contact $s \in N_k(g) \setminus i$, $k \neq j$.

In contrast, if $j, k \in N_i(g)$ and $jk \in E$ (i.e. j and k form a three-cycle with i), the information flow from j to i depends on both j's and k's status and, similarly, that from k depends on k's and j's status. More precisely, the probability that j transmits information about a job opening to i depends on whether j is employed and possesses such information and on whether k also needs a job. If k does, k will provide no information to i and this is independent of cycles. However, since j and k are friends, k now competes with i for information from j,

[23] Detailed calculations may be found in Appendix ??.

decreasing i's probability of receiving information from j. Such *indirect* effect of the status of k on $P^i_j(g)$–and thus on $P^i(g)$–is a direct consequence of the cycle $\{i,j,k\}$. This indirect effect is missing when j and k do not belong to any short cycle with i.

Similar dependence appears when j and k belong to a four-cycle together with i. If a four-cycle $\{i,j,s,k\}$ is present, the lack of independence in information flows from j and k to i comes from the fact that the probabilities of receiving a job offer from each of them depend on a common variable, the status of individual s.

We formalize these statements using the concept of stochastic affiliation:[24,25]

Proposition 1 (Affiliation) *Assume $y^2_i(g)=1$ and consider $j,k \in N_i(g)$.*
(a) $I^i_j(g)$ and $I^i_k(g)$ are strictly affiliated if either $S^i_S(g)=\{i,j,k\}$ or $S^i_S(g)=\{i,j,s,k\}$ for one $s \neq i,j,k$.
(b) Conditional on the status of j's and k's neighbors who form three or four-cycles with i, $I^i_j(g)$ and $I^i_k(g)$ are conditionally independent.

Strict affiliation in information flows from node i's potential providers implies that, conditional on receiving an offer from a provider j with high probability, the probability that i receives another offer from a provider k increases if the two providers are connected directly or through another node $s \neq i$.

We illustrate the intuition behind the proof using an example. Consider a triangle network g_b in Figure 2 composed of individuals i,j, and k and assume that node i is unemployed. Denote $f(I^i_j(g_b),I^i_k(g_b))$ the joint density function of $I^i_j(g_b)$ and $I^i_k(g_b)$. First, consider that $j \in N_i(g)$ is employed and hears about an open position; this happens with probability α. Only j will pass information to i, but not k, when k is employed but does not have information about other jobs, an event that has probability $1-\alpha-\beta$. Node j transmits a job offer to i with probability 1 in such a case. In contrast, i only receives a job offer from j with probability $\frac{1}{2}$ if k needs a job and thus competes with i for j's information, an event that has probability β. As a result and due to the network symmetry,

$$f(1,0)=f(0,1)=\alpha\beta\frac{1}{2}+\alpha(1-\alpha-\beta).$$

Second, i receives no offer from any of her neighbors if either none of them has any information to pass or one of them does but passes it along to a competitor, leading to

$$f(0,0)=(1-\alpha)^2+2\alpha\beta\frac{1}{2}$$

Last, a node does not lose the job and hears about a vacancy with probability α. Therefore, i might receive two offers from both j and k with probability $f(1,1)=\alpha^2$.

Note that

$$f(1,1)*f(0,0)=\alpha^2\Big((1-\alpha)^2+\alpha\beta\Big)>\alpha^2\Big((1-\alpha-\beta)+\beta\frac{1}{2}\Big)^2=f(0,1)*f(1,0)$$
$$(1-\alpha)^2+\beta\alpha>(1-\alpha)^2-\beta(1-\alpha-\frac{1}{4}\beta),$$

which implies that $I^i_j(g_b)$ and $I^i_k(g_b)$ are strictly affiliated.

To illustrate the case of a four-cycle, consider the four-node network g_c in Figure 2 and take the perspective of a node i with $N_i(g_c)=\{j,k\}$ and $N^2_i(g_c)=\{s\}$. In this example, it cannot happen that i receives an offer with probability 1 from j and with probability $\frac{1}{2}$ from k, or viceversa. If $y^2_s(g_c)=0$ (i.e. s does not need a job), both potential providers j and k can only pass a job along to i with probability 1. Conversely, if $y^2_s(g_c)=1$, both potential providers j and k will pass information to i with a lower probability $\frac{1}{2}$. Hence, the probabilities that each of the two neighbors of i pass information to her are not independent.

[24] This concept was introduced to economics by Milgrom and Weber (1982). Two random variables are affiliated if, conditional on observing a low (high) value of one variable, the probability of observing a low (high) value of the other variable increases. Formally, two random variables X and Y are affiliated if for all $x<x'$ and for all $y<y'$:

$$f(x',y)*f(x.y')\leq f(x',y')*f(x,y)$$

where *f(x,y)* is the joint density function of variables X and Y. Since two independent random variables are affiliated according to the above expression, we say that two variables are *strictly* affiliated if the condition holds with strict inequality.

[25] All proofs are relegated to Appendix **??**.

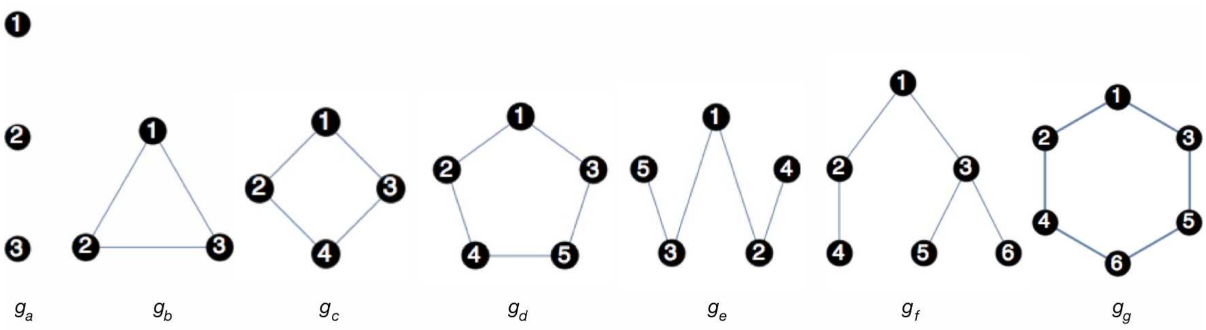


Figure 2: Networks with different local connectivity and clustering patterns: empty network with $n = 3$, triangle, square, pentagon, two trees, and a hexagon.

Part (b) of Proposition 1 shows that the affiliation in the information flows is the only reason for such dependence. Once we condition on the status in the corresponding neighborhoods, the information flows become independent.

3.2 Effects of cycles on employment

Proposition 1 shows that network cycles generate affiliation in information flows but provides no prediction concerning the direction of the effect. Calvó-Armengol (2004) shows that, within cycle-free networks, direct contacts are beneficial whereas second-order contacts are detrimental for the individual probability of employment. However, example 1 indicates that this does not necessarily hold in networks that contain short cycles. Let us denote $\left(n_i(g); \{n_j(g)\}_{\forall j \in N_i(g)}\right)$ as the *joint degree distribution* of i. The following proposition characterizes the effects of short cycles, keeping constant the joint degree distribution (i.e. the number of direct contacts and the number of contacts of her contacts for each i):

Proposition 2 (Effect of cycles) *Let $g = (N, E)$ and $g^x = (N^x, E^x)$, $x \in \{t, s\}$, be three networks, such that $N \subseteq N^x$, $\left(n_i(g); \{n_j(g)\}_{\forall j \in N_i(g)}\right) = \left(n_i(g^x); \{n_j(g^x)\}_{\forall j \in N_i(g^x)}\right)$ for $\forall i \in N$ and $x \in \{t, s\}$, $S_S(g^t) = S_S(g) \cup \{i, j, k\}$ and $S_S(g^s) = S_S(g) \cup \{i, j, z, k\}$ for some $i, j, k, z \in N$. Then,*

(i) $P^h(g) > P^h(g^t)$ *for* $h \in \{i, j, k\}$ *and* $P^f(g) = P^f(g^t)$ *for all* $f \in N \setminus \{i, j, k\}$.

(ii) $P^h(g) > P^h(g^s)$ *for* $h \in \{i, j, k, z\}$ *and* $P^f(g) = P^f(g^s)$ *for all* $f \in N \setminus \{i, j, k, z\}$.

(iii) $P^h(g^s) > P^h(g^t)$ *for* $h \in \{i, j, k\}$ *and* $P^h(g^s) = P^h(g^t)$ *for* $f \in N \setminus \{i, j, k, z\}$.

Figure 3 provides an example of the networks in Proposition 2. Networks g, g^t and g^s have the same *joint* distribution of degree and second-order degree and the set of cycles is the same with one exception: g^t (g^s) has one additional triangle (square) compared to g.[26]

Part (i) indicates that if an individual has the same degree and her neighbors also have the same degrees in both networks but she is involved in one triangle more in g^t than in g (e.g. the case of nodes 1, 2, and 3 in Figure 3), she is less likely to get a job offer through her network contacts in g^t. In contrast, the remaining nodes (i.e. nodes $4 - 7$ in Figure 3) are unaffected by the difference between g and g^t. Part (ii) shows that the same holds for a square but, as Part (iii) indicates, the impact of a triangle is larger that that of a square.

[26]Note that keeping the second order degree constant implies that the number of neighbors of i's contacts is constant; thus, whenever the distribution of degree and second order degree is the same in the different networks, the conditions of Proposition 2 hold.

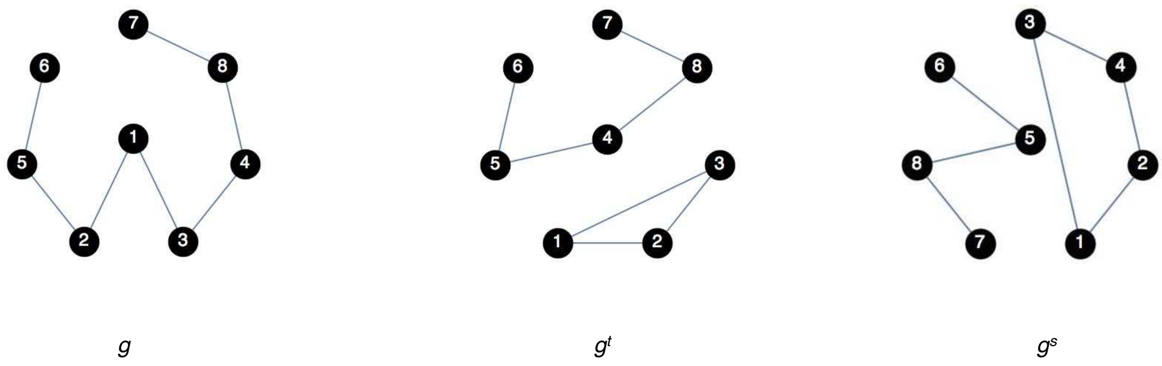


Figure 3: Example of g, g^t and g^s from Proposition 2.

A direct consequence of Proposition 2 for labor-market outcomes is that under the conditions of the proposition, the rate of employment of i is lower if i belongs to a triangle or square in g^t or g^s, respectively: $E_i(g) > E_i(g^s) > E_i(g^t)$. That is, i is more likely to be unemployed in $g^x, x \in \{s,t\}$ than in g and more so if the additional cycle is a triangle rather than a square, while $E_i(g) = E_i(g^s) = E_i(g^t)$ for the nodes that are not involved in any of the additional cycles.

To illustrate the intuition behind the proposition, let us compare a tree network g, in which $S_S(g) = \emptyset$, with a network g^t such that $S_S(g^t) = \{i, j, k\}$. Let $N_i(g) = N_i(g^t) = \{j, k, v, \ldots, z\}$. Assume $ik, ij \in E \cap E^t$, and $jk \in E^t$ but $jk \notin E$. If $i \in N$ is in status 2 (i.e. i needs a job), the probability that she does not receive any offer from $j \in N_i(g)$ is $R^i_j(g) = \alpha q_j(n_j) + (1-\alpha)$. In the following illustration, to simplify notation we omit the superscript i and the dependence on g to write R_j.

Since $S_S(g) = \emptyset$, the probability that i does not receive any offer from any neighbor in g is simply the product of the individual probabilities over all i's neighbors. Therefore, the probability that i receives at least one offer from her contacts is (see Proposition 1 in Calvó-Armengol, 2004):

$$P^i(g) = 1 - R_j R_k R_v \ldots R_z = 1 - \prod_{h \in N_i(g)} R_h. \tag{5}$$

Since the only difference between g and g^t relevant for i is that $jk \in E$, Proposition 1 shows $I^i_j(g^t)$ and $I^i_k(g^t)$ are affiliated. As a result, $R_{jk} \neq R_j R_k$, where R_{jk} is the probability that i does not receive any offer neither from j nor from k. Consequently,

$$P^i(g^t) = 1 - R_{jk} R_v \ldots R_n = 1 - R_{jk} \prod_{h \in N_i(g) \setminus \{j,k\}} R_h. \tag{6}$$

Note from (5) and (6) that $P^i(g) > P^i(g^t)$ if $R_{jk} > R_j R_k$.

Table 1 presents the probabilities R_{jk} conditional on the status of j and k for networks g and g^t under the conditions of Proposition 2.[27] The first column of the table lists the four possible combinations of the status of j and k: both j and k are in status 1 and thus they are potential providers of i (row 1), two cases in which one of them is a provider while the other one is not (rows 2 and 3), and the fourth situation where none of them is a provider of i (row 4). The second and third columns, denoted respectively as g and g^t, contain the probabilities that i does not receive any offer neither from j nor k under each scenario in any of the two networks. The probabilities R_{jk} are obtained by simply adding up the four expressions in the corresponding columns.[28]

The last column (labeled $g^t - g$) in Table 1 is the difference between these probabilities (R_{jk} across the two networks. If we add up the four rows of this last column, we obtain

$$R_{jk} - R_j R_k = \alpha^2 \Big[\Big(q_j(n_j) - q_j(n_j - 1) \Big) \Big(1 - q_k(n_k - 1) \Big) + \Big(q_k(n_k) - q_k(n_k - 1) \Big) \Big(1 - q_j(n_j) \Big) \Big] > 0. \tag{7}$$

[27] Appendix **??** contains the detailed computation of the probabilities in Table 1.

[28] Note that $R_{jk}(g) = R_j(g) R_k(g)$ but $R_{jk}(g^t) \neq R_j(g^t) R_k(g^t)$.

R_{jk}/status	g^t	g	$g^t - g$
j, k providers	$\alpha^2 q_j(n_j - 1) q_k(n_k - 1)$	$\alpha^2 q_j(n_j) q_k(n_k)$	$\alpha^2 \left[q_j(n_j - 1) q_k(n_k - 1) - q_j(n_j) q_k(n_k) \right]$
j is provider, k not	$\alpha q_j(n_j) - \alpha^2 q_j(n_j - 1)$	$\alpha(1-\alpha) q_j(n_j)$	$\alpha^2 \left[q_j(n_j) - q_j(n_j - 1) \right]$
k provider, j not	$\alpha q_k(n_k) - \alpha^2 q_k(n_k - 1)$	$\alpha(1-\alpha) q_k(n_k)$	$\alpha^2 \left[q_k(n_k) - q_k(n_k - 1) \right]$
j, k not providers	$(1-\alpha)^2$	$(1-\alpha)^2$	0

Table 1: Probability that i does not receive any offer from $j, k \in N_i(g)$ in g^t and g.

by Claim 1. Therefore, $P^i(g) > P^i(g^t)$.

Table 1 further illustrates how the affiliation affects the information flows from j and k to i. In g^t, the affiliation increases the likelihood that i receives two offers from both j and k or none, while decreasing the probability of receiving just one offer. When j and k are providers (row 1 in Table 1), $R_{jk} - R_j R_k < 0$ by Claim 1. In contrast, two events in Table 1 (one neighbor can pass information to i while the other cannot, corresponding to cases in rows 2 and 3) yield a lower probability of not getting at least one offer in g^t than in g, $R_{jk} - R_j R_k > 0$. Expression (7) shows that the opposing effects are non-neutral and the aggregate effect on i is negative (higher probability of not getting at least one offer). In g^t, i receives two offers simultaneously from j and k more often but she can accept only one of the jobs while the other one remains unfilled. Therefore, short network cycles, by inducing affiliation, increase the likelihood of mismatch between candidates and jobs; affiliation decreases the employment prospects of people in closely-knit neighborhoods. In economic terms, transitivity in relationships and overlapping in the neighborhoods of different individuals, through the effect of affiliation of information flows, prevent an efficient diffusion and generate labor market frictions.[29]

The above argument extends naturally when the starting network g is not a tree, and to cycles of length four. Regarding the former, $P^i(g)$ and $P^i(g^t)$ would be more complex than a product of R_h's if $S(g) \neq \emptyset$, but the comparison between the corresponding expressions (5) and (6) would again reduce to the comparison of the probabilities $R_{jk}(g)$ and $R_j(g)R_k(g)$. As for four-cycles, the argument is the same, except that the affiliation, and the fact that $R_{jk}(g^s) \neq R_j(g^s)R_k(g^s)$, stems from the status of contacts of neighbors such as $s \in N_{jk}(g^s)$. If s is employed, her status makes it more likely that *both* j and k share a job information with i and when s is unemployed it is more likely that i does not receive any offer from any of them.

Proposition 2 has several implications and raises a few issues that are worth stressing here. First, note that the above comparison holds both within and across networks. That is, it does not matter whether we compare two individuals across two networks as in Proposition 2 or two individuals within the same network. The reason is that, in the static, one-period version of the model, only the local neighborhood matters. More precisely, the employment probability of an individual is determined only by the status of agents in her first- and second-order neighborhood and by the geometry of these local neighborhoods. Therefore, any two nodes with the same degree, second-order degree, and, say, the same numbers of squares but different number of triangles could also be compared. The following two remarks clarify this point:

Remark 1 *In the one-period model, m-cycles do not affect employment prospects for any $m > 4$.*

Remark 2 *In the one-period model, $I_g^i(g)$ is unaffected by cycles that do not contain i.*

It is important to note that none of the above two remarks would hold in a dynamic model (see Section 4, where we analyze the dynamics).

Second, what can we say about the unemployment at the network level? Proposition 2 can be interpreted as an individual as well as a network-level result. Since the employment prospects of agents only differ in the probability $P^i(g)$, the next statement directly follows from Proposition 2:

[29] Note that this effect is stronger that a simple diversification argument. A higher probability of receiving two offers or none, and a lower probability of having just one offer could be equivalent if the individual could benefit more from receiving two offers than from receiving just one. However, two offers are equivalent to just one offer because the individual can only have one job. Thus, the negative effect is stronger, because people cannot fully enjoy the advantage of receiving multiple job offers. In other words, even a slightly risk loving individual would, *ceteris paribus*, prefer not to be involved in cycles.

Corollary 1 (Network level) *Consider networks g, g^t, and g^s from Proposition 2 with $N = N^t = N^s$. Then, the unemployment rate in each network is such that:*

$$U(g^t) > U(g^s) > U(g). \tag{8}$$

In words, our results allow us to compare two networked societies in terms of unemployment as long as the assumptions in Proposition 2 hold: when two networks have the same *joint degree distribution* and they only differ in one short cycle, the network that contain the cycle has higher unemployment rate and the effect is stronger for a triangle than for a square.

The above discussion raises a natural question: can we compare two societies if we relax some of the assumptions of Proposition 2? The answer is no. The following example shows that having the same *joint degree distribution*, $\left(n_i(g); \{n_j(g)\}_{\forall j \in N_i(g)}\right)$, in both networks is a necessary condition. If we relax this assumption, for example by only maintaining constant the degree distribution, Proposition 2 may fail to hold.

Figure 4: Networks in Example 2.

Example 2 (Equal joint degree distribution as a necessary condition) *Consider networks g and g' in Figure 4. Both networks have the same degree distribution, but the joint distribution differs across them. At the individual level, being involved in more cycles does not necessarily translate in lower unemployment if the number of direct and indirect contacts is not held constant. For instance, node 1 belongs to the cycle {1,2,3} in network g' while she is not involved in any cycle in network g. Nevertheless, $P^1(g') = 0.1334 > P^1(g) = 0.1285$ for $a = 0.1$ and $b = 0.2$.*[30] *The reason is that node 1 competes with a smaller number of agents in g' than in g for information and this benefits her more than the harm that comes from the affiliation induced by the triangle. From the perspective of the entire network, although $S(g') \setminus S(g)$ = {1,2,3}, $E(g') = 0.84013 > E(g) = 0.84012$ for $a = 0.1$ and $b = 0.2$. Hence, having less cycles does not guarantee lower unemployment if the joint distribution of the direct and two-links-away friends is not held fixed.* □

Example 2 illustrates that, if two networks have the same distributions of degrees and neighbors' degrees but a distinct joint distribution of both variables, the (un)employment rates cannot be generally ranked according to Corollary 1. This is an important observation: Example 2 shows that Proposition 2 and Corollary 1 cannot be generalized by relaxing the assumption of holding fixed the joint distribution of connectivity and second-order connectivity.

Next we compare specific networks with different clustering coefficient.

Corollary 2 (Vertex-transitive networks) *Consider two vertex-transitive networks g and g', with $\left(n_i(g); \{n_j(g)\}_{\forall j \in N_i(g)}\right) = \left(n_i(g'); \{n_j(g')\}_{\forall j \in N_i(g')}\right)$ for $\forall i \in N$. If $C_i(g') \geq C_i(g)$, $C_i^4(g') \geq C_i^4(g)$ for all i, and at least one of them is satisfied with strict inequality, then $U(g') > U(g)$.*

[30] See Section ?? for the derivation of all the probabilities.

The corollary is a direct consequence of Proposition 2 and the definition of vertex-transitive networks (see Section 2.1). Remember that a network is vertex-transitive if all the nodes occupy identical positions. Hence, they have the same degree, second-order degree, clustering coefficient, global centrality etc. As a consequence, if we compare two such networks that have the same number of neighbors and number of neighbors' contacts, we can compare their unemployment rates in line with both Proposition 2 and Corollary 1. To provide an illustration, Corollary 2 predicts that unemployment increases as we move from the first to the third network in Figure 1.

According to Proposition 2 what matters for labor prospects is the number of short cycles, while we refer to the clustering coefficients in Corollary 2. Due to the popularity of the clustering coefficient, one may ask why we do not link unemployment directly to the coefficient in Proposition 2? The main reason is that the relationship between the number of cycles and the clustering coefficients is not one-to-one. As an illustration, consider node i in networks g_a - g_c in Figure 5. Although $C_i(g_a) = C_i(g_b) = \frac{1}{3}$, $P^i(g_a) \neq P^i(g_b)$ because the two triangles in g_b additionally form a four-cycle that affects i's the information flows and thus the employment prospects of those involved in the four-cycle. Similarly, $C_i^4(g_b) = C_i^4(g_c)$ but $P^i(g_b) \neq P^i(g_c)$. In vertex-transitive networks to which we refer in Corollary 2, differences in the number of cycles in which agents are involved are controlled for, since all agents occupy identical positions. The short cycles in their neighborhoods are then distributed equally. However, more complex architectures contain more complex interactions between the two clustering coefficients and cycles of different lengths, depending on who is involved in which cycle. Therefore, we express our main results in terms of additional network cycles to illustrate their *ceteris paribus* effect. A result a là Proposition 2 that would link employment to the clustering coefficient should account for both the number of cycles and their distribution across agents.

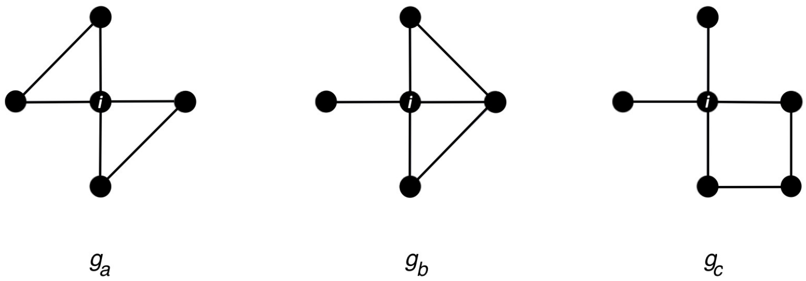


Figure 5: Relation between short cycles and the clustering coefficients

4 Dynamic analysis

In this section, we examine the long-run consequences of network cycles in labor markets. Calvó-Armengol and Jackson (2004) showed in a similar model that the unemployment rate is positively correlated across time periods and path-connected agents, and that there is duration dependence. They further provide numerous examples illustrating that the network topology shapes the long-run labor-market performance. In this section, we separate the impact of short cycles on long-run employment and inequality patterns from that of other network features.

In our dynamic setup time evolves in discrete steps denoted by t. At the beginning of each period t, each worker may be employed or unemployed, depending on her employment status in period $t-1$. If she starts employed, she has a probability $b \in (0,1)$ of losing the job. Afterwards, all workers (employed and unemployed) hear about a vacancy with probability $a \in (0,1)$. As in the static model, losing a job and hearing about a vacancy are independently distributed and independent across individuals and periods. Afterwards, unemployed individuals who have heard about a job accept it immediately, while those employed pass it along to one of their unemployed connections at random.

We assume that at $t = 1$, all nodes start being employed.[31] The only difference between the first and the subsequent periods is the initial employment status; in subsequent periods, not necessarily all individuals are employed at the beginning of the period, but they may bring unemployment from the previous period.

Let the *state* at the beginning of period t be the vector $s_t = (E_{1t}, \ldots, E_{nt})$, where $E_{it} = 0$ if node i is unemployed when period t starts and $E_{it} = 1$ if i is employed. Each state s_t has an associated employment rate $E_t = \frac{\sum E_{it}}{n}$ and unemployment rate $U_t = 1 - \frac{\sum E_{it}}{n}$. From $t = 2$ on, the state s_t will be determined by the parameters a and b, the network architecture g that channels the flow of information about job openings and the employment state of the previous period s_{t-1}; the rest of the history is irrelevant. This constitutes a Markov chain.

Since the state space is finite, we can represent the transition distribution probability through the transition matrix P with element (i, j) given by $p_{ij} = \Pr(S_{t+1} = j \mid S_t = i)$. Each row of P contains the probabilities of each possible state s_{t+1}, i.e. all possible combinations of employment and unemployment for all nodes given that they started the period in state s_t.

In this framework, the Markov chain is time-homogeneous, so that the transition matrix P is the same in each period, and the t-period transition probability can be computed as the t-th power of the transition matrix, P^t. As a result, the dynamics can be modeled as a finite-state irreducible and aperiodic Markov chain and the process converges to a unique limit distribution (Young, 1993). Note that the initial state becomes irrelevant after a certain number of periods and the limit probabilities over states only depend on the parameters of the model, a and b, as well as on the network architecture g. We are interested in the steady-state probability distribution over all possible employment states: $lim_{t\to\infty} P^t = \Pi$, where Π is a matrix in which each row is the limit distribution of the process and yields the probability of each possible state s. From this limit probability distribution over states, we first ask whether the effects of short cycles identified in Section 3 persist in the long run. In addition, we study other moments of the distribution, the space correlations and the persistence of employment, focusing on how they change as we systematically manipulate the close-knittedness of network neighborhoods.

We start by computing the exact limit distribution of the proposed Markov chain for a few simple networks and all values of $a, b \in (0, 1)$ (subsection 4.1). However, certain statistics do not have a general closed-form solutions, even in the case of very simple networks, and characterizing the limit distributions for larger and more complex network structures is not feasible. As a consequence, the rest of the section relies on Monte Carlo methods to study how network cycles shape the steady state (un)employment patterns, in a set of selected networks and for given values of a and b. We present the results of the simulations for $a = b = 0.1$. The conclusions are qualitatively robust to alternative parameter constellations.[32]

To facilitate the comparison of the long-run dynamics across networks, we build networked *economies* with different architectures but the same number of nodes. For each network architecture we simulate 100 economies for 10,000 periods for each parameter constellation and, to analyze the steady state distribution, we report the unemployment patterns of the last 1,000 periods.[33]

We present our results in three subsections. Subsection 4.1 analyzes whether the length of network cycles matters for long-run employment. Subsection 4.2 studies how network cycles shape long-run inequality within networks, and Subsection 4.3 makes comparisons across homogeneous networks.

4.1 Long-run unemployment in cycle networks

In this subsection, we characterize the steady state distributions of the Markov chain for a few cycle networks (see Section 2 for a definition). Consider the networks g_a through g_d in Figure 2 (an empty network with $n = 3$, a triangle, a square, and a pentagon, respectively). The one-period analysis shows that three- and four-cycles matter and the former have a larger impact on employment than the latter. Here, we show that this result

[31] The model converges to a unique steady-state distribution, independently of the initial state of the system. The initial employment state is thus irrelevant after convergence.

[32] The analysis for different values is available from the authors upon request.

[33] We performed robustness checks to see how fast the model converges and it actually converges to the steady state relatively fast. We are thus confident that running the model for 10,000 periods and selecting the last 1,000 provides a good approximation of the steady-state distributions.

persists in the long run. The dynamic model further allows us to test whether the influence of cycles extends to cycles of longer lengths and whether the effect vanishes as we increase the cycle length.

The networks g_a to g_d are simple and the positions of all nodes symmetric in each of them, which allow us to compute the transition matrices, the limit state distributions as a function of a and b, and the associated average steady-state employment rates for each node and network.[34] Figure 6 plots the differences in employment rates between a few pairs of networks as a function of the parameters (a, b). The figure reveals that the steady-state employment rate in networks g_a to g_d can be ranked as $E[g_a] < E[g_b] < E[g_c] < E[g_d]$ for any $a, b \in (0, 1)$. That is, the steady state employment rates scale up with the length of the cycle (from three to five) in each network. This points to the persistence in the long run of the negative role of short cycles that we identified in the static model. In the following subsections we corroborate this negative effect using alternative manipulations of network close-knittedness. Moreover, the above ranking of the average employment rate shows that the impact extends to five-cycles and is decreasing with the cycle length.

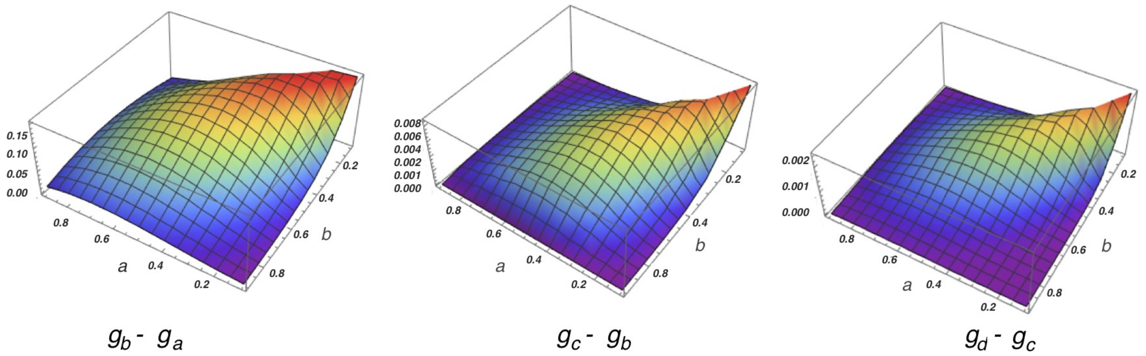


Figure 6: Differences in steady-state employment rates in networks g_a, g_b, g_c, and g_d from Figure 2.

Although we are able to obtain explicitly the limit distributions as a function of a and b for some of the networks in Figure 2, certain statistics cannot be computed explicitly and the complexity of the limit distributions increases dramatically with the network size. Thus, in the following we provide numerical experiments for $a = b = 0.1$ in cycle networks. To that aim, we complement the networks g_a to g_d, analyzed above, with the hexagon network g_g in Figure 2. Table 2 reports the long-run labor-market outcomes from the simulation exercise. Each cell reports the average steady-state outcome in the last 1,000 periods (out of the simulated 10,000) from 100 economies composed of 60 individuals each.

We report four types of outcomes in Table 2 (as well as in Tables 3, **??**, and **??**):

A. *Employment statistics* report the number of nodes in each economy, the number of economies and the total number of cycles,[35] the average employment rate, its standard deviation and coefficient of variation.

B. *Time and spatial correlations* report the serial correlations in average employment and the simple-matching coefficients (SMC) in employment of connected and two-links-away nodes.[36]

C. *Transition rates* provide information about the changes and persistence of the labor-market status of agents.[37]

D. *Kolmogorov-Smirnov, Wilcoxon, and Fligner-Killeen tests* report the statistics (and p-values) of non-parametric

[34] We include the empty network for comparison purposes.

[35] Each economy contains many cycles; for example, with 60 nodes an economy contains 20 triangles or 15 squares or 12 pentagons or 10 hexagons.

[36] The simple-matching coefficient reports the fraction of network links in which the two nodes involved share the employment status. For instance, a value of 0.6 means that 60% of connected members in all simulated economies in the last 1,000 periods were either both employed or both unemployed.

[37] We label EE the fraction of people who keep their job from one period to the next, UU is the fraction of those who remained unemployed, EU is the fraction of the population that were employed in one period but lost their job in the next, while UE is the reverse case. For example, $EE = 0.80$ means that that 80% of people across all the simulated economies, in the last 1,000 periods, kept their jobs from one period to the next. The other cases are interpreted accordingly.

tests of equality of distributions, means, and variances of employment, respectively, across all pairs of networks after convergence.

	Networks				
	Empty ($n = 3$)	Triangle	Square	Pentagon	Hexagon
A. Employment statistics					
Nodes in each economy:	60	60	60	60	60
Number of economies:	100	100	100	100	100
Total number of cycles:	-	2000	1500	1200	1000
Employment rate (E):	0.5275	0.6812	0.6892	0.6922	0.6917
St. Dev. of E:	0.0648	0.0696	0.0698	0.0694	0.0695
Coef. Variation of E:	0.1229	0.1022	0.1013	0.1003	0.1005
B. Time and spatial correlations					
Correlation(E_{t-1}, E_t):	0.8119	0.8039	0.8043	0.7993	0.8001
Correlation(E_{t-2}, E_t):	0.6596	0.6496	0.6493	0.6424	0.6433
Average SMC (1st neighbors):	NA	0.6380	0.6263	0.6232	0.6229
Average SMC (2nd neighbors):	NA	NA	0.6130	0.5965	0.5965
C. Transition rates					
Fraction continue employed EE:	0.4801	0.6267	0.6340	0.6367	0.6362
Fraction continue unemployed UU:	0.4251	0.2644	0.2557	0.2522	0.2527
Fraction lost job EU:	0.0474	0.0544	0.0552	0.0556	0.0556
Fraction found job UE:	0.0474	0.0545	0.0552	0.0555	0.0556
Conditional: EE/(EU+EE)	0.9101	0.9200	0.9199	0.9198	0.9197
Conditional: UU/(UE+UU)	0.8997	0.8292	0.8225	0.8195	0.8198
D. Kolmogorov-Smirnov, Wilcoxon, and Fligner-Killeen tests					
(equality of distribution; mean; variance)		D (p)	W (p)	χ^2 (p)	
Empty vs. triangle		0.7460 (0)	549147871 (0)	52.8604 (0.0440)	
Empty vs. square		0.7674 (0)	467889490 (0)	37.8294 (0.3414)	
Empty vs. pentagon		0.7782 (0)	432566331 (0)	33.9011 (0.4725)	
Empty vs. hexagon		0.7779 (0)	440002393 (0)	38.5586 (0.3118)	
Triangle vs. square		0.0477 (0)	4676213322 (0)	47.9243 (0.0715)	
Triangle vs. pentagon		0.0638 (0)	4553235062 (0)	43.5134 (0.1271)	
Triangle vs. hexagon		0.0596 (0)	4570659747 (0)	42.7190 (0.1734)	
Square vs. pentagon		0.0166 (0)	4877763199 (0)	33.3234 (0.5006)	
Square vs. hexagon		0.0159 (0)	4895893172 (0)	54.7886 (0.0178)	
Pentagon vs. hexagon		0.0042 (0.3354)	5018302682 (0.1552)	32.3255 (0.5979)	

Table 2: Long-run labor market statistics in cycle networks.

Table 2 corroborates our previous result for the static case: $E[g_a] < E[g_b] < E[g_c] < E[g_d]$. The tests for equality of distribution and mean show that the labor market outcomes are all statistically different from each other and therefore the type of cycle prevalent in the economy is relevant ($p < 0.00001$; Kolmogorov-Smirnov and Wilcoxon rank-sum tests); the variances of the average employment do not differ systematically though.[38] The differences in the employment rates are relatively small between the triangle, square, and pentagon economies. While the increase is over 15% from the empty network to the triangle, it reduces to 0.79% from the triangle to the square, and 0.31% from the square to the pentagon. Quantitatively speaking, in an economy composed of 1,000,000 workers, for the selected parameters values ($a = b = 0.1$), there would be roughly 7,975 and 11,079 unemployed individuals more in the triangle economy in the long-run, as compared to the square and pentagon economies, respectively. Figure 7 reveals that the different economies can be ranked in the sense of the first-order stochastic dominance: the limit distribution of employment in the pentagon economy first-order stochastically dominates (FOSD, hereafter) the square economy, which FOSD the triangle economy.[39]

Although the long-run employment distributions of the pentagon and hexagon economies FOSD all the remaining ones, these two economies generate very similar employment distributions. In fact, we cannot reject

[38]The non-parametric Fligner-Killeen tests of equality of variances only allow to reject the equality at 5% in two cases, but the ranking is not systematic.

[39]To focus on cycles, the empty network is omitted in Figure 7 but the ranking in the sense of the first-order stochastic dominance extends to the empty network.

the equality of the distributions or their moments. This suggests that the economically relevant and statistically strong long-run impact of network cycles is limited to relatively short network cycles.

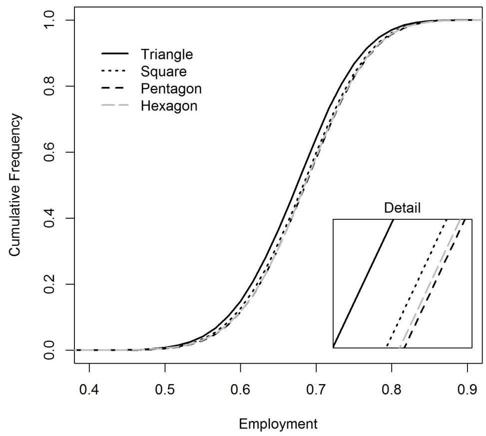


Figure 7: Cumulative density function of the employment rate in cycle networks.

The simulated limit distributions enable us to analyze other long-run effects of network cycles that go beyond the average employment rates of the cycle economies. For instance, one important macroeconomic question concerns employment fluctuations. We thus ask how network cycles affect the volatility and structure of employment. As mentioned above, we find no systematic effects of network cycles on employment volatility across our comparison networks. Even though we confirm there are time and network correlations in employment and that the correlations decrease with time and network distance, we detect virtually no systematic impact of network-cycle length on the values of all these correlations.

In sum, the numerical experiment performed in this section confirms and complements our previous theoretical results. Network cycles decrease employment prospects even in the long run and their impact diminishes with the length of the cycle. Furthermore, although relatively longer cycles could theoretically induce associations in information flows once the model is repeated, we find that statistically strong and economically relevant long-run effects of network cycles are limited to short cycles, such as triangles and squares and to a lesser extent five-cycles. The effects do not seem to go beyond five-cycles after convergence.[40] In the following sections, we limit our attention to the role of triangles.[41]

4.2 Network cycles and long-run inequality

In the previous subsection, all nodes occupy symmetric positions, which precludes within-network comparisons. However, an important question is whether network cycles can be a source of long-run inequality. In this section, we present the results of two experiments. First, we investigate whether triangles affect inequality persistently even if the nodes with low and high clustering are not segregated in the network. Then, we analyze the impact of the clustering coefficient in an economy simulated on a real-life friendship network.

Different clustering in an integrated network. We explore a network that integrates individuals with higher and lower clustering. Specifically, we analyze the long-run effects of short cycles in the network depicted in Figure 8. The network is composed of 28 individuals with equally sized first- and second-order neighborhoods

[40] This conclusion only holds generally in qualitative terms. For other values of a and b, the exact cycles lengths for which differences vanish may change.

[41] The effects of squares and pentagons in the reported simulation exercises are again qualitatively similar but quantitatively weaker than those of triangles.

($n_i(g) = 3$ and $n_i^2(g) = 6$ for each i) but differing clustering patterns. In particular, 7 nodes have clustering equal to zero, whereas 21 agents belong to one three-cycle (i.e. have clustering equal to 1/3).

More importantly for our purpose, the low- and high-clustering individuals are not segregated in the network: each low-clustering individual is only connected with high-clustering individuals and each high-clustering node is connected to one low-clustering and two high-clustering agents. Hence, this network enables us to study whether short network cycles have any implications on long-run inequality even if people are not segregated by the density of their neighborhoods. Such "no-segregation" condition is important because of the steady-state spatial correlations in employment status across directly and indirectly connected individuals, shown in Calvó-Armengol and Jackson (2004). Such long-run employment correlations across network paths reduce the inequality across nodes in the same component and may thus potentially eliminate the negative effect of short cycles if high- and low-clustering individuals are close to each other in the network. We therefore test whether differing clustering patterns can still generate inequality even when nodes are not segregated by clustering.

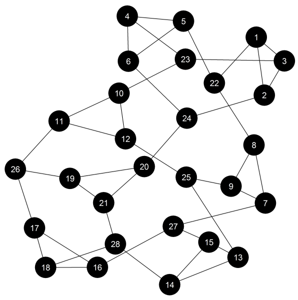


Figure 8: Regular network with $n_i(g) = 3$ and $n_i^2(g) = 6$ for each i, and two types of nodes regarding their close-knittedness–high- and low-clustering agents–who are mutually integrated.

Table 3 summarizes the long-run labor-market outcomes for the whole network (column *All*) as well as disaggregated for the low- and high-clustering individuals. The structure of the table is the same as in Table 2. To ensure the comparability of all statistics across the two node types, we report the results of 100 independent realizations for each type. Due to the differing number of the two types of nodes in the network, we simulated 300 economies/networks and report the results of 100 randomly chosen economies for the high-clustering individuals and of all the 300 economies for the low-clustering agents. This way, we compare a total of 2,100 low- vs. 2,100 high-clustering individuals in the last 1,000 periods of the simulated Markov chain.

From Table 3, we observe some differences between the two types in their long-run employment prospects. Low-clustering individuals are somehow more likely to be employed and enjoy smaller employment volatility. However, these differences are small and statistically non-significant in the Wilcoxon signed rank and Fligner-Killeen tests ($p = 0.1137$ and 0.2587 for the mean and the variance, respectively). The small size of the differences observed in mean and variance between low and high clustering nodes is driven by the spatial correlation and the fact that there is no spatial segregation between the two types. Nevertheless, we reject the equality of the two distributions at any reasonable significance level, using the Kolmogorov-Smirnov test of equality of the distributions ($p < 0.00001$). Figure 9(a) visualizes the comparison. We observe that the steady-state employment distribution of the low-clustering individuals second-order stochastically dominates the employment distribution of the non-zero-clustering nodes.

The two types of nodes further differ in the persistence of the employment status. Note that short cycles increase the serial correlations of employment (see the reported 95% confidence intervals of $\mathrm{Cor}(E_t, E_{t-1})$). That is, the correlation of employment between two consecutive periods is higher if the node is embedded in a more dense neighborhood. As a result, the employment cycles exhibit a different structure for each type. Figure 9(b) provides an example of employment cycles of low and high clustering nodes in one of the simulated economies. It illustrates that high clustering maintains the employment state more stable across consecutive periods but, once

	Type of node		
	All	Low clust.	High clust.
A. Employment Statistics			
Num. nodes:	28	7	21
Num. economies/networks:	300	300	100
Employment rate (E):	0.7314	0.7341	0.7314
St. deviation of E:	0.1023	0.0979	0.1119
Coef. variation of E:	0.1399	0.1334	0.1530
B. Time and spatial correlations			
Cor(E_{t-1}, E_t):	0.7951 (p=0)	0.7180	0.7763
95% confidence intervals of Cor(E_{t-1}, E_t):		[0.7150; 0.7210]	[0.7738; 0.7787]
Cor(E_{t-2}, E_t):	0.6343	0.5197	0.6059
Simple-Matching Coef. (1st neighb.):	0.6446		
Simple-Matching Coef. (2nd neighb.):	0.6205		
C. Transition rates			
EE:	0.6750	0.6773	0.6743
UU:	0.2123	0.2091	0.2133
EU:	0.0563	0.0568	0.0562
UE:	0.0563	0.0568	0.0562
Conditional: EE/(EU+EE)	0.9230	0.9226	0.9231
Conditional: UU/(UU+UE)	0.0563	0.7864	0.7915
D. Kolmogorov-Smirnov, Wilcoxon, and Fligner-Killeen tests			
(equality of distribution; mean; variance)	D (p)	W (p)	χ^2 (p)
Low vs. high clustering	0.03461 (0)	5020228825 (0.1137)	20.309 (0.2587)

Table 3: Steady-state labor-market statistics in the network from Figure 8 (last 1,000 periods).

we escape a state, the troughs and peaks of employment cycles may be lower and higher, respectively, in close-knit network environments. These features result from a combination of different network effects. Since short network cycles increase spatial correlation, connected individuals are more likely to be in the same state across periods and network links, and the effect "drags" the employment of most members of a clustered community up or down, toward a new common employment status.

The higher employment persistence among high-clustering individuals also affects labor-market transitions. While high-clustering individuals are less likely to preserve their employment, they are slightly more likely to remain unemployed across periods.

Summarizing, network integration of low and high clustering nodes quantitatively and qualitatively attenuates but does not eliminate the detrimental impact of short cycles. As a consequence, a policy aiming at the integration of communities with differing clustering patterns does not necessarily eliminate all labor-market disadvantages generated by network close-knittedness.

Clustering in a real-world network. In contrast to the network depicted in Figure 8, typical social networks may show a certain degree of segregation of clustering patterns and perhaps some correlation between clustering and connectivity, among other features. For that reason, we now turn to a real-life friendship network to simulate our model and study the long-run effects of clustering. We use the network elicited in Brañas et al. (2010). First, this particular network is not too large and thus computationally it is not too demanding. And, second, it exhibits typical features of real-life social networks, including a large variability in the joint distribution of degrees, second-order degrees and clustering, positive assortativity, and negative clustering-degree correlation (see Brañas et al. (2010) for details). We use the giant component of their network with $n = 76$, depicted in Figure **??** in Appendix **??**, and again simulate 100 independent networked economies over 10,000 periods. The average steady-state employment rate is 72.37% ($E = 0.7237$; $sd(E) = 0.06$) and it exhibits large serial and spatial correlations.

Our main interest is to determine how these patterns at the individual level correlate with network positioning. Table 4 reports three regressions, one for each relevant labor market outcome: individual employment (E_i; column (1)), its standard deviation (sd(E_i); column (2)) and time correlation ($Cor(E_{it}, E_{it-1})$; column (3)). We consider the three dependent variables for each network member in the last 1,000 periods. The independent variables are the individual's first- and second-order degrees and clustering coefficient. In columns (2) and (3),

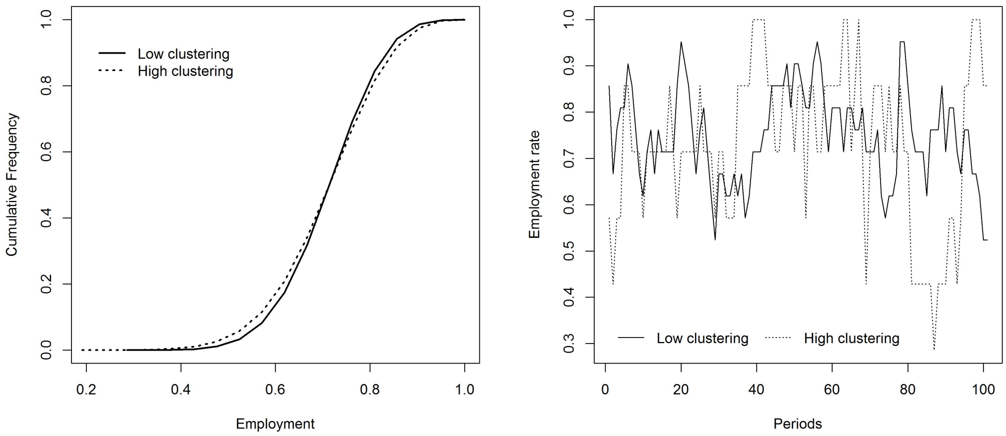


Figure 9: Cumulative density function of employment and employment fluctuations with two types of nodes (Figure 8).

we further control for the average employment of the node. To provide a clean effect of the clustering coefficient, the regressions only include individuals with $n_i(g) \geq 2$ so that they have a well defined clustering. Standard errors are clustered at the network level (the smallest independent unit in the simulated data). control for the average individual employment in the steady state

The results corroborate the theoretical hypotheses. The average employment, its volatility, and the serial correlations all change systematically with individual degree, second-order degree, and the clustering coefficient. The estimates are statistically strong ($p < 0.00001$). Most importantly, holding the first- and second-order degree constant, the clustering coefficient decreases individual employment and increases simultaneously its volatility and autocorrelation.[42]

4.3 Vertex-transitive networks

In this section, we propose an exercise that resembles a first-order stochastic dominance shift of the distribution of the clustering coefficient. To that aim, we design a series of vertex-transitive networks a là Figure 1, in which all nodes occupy identical positions. More precisely, we hold the first- and second-order degree distributions constant across the networks to the extent possible but vary systematically the number of triangles each node is embedded in. We perform this exercise for degree-three and degree-four networks. Figures **??** and **??** in Appendix **??** illustrate the networks under our comparisons; Tables **??** and **??** deliver the steady-state labor-market statistics.

Figure 10 summarizes the main findings of this subsection. Independently of whether we focus on $n_i(g) = 3$ or 4 for each node, the steady-state probability distributions of employment dominate each other as we decrease the number of triangles, in which each individual is involved. In all cases, the distributions and the means are significantly different ($p < 0.00001$), while the variances do not differ systematically. In quantitative terms, the steady-state employment rate decreases from 73.62% to 72.95% under $n_i(g) = 3$ as we mode from a cycle-free network to the case when each node is involved in exactly one triangle. This would correspond to over 6,660 unemployed in a one-million-people economy. The figure further decreases to 71.57% in a network with $C(g) = 1$, corresponding to 20,525 more unemployed with respect to the zero-clustering network. The average employment is naturally higher in networks where $n_i(g) = 4$ for each node, but the ranking with respect to triangles is

[42]Tables **??** - **??** in Appendix **??** illustrate the importance of our *ceteris paribus* condition. The estimated effects of the regressors frequently switch signs depending on their combination in the model. Moreover, the tables show that the results are robust to controlling for global centrality of each node.

	Dependent variable:		
	E_i	sd(E_i)	Cor(E_{it}, E_{it-1})
	(1)	(2)	(3)
Degree	0.029***	−0.001***	−0.011***
	(0.0003)	(0.0001)	(0.0004)
Second-order degree	−0.008***	0.001***	0.004***
	(0.0004)	(0.0001)	(0.0004)
Clustering coef.	−0.006***	0.001***	0.012***
	(0.001)	(0.0003)	(0.001)
Average E_{it}		−0.535***	−0.383***
		(0.002)	(0.010)
Constant	0.662***	0.830***	1.005***
	(0.002)	(0.002)	(0.007)
Observations	6,300	6,300	6,300
R^2	0.648	0.964	0.653
Adjusted R^2	0.648	0.964	0.653
Residual Std. Error	0.035 (df = 6296)	0.006 (df = 6295)	0.028 (df = 6295)
F Statistic	3,866.398*** (df = 3; 6296)	41,584.200*** (df = 4; 6295)	2,962.656*** (df = 4; 6295)

Note: robust st. errors clustered at network level in parentheses. *p<0.1; **p<0.05; ***p<0.01

Table 4: Real-world network: OLS. People with $n_i(g) \geq 2$ who have well defined clustering coefficient.

preserved: the employment rates are 76.5%, 75.88%, 75.36%, and 73.8% as we move from the zero-clustering network to a fully clustered architecture in our four networks under study. We thus conclude that, *ceteris paribus*, first-order stochastic dominance shifts of the distribution of the clustering coefficient organize the distribution of employment in the sense of the first-order stochastic dominance.

In line with the previous sections, short network cycles again induce larger serial correlations in the steady-state employment status in the networks analyzed here. As a consequence, since the variance is similar across the networks but the time correlations increase steadily with the number of triangles, the peaks and troughs of the employment cycles are somehow higher and lower as we increase the clustering of the networks under study. This–jointly with the lower employment prospects in more close-knit networks–again affects labor-market transitions: the likelihood of maintaining a job is virtually unaffected across the networks as we increase their close-knittedness. In contrast, the probability of remaining unemployed between two consecutive periods increases steadily. Correlations in employment of linked people increase in the limit distributions but they decrease between two-links-away individuals. All these observations corroborate the conclusions from the previous subsections.

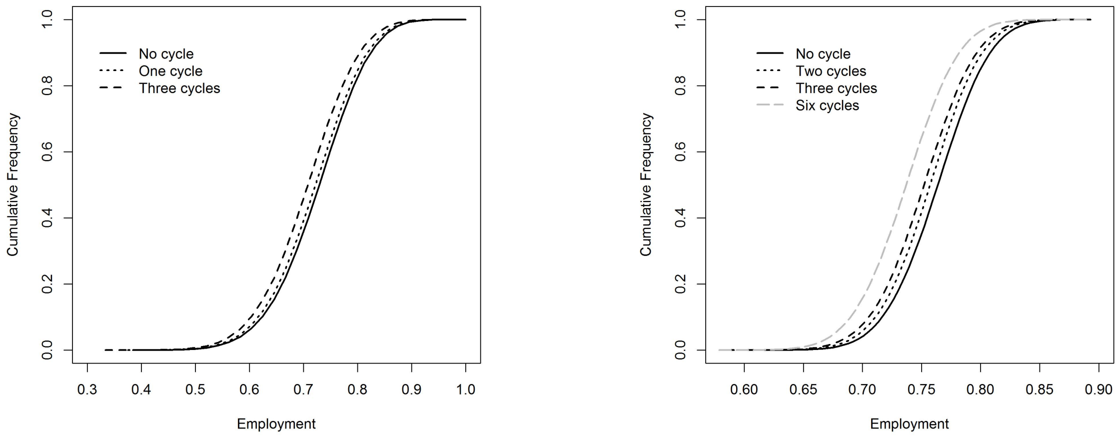
5 Wages

In this section, we briefly analyze the impact of three-cycles on wages. To that purpose, we analyze an extension of the static model from Section 2, in which each job offer comes with a wage.

5.1 Information transmission with wages

Let $W_i(g)$ be a random variable denoting the wage of the position occupied by individual i in network g. For simplicity, we assume that there are two wage levels in the economy: low-paying positions with wage w_0 and high-paying positions with wage $w_1 > w_0$. Initially, all people are employed in a high-paying job.[43] Again, each worker may lose her job with probability $b \in (0, 1)$. Then, each individual hears about a low- or a high-paying

[43] Once again, this assumption is inconsequential. All the results are qualitatively robust as long as all people occupy a high-paying job with the same probability.



(a) $n_i(g) = 3$ for each $i \in N$ (b) $n_i(g) = 4$ for each $i \in N$

Figure 10: Cumulative distributions of employment in vertex-transitive networks.

job with probabilities a_0 and a_1, respectively, with $a_0 + a_1 = a \in (0, 1)$. At this stage, each worker can find herself in one out of six situations (status):

Status 1: with probability $\alpha_0 = a_0(1-b)$, she is employed in a high-paying job and possesses information about a low-paying job,

Status 2: with probability $\alpha_1 = a_1(1-b)$, she is employed in a high-paying job and possesses information about a high-paying job,

Status 3: with probability $\beta = b(1 - a_0 - a_1)$, she is unemployed and has no information about any vacancy,

Status 4: with probability $\delta_0 = a_0 b$, she is unemployed but has heard about a low-paying job,

Status 5: with probability $\delta_1 = a_1 b$, she is unemployed but has heard about a high-paying job,

Status 6: with probability $\gamma = (1 - b)(1 - a_0 - a_1)$, she is employed with no offer to pass to her contacts.

where α=α_0+α_1. Let us write $\bar{y}_i^s(g) = 1$ if agent i is in status s.[44] At this stage, unemployed workers who learn about a vacancy (agents in status 4 or 5) immediately accept the offer, regardless of whether the job is high- or low-paying. Employed workers who learn about a low-paying job (status 1) or a high-paying job (status 2) pass the offer uniformly at random onto one of their unemployed contacts (in status 3) who accepts the offer.

Individuals in status 1 and 2 are potential providers; we call them *low providers* and *high providers*, respectively. Second-order neighbors in status 3 will be called *competitors*. As in Section 2, it is possible that an unemployed individual receives multiple offers simultaneously. In such a case, she accepts the job with the highest wage, while the other positions remain unfilled.

5.2 The incidence of triangles on wages

We first show that adding a triangle to a network a là Proposition 2 reduces the expected wage of the nodes involved in the triangle:

[44]Bear in mind that there are four status in Sections 2 - 4, while there are six of them in this extended version of the model.

Proposition 3 (Wage) *Consider networks g and g^t defined in Proposition 2. Then, $E[W_h(g)] > E[W_h(g^t)]$ for $h \in \{i, j, k\}$ and $E[W_z(g)] = E[W_z(g^t)]$ for all $z \neq i, j, k$.*

Proposition 3 complements Proposition 2 by showing that lower employment prospects of clustered individuals and networks translate into lower expected wages. However, this result raises a question: Is this finding driven by the unemployment channel or does higher clustering affect wages through additional mechanisms? To answer this question, the following proposition focuses on the expected wage conditional on ending up employed and asks whether close-knit neighborhoods benefit or hurt employed individuals:

Proposition 4 (Conditional wage) *Consider networks g and g^t defined in Proposition 2. Then, $E[W_h(g^t) \mid E_h(g^t) = 1] > E[W_h(g) \mid E_h(g) = 1]$ for $h \in \{i, j, k\}$ and $E[W_z(g^t) \mid E_h(g^t) = 1] = E[W_z(g) \mid E_h(g) = 1]$ for all $z \notin \{i, j, k\}$.*

We know from Propostion 2 that forming part of short network cycles decreases one's employment probability. Therefore, it is not surprising that in turn decreases one's expected wage. Nevertheless, Proposition 4 shows that the negative effect is driven by the unemployment channel. If we compare the wages of two *employed* individuals whose local positioning only differs in the cohesion of their networks, the presence of triangles actually benefits people. The intuition behind this findings is closely related to intuition behind Proposition 2. The lack of independence of information flows from different neighbors persists, leading to higher probability of receiving multiple offers. However, while multiple offers do not increase one's employment likelihood because each agent can only accept one job, they do increase the probability of hearing about at least one high-paying job. As a result, receiving multiple offers is not redundant any longer and the expected wage conditional on being employed is higher in clustered neighborhoods.

This result provides an additional channel for how network cycles contribute to the persistence and widening of income inequalities across communities and over time periods. Well-off communities or economic crises with high employment rates are benefited by close-knittedness, while bad neighborhoods and periods of economic unease that suffer from high unemployment rates are actually hurt by the same network feature.

6 Conclusions

This paper analyzes systematically the role of short network cycles in labor market outcomes. We show formally that densely-knit neighborhoods lead to the affiliation in information diffusion with important micro- and macroeconomic consequences on expected unemployment rates, wages, inequality, and employment fluctuations. In particular, network cycles lead to lower expected employment rates both at the individual and the population level and both in the short and long run. Moreover, clustering leads to employment fluctuations with higher volatility and more persistence (higher time correlation). Clustering results also in lower expected wages. This effect is, however, driven by the lower probability of employment; for employed workers, expected wages are higher if they belong to short cycles. The reason is that detected affiliation can benefit workers because they may benefit from receiving multiple offers by selecting better-paying jobs. The main direction for future research stemming from our paper is the empirical test of the theoretical results.

References

[1] Acemoglu, D., Ozdaglar, A., & Yildiz, E. (2011). "Diffusion of innovations in social networks". In *2011 50th IEEE Conference on Decision and Control and European Control Conference: 2329-2334*. IEEE.

[2] Baer, M. (2010). "The strength-of-weak-ties perspective on creativity: a comprehensive examination and extension". *Journal of Applied Psychology* 95(3): 592.

[3] Barr, T., Bojilov, R., & Munasinghe, L. (2019). "Referrals and Search Efficiency: Who Learns What and When?". *Journal of Labor Economics* 37(4): 1267-1300.

[4] Bayer, P., Stephen L.R. & Topa, G. (2008). "Place of Work and Place of Residence: Informal Hiring Networks and Labor Market Outcomes". *Journal of Political Economy* 116(6): 1150-1196.

[5] Beaman, L.A. (2011). "Social networks and the dynamics of labor market outcomes: Evidence from refugees resettled in the US". *The Review of Economic Studies* 79(1): 128-161.

[6] Beaman, L.A. (2016) “"Social networks and the labor market",” in *The Oxford Handbook on the Economics of Networks*, edited by Bramoullé, Yann, Andrea Galeotti and Brian Rogers, Oxford: Oxford University Press.

[7] Beaman, L., Keleher, N. & Magruder, J. (2018). "Do job networks disadvantage women? Evidence from a recruitment experiment in Malawi". *Journal of Labor Economics* 36(1): 121-157.

[8] Beaman, L., & Magruder, J. (2012). "Who gets the job referral? Evidence from a social networks experiment". *American Economic Review* 102(7): 3574-93.

[9] Bentolila, S., Michelacci. C. & Suarez. J. (2010). "Social contacts and occupational Choice". *Economica* 77(305): 20-45.

[10] Bian, Y., Huang, X. & Zhang, L. (2015). "Information and favoritism: The network effect on wage income in China". *Social Networks* 40: 129-138.

[11] Blau, P. M. (1986). *Exchange and power in social life*. Transaction Publishers.

[12] Bloch, F., Genicot, G. & Ray, D. (2008). "Informal insurance in social networks". *Journal of Economic Theory* 143(1): 36-58.

[13] Boorman, S. A. (1975). "A combinatorial optimization model for transmission of job information through contact networks". *Bell Journal of Economics* 6(1): 216-249.

[14] Brañas-Garza, P., Cobo-Reyes, R., Espinosa, M. P., Jiménez, N., Kovářík, J., & Ponti, G. (2010). "Altruism and social integration". *Games and Economic Behavior* 69(2), 249-257.

[15] Brown, M., Setren, E. & Topa, G. (2016). "Do informal referrals lead to better matches? Evidence from a firm's employee referral system". *Journal of Labor Economics*, 34(1): 161-209.

[16] Burgess, S., Lane, J., & Stevens, D. (2001). "Churning dynamics: an analysis of hires and separations at the employer level". *Journal of Labor Economics* 8: 1-14.

[17] Burt, D. R. (2001). "Bandwidth and echo: Trust, information, and gossip in social networks", in *Networks and Markets: Contributions from Economics and Sociology*, ed. Alessandra Casella and James E. Rauch, Russell Sage Foundation, 2001.

[18] Burt, R. S. (2009). *Structural holes: The social structure of competition*. Harvard University Press.

[19] Burks, S. V., Cowgill, B., Hoffman, M. & Housman, M. (2015). "The value of hiring through employee referrals". *The Quarterly Journal of Economics* 130(2): 805-839.

[20] Calvó-Armengol, A. (2004). "Job contact networks". *Journal of Economic Theory* 115(1): 191-206.

[21] Calvó-Armengol, A. & Jackson, M. O. (2004). "The effects of social networks on employment and inequality". *American Economic Review* 94(3): 426-454.

[22] Calvó-Armengol, A., & Jackson, M. O. (2007). "Networks in labor markets: Wage and employment dynamics and inequality". *Journal of economic theory* 132(1): 27-46.

[23] Campbell, A. (2013). "Word-of-mouth communication and percolation in social networks". *American Economic Review* 103(6): 2466-98.

[24] Cappellari, L. & Tatsiramos, K. (2015). "With a little help from my friends? Quality of social networks, job finding and job match quality". *European Economic Review* 78(C): 55-75.

[25] Centola, D. (2010). "The spread of behavior in an online social network experiment". *Science* 329(5996): 1194-1197.

[26] Chiang, W.-K. & Chen, R.J. (1995). "The (n,k)-Star Graph: A Generalized Star Graph". *Information Processing Letters* 56: 259-264.

[27] Coleman, J. S. (1988). "Social capital in the creation of human capital". *American Journal of Sociology* 94: S95-S120.

[28] Conley, T. G. and Topa, G. (2002). "Socio-economic distance and spatial patterns in unemployment". *Journal of Applied Econometrics* 17(4): 303-327.

[29] Corcoran, M., Datcher, L., & Duncan G. J. (1980). "Information and Influence Networks in Labor Markets". in *Five Thousand American Families: American Journal of Sociology*, vol. 8, edited by Greg J. Duncan and James N.

[30] Diamond, P. (1981). "Mobility Costs, Frictional Unemployment, and Efficiency". *Journal of Political Economy* 89, 798-812.

[31] Dustmann, C., Glitz, A., Schőnberg, U., & Brücker, H. (2015). "Referral-based job search networks". *The Review of Economic Studies* 83(2): 514-546.

[32] Franzen, A. & Hangartner D. (2006). "Social Networks and labor Market Outcomes: The Non-Monetary Benefits of Social Capital". *European Sociological Review* 22(4): 353-368.

[33] Friedkin, N. (1980). "A test of structural features of Granovetter's strength of weak ties theory". *Social networks* 2(4): 411-422.

[34] Galeotti, A., and Merlino, L. P. (2014). "Endogenous job contact networks". *International Economic Review* 55(4): 1201-1226.

[35] Gee, L. K., Jones, J., & Burke, M. (2017). "Social Networks and labor markets: How strong ties relate to job finding on Facebook's social network". *Journal of Labor Economics* 35(2): 485-518.

[36] Goyal S. (2007). *Connections: an introduction to the economics of networks*. Princeton University Press.

[37] Granovetter, M. (1973). "The strength of weak ties". *American Journal of Sociology* 78(6): 1360-1380.

[38] Granovetter, M. (2005). "The impact of social structure on economic outcomes". *Journal of Economic Perspectives* 19(1): 33-50.

[39] Granovetter, M. (2018). *Getting a job: A study of contacts and careers*. University of Chicago Press.

[40] Hansen, M. T. (1999). "The search-transfer problem: The role of weak ties in sharing knowledge across organization subunits". *Administrative Science Quarterly* 44(1): 82-111.

[41] Ioannides, Y. M., & Datcher Loury, L. (2004). "Job information networks, neighborhood effects, and inequality". *Journal of Economic Literature* 42(4): 1056-1093.

[42] Jackson, M. O. (2010). *Social and economic networks*. Princeton University Press.

[43] Jackson, M. O., Rodríguez-Barraquer, T. & Tan, X. (2012). "Social capital and social quilts: Network patterns of favor exchange". *The American Economic Review* 102(5): 1857-1897.

[44] Jackson, M. O. & Rogers, B. W. (2007). "Meeting strangers and friends of friends: How random are social networks?". *The American Economic Review* 97(3): 890-915.

[45] Jackson, M. O., Rogers, B. W., & Zenou, Y. (2017). "The economic consequences of social-network structure". *Journal of Economic Literature* 55(1): 49-95.

[46] Killworth, P. & Bernard, H. (1974). "CATIJ: A new sociometric and its application to a prison living unit". *Human Organization* 33(4): 335-350.

[47] Kydland, F. E. (1995). "Business Cycles and Aggregate Labor Market Fluctuations". In *Frontiers of Business Cycle Research*, TF Cooley (Ed.).

[48] Lind, P. G., González, M. C. & Herrmann, H. J. (2005). "Cycles and clustering in bipartite networks". *Physical Review E* 72(5): 056127.

[49] Lippert, S., & Spagnolo, G. (2011). "Networks of relations and word-of-mouth communication". *Games and Economic Behavior* 72(1): 202-217.

[50] Louch, H. (2000). "Personal network integration: transitivity and homophily in strong-tie relations". *Social Networks* 22(2000): 45-64.

[51] Maarten H. & Wolbers, J. (2000). "The Effects of Level of Education on Mobility between Employment and Unemployment in the Netherlands". *European Sociological Review* 16(2): 185-200.

[52] Marsden, P. V. & Campbell, K. E. (1984). "Measuring tie strength". *Social Forces* 63(2): 482-501.

[53] Marsden, P. V. & Campbell, K. E. (1990). "Recruitment and selection processes: The organizational side of job searches". In R. L. Breiger (Ed.), *Social mobility and social structure* (pp. 59-79). New York: Cambridge University Press.

[54] Marmaros, D. & Sacerdote, B. (2002). "Peer and Social Networks in Job Search". *European Economic Review* 46: 870-879.

[55] Milgrom, P. R., & Weber, R. J. (1982). "A theory of auctions and competitive bidding". Econometrica: *Journal of the Econometric Society*, 1089-1122.

[56] Montgomery, J. (1991). "Social Networks and Labor-Market Outcomes: Toward an Economic Analysis". *The American Economic Review* 81: 1408-1418.

[57] Montgomery, J. D. (1992). "Job search and network composition: Implications of the strength-of-weak-ties hypothesis". *American Sociological Review* 57(5): 586-596.

[58] Mortensen, D. T., & Pissarides, C. A. (1999). "Job reallocation, employment fluctuations and unemployment". *Handbook of Macroeconomics*, 1, 1171-1228.

[59] Munshi, K. (2003). "Networks in the modern economy: Mexican migrants in the US labor market". *The Quarterly Journal of Economics* 18(2), 549-599.

[60] Myers C. & Shultz G.P. (1951). *The dynamics of a labor market.* New York: Prentice Hall.

[61] Pallais, A., & Sands, E. G. (2016). "Why the referential treatment? Evidence from field experiments on referrals". *Journal of Political Economy* 124(6): 1793-1828.

[62] Pissarides, C. A. (1985). "Short-run equilibrium dynamics of unemployment, vacancies, and real wages". *The American Economic Review* 75(4): 676-690.

[63] Pellizzari, M. (2010). "Do friends and relatives really help in getting a good job?" *Industrial and Labor Relations Review* 63 (3): 494-510.

[64] Ruef, M. (2002). "Strong ties, weak ties and islands: structural and cultural predictors of organizational innovation". *Industrial and Corporate Change.* 11(3): 427-449.

[65] Topa, G. (2001). "Social interactions, local spillovers and unemployment". *The Review of Economic Studies* 68(2): 261-295.

[66] Vega-Redondo, F. (2007). *Complex social networks*. Cambridge University Press.

[67] Weisstein, Eric W. "Vertex-Transitive Graph." From MathWorld–A Wolfram Web Resource. Retreived from: https://mathworld.wolfram.com/Vertex-TransitiveGraph.html.

[68] Young, H. P. (1993). "The evolution of conventions". *Econometrica* 61(1): 57-84.

7 Appendix

All proofs can be found in the following link, by clicking on "Are close-knit networks good for employment?":

https://sites.google.com/view/sofiaruizpalazuelos/research