



Paper to be presented at the Summer Conference 2010

on

"Opening Up Innovation:
Strategy, Organization and Technology"

at

Imperial College London Business School, June 16 - 18, 2010

NETWORKS OF INVENTORS

Michael David König

ETH Zurich
mkoenig@ethz.ch

Moritz Müller

ETH Zurich
m-mueller@ethz.ch

Christian Pich

ETH Zurich
cpich@ethz.ch

Sidonia von Proff

University of Marburg
vonproff@uni-marburg.de

Abstract:
The paper at hand presents a micro-founded network model which reproduces macro characteristics that correspond well with the real-world inventor network of US patents between 1975 and 2009. The focus lies on explaining how prevalent local (among co-inventors of co-inventors) and global (random and preferential) search strategies for collaborators are. The relationship between the two strategies has changed during the last decades, as the empirical data reveal. In particular, we find that links between inventors are established mainly by performing a local search and this effect has become more prevalent over time.

Networks of Inventors

Michael D. König^a, Moritz Müller^a, Christian Pich^a, Sidonia von Proff^{a,b}

^a*Chair of Systems Design, Department of Management, Technology and Economics,
ETH Zurich, Kreuzplatz 5, CH-8032 Zurich, Switzerland*

^b*Chair of Economic Geography and Location Research, Department of Geography,
Philipps-University Marburg, Deutschhausstrasse 10, D-35032 Marburg, Germany*

Abstract

The paper at hand presents a micro-founded network model which reproduces macro characteristics that correspond well with the real-world inventor network of US patents between 1975 and 2009. Our focus lies on explaining how prevalent local (among co-inventors of co-inventors) and global search strategies for collaborators are. The relationship between the two strategies has changed during the last decades, as the empirical data reveal. In particular, we find that links between inventors are established mainly by performing a local search and this effect has become more prevalent over time.

Key words: innovation networks, network formation

JEL: C63, D83, D85, L22

1. Introduction

The analysis of social networks, established by social scientists,¹ has early been applied to the study of innovation diffusion [Abrahamson and Rosenkopf, 1997; Granovetter, 1978]. Social network analysis can answer questions on the breadth and speed of certain innovations. Starting one step earlier, on the level of inventions, similar research questions arise: where does the knowledge incorporated in inventions come from? Different pieces of knowledge provide the basis for new combinations of ideas, as Schumpeter called innovations, and they are put together from teams of inventors. At the same time, innovating becomes more and more interdisciplinary and the team sizes of patents as well as publications increase [Wuchty et al., 2007].

Inventor networks are social networks and the same tools for their analysis as for other social network phenomena can be applied here. Since patents

Email addresses: mkoenig@ethz.ch (Michael D. König),
mueller@cournot.u-strasbg.fr (Moritz Müller), cpich@ethz.ch (Christian Pich),
vonproff@staff.uni-marburg.de (Sidonia von Proff)

¹See Wasserman and Faust [1994] for an introduction and overview.

are a prerequisite for many types of innovations, patent network analysis is an adequate (albeit imperfect) approach to investigate innovation collaboration. Patent documents display plenty of information and they are relatively easy to access. By investigating who invents with whom (co-inventors on each patent) a collaboration network can be constructed.

The resulting network is a snapshot of collaborations at the end of a certain time period. The question of how the network has evolved to a certain structure have to be answered by using both, theoretical and empirical methods. Models of network evolution are the theoretical background for this purpose. A few attempts have been made to compare the outcome networks generated by the theoretical models with an empirical network, mostly by means of a simulation techniques [see e.g. [Toivonen et al., 2009](#), for an analysis of various social network models]. However, existing research restricts most often to either descriptive statistics of real-world networks [e.g. [Gay and Dousset, 2005](#)], or theoretical studies with or without a simulation analysis [e.g. [Marsili et al., 2004](#)]. Some studies go deeper, but use a regional limited database [cf. [Fleming et al., 2007](#)].

The paper at hand takes into account the three strands of analysis: empiricism, theory, and simulation. We use the large database of US patents provided by the Harvard Dataverse Network Project in 2009 [[Lai et al., 2009](#)]. This database extends the NBER patent database [[Hall et al., 2001](#)]. Only a few studies are based on these data up to now. The theoretical model that we develop builds on previous ones in the literature, most notably the growing network model of [Jackson and Rogers \[2007\]](#), where local and global search mechanism are combined, the preferential attachment model by [Barabasi and Albert \[1999\]](#), and the local search model of [Ehrhardt et al. \[2006\]](#); [Marsili et al. \[2004\]](#) for a network with a fixed, predefined size. This new model generates networks which are astonishingly consistent with the empirical inventor network, generated from the patent data. This indicates that the mechanisms underlying the formation of the inventor network from patents filed in the US seem to be consistent with the model.

Caveats are that patent networks represent only a certain part of a collaboration network (depending on the possibility and the wish to protect an invention, which are industry-specific) and that patent documents are not exact, i.e. may have mistakes and the inventors listed have to be matched with real persons. Nevertheless, important insights can be gained by analyzing them and the paper at hand contributes to a small but increasing strand of research which focuses on the quantitative side of innovation network analysis.

The remainder of the paper is structured as follows. In [Section 2](#), the relevant state of the art research is summarized. [Section 3](#) contains the theoretical model and discusses its assumptions. The empirical analysis of the model is presented in [Section 4](#). [Section 4.1](#) describes the US patent data and the corresponding collaboration network. [Section 4.2](#) contains the

estimation of the model parameters. A discussion of the results follows in Section 4.3, before Section 5 concludes.

2. Background

The importance of collaboration in R&D activities and, within that, the necessity to have face-to-face contacts, is known from innovation economics literature [cf. Sorenson, 2003; Storper and Venables, 2004]. Co-inventors spend quite a lot of time together and have a great chance to exchange knowledge exceeding the current research project. This is supported by existing patent literature that found significant information flows between co-inventors [cf. Fleming and Frenken, 2007]. Since individual persons are dispersed in (physical and social) space the influence of distance on collaboration is of interest and how individuals find their collaboration partners.

There are different measures of distance such as social, geographical, or cognitive distance [cf. Boschma, 2005]. Several previous studies focused on explaining citation patterns with distance measures [cf. Breschi and Lissoni, 2005; Jaffe et al., 1993; Singh, 2005]. For example, Breschi and Lissoni [2005] analyzed empirically an Italian inventor network by looking how the social network (co-inventorship) corresponds to citation patterns. They found that indirect social connections help to overcome geographical distance for collaborations. Moreover, collaboration and, thus, the participation in an R&D network affect the performance of companies. Therefore it is of interest, how the connections are established. The partner must have the right competences and some trust is required in order to cope with the uncertainty of R&D projects. It is impossible to get perfect information about potential partners and the search strategy must keep track of search costs. Ties are often repeated when they have been successful and the next step is to search among the partners' partners. In addition, there is a random part in the search. Collaborators find each other by meeting on conferences or fairs, by random personal acquaintances, by job mobility between companies, and more. The result is a patent network with small world properties: high clustering and rather short average distance. These findings are supported by the works of Newman [2001, 2004] when analyzing scientific co-authorship networks and Fleming and Frenken [2007]; Fleming et al. [2007]; Fleming and Marx [2006] analyzing patent data.

The question is now, which search strategy is more effective for and more prevalent in innovation processes: the search among partners' partners or a global search mechanism. Among others, Burt [2004] argues that having structural holes (i.e. unconnected neighbors) in the network is the key point to exchanging new information between groups while being economical. If there are no links to far-distant parts of a network but only local links, then the risk of a lock-in increases. Proponents of the social capital theory in contrast argue that a densely connected network fosters efficient collaboration

[e.g. Walker et al., 1997]. The density of links in such clusters make knowledge exchange fast and easy. These two different points of view, structural holes vs. social capital, can be reconciled by attributing their importance to different stages in the industry life cycle. Consider the evolution of an R&D collaboration network [Cowan and Jonard, 2009]. In young industries, connecting different components and collecting new ideas from different parts of the network can be an advantage, while in a more mature industry stronger ties are fruitful for exploiting the technologies [cf. König, 2008, Chapter 1.9]. A similar concept is the notion of knowledge specialization versus knowledge brokerage, or exploitation versus exploration (see Carnabuci [2005], Chapter 6). In an empirical study with US patents and patent citations, Carnabuci [2005] finds that both types of behavior exist and they are related to the stadium of the industry life cycle.

If the point of view is overarching and not industry-specific, there are industries in different phases of the industry life cycle and a co-inventor analysis gives information about the average collaboration behavior. Jackson and Rogers [2007] investigated how often links are globally formed in comparison to network-based (local) links. The resulting model exhibits several features common for social networks, namely short average distance between nodes in the network, high clustering, skewed degree distributions, assortativity, and a negative correlation of clustering with degree. The empirical analysis in Jackson and Rogers [2007] shows large differences between networks regarding the network-based share of link formation. Unsurprisingly, the link formation process of High School romances and co-authors of scientific publications differs. However, their model assumptions (on the microlevel of why links are formed) are not validated except for a comparison whether the overall model suits to the empirical data (on the macro-level of the aggregate network topology). Based on this model, Kovarik and van der Leij [2009] explain link formation with risk aversion, while Campbell [2008] does so with signaling willingness to cooperate. Overall, the literature on how inventors form links is remarkably sparse and especially lacking a dynamic view. To the authors' knowledge, there are no empirical studies on the background of link formation, which have to take into account psychological, sociological, and economic reasons at the same time. We argue that the importance of face-to-face contacts or spatial proximity (as explained above) should make the local search strategy more prevalent. Links to distant parts of the network are important but exist much less often.

Hypothesis 1. *Links between inventors are established mainly by performing a network-based search.*

However, the search strategies need not be constant over time. A static view neglects that a network is the outcome of interaction among the agents and evolves over time. Whenever separated components of a network become connected a lot of new indirect links emerge and network properties

will change. Thus, the underlying link formation mechanism needs further investigation. The challenge in link formation research is that there exist on the one hand microeconomic concepts, which can hardly be tested and on the other hand there are physics-based statistical models which are supported by empirics but which have no real behavioral foundation. There are a few attempts to close this research gap and the paper at hand contributes to that line of research.

[Fleming and Frenken \[2007\]](#) compared two patent networks, which differed in their small world characteristics. In a static-comparative analysis they investigate the development of the networks over time. The inventor network of the Boston region had a much smaller largest component at the end than the inventor networks in the Silicon Valley even though they had comparable basic conditions. They found IBM to be the central actor in the Silicon Valley linked to other companies and universities by former employees. Thus, former employees of a central company have established the critical links between subnetworks in the Silicon Valley network. Job mobility has been an important part of the link formation process.

If the networks-based search was so important for the development of a large network component in Silicon Valley the question arises whether local search has become more important during the last decades. Two forces pull in opposite directions. On one hand, improved information and communication technologies facilitate cooperation over distance. At the same time, passenger transportation costs have decreased and air travel connections have improved. However, there are several studies available that cannot find an increase in innovation cooperation over distance [cf. [Dettmann and von Proff, 2008](#); [Howells, 1995](#)]. On the other hand, globalization and the internet can lead to an information overload. For example, the amount of scientific publications makes it impossible to be aware of all specialists in your own field of research. In order to cope with the abundant information it is most likely that individuals rely on a recommender system in the way that they ask among their colleagues about potential collaborators. This would lead to an increase in local search, i.e. link formation among colleagues' colleagues. We argue that the latter effect of local search is dominating the first effect of global search for innovation collaborations, and even more so, as the knowledge based economy generates an ever growing knowledge base. In line with this point of view, we propose the following hypothesis:

Hypothesis 2. *Network-based search has become more prevalent during the last decades.*

In the following sections we introduce a formal model of how collaboration networks of inventors are formed and then proceed by estimating this model using the one-mode projection network of inventors generated from

USPTO patent data.² These estimates and their significance will allow us to test the above mentioned hypotheses.

3. The Model

In this section we introduce a model for the process in which collaborations in an inventor network are formed. Building on the discussion in the previous section, this model explicitly takes into account the fact that an inventor can find co-inventors through global and local search in the existing network of inventors.

More formally, we consider the following network formation process $(G(s))_{s=0}^S$ for some $S > 0$. $G(0)$ is the initial network. At every step $s = 1, 2, \dots, S$ links are created either (i) by a new inventor that enters the network with probability $p \in [0, 1]$ or (ii) with probability $1 - p$ by an incumbent inventor already existing in the network $G(s) = (N(s), E(s))$, consisting of a set of $N(s)$ inventors and a set $E(s)$ of edges between them.

- (i) With probability p a new inventor enters. The entering inventor $i \notin N(s)$ selects $m_g \geq 0$ inventors from the network $G(s)$ at random. We denote the set of these inventors by $N_i^{G,E}(s)$, with $|N_i^{G,E}(s)| = m_g$. For each inventor $j \in N_i^{G,E}(s)$ a link is created with probability

$$\begin{aligned} p_{G,E} &= \mathbb{P}\left(G(s+1) = G(s) + ij | G(s), j \in N_i^{G,E}(s)\right) \\ &= \frac{e^{\alpha_1^E + \alpha_2^E d_j(s) + \alpha_3^E t}}{1 + e^{\alpha_1^E + \alpha_2^E d_j(s) + \alpha_3^E t}}. \end{aligned} \quad (1)$$

Here $t = 1, \dots, T$ corresponds to a slowly varying time frame that is constant over the steps $s = 1, 2, \dots, S$. Equation (1) captures a *global search* mechanism. The entering inventor i then selects $m_L \geq 0$ inventors from the union of the neighborhoods $\bigcup_{j \in N_i^{G,E}(s)} N_j(s)$ of the inventors $j \in N_i^{G,E}(s)$ uniformly at random. Let us denote the set of these inventors by $N_i^{L,E}(s)$, with $|N_i^{L,E}(s)| = m_L$. For each inventor $k \in N_i^{L,E}(s)$ a link is formed with probability

$$\begin{aligned} p_{L,E} &= \mathbb{P}\left(G(s+1) = G(s) + ik | G(s), k \in N_i^{L,E}(s)\right) \\ &= \frac{e^{\beta_1^E + \beta_2^E d_k(s) + \beta_3^E t}}{1 + e^{\beta_1^E + \beta_2^E d_k(s) + \beta_3^E t}}. \end{aligned} \quad (2)$$

Equation (2) corresponds to a *local search* mechanism.

²The bipartite network of patents and inventors corresponds to a two-mode network and the projection of this network on the inventors is a one-mode network [Wasserman and Faust, 1994]. See Grewal et al. [2006] for a similar analysis.

- (ii) With probability $1 - p$, incumbent inventors, that have been added already to the network, can form links. We assume that an incumbent inventor $i \in N(s)$ is selected with probability $e^{\rho d_i(s)} / \sum_{j \in N(t)} e^{\rho d_j(s)}$, $\rho \geq 0$. This assumption considers that inventors with a large number of collaborations are more active and thus more likely to initiate a new collaboration. Inventor i then selects m_G inventors $l \in N(s) \setminus \{i\}$ uniform at random. Let the set of inventors selected in this way be $N_i^{G,I}(s)$, with $|N_i^{G,I}(s)| = m_G$. For each inventor $l \in N_i^{G,I}(s)$ a link is created with probability

$$\begin{aligned} p_{G,I} &= \mathbb{P}\left(G(s+1) = G(s) + il \mid G(s), l \in N_i^{G,I}(s)\right) \\ &= \frac{e^{\alpha_1^I + \alpha_2^I d_i(s) + \alpha_3^I t}}{1 + e^{\alpha_1^I + \alpha_2^I d_i(s) + \alpha_3^I t}}. \end{aligned} \quad (3)$$

Equation (3) corresponds to a *global search* mechanism. Moreover, inventor i selects m_L inventors in her second-order neighborhood $N_i^{(2)}(s)$. Let the set of these inventors be $N_i^{L,I}(s) \subseteq N_i^{(2)}(s) \cup \bigcup_{j \in N_i^{G,I}(s)} N_j(s)$, such that $|N_i^{L,I}(s)| = m_L$. For each $k \in N_i^{L,I}(s)$ a link is created with probability

$$\begin{aligned} p_{L,I} &= \mathbb{P}\left(G(s+1) = G(s) + ik \mid G(s), k \in N_i^{L,I}(s)\right) \\ &= \frac{e^{\beta_1^I + \beta_2^I d_k(s) + \beta_3^I t}}{1 + e^{\beta_1^I + \beta_2^I d_k(s) + \beta_3^I t}}. \end{aligned} \quad (4)$$

Equation (4) corresponds to a *local search* mechanism. Through this local search mechanism a total of m_L links are created in expectation.

An illustration of the above mechanism of link formation can be seen in Figure 1. It is interesting to note that for the case of $p = 1$, $\alpha_2^E = \alpha_3^E = \beta_2^E = \beta_3^E = 0$ we obtain the model of Jackson and Rogers [2007]. Moreover, in the case of $p = 0$ the model of Ehrhardt et al. [2006]; Marsili et al. [2004] can be nested into our framework. We thus propose a general model of network formation into which previous prominent models in the literature can be casted.

4. Data Analysis

In the following sections we estimate the model that we have introduced in the previous section. In Section 4.1 the data set is presented and a number of descriptive statistics are shown and discussed. Next, in Section 4.2 the estimation strategy is introduced. The estimation results are then presented in Section 4.3.

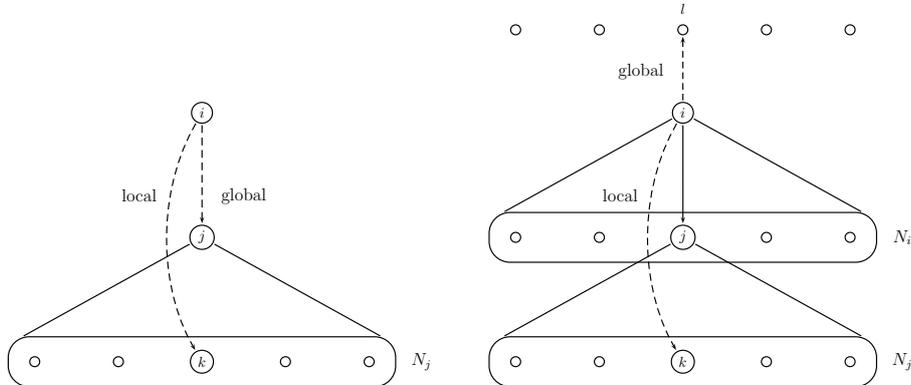


Figure 1: Illustration of the link formation process. (Left) A link is created by an entering inventor i through global search to inventor $j \in N_i^{G,E}$. Moreover, a link is created to a neighboring inventor $k \in N_i^{L,E}$ of j , $k \in N_j$, via local search. (Right) A link is created by an incumbent inventor i to an inventor $l \in N_i^{G,I}$ through global search. Moreover, a link is created to a second-order neighbor $k \in N_i^{L,I}$ of i , i.e. $k \in N_i^{(2)}$ with $k \in N_j$ and $j \in N_i$, via local search.

4.1. Sample

We have analyzed the USPTO patent data from 1975 to 2009 as provided by Lai et al. [2009]. The network of inventors has been constructed by creating a link between any pair of inventors that has appeared together on a patent. The resulting network is undirected.

Some summary statistics can be seen in Figure 2. We have focused on four specific ones, since these have been used in characterizing the topological properties of networks in a wide range of related areas [Jackson, 2008]. In the figure we show the degree distribution $p(d)$, the clustering-degree correlation $C(d)$, average nearest neighbor connectivity $d_{nn}(d)$ and component size distribution $p(s)$ [Wasserman and Faust, 1994]. The degree distribution $p(d)$ exhibits a power law tail showing that the network of inventors is highly asymmetric with only a few inventors having a large number of collaborations. We also see that the average degree is increasing due to the addition of links to the network over time. The clustering coefficient $C(d)$, measuring the tendency of co-inventors to have appeared on a patent together, is relatively high and exhibits a decreasing trend with increasing degree. This indicates the fact that highly connected inventors tend to be connected, on average, with inventors who have not collaborated with each other. The nearest neighbor connectivity $d_{nn}(d)$ is defined as the average degree of the neighbors of an inventor with degree d [Pastor-Satorras et al., 2001]. It turns out that inventor networks are assortative, that is, high de-

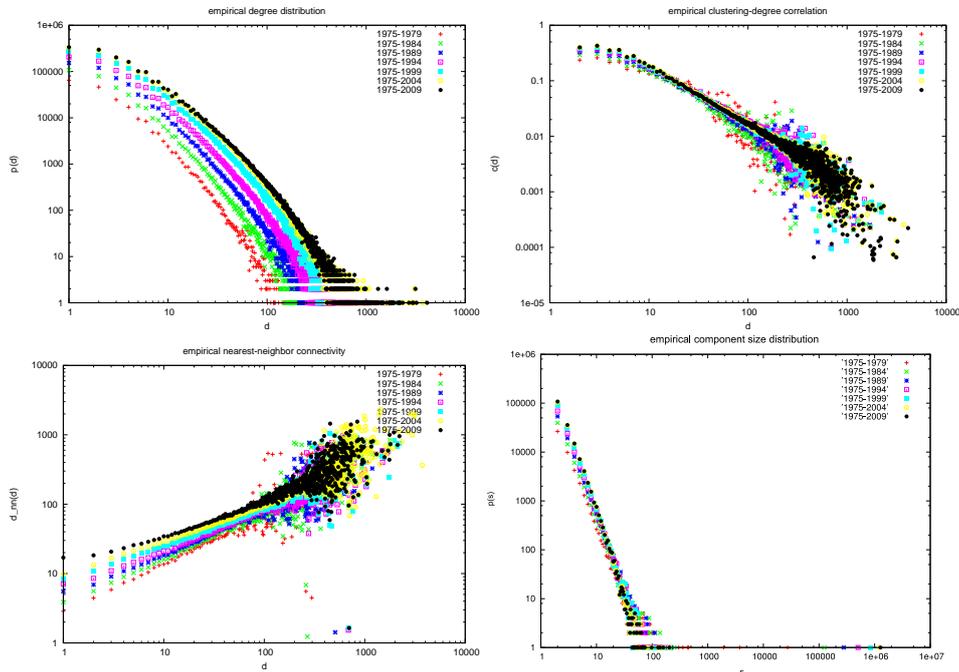


Figure 2: Empirical degree distribution $p(d)$, clustering-degree correlation $C(d)$, average nearest neighbor connectivity $d_{nn}(d)$ and component size distribution $p(s)$ constructed from the US patent network. All network statistics seem to follow a stable power law.

gree inventors are connected to other high degree inventors while low degree inventors maintain collaborations with other low degree inventors [Newman, 2004]. Finally, there exists a large connected component and a number of smaller components in the inventor network. The distribution of the sizes of the smaller connected components can well be described with a power law. It is interesting to note from these descriptive statistics, that over a wide range of values, the distributions can all be described by power laws. Moreover the power law behavior of the distributions seems to be stable over time.

For the purpose of parameter estimation we extracted only those patents which fall into the USPTO patent category chemicals. In this way we retrieved 397018 from a total of 3684228 patents in all categories available. The network of inventors in this class comprises of 294063 nodes and 532367 links.

4.2. Estimation Strategy

The model we have introduced in Section 3 describes the formation of a network of inventors by to the creation of bilateral relations of entrants as well as incumbents. Unfortunately, we do not observe the actual process of

how inventors get to know each other and form collaborative ties over time. However, what we do observe is the outcome of this latent partner search process in terms of the patents filed at the patent office, and from these patents, the patent collaboration network after a certain period of time. We take this collaboration network as a snapshot of the current co-patenting network of inventors and compare it with the outcome network of our network formation process. Therefore, we indicate the path of network growth by a series of snapshots of the observed co-patenting network. Because the model describes the evolution of the network as a Markov chain, it yields expectations on a network in one time step, given a network in the previous time step. Thus, our unit of observation is the change occurring from one network snapshot to the next. The objective is to obtain an estimate for which the model generates a network growth path which is close to the observed one.

We obtain this estimate by the Generalized Method of Moments (GMM). Our approach is strongly inspired by [Snijders, 2001], who also estimates a Markov chain model of network formation on a series of network snapshots. However, one difference is that we directly estimate our discrete Markov chain model of network formation whereas [Snijders, 2001] makes his continuous. Whether the model is continuous or discrete does not make a fundamental difference. More important is the common Markov chain assumption that the network forms in a series of subsequent elementary events and only the last state of the network formation history matters for the next step. This assumption is crucial for estimation because it implies that a process of network formation which spans a longer time period can be perceived as a sequence of shorter, independent network formation processes. This way a longitudinal sample of a network can be perceived as a collection of several, independent observations. This allows for estimating the models using the Method of Moments (MoM), as in [Snijders, 2001], or the GMM, as in our work.

Both the GMM and MoM yield estimates such that the population moments implied by the model correspond to the observed sample moments. In the case at hand the sample moments are changes of the co-patenting network statistics from one time step to the other and the population moments are the predictions of these changes by the model. For example, one moment used in our estimation are changes of the degree distribution. In order to identify the parameters of the model, one needs to specify at least as many moment conditions as parameters. The difference between MoM and GMM is that for MoM the number of moment conditions equals the number of parameters whereas GMM allows for more moment conditions than there are parameters. We adopt the GMM as it allows to incorporate moment conditions on more network aspects and implications of the model.

The GMM estimator $\hat{\Theta}$ minimizes an objective function $Q(\Theta)$ which is a weighted sum of moment functions \mathbf{h}_t over all observations $t = 1, \dots, T$,

i.e.

$$\hat{\Theta} = \underset{\Theta}{\operatorname{argmax}} Q_M(\Theta) = \underset{\Theta}{\operatorname{argmax}} \left[\frac{1}{T} \sum_{t=1}^T \mathbf{h}_t \right]' \mathbf{W} \left[\frac{1}{T} \sum_{t=1}^T \mathbf{h}_t \right], \quad (5)$$

where \mathbf{W} is the weighting matrix. Each moment vector function \mathbf{h}_t , relates one observation to the model parameters w.r.t. the specified moments. The idea that the sample moments equal the population moments finds its expression in the assumption that the expectation of the moment functions is zero ($E[\mathbf{h}_t] = 0$).

Our sample is a series of network snapshots ($G(0), G(1), \dots, G(t) \dots G(T)$) and in this sample, the principle unit of observation is the change occurring from one network, $G(t-1)$, to the next network $G(t)$. Changes of the corresponding network statistics ($\Delta S_t = S_t - S_{t-1}$) describe *how* the network changes and, therefore, are appropriate moments for the estimation. The estimation is conditional on the number of entrants. Therefore, given the prior network $G(t-1)$ and parameter estimates $\hat{\Theta}$, the model yields expected network statistics, \hat{S}_t , of a network as large as the observed network in terms of size [see also [Snijders, 2001](#)]. The expected network statistics are used to indicate the expected change (i.e. $\Delta \hat{S}_t = \hat{S}_t - S_{t-1}$). The prediction error of the model is then simply the difference of the expected and the observed change statistics ($u_t = \Delta \hat{S}_t - \Delta S_t$). In our estimation, prediction errors enter in two ways. Firstly, the error is minimized by introducing the squared error as moment function ($h_t = u_t^2$). Secondly, as the model describes a Markov chain in which errors are uncorrelated over time, a further set of moment functions is obtained by multiplying each error with its lag ($h_t = u_t u_{t-1}$). Preventing correlation of lagged errors is advantageous because it further releases the effect of time on the modes of partner choice.

Network characteristics of interest have been discussed in the previous subsection. The shape of the degree distribution, clustering distribution, and nearest neighbor connectivity are all likely to affect many network processes; most notably the diffusion and accumulation of knowledge [[Rogers, 1995](#)]. Our model closely relates to these distributions because the specific relation of global and local partner search in combination with preferential attachment is able to generate specific shapes of the distributions. Technically, this is important for identification of the model parameters because the expected moments need to vary with the parameters. The remaining issue is to choose those statistics which indicate how the network grows from one instance to the next. Because the statistics should be roughly on the same level for estimation, we replace the clustering coefficient by the number of triangles and the nearest neighbor connectivity by the size of the 2nd order neighborhood. For each network measure an empirical and a theoretical (expected) histogram is constructed which displays the average change of a network measure by degree classes. Each column of the histograms represents one moment and one error is the difference between a column of

the empirical and the expected histogram. Because the degree distribution is highly skewed, the binning of the histograms is logarithmic. This has the effect that the bins become larger for higher degrees and, thereby, each bin includes a comparable number of nodes. The estimation results presented in the next section are based on five bins which include nodes of degree (0), (1, 2), (3, ..., 7), (8 ... 20), and (20 ... max).

Stochastic approximation. The GMM estimate $\hat{\Theta}$ is found by minimizing the quality function $Q_M(\Theta)$, as described in Equation 5. In order to actually calculate the value of $Q_M(\Theta)$, one would need to evaluate the moment functions, \mathbf{h}_t . These moment functions involve expected network statistics $\mathbb{E}[S(G(t))]$, given the parameter vector Θ and the observed network in the prior period $G(t-1)$, i.e. $\mathbf{h}_t = f(\mathbb{E}[S(G(t))|\Theta, G(t-1)])$. Because there exists no closed-form solution for $\mathbb{E}[S(G(t))]$, we obtain the expectation using a Monte Carlo simulation of the Markov chain described by the model in Section 3. We then apply an appropriate search algorithm in the parameter space of Θ in order to find the minimum of the quality function $Q_M(\Theta)$.

Note that the quality function $Q_M(\Theta)$ generated by Monte Carlo simulations is itself a random variable. The average over an infinite number of simulations would be needed to compute the actual expectation of $Q_M(\Theta)$. However, since our network formation process is a Markov chain, we can use the fact that a realization (simulation) of the process comes arbitrarily close to the expected value if the network is large enough [see e.g. [Benaim and Weibull, 2003](#)]. The networks we are analyzing are large (see Section 4.1) and thus we can take the value of the quality function $Q_M(\Theta)$ of a single simulation for a given parameter set Θ as a sufficiently close approximation to the actual expected value. This makes our numerical estimation remarkably faster.

Since no exact solution of $\mathbb{E}[S(G_t)]$ is available, the quality function $Q_M(\Theta)$ is stochastic and we need to rely on Monte Carlo simulations to find its minimum using stochastic approximation techniques [[Pflug, 1996](#)]. Stochastic approximation provides the toolbox for optimizing stochastic functions, that is functions including a random component as it is given by the quality function $Q_M(\Theta)$. In order to approximate the solution to equation 5, we rely on the gradient-free stochastic perturbation algorithm proposed by [Spall \[2003, ff. 150\]](#). This method is similar to a steepest descent method such as for example the Newton-Raphson algorithm with fast convergence in stochastic settings. For a general overview on stochastic approximation see also [Pflug \[1996\]](#).

4.3. Results

Estimation restrictions. Preliminary results have been obtained for a restricted sample as well as a restricted set of parameters. The restricted

sample is based on all chemical patents³ filed by the USPTO with two or more inventors. The co-patenting network is accumulated from 1975 to 1995. The subsequent period from 1995 to 2000 is used for estimation which yields 5 time periods. This sample is used to obtain first estimates of the time effect on local and global network search. However, assuming that the decision making of entrants and incumbents does not differ, we force the parameters to be equal for both. Thus, for Equations (1) to (4) we set $\alpha_k = \alpha_k^E = \alpha_k^I$ and $\beta_k = \beta_k^E = \beta_k^I$, for $k = 1, 2, 3$. Furthermore, the parameter ρ , which models the probability of selecting an incumbent to initialize collaborations, is fixed at a value of 1. This choice leads to selection probabilities being proportional to the degree of the inventor. Finally, in order to reduce computational load, we chose to make the sets of potential collaborators small and set $m = m_L = m_G = 10$. In the near future, the parameters which model the decision of the inventors (i.e. the parameters α 's, β 's, and ρ) are going to be estimated separately and the sets of potential collaborators (i.e. the m_G, m_L) are going to be increased in size until their effect on the simulation results becomes zero.

Preliminary results. The results give a clear indication in favor of the hypotheses that local search becomes more prevalent over time. Table 4.3 presents the estimation results for two models. In both models we observe a significant negative time trend for global attachment (α_3) and a positive time trend for local attachment (β_3) (significant in model A). Considering the standard errors, in both models the local and global time trends are clearly different from each other. Local search increases relative to global search over time.

The estimates also reveal theoretical insights on models of network formation. In model A, the preferential attachment effect is set to zero and in model B it is left free for estimation. For interpretation of the negative sign of the preferential attachment coefficient, consider that incumbents are chosen to initialize a collaboration proportional to their degree. Therefore, a negative sign does not necessarily imply that there is no preferential attachment effect in model B. It rather suggests that selection proportional to degree is too strong. Nevertheless, we clearly see that the introduction of a preferential attachment effect has a very strong effect on the probability that an entry step occurs. This means that in models of network growth the entry of new nodes in the network and preferential attachment among incumbents potentially trade off and are likely to generate similar network structures. Actually both models yield predictions which fit very well the empirical distribution as shown in figure 3.

³Patents in main class one of the US patent classification

Table 1: Estimation of the time effect on global (α_k) and local (β_k) search moderated by preferential attachment

		Model A	Model B
Intercept	(α_1)	-1.160*** (0.038)	-1.048*** (0.057)
Pref. Attach.	(α_2)	!0	-0.470*** (0.051)
Time	(α_3)	-0.163*** (0.037)	-0.282*** (0.027)
Intercept	(β_1)	-1.322*** (0.208)	-1.519*** (0.137)
Pref. Attach.	(β_2)	!0	-0.440*** (0.064)
Time	(β_3)	0.602*** (0.135)	0.256 (0.204)
Entry step	(p)	1.000*** (0.127)	0.362** (0.178)
No. of time periods T		5	5
Quality function Q		24	20

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$ result from two-tailed z-tests of coefficients being zero. If '!'0', parameter is set to zero.

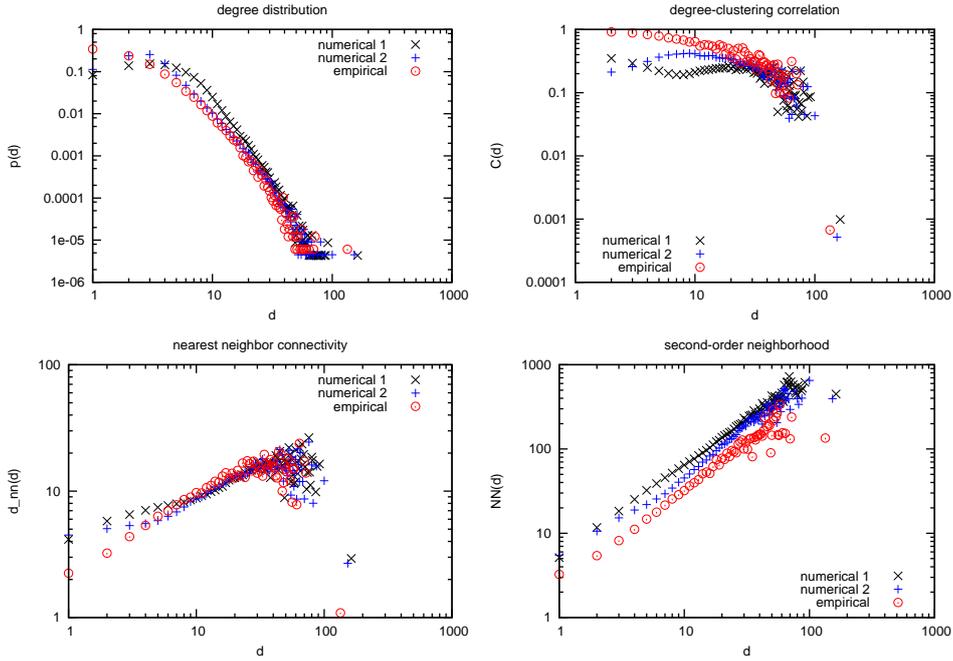


Figure 3: Comparison of the empirical statistics for the degree distribution $p(d)$, clustering-degree correlation $C(d)$, average nearest neighbor connectivity $d_{nn}(d)$ and size of the second-order neighborhood $N^{(2)}$. The empirical distributions correspond to the inventor network which evolved from 1995 to 2000. The theoretical distributions correspond to numerical simulations with the estimated parameter sets for Model A (without preferential attachment effect) and Model B (with preferential attachment) for the same time period.

5. Discussion and Conclusion

In this paper we introduced an extension of the network growth model of Jackson and Rogers [2007]. Our model allows for link formation among incumbents and models explicitly the agent's collaboration decisions. The introduction of a preferential attachment effect in the collaboration decision of the agents shows that entries and link formation among incumbents are capable of affecting the shape of the distributions of the network statistic in similar ways. The estimates of the model indicate that the evolution of inventors networks probably needs to be explained by a combination of both. Most importantly, this working paper provides a first intuition on how the relative importance of local and global search changes over time. Whereas the importance of global search seems to decrease, local network search seems to become more important. In future work we will investigate in more detail the temporal change in the parameters governing the influence of local vs. global search.

References

- Abrahamson, E. and Rosenkopf, L. (1997). Social network effects on the extent of innovation diffusion: A computer simulation. *Organization Science*, 8(3):289–309.
- Barabasi, A.-L. and Albert, R. (1999). Emergence of scaling in random networks. *Science*, 286:509–512.
- Benaim, M. and Weibull, J. (2003). Deterministic approximation of stochastic evolution in games. *Econometrica*, pages 873–903.
- Boschma, R. (2005). Proximity and innovation: A critical assessment. *Regional Studies*, 39(1):61–74.
- Breschi, S. and Lissoni, F. (2005). *Handbook of Quantitative Science and Technology Research*, chapter Knowledge Networks from Patent Data, pages 613–643. Springer Netherlands.
- Burt, R. (2004). Structural Holes and Good Ideas. *American Journal of Sociology*, 110(2):349–399.
- Campbell, A. (2008). Signaling in social networks. mimeo.
- Carnabuci, G. (2005). *A Theory of Knowledge Growth*. Universiteit van Amsterdam. Dissertation.
- Cowan, R. and Jonard, N. (2009). *Innovation Networks*, chapter Structural Holes, Innovation and the Distribution of Ideas, pages 127–144. Springer.
- Dettmann, A. and von Proff, S. (2008). Inventor collaboration over distance a comparison of academic and corporate patents. mimeo.
- Ehrhardt, G., Marsili, M., and Vega-Redondo, F. (2006). Diffusion and growth in an evolving network. *International Journal of Game Theory*, 34(3).
- Fleming, L. and Frenken, K. (2007). The evolution of inventor networks in the silicon valley and boston regions. *Advances in Complex Systems*, 10(1):53–71.
- Fleming, L., King, C., and Juda, Adam, I. (2007). Small worlds and regional innovation. *Organization Science*, 18(6):938–954.
- Fleming, L. and Marx, M. (2006). Managing creativity in small worlds. *California Management Review*, 48(4):6–27.
- Gay, B. and Dousset, B. (2005). Innovation and network structural dynamics: Study of the alliance network of a major sector of the biotechnology industry. *Research Policy*, pages 1457–1475.
- Granovetter, M. (1978). Threshold models of collective behavior. *American Journal of Sociology*, 83(6):1420.
- Grewal, R., Lilien, G. L., and Mallapragada, G. (2006). Location, location, location: How network embeddedness affects project success in open source systems. *Management Science*, 52(7):1043–1056.

- Hall, B., Jaffe, A., and Trajtenberg, M. (2001). The NBER Patent Citation Data File: Lessons, Insights and Methodological Tools. *NBER Working Paper*.
- Howells, J. R. (1995). Going global: The use of ict networks in research and development. *Research Policy*, 24(2):169–184.
- Jackson, M. (2008). *Social and Economic Networks*. Princeton University Press.
- Jackson, M. O. and Rogers, B. W. (2007). Meeting strangers and friends of friends: How random are social networks? *American Economic Review*, 97(3):890–915.
- Jaffe, A. B., Trajtenberg, M., and Henderson, R. (1993). Geographic localization of knowledge spillovers as evidenced by patent citations. *The Quarterly Journal of Economics*, 108(3):577–598.
- König, Michael, D. (2008). *Dynamic R&D Networks: the efficiency and evolution of interfirm collaboration networks*. PhD thesis, Department of Management, Technology and Economics, ETH Zurich, Switzerland.
- Kovarik, J. and van der Leij, M. J. (2009). Risk aversion and networks: Microfoundations for network formation. mimeo.
- Lai, R., D’Amour, A., and Fleming, L. (2009). The careers and co-authorship networks of U.S. patent-holders, since 1975. Harvard Business School, Harvard Institute for Quantitative Social Science.
- Marsili, M., Vega-Redondo, F., and Slanina, F. (2004). The rise and fall of a networked society: a formal model. *Proceedings of the National Academy of Sciences*, 101:1439–1442.
- Newman, M. E. J. (2001). The structure of scientific collaboration networks. *Proceedings of the National Academy of Sciences of the United States of America*, 98(2):404–409.
- Newman, M. E. J. (2004). Coauthorship networks and patterns of scientific collaboration. *Proceedings of the National Academy of Sciences of the United States of America*, 101(Suppl 1):5200–5205.
- Pastor-Satorras, R., Vazquez, A., and Vespignani, A. (2001). Dynamical and correlation properties of the internet. *Physical Review Letters*, 87.
- Pflug, G. (1996). *Optimization of stochastic models: the interface between simulation and optimization*. Kluwer Academic Publishers.
- Rogers, E. (1995). *Diffusion of innovations*. Free Press.
- Singh, J. (2005). Collaborative networks as determinants of knowledge diffusion patterns. *Management Science*, 51(5):756–770.
- Snijders, T. (2001). The Statistical Evaluation of Social Network Dynamics. *Sociological Methodology*, 31(1):361–395.
- Sorenson, O. (2003). Social networks and industrial geography. *Journal of Evolutionary Economics*, 13(5):513–527.
- Spall, J. (2003). *Introduction to stochastic search and optimization: estimation, simulation, and control*. John Wiley and Sons.
- Storper, M. and Venables, A. J. (2004). Buzz: face-to-face contact and the urban economy. *J Econ Geogr*, 4(4):351–370.
- Toivonen, R., Kovanen, L., Kivel, M., Onnela, J.-P., Saramki, J., and Kaski, K. (2009). A comparative study of social network models: Network evolution models and nodal attribute models. *Social Networks*, 31(4):240–254.
- Walker, G., Kogut, B., and Shan, W. (1997). Social capital, structural holes and the formation of an industry network. *Organization Science*, 8(2):109–125.
- Wasserman, S. and Faust, K. (1994). *Social Network Analysis: Methods and Applications*. Cambridge University Press.
- Wuchty, S., Jones, B. F., and Uzzi, B. (2007). The Increasing Dominance of Teams in Production of Knowledge. *Science*, 316(5827):1036–1039.