DE GRUYTER Rev Netw Econ 2020; 19(2): 85–114

Neil Gandal*, Nadav Kunievsky and Lee Branstetter

Network-Mediated Knowledge Spillovers in ICT/Information Security

https://doi.org/10.1515/rne-2020-0034
Received June 20, 2020; accepted October 15, 2020

Abstract: A large literature has used patent data to measure knowledge spillovers across inventions but few papers have explicitly measured the impact of the collaboration networks formed by inventors on the quality of invention. This paper develops a method to measure the impact of collaboration networks of inventors on invention quality. We apply this methodology to the information and communication technology (ICT) and information security sectors in Israel and find that the quality of Israeli inventions are systematically linked to the structure of the collaborative network in these sectors. We are very grateful to the editor Lukasz Grzybowski and an anonymous referee for very helpful comments and suggestions that significantly improved the paper. We thank the Maurice Falk Institute for Economic Research in Israel, Start-Up Nation Central, the U. S. National Science Foundation (SciSIP grants 1360165 and 1360170), and Portugal's Foundation for Science and Technology for financial support of this research. Lee Branstetter's work on this project was supported by the National Science Foundation and we thank Britta Glennon for excellent research assistance. We are also grateful to Tim Bresnahan, Eugene Kandel, Imke Reimers, and seminar/conference participants at the 19th CEPR IO conference, the 10th Paris conference on Digital Economics, Collegio Carlo Alberto, Hebrew University, Stanford University, Tel Aviv University, ad UC-Berkeley and for helpful comments and suggestions.

Keywords: ICT, information security, knowledge spillovers, networks, patent citations

***Corresponding author: Neil Gandal**, Tel Aviv University & CEPR, Tel Aviv-Yafo, Israel, E-mail: gandal@tauex.tau.ac.il. https://orcid.org/0000-0002-0031-569X
Nadav Kunievsky, University of Chicago, Chicago, Illinois, USA
Lee Branstetter, Carnegie Mellon University, Pittsburgh, Pennsylvania, USA

1 Introduction

High-tech R&D is typically done by teams. Working in teams necessarily involves exchanging ideas and sharing information. Participants of such research teams carry this knowledge to other teams and other projects in which they are involved or become involved, and knowledge can continue to flow between former collaborators even after they move across regions or to different firms and cease direct collaboration (Agrawal, Cockburn, and McHale 2006; Almeida, Song, and Wu 2001). The networks traced out by collaborations can become a key mechanism through which knowledge flows (Akcigit et al. 2018). Interestingly, though a great deal of research has focused on measuring knowledge spillovers in patents, over time and space, fewer papers have linked the structure of the networks formed by inventors' prior collaborations, and the knowledge spillovers that may flow through these networks, to the quality of patents.

Knowledge spillovers lie at the heart of modern theories of endogenous growth (Acemoglu 2009; Romer 1986, 1990), international trade (Branstetter and Saggi 2011; Grossman and Helpman 1991); international investment (Keller and Yeaple 2013), and economic development (Jones 2014). The late Zvi Griliches and several generations of his students, including Jaffe and Trajtenberg (2002), introduced a series of econometric techniques for empirically measuring the strength of these spillovers across time and space, using patents and patent citations. A large and growing literature has deployed these techniques across a wide range of technological domains, organizational categories, and countries, strongly affirming the existence and importance of knowledge spillovers.[1] Despite this extensive literature, the exact mechanisms through which knowledge spillovers are propagated, their relative importance in mediating these knowledge flows, and the effects of these spillovers on the quality of inventions remain imperfectly understood.

Some early research (Griliches 1979, 1992; Keller 1998) presumed that at least some spillovers might flow through contact in the marketplace with products or services embodying new technology. Other firms might reverse-engineer and build on this technology without ever forging any direct contact between their R&D engineers and those of the firm that created the original product. While this kind of spillover is certainly possible, in modern technology-intensive industries, spillovers are also likely to occur through more direct interaction between individuals

1 The empirical literature on knowledge spillovers is quite extensive, and we lack the space to review it fully. Scherer (1982), Jaffe (1986), Bernstein and Nadiri (1988), and Irwin and Klenow (1994) authored influential early studies, and Griliches (1992) provided a survey of early empirical work. Keller (2004) provides a review of the empirical literature focused on international knowledge spillovers, which is not the focus of the current paper.

who work together and exchange ideas and information. In this paper, we wish to examine whether the direct interaction among researchers increases knowledge spillovers and raises invention quality.

We examine this issue in the context of the information and communication technology (ICT)/information security (IS) industry in Israel. Israeli firms occupy an especially prominent role this sector. No rigorous statistical study has yet examined whether networks in this sector have actually played an important role in the rise of Israel as a center of ICT/IS innovation. In the paper, we use a simple model to examine the existence and importance of collaborator network-mediated knowledge spillovers. Like the other papers in this literature, we assume that the quality of a patented invention is closely related to its count of forward citations.

In order to apply the model, we have to address the issue that patent networks form sequentially and therefore play a dual role in expanding the number of citations received by a given patent. First, the denser and richer the patent network available to the inventors is at the time of the patent application, the more access the inventors have to useful knowledge obtained through their prior collaborations, both direct and indirect. This enhances the quality and value of invention *i*, and hence leads to more citations. We refer to this effect as the "ex-ante" knowledge spillover. After invention *i* is generated, the evolving network propagates knowledge of this useful invention (and the technical innovations it contains) to other inventor teams working on related technologies, leading to more citations over time. It is also possible that successful invention shapes the pattern of subsequent collaborations, as other inventors seek to partner with the creators of especially impactful inventions. In other words, invention quality could drive network density and richness rather than the other way around. We refer to these impacts as the "ex post" effect. We are primarily interested in the first, "ex ante" effect of network density on invention quality.

To make this point clear, the "ex-ante" network is the network that exists when the patent application is submitted. This means that in order to calculate the "ex-ante" network for each patent, we "remove" all patents that did not exist when the patent application was submitted and we calculate the network for the relevant patent exactly at that point in time. We do this patent by patent, which means that each patent has a different "ex ante" network. Logically, only patents that had been granted at the time that a particular patent is issued can provide a knowledge spillover to that particular patent. That is, the "ex ante" network contains the knowledge that could have plausibly affected the particular patent/invention. Thus, the "network" effect we find is indeed caused by the ex-ante network.

This achieves several important goals: Primarily it allows us to separate the "ex ante" effect from the "ex post" effect. The "ex post" is the network that is in place at the end of our data period. Additionally, this method solves the issue of the potential endogeneity of the "ex-post" network. This is because of the potential for reverse causality between invention quality and centrality as measured by the ex post network. We discuss this point in detail when we conduct the analysis. Finally, and this is a useful property of computing different "ex ante" networks for each patent, there is very little correlation among the different degree centrality network measures (degree, closeness, and betweenness) that we define and discuss in the analysis.

Using data from U.S. PTO patent grants in information security, we find that the quality of Israeli ICT/information security inventions is systematically linked to the structure of the collaborative network that existed among Israeli engineers at the time of the patent application. In particular, we find positive and statistically significant evidence of "ex-ante" direct and indirect knowledge spillovers among Israeli inventors in ICT/IS that raise the quality of invention. This research highlights the importance of direct interaction among inventors as a conduit for flows of frontier scientific knowledge, and, through these flows, the realization of higher invention quality. To illustrate the importance of the discussion above, had we "incorrectly" employed the "ex-post" network (at the end of our data window), rather than the "ex-ante" network in the analysis, we would have overestimated the extent of the knowledge spillovers by approximately 50 percent.

Our results suggest that the knowledge spillovers propagated through Israeli inventor networks in ICT/IS help explain the high quality and significant impact of Israeli invention in the information security domain.[2] Thus, national institutions and history can shape the density and effectiveness of inventor networks.

1.1 Literature Review

Our paper is related to two strands of literature. The first strand, pioneered by Trajtenberg (1990), uses patent citations as measures of the quality of innovations and as measures of knowledge spillovers across inventions. More important inventions tend to be cited more frequently by subsequent patents, in the same way that important and influential papers receive more citations from later scholarship. Empirical techniques initially developed by Jaffe, Trajtenberg, and Henderson (1993) and reviewed in Jaffe and Trajtenberg (2002) use patent citations to measure knowledge spillovers across time and space. As this literature evolved, a growing number of papers sought to directly measure social, contractual, or institutional

2 Like ICT/IS, Israeli Fin-Tech or Med-Tech sectors require computer science expertise and programming skills, areas in which Israel should have a comparative advantage. We find no evidence of such spillovers in either the Israeli Fin-Tech or Med-Tech sectors.

connections between inventors that might mediate knowledge spillovers between them. Branstetter (2001, 2006), Singh (2008), Berry (2014), and Alcacer and Zhao (2012), among others, built on the techniques of Jaffe, Trajtenberg, and Henderson, and used them to measure the degree to which multinationals can enhance flows of knowledge spillovers across national boundaries by creating R&D facilities abroad. Gomes-Casseres, Hagedoorn, and Jaffe (2006) and Branstetter and Sakakibara (2002) have used patent and citation data to measure the impact of formal interfirm research collaboration on knowledge spillovers. Almeida, Song, and Wu (2001) and Agrawal, Cockburn, and McHale (2006), among many others, have sought to measure the impact of the movement of specific individual inventors across organizational boundaries on knowledge spillovers between them. Few previous studies in the economics literature have examined the impact of inventors' collaboration networks traced out by coinventions (that is, inventors appearing together on the same patent document) on knowledge flows and invention quality.[3]

This omission in the innovation literature is striking given the significant attention placed on collaboration networks in other, closely related social science literatures. Recent studies have examined the relationship between network structure and behavior (e.g. Ballester, Calvó-Armengol, and Zenou 2006; Calvo-Armengol and Jackson 2004; Goyal, van der Leij and Moraga-Gonzalez, 2006; Jackson and Yariv 2007; Karlan et al. 2009) and the relationship between network structure and performance (Ahuja 2000; Calvó-Armengol, Patacchini, and Zenou 2009, Fershtman and Gandal 2011; Gandal and Stettner 2016).[4]

Schilling and Phelps (2007), and a related stream of papers, examine alliance structures among firms, showing that firms with denser alliance networks generate more patent applications. Analysis is at the firm level, not the inventor level, and it is firm alliance structures, not coinventions, that forms the basis for the network. Patent counts are used as the innovative outcome, but are not adjusted for

3 Breschi and Lissoni (2009) provide one exception. Their question and approach differs from ours. They are primarily interested in distinguishing knowledge flows that are due to (1) local proximity versus those due to (2) inventors who move from firm to firm locally. While they build a co-invention network, they do not formally use the properties of the network in the analysis, and do not link structural characteristic of the network to the quality of patents. Nerkar and Paruchuri (2005) and Singh (2005) distinguish between ex ante and ex post inventor networks. They do so in studies focused on the role of networks in propagating knowledge flows between inventors in a single U.S. firm (Nerkar and Paruchuri) or between U.S.-based inventors in different firms and regions (Singh). Neither paper directly examines invention quality, so the reverse causality problem we seek to resolve does not arise in these earlier papers.

4 There is a separate economic literature on innovation and technology adoption in industries characterized by network effects in final demand (e.g. Kretschmer 2008). While these effects likely impact some of the firms in our data set, they are not the focus of our study.

citations. Paruchuri's (2010) work is somewhat closer to the approach taken in this paper. Paruchuri examines the relationship between the centrality of inventor in a pre-existing ("ex ante") intrafirm coinventor network and the placement of the firm in an interfirm research alliance network. This paper does consider invention quality, but looks only at patent citations received only by inventors in the same firm (self-citations) and limits its purview to only eight firms in the pharmaceutical industry.

Perhaps the closest prior work to our own is Fleming, King, and Juda (2007), who examine regional innovation networks within the United States. They focus most intensively on Silicon Valley and Boston, using data from the late 1970s through the late 1990s, and explore the hypothesis that so-called "small world" network structures contribute to innovative outcomes. In general, they find much less evidence for this than expected. These authors find that the number of patented inventions generated by the regional cluster is positively related to the size of the giant component and negatively to path length within the component, but analysis is conducted at the cluster level rather than at the patent level. In addition, the connection between path length and patent output disappears when patents are weighted by citations.

This paper seeks to fill a gap in the literature by assessing the degree to which collaboration networks, as traced out by pre-existing instances of "coinvention" by inventors named in patent documents, both shape the pattern of knowledge spillovers *and* influence the quality of individual inventions. We focus on a particular technological domain, but allow our coinvention networks to span organizational boundaries.

1.2 Our Analysis and Results

In this paper, we use data on the inventors that appear in patent documents to trace out and construct a two-mode network: (I) a Patent network and (II) an Inventor network. In the case of the patent network, the nodes are patents, and two patents are linked if there are common inventors who work on both patents. In the case of the inventor network, the nodes of this network are the inventors themselves. There is a link between two inventors if they are co-inventors of the same patent (In section 2 below we provide a simple example to distinguish these two networks).

We examine the patent network and the inventor (collaboration) network of inventors creating technologies in the domain of information security, broadly defined. Our broad definition includes patents in ICT patent classes that the USPTO and researchers working in the field have defined as information security related classes; these are listed in detail in the Appendix and discussed later in the paper.

For each patent, we calculate its proximity to other patents in the network, where the links are through inventors. We then calculate the centrality of these patents within the patent network, in a manner defined below. Similarly, we calculate the centrality of inventors within the inventor network.

We then regress patent invention quality, measured by the total number of forward citations, on network centrality measures within the patent network at the time when the patent application was submitted. We control for other characteristics of the patent. We find that in the case of Israel, the network centrality measure "closeness" (which is defined below) is significantly associated with the variation in patent quality. This result provides evidence of direct and indirect knowledge spillovers, propagated through inventor networks, which raise the quality of invention.

We use instances of the same inventors appearing together in a patent document to trace out the networks through which knowledge spillovers will be presumed to flow. Of course, this definition necessarily omits instances of collaboration or communication that are not reflected in the patent documentation "paper trail". While acknowledging this point, we argue that unmeasured communication and interaction are likely to be highly correlated in space and time with the data that we do observe in the patent data record.

1.3 Israel's Emergence as a Global Center of Innovation in ICT/ Information Security

Our primary focus is on Israel, which is recognized as one of the most innovative countries in the world. Widely cited indices of national innovative capacity, such as the Bloomberg Index of Innovation or the Global Competitiveness Index compiled by the World Economic Forum, regularly rank Israel among the world's top five innovating countries, despite its small size.[5] Reflecting this technological strength, the country has become a major global center for high-tech entrepreneurship. Excluding the U.S., only China has more firms listed on the NASDAQ stock exchange.[6] Leading players in the global IT sector, such as Intel, IBM, Google, Motorola, Apple, Microsoft, and many others have set up research centers in Israel, hoping to harvest local talent and knowledge. Israeli companies today

5 See "The Bloomberg Innovation Index", http://www.bloomberg.com/graphics/2015-innovative-countries/ (accessed 17/12/2016) and "Global Competitiveness Report 2015–2016 – Reports – World Economic Forum", http://reports.weforum.org/global-competitiveness-report-2015-2016/economies/#economy=ISR (accessed 17/12/2016).

6 "Companies in Israel – Nasdaq.com", http://www.nasdaq.com/screening/companies-by-region.aspx?region=Middle+East&country=Israel (accessed 17/12/2016).

play a key role in shaping the global IT industry – from chips to the end user applications. Israeli firms occupy an especially prominent role in information security, which is one of the largest and fastest growing sub-sectors of ICT.

Popular explanations of Israel's technological ascendancy characterize Israel's size as a strength, asserting that the small nation is characterized by tightly connected networks, through which knowledge spillovers can easily flow. Elite Israel Defense Force (IDF) units, such as the well-known Unit 8200, are believed to play an important role in seeding successful startups in Israel by creating a connected network of programmers.[7] Unit 8200, and similar units, effectively nudge a fraction of their most gifted alumni into high-tech entrepreneurship in ICT and related domains. Once they leave the military, 8200 veterans use the network of 8200 veterans to found start-ups and develop technologies based in part on their experience and connections in the military.[8] The theme of knowledge spillovers from connected networks of former members of the military intelligence corps runs through the book *Start-Up Nation* (Senor and Singer 2009) and other sources, but no rigorous statistical study has yet confirmed that these networks have actually played an important role in the rise of Israel as a center of ICT innovation.

In this paper, we do not address the role of particular military units in fostering Israeli networks of information technology developers. However, we undertake what is, to the best of our knowledge, the first empirical effort to measure these networks, as they are traced out in patent data, and ascertain the degree to which network density affects the quality of Israeli invention. To capture information security inventions, we include all patents granted within a broad range of ICT patent classes that have been identified by the USPTO and previous researchers as containing information security patents. These classes are reasonably broad, and contain within them many patents that are not strictly information security inventions, *per se*. We deliberately used a broad definition, in order to be reasonably confident that we obtained all relevant inventions. Narrowly defined fields have limited numbers of patents, making econometric work more challenging. Finally, Israel is very different from the other countries because a large proportion of its patents in the ICT/Information Security sector (47 percent) are assigned to US firms. No other country with significant numbers of patents in this sector has more

7 Unit 8200, a military intelligence unit focusing on signal intelligence and code decryption, is the largest unit in the Israel Defense Forces, comprising several thousand soldiers. It is comparable in its function to the United States' National Security Agency. See Idan Tendler, "From the Israeli Army Unit 8200 to Silicon Valley," 23 March 2015, available at https://techcrunch.com/2015/03/20/from-the-8200-to-silicon-valley/.

8 "70 percent of successful Israeli startups are led by 8200 graduates," says NBIC Director Fadi Swidan, from "High-tech elites to nurture Arab-Israeli startups,"17.4.2016, available at http://www.israel21c.org/high-tech-elites-to-nurture-arab-israeli-startups/.

than 17 percent US assignees, and most of the countries have less than five percent or fewer US assignees.

We now briefly examine ICT/information security patents by patent class for several countries and for the state of California, which is considered to be on the forefront of knowledge in ICT/Information Security (as well as other areas).[9] The percent of patents in each of the ICT/Information Security patent classes is shown in Table 1 for California, Israel, Japan, and Korea (All tables are at the end of the paper).

In the case of Korea, almost 60% of the patents are from categories 365 and 455. These are the two largest categories for Japan as well and account for 36% of the patents in that country. These are hardware-oriented patent classes, only parts of which are strongly related to information security. The distribution of patents across classes in Israel and California looks quite different from Japan and Korea. The percentage of patents in classes 365 and 455 are 21 and 24% respectively. Excluding class 455, which is very broad and has a large number of patents, the

Table 1: ICT/Information security patents class distribution, by country.

Patent class	(1) Israel	(2) California	(3) Japan	(4) South Korea	(5) Taiwan	(6) Canada	(7) Finland	(8) Germany	(9) France
326	2.03%	5.06%	3.46%	3.36%	3.71%	1.86%	0.08%	2.69%	2.83%
340	7.11%	5.63%	8.61%	4.05%	11.87%	11.80%	3.90%	19.50%	13.31%
365	7.24%	10.40%	20.88%	34.94%	22.31%	3.86%	0.33%	9.24%	8.96%
380	2.18%	1.71%	2.12%	1.64%	0.91%	3.77%	2.28%	2.01%	5.41%
455	14.46%	13.41%	15.56%	24.72%	15.87%	27.27%	60.37%	17.89%	19.22%
704	3.77%	3.11%	4.75%	2.99%	2.06%	4.33%	4.76%	5.06%	4.11%
705	4.23%	8.11%	2.91%	0.97%	1.23%	5.78%	2.23%	5.68%	3.61%
708	2.64%	2.15%	2.52%	1.22%	2.08%	1.26%	0.72%	2.76%	4.07%
709	13.63%	15.05%	7.52%	4.17%	2.89%	15.24%	11.13%	8.17%	8.93%
710	5.84%	6.24%	5.51%	2.95%	8.68%	3.26%	1.14%	5.12%	4.41%
711	11.40%	8.83%	8.50%	3.91%	7.36%	3.74%	1.31%	3.56%	4.79%
713	7.37%	6.84%	6.06%	4.51%	9.66%	7.36%	4.70%	6.02%	8.43%
714	11.34%	8.71%	9.19%	8.47%	9.32%	5.25%	3.12%	10.05%	8.52%
726	6.76%	4.74%	2.41%	2.10%	2.06%	5.22%	3.92%	2.24%	3.39%

The table presents the patent class distribution in Israel and other leading countries, for patents in Information Security/ICT area, granted in 1985 onward.

9 We cannot conduct the formal analysis for other countries because it is very difficult to "disambiguate" similar names of inventors. In the case of Israel, we solved this issue by examining each individual inventor.

largest three patent classes for both Israel and California are classes 709, 711, and 714. Patent class 709 covers Electrical Computers and Digital Processing Systems: Multicomputer Data Transferring.Patent Class 711 covers Electrical Computers and Digital Processing Systems: Memory. Patent Class 714 covers Error Detection/Correction and Fault Detection/Recovery. These classes are more oriented to software than patent classes 365 (Static Information Storage and Retrieval) and Patent Class 455 (Telecommunications).

When we look at the percent of patents in the "700 classes," less the percent of patents in the other classes containing ICT/information security patents, we see an interesting bifurcation. Israel, Canada, and California have many more patents in the "software" classes, while Korea, Taiwan, and Finland have many more patents in the hardware-oriented classes containing ICT/information security patents. Germany, France, and Japan are in the middle. See Table 2.

2 Theoretical Foundations for Network-Mediated Knowledge Spillovers

Network-mediated knowledge spillovers can be either direct or indirect. In the case of network-mediated spillovers between patented inventions, *direct* spillovers occur when two patented inventions have a common inventor who transfers knowledge from one patent to another. That is, an inventor takes the knowledge that he/she acquired while working on a previously patented invention and implements it in another invention. However, knowledge may also flow between invention teams even if they are not directly connected by a common inventor. The indirect route occurs whenever an inventor learns something from participating in

Table 2: "700 classes versus other classes", by country.

	(1) Israel	(2) California	(3) Japan	(4) South Korea	Taiwan	Canada	Finland	Germany	France
% of 700 classes	0.67%	0.65%	0.50%	0.32%	0.45%	0.52%	0.33%	0.49%	0.50%
% of other classes	0.33%	0.35%	0.50%	0.68%	0.55%	0.48%	0.67%	0.51%	0.50%
Difference	0.34%	0.29%	0.00%	−0.36%	−0.09%	0.04%	−0.34%	−0.03%	0.01%

The table presents the difference in the share of patents from the "700 classes" and other classes which fall under the category of Information Security/ICT and were granted in 1985 onward.

one invention, takes the knowledge to a second invention and "shares" it with another inventor on that invention team, who, in turn, uses it when she works on a third invention. In such a scenario, knowledge flows from the first patent to the third patent, even though they do not have any inventors in common. Clearly, such indirect spillovers may be subject to decay depending on the distance (the number of the indirect links) between the patents.

Fershtman and Gandal (2011) show theoretically that when there are project spillovers that decrease with decay, there should be a positive correlation between project success and project *closeness centrality*, which is defined as the inverse of the sum of all distances between the project and all other projects. Closeness centrality measures how far each project is from all the other projects in the network.

2.1 An Example Constructing the Patent and Inventor Networks

Before we proceed, the example below shows how to construct the patent and inventor network. Suppose that there are six inventors and five patents with the following patent-inventor data:

Patents	Inventors
Patent 1	Polly & Cindy
Patent 2	Steve
Patent 3	Thomas, Elizabeth, & Jack
Patent 4	Polly & Jack
Patent 5	Steve & Jack

The first network in Diagram 1 below shows the two-mode network with both patents and innovators. The second network shows the "Inventor Network," where two inventors are connected if they work on a patent together. The third network is the "Patent Network." Two patents are connected if they have an inventor in common.

In the inventor network, "Jack" is the most central and he is directly connected to all other inventors except Cindy. In the patent network, both patents 4 and 5 are directly connected to three other patents. Although patents 1 and 3 are not connected, knowledge can indirectly flow between those patents via patent 4. This is because Polly works on both patents 1 and 4, while Jack works on patents 4 and 3.

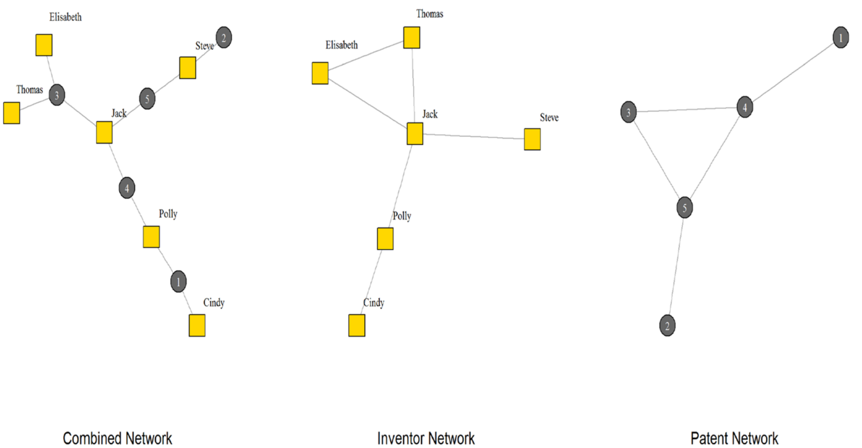

Diagram 1: A two-mode network and corresponding patent networks.

2.2 A Formal Model for Exploring Network-Mediated Knowledge Spillovers

As discussed, the academic literature has frequently used forward patent citations as a measure of invention quality. Following this convention, we assume that the quality (denoted S_i) of each patent "i" is closely related to its count of forward citations, i.e. the citations received from subsequently granted patents. As is typical, we exclude self-citations (both to assignees and to inventors).

We write:

$$S_i = X_i\omega + \varepsilon_i \tag{1}$$

where X_i is a vector of observable patent characteristics, ω is a vector of parameters to be estimated, and ε_i is an error term.

We define two patents to be linked if they have an inventor in common. We focus on national networks. A patent is defined to be from a country if all its inventors are residents of said country, i.e. all inventors have an address in that country on a given patent document. This means, for example, that an Israeli working in the Silicon Valley lab of her multinational employer would be considered "American" for our purposes, because she is a resident of the U.S. at the time of the patent application.

There are three common measures of centrality that measure different types of spillovers: degree, closeness, and betweenness. Degree centrality measures the number of direct connections of the node (a patent in our setting). Closeness centrality inversely measures how far the node is from all other nodes in the connected component. Betweenness centrality measures the number of shortest paths that pass through the node (patent).

If (only) *degree centrality* is significant in explaining the success of a patent, then there are direct knowledge spillovers from directly connected nodes, but no indirect spillovers.

If *closeness* is significant in explaining the success of a patent, then there are both direct and indirect knowledge spillovers from directly and indirectly connected nodes, and the spillovers decay with distance between the patents.

If *betweenness* is significant in explaining the success of a patent, then there are knowledge spillovers from being in the center of the information flow.[10]

We find that for the case of Israel, only closeness centrality is associated with patent success. This suggests that there are both direct and indirect knowledge spillovers that decay with distance.

Given that the other measures are not associated with success, we focus on closeness centrality in the modified version of the Fershtman and Gandal (2011) model we now describe. The model assumes that each patent i may enjoy positive spillovers from patents that are directly connected and patents that are indirectly connected, but that these spillovers are subject to decay that increases as the distance between the patents – that is, the number of intervening connections – in the patent network increases. Formally when the distance between patent i and j is $d\,(i, j)$, we assume that the quality of each patent varies with $Y/\Sigma_j d\,(i,j)$ where γ is the magnitude of the spillover.[11]

Under this assumption, the quality of each patent i can be written

$$S_i = X_i\omega + \frac{\gamma}{\Sigma_j d\,(i,j)} + \varepsilon i. \tag{2}$$

Formally, closeness centrality is the inverse of the sum of all the (shortest) distances between a focal patent and all other patents multiplied by the number of other patents. Closeness centrality measures how far each patent is from all the other patents in a network and is calculated as:

10 As we noted in the introduction – and this is simply a property of different "ex ante" networks for each patent – there is very little correlation among the different degree centrality network measures (degree, closeness, and betweenness). Indeed, the correlation coefficients between the three network centrality measures (degree, closeness, and betweenness) are all less than 0.15. When we "incorrectly" use the "ex post" network (rather than the "ex ante" network) in the analysis, those three measures are indeed highly correlated.

11 For two patents that are directly connected (that is, share an inventor in common), $d\,(i, j) = 1$. For two patents that are indirectly linked via a third patent, $d\,(i, j) = 2$.

$$C_i \equiv \frac{(N-1)}{\sum_{j \in N} d(i,j)}, \tag{3}$$

where N is the number of patents and $d(i, j)$ is the shortest distance between Israeli patents i and j, as measured by the network traced out in patent documents. Patents that indirectly link to a large number of other patents have a higher closeness centrality measure than patents near or at the edge of a network. (See Freeman (1979), pp. 225–226).

Using (3), the expression for closeness centrality, patent i's success can be rewritten as

$$S_i = X_i \omega + \gamma \frac{C_i}{N-1} + \epsilon_i \tag{4}$$

Hence, for each patent (denoted "i"), we calculate the cited patent's "country network" closeness centrality. By construction, we only consider the possibility of *intra*national knowledge spillovers, because our networks are based on co-inventions between inventors who "meet" in the same national territory.[12] The closeness centrality measure is only defined within groups of patents that are actually connected to each other by common inventors. For that reason, following the usual practice in the network literature, we will focus our analysis on the largest single group of patents within Israel (and our other sample countries) that are connected to a common network. This is referred to in the literature as the "giant component."

2.2.1 Endogeneity/Causality: "Ex Ante" Versus "Ex Post" Network

Importantly, we need to address the endogeneity issue associated with network formation. High quality patents will also attract large numbers of citations from subsequently granted patents. As the quality of these patents are recognized, more inventors may want to collaborate with the inventors or with their collaborators. This raises the possibility that the causal linkage between network density and invention quality runs in both directions, with higher quality patents effectively growing a denser network around them after they are invented.

---

12 This does not imply the assumed absence of international spillovers but rather the difficulty of tracking inventor networks across countries and our interest in measuring the impact of intra-national networks, especially in Israel, on invention quality. To the extent that unmeasured international collaborations raise the quality of invention, our approach is likely to generate a downward-biased estimate of the impact of Israeli inventor networks on inventions quality.

To address this issue, we need to distinguish between the "ex-ante" network that was in effect when the application for the patent was filed, and the "ex-post" network that exists at the end of our data window. To do this, we create, for each patent, the coinvention network that exists at the time of the patent filing — meaning that there is a different network for each patent. Logically, this "ex ante" network is the network that could have plausibly raised the quality of the invention.

Using the "ex-ante" network helps resolve the endogeneity issue, in part because of the difficulty inventors face in forming coinvention links in a forward-looking manner. In a relatively new and fast-moving domain like information security, it can be hard to anticipate which inventors will create high quality patents in the future. The ability to form linkages is also constrained by organizational boundaries. Patents are generally owned by a single firm and all the inventors listed on a particular patent work typically work at that firm. Few inventors would choose to leave a firm just to apply for a future patent of uncertain quality with someone else at another firm. Hence, we believe this methodology helps address the endogeneity issue.[13]

We take a number of other steps described in more detail later in the paper to strengthen our causal inference. If a disproportionate number of high quality patents are (co)invented by a small number of superstar inventors, and these inventors also possess dense collaboration networks, even ex ante, then we could mistakenly associate the quality of these inventions with the density of the networks rather than the unique inventive capabilities of the superstars. Including a dummy variable for the presence of a "top 1 percent" inventor on a patent does not qualitatively change our results.[14] The measured closeness of a patent will also be (mechanically) related to the number of inventors listed on the patent, and we directly control for this in all regressions. The position of a patent within the network could be related to its technological proximity to prior work – to control for this, we incorporate the number of backward citations listed in the patent document. The quality of a patent is also likely to be related to the level of R&D expenditure that went into it, and, as we have noted several times, large U.S.-based multinationals are a conspicuous presence in the Israeli information security

13 Indeed, when we "incorrectly" use the "ex-post" network that exists at the end of our data window, rather than the "ex-ante" network in the analysis, we find that the coefficient on closeness is approximately 50% higher, highlighting the potential endogeneity of the "ex-post" network.

14 The robustness of our results to "superstar effects" also helps guard against the possibility that unmeasured characteristics like individual "power" or status that are highly correlated with individual networks are driving our results.

sector, collectively investing hundreds of millions of dollars in acquisitions and R&D.[15] These large multinationals have created unusually large teams of well-networked inventors and provided them with unusually large amounts of money – thereby creating high quality patents whose quality could inappropriately ascribed to the density of the networks possessed by the inventors. We address this by dropping *all* patents assigned to U.S. MNCs in our sample, and our empirical results actually get stronger.

3 Data and Empirical Analysis

3.1 Defining and Delimiting Our Patent Populations

We now turn to our empirical analysis. In order to begin, we need to define the relevant i patent classes. As we have already noted, from detailed examination of United States Patent and Trademark Office (USPTO) patent class descriptions and the work of prior researchers, we were able to determine the patent classes relevant for information security innovations, broadly defined. These ICT patent classes are shown in the Appendix.[16]

We then collected data from the USPTO on all patents granted in the relevant patent classes. Our data include the number of forward citations, backward citations (citations made to previously granted patents), grant year, location of inventor (hence we know whether the inventor(s) are Israeli), patent class, the number of claims, the number of subclasses, the number of inventors, and the assignee (owner) of the patent. We use all of these variables in the analysis.

The number of U.S. patents by country in the relevant patent classes for the years 1985–2014 is given in Table 3. Since there were relatively few information security patents before 1985, we start with that grant year. In the 1985–2014 period, the USPTO issued approximately 340,000 patents in our target patent classes in which all inventors are from the same country. The table shows that more than 50% of the patents were issued between 2005–2014.[17]

15 Four large U.S.-based MNCs own more than a quarter of the patents in our sample. The role of these American acquirers is controversial, even among Israelis otherwise committed to globalization and open markets. One famous Israeli economist is known to have complained that, “Israel does not export software, it exports software companies.”

16 See https://www.uspto.gov/web/patents/classification/uspc726/defs726.htm, accessed 25 June 2017. We included class 709, which does not appear as a relevant patent class in the USPTO document, but according to Arora and Nandakumar (2012), should be included in the information security sector. Nothing changes if we eliminate that class.

17 Patents with missing data account for less than 5% of all patents (and 3% for Israeli patents).

Table 3: USPTO ICT/Information security patents by country for 1985–2014.

	(1) Number of patents	(2) Share of patents
Israel	4431	1%
South Korea	17,799	5%
Taiwan	8200	2%
Japan	64,618	19%
Canada	8057	2%
Finland	3497	1%
Germany	10,472	3%
France	6191	2%
USA	190,392	56%
Other countries	25,871	8%
Total	339,528	100%

The table presents the number of patents that originated in the respective country, between 1985 and 2014, and are listed in the USPTO database. We identify a patent as one that was originated in a specific country if all its inventors' home addresses were listed under that country, according the USPTO data.

Because we construct the patent network (for each patent) at the time the patent was applied for, we need to have a large enough existing giant component of connected patents already in existence. A giant component formed in Israel at the end of 2006. Before that time, there were several smaller components. This is typical of network formation. In the case of Israel, this means we can include patents that were applied for beginning in 2007.

In our database, we have patents issued through 2014. Figures 1 and 2 show the formation and development of the Israeli network and its giant component over time. Complete data exist for 881 USPTO patents with Israeli inventors in this period. That is, for these patents, all inventors had an address in Israel. We exclude patents with both Israeli inventors and inventors from other countries (primarily the US) from the main analysis, since we want to focus on the local network. The number of Israeli patents is small relative to the total number of patents in the relevant classes with all inventors based in the same country. Table 4 shows that Israeli patents as a proportion of all patents granted by the USPTO in these classes increased steadily over the 1985–2014 period, but remained a small percentage of the total. The conventional wisdom regarding Israeli patents in these classes is that they stand out in terms of quality rather than quantity.[18]

18 It is also possible – and, in fact, likely –that our data include many patents that are not information security patents, strictly defined, and that the Israeli share of a more narrowly defined set of information security patents would be much higher. We chose to err on the side of being reasonably comprehensive in our definition of information security patents.

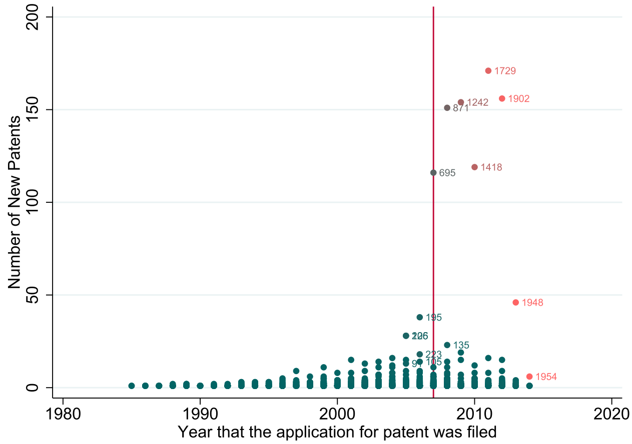

Figure 1: Introduction of patents to the Israeli network.
Note: Each dot represents a component. The *Y* axis shows the number of new patents added to a specific component at a certain year. The colors represent the size of the component, where blue dots represent small components and redder ones bigger components. The number shows the exact size of big components.

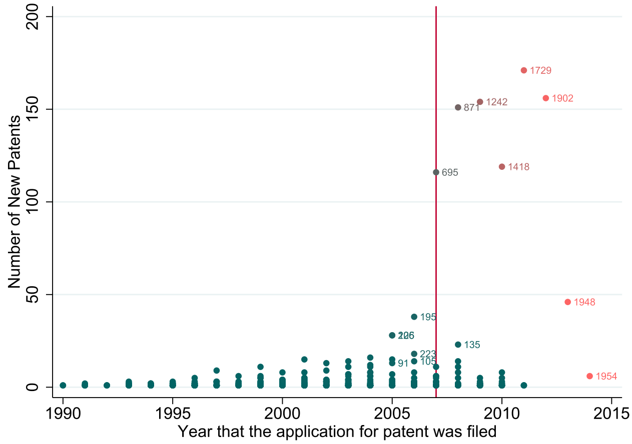

Figure 2: Introduction of patents to the Israeli giant component.
Note: Each dot represents a component. The *Y* axis shows the number of new patents added to a specific component at a certain year. The colors represent the size of the component, where blue dots represent small components and redder ones bigger components. The number shows the exact size of big components.

Table 4: Israeli ICT/Information security patents 1985–2014.

	(1) # of all patents	(2) # of Israeli patents	(3) % of Israeli patents
1985–1989	11,253	32	0.28%
1990–1994	16,417	71	0.43%
1995–1999	36,492	256	0.70%
2000–2004	54,745	554	1.01%
2005–2009	82,732	980	1.18%
2010–2014	137,889	2538	1.84%
Total	339,528	4431	1.31%

Column 1 presents the number of patents issued by all countries and were granted at each five-year period, between 1985 and 2014. Column 2 presents the number of Israeli patents that were granted at the same period. Column (3) shows the percentage of the Israeli patents out of all the patents issued at the same period. We identify a patent as one that was originated in a specific country if all inventors' home addresses were listed in that country, according the USPTO data.

3.2 Construction of the Patent Network

We construct the network of Israeli patents by defining two patents to be linked if they have an inventor in common. Thus, we link patents via the recorded names of inventors. Although the USPTO data are reasonably thorough, the empirical literature has noted the challenges that arise in the "disambiguation" of similar names (Marx, Singh, and Fleming 2015; Trajtenberg, Shif, and Melamed 2009; Ventura, Nugent, and Fuchs 2015). For the purposes of our study, we think of the use of recorded inventor names in USPTO data as raising two measurement challenges, which we refer to as "false positives" and "false negatives."

A **false positive** means that we identify a connection between two patents in the network, where this connection does not actually exist. A false positive occurs if two (or more) separate inventors have the same name. In order to reduce the potential for false positives, we drop inventors with 100 or more patents.[19] Inventor names with a very large number of patents attached to them could, in fact, reflect multiple inventors, and inclusion of such inventors could lead to substantial mismeasurement. In the case of the Israeli network, we individually examined the names of all patent holders with more than 20 patents – and did not find a single case of a false positive. We are thus confident that our results are not driven by false positives in the Israeli data.

A **false negative** means we do not find a connection between two patents due to different spelling, or typing mistakes of the inventors' names. In order to reduce the

19 We note, however, that the qualitative nature of our results is not affected whether we retain or drop inventors with more than 100 patents. There are no such inventors in the Israeli network in any case.

probability of false negatives, we standardize all inventor names in the following ways. First, we used only lower case letters for the names. Second, we removed leading and following spaces. Third, we replaced all "-" symbols with spaces between names. Finally, we removed all punctuation symbols, such as parenthesis, commas etc. This standardization should help minimize the false negatives in our data. To the extent that they remain, and that our network omits important connections, we are underestimating the extent of the network and its potential impact on invention quality. Descriptive statistics for the Israeli network appear in Table 5.

Israel is unique among countries in that many of its patents have US assignees. Fully 47% of the 881 Israeli patents in the giant component that were applied for beginning in 2007 have US assignees.[20] For comparison, no other country has more than 17% "US Assignees" in these patent classes (applied for beginning in 2007), and most have less than 5% US assignee patents. Hence, by this measure, Israel is "off the charts." The high frequency of US assignees reflects the unique history of Israel's high-tech sector. US multinationals established research, design, and production facilities in Israel at the inception of the Israeli ICT industry's

Table 5: Descriptive statistics – Israel.

	N	Mean	Std. Dev	Minimum	Maximum
Forward citations	881	1.70	6.05	0	66
Forward citations – "no self citations"	881	1.42	5.74	0	64
Grant year	881	2012.62	1.36	2008	2014
Application year	881	2009.78	1.83	2007	2014
# of inventors	881	2.75	1.49	1	11
Backward citations	881	36.15	82.85	0	547
US assignee	881	0.47	0.50	0	1
# of claims	881	22.13	11.59	3	153
# of subclasses	881	3.70	2.77	1	26
Closeness/(N–1)	881	0.000105	0.000051	0.000034	0.000236

The table presents descriptive statistics for all the Israeli patents that are in the Israeli Giant component in 2014, and were were applied between 2007 and 2014. Forward citations include the number of citations a patent receives. Forward citations – "No self cites", includes all the citations a patent receives, excluding citation made by patents from the same inventors or that were made by the same assignee. Grant year is the year the patent was approved by the USPTO. Number of inventors are the number of inventors listed as the patent inventors. Closeness is the patent closeness centrality measure, in the patent network formed until the patent application year. Backward citations are the number of patents that were cited by the patent. US Assignee is an indicator variable that indicates whether the patent was applied by a US assignee. We identify a patent as one that was originated in Israel if all its inventors' home addresses were listed under Israel, according the USPTO data.

20 Since the data are from the USPTO, we know whether the assignees are US or foreign entities. In the case of Israel, virtually all non-Israeli assignees are US assignees.

development, and have continued to play an important role. These subsidiaries were often led by Israelis returning home to Israel after years — even decades — of distinguished engineering leadership inside US-based companies. Today, many Israeli start-ups are eventually acquired by US firms, and the purchase moves ownership of Israeli patents to a U.S. entity.

3.3 Measuring Spillovers via Connected Networks

In this section, we estimate Eq. (4) which we repeat below:

$$S_i = X_i\omega + \gamma \frac{C_i}{N-1} + \epsilon_i \tag{4}$$

Recall that S_i, the number of forward citations received by a given patent, is our measure of quality. We exclude self-citations and citations made by patents from the same assignee and the same inventor.

We further assume that the number of forward citations received by patent i depends on a vector of observable factors, denoted X_i. These include characteristics of the patent: backward citations (citations made to previously granted patents), grant year, location of inventor (hence we know whether the inventor(s) are Israeli), patent class, the number of claims, the number of subclasses, the number of inventors, and the assignee (owner) of the patent.

C_i is the *closeness centrality* of patent i in the Israeli network and γ is the parameter associated closeness.

Recall that patent networks play a complex role in expanding the number of citations received by a given patent. First, existing patent networks, as measured by closeness at the time of the patent application, provide the inventors of a given patent access to useful knowledge that enhances the quality and value of invention i, and hence lead to more citations. Second, after invention i is generated, the growing network propagates knowledge of this useful invention (and the technical innovations it contains) to other inventor teams working on related technologies, leading to more citations over time.

Fortunately, we can disentangle these separate effects by constructing a network for each patent at the time the patent was applied for. Using the existing networks for each patent, we can estimate (4) to measure the "ex-ante" effect. Although this makes the empirical work computationally intensive, it is necessary in order to examine whether inventions benefit from the network that was in place when the patent application was filed.

Citations are highly skewed; additionally, some of the independent variables (like the number of inventors) are also highly skewed. Hence, it makes sense to use logarithms and use the log/log specification. The term "ln" before the variable means natural log. The dependent variable used in the regressions in Table 6 is the natural log of forward citations excluding citations from the same inventor and assignee. Since some patents receive no forward citations and the natural log of zero is undefined, following a common practice in the patents literature, we will add one to the number of forward citations and take the natural log of this transformed variable.

3.4 Measuring the "Ex-Ante" Network Spillover Effect

The independent variables are the number of inventors on each patent, the number of backward citations, the number of claims, the number of subclasses, and the closeness of the patent, where we measure closeness at the time when the patent application is filed. We control for grant year in every regression.[21]

Column 1 in Table 6 shows the results for the Israeli patents. The estimated coefficient on closeness (γ) is positive and significant (0.17, $t = 3.25$***), suggesting that there are knowledge spillovers from ex-ante "connections" in the giant component.

As far as the other independent variables are concerned, similar to other studies we find that the estimated coefficients on the number of backward citations, the number of inventors, and the number of claims are statistically significant in all the specifications in Table 6. The estimated coefficient on the number of subclasses is positive, but not statistically significant.

3.5 Robustness Analysis

- In columns 2 and 3, we repeat the analysis in column 1 for US and Israeli assignees separately. We find that the estimated coefficient on closeness (γ) is positive and significant for both groups (0.17, $t = 2.08$** for Israeli assignees, coefficient = 0.26, $t = 4.24$*** for US assignees), again suggesting that there are knowledge spillovers from "connections" in the giant component.[22] The estimated coefficient on backward citations is positive and significant in all cases, while the estimated coefficient on the number of innovators is significant for the full sample and for "Israeli assignees."
- Collectively, four large American firms (Apple, Google, IBM, and Intel) hold 28 percent of the "Israeli" patents in the data set. In this sense, Israel is very

21 When conducting robustness results, we also include dummy variables for patent classes. Again, our main results are unchanged: (the estimate for γ is 0.16, $t = 3.09$***).

22 The coefficient for American assignees seems significantly larger than that of Israeli assignees, but the differences are not statistically significant.

Table 6: Regression results – Israel.

	(1) Network at the time of patent application	(2) Network at the time of patent application – US assignees	(3) Network at the time of patent application Israeli assignees	(4) Network at the time of patent application without large assignees	(5) Network at the time of patent application with Super Star Ind.
ln(Closeness)	0.169***	0.261***	0.180**	0.179***	0.171***
	(0.052)	(0.062)	(0.082)	(0.061)	(0.054)
ln(# of inventors)	0.061*	0.008	0.142**	0.062	0.061*
	(0.037)	(0.039)	(0.062)	(0.046)	(0.037)
ln(Backward cites)	0.078***	0.027	0.127***	0.097***	0.078***
	(0.015)	(0.016)	(0.026)	(0.020)	(0.016)
ln(# of claims)	0.220***	0.120**	0.247***	0.218***	0.220***
	(0.043)	(0.058)	(0.060)	(0.050)	(0.043)
ln(# of subclass)	0.038	0.031	0.040	0.050	0.038
	(0.032)	(0.031)	(0.055)	(0.038)	(0.032)
Super Star					−0.006
					(0.044)
N	881	416	464	681	881
Adj. R^2	0.387	0.311	0.385	0.393	0.387
Grant year dummies	Yes	Yes	Yes	Yes	Yes

The dependent variable is the natural log of one plus the number of forward citations. While counting forward citations, we exclude citations made by the patent's inventors' other patents, and citations made by other patents that are listed under the patent's assignee. Number of inventors is the number of inventors listed in the USPTO data. The variables are as explained in the text. In column (1) we regress on all Israeli patents applied for after 2007 and we define the Degree and Closeness of a patent by looking at the network that existed up until the patent application year. In column (2) we regress the same specification only on patents with US Assignee. In column (3) we repeat the regression only for patents with non-US assignee. In column (4) we omit patents that were applied by IBM, Apple, Google and Microsoft and In column (5) we control for "Super Stars". Standard errors are in parentheses.
(*= significant at 10% level, **= significant at 5% level, ***= significant at 1% level).

different from all other countries: they have very small percentages of US assignees. When we exclude patents assigned to these major firms (column 4, Table 6), the estimate of γ remains positive and highly significant (0.18, $t = 2.83$***), and similar to that in column 1 in Table 6. Hence, the results are not affected by excluding the very large firms from the analysis.

- When we include four centrality measures (degree, closeness, betweenness as well as eigenvector centrality) together in the same specification, it does not change the coefficient on closeness very much. It has a value of 0.20 (instead of 0.17 when we only include closeness) and a standard error of 0.054 (instead of 0.053 when we only include closeness). Also, only the coefficient on closeness is positive and statistically significant. The estimated coefficients on degree and betweenness are negative and not statistically significant. The coefficient on eigenvector centrality is virtually zero (a t-statistic of 0.10). In Table A1 in the Appendix, we provide regressions with the other centrality measures separately and as well as a regression with all of the centrality measures.
- Forward citations are highly skewed, logging forward citations and using it as the dependent variable is the standard approach in this literature. We believe that this is the best approach in our setting. Nevertheless, also estimated a negative binomial regression model (an event count model). We find that the estimated coefficient of closeness is also statistically significant (a z-stat = 3.15 with the negative binomial distribution vs. a t-stat = 3.27 with the dependent variable in our preferred model). However, the negative binomial distribution is a much poorer fit than our preferred model. The negative binomial has a Pseudo R-squared value of 0.19 while our preferred model has an adjusted R-squared of 0.38.

3.6 Additional Robustness Analysis: Employing Characteristics from the Innovator Network

In addition to the patent network generated by connections among inventors, there is also a related inventor network. Indeed, as we noted, our data form a two-mode-network: (I) patents and (II) inventors. The two-mode-network can be partitioned into two types of nodes, e.g. patents and inventors. We can then use the two-mode network to construct two different one-mode networks: (i) the patent network and (ii) inventor network. Here we add the inventor network to the analysis, where, in the inventor network, two inventors are connected if they work together on a patent. The nodes of the inventor network are innovators and the nodes of the patent network are patents.

We include the inventor network by introducing a dummy variable that equals one for inventors who are ranked in the top one percent of all inventors in the country in terms of the number of patents the innovator holds. This dummy variable (“Super Star”) takes on the value one if the patent has a top one-percent innovator on the patent and zero otherwise. This controls for inventor quality. When constructing the “Star” variable, we make these calculations at the end of time, reflecting the notion that inventor quality is inherent. Using the top one percent is ideal because in the giant component, roughly half (about 45 percent) of the patents have such an inventor. In the Israeli patent data, 77% of the inventors have one or two patents, while 10% have more than five patents.

It is interesting to examine whether (controlling for network structure) such “stars” affect the success of the patent. We find that in the case of Israel, beyond the effect it has on the network, the presence of such stars does not affect the success of the patent. The coefficient of a dummy variable signifying the existence of a top 1% inventor on the inventor team of a patents is quite small and statistically insignificant (The coefficient estimate is only –0.004, and the t-statistic is –0.11). The estimate of γ is unaffected by inclusion of this variable. The estimated coefficient on γ remains positive and statistically significant (0.17, $t = 3.16^{***}$). See column 5, table 6. These results suggest that our measure of closeness centrality is not merely a proxy for the presence on the inventor team of highly accomplished individual inventors. Instead, they are fully consistent with the view that much of what makes highly accomplished inventors valuable members of a research team is the knowledge spillovers — direct and indirect — that they bring into a research collaboration.

4 Conclusions

For nearly a quarter century, researchers have used patent citation data to trace out knowledge spillovers across inventions, organizations, and regions. From the inception of this literature, researchers have recognized the potential importance of direct interaction between inventors, but relatively few studies have sought to measure inventor networks explicitly, and fewer still have sought to quantify the degree to which these networks raise invention quality.

Drawing inspiration from related work on open source software projects, this study seeks to advance the literature by using the pattern of inventor interaction traced out in patent documents to create measures of inventor networks; we go on to empirically measure the association between the location of a patent within this network and the quality of invention as measured by forward citations. We apply these techniques in an interesting context – ICT/information security technology in Israel. This is a domain in which Israeli inventors have recently emerged as globally important creators of new technology. Industry accounts suggest that the rapid rise of Israeli firms to this position of global prominence has been driven, in part, by the unusually strong networks that characterize Israeli inventors operating in this domain. These networks are believed to help produce better inventions, and then rapidly convey the new technologies embodied in these inventions to subsequent inventor teams. Despite wide acceptance of this conventional wisdom, no empirical research has yet convincingly related Israeli invention quality to Israeli inventor networks. This paper presents empirical evidence supporting and extending this conventional wisdom. We find that the quality of Israeli inventions is systematically related to the location of these patents within the Israeli invention network.

Appendix Patent classes

A1: Relevant patent classes for ICT/Information security:

Table A1: Different centrality measures.

	(1) log(Fwd. citations + 1)	(2) log(Fwd. citations + 1)	(3) log(Fwd. citations + 1)	(4) log(Fwd. citations + 1)
ln(Eigenvector centrality + 1)	−0.102			0.019
	(0.795)			(0.897)
ln(Betweenness + 1)		−0.010		−0.010
		(0.006)		(0.007)
ln(Degrees)			−0.019	−0.027
			(0.018)	(0.022)
ln(Closeness)				0.203***
				(0.054)
ln(# of inventors)	0.073**	0.110**	0.079**	0.106**
	(0.037)	(0.044)	(0.037)	(0.044)
ln(Backward cites)	0.078***	0.081***	0.083***	0.087***
	(0.015)	(0.015)	(0.016)	(0.016)
ln(# of claims)	0.220***	0.219***	0.219***	0.216***
	(0.043)	(0.043)	(0.043)	(0.043)
ln(# of subclass)	0.040	0.043	0.039	0.038
	(0.032)	(0.032)	(0.032)	(0.032)
N	881	881	881	881
Adj. R^2	0.379	0.381	0.380	0.380
Grant year dummies	Yes	Yes	Yes	Yes

The table presents regression results using the specification used in table 6, column 1, where instead of using closeness as a centrality measure, we use other measures. The dependent variable is the natural log of one plus the number of forward citations. While counting forward citations, we exclude citations made by the patent's inventors' other patents, and citations made by other patents that are listed under the patent's assignee. In column (1) we use Eigenvector centrality measure. In column (2) we use Betweenness centrality measure. In column (3) we use the number of degrees. In column (4) we include all four centrality measures. Standard errors are in parentheses.
(*= significant at 10% level, **= significant at 5% level, ***= significant at 1% level).

326,	Electronic digital logic circuitry, subclass 8 for digital logic circuits acting to disable or prevent access to stored data or designated integrated circuit structure.
340,	Communications: Electrical, subclasses 5.2 through 5.74, for authorization control without significant data process features claimed, particularly subclasses 5.22–5.25 for programmable or code learning authorization control; and subclasses 5.8–5.86 for intelligence comparison for authentication.
365,	Static information storage and retrieval, subclass 185.04 for floating gate memory device having ability for securing data signal from being erased from memory cells.
380,	Cryptography, subclasses 200 through 242 for video with data encryption; subclasses 243–246 for facsimile encryption; subclasses 247–250 for cellular telephone cryptographic authentication; subclass 251 for electronic game using cryptography; subclasses 255–276 for communication using cryptography; subclasses 277–47 for key management; and subclasses 287–53 for electrical signal modification with digital signal handling.
455,	Telecommunications, subclass 410 for security or fraud prevention in a radiotelephone system.
704,	Data processing: Speech signal processing, linguistics, language translation, and audio compression/Decompression, subclass 273 for an application of speech processing in a security system.
705,	Data processing: Financial, business practice, management, or cost/price determination, subclass 18 for security in an electronic cash register or point of sale terminal having password entry mode, and subclass 44 for authorization or authentication in a credit transaction or loan processing system.
708,	Electrical computers: Arithmetic processing and calculating, subclass 135 for electrical digital calculating computer with specialized input for security.
709,	Electrical computers and digital processing systems: Multicomputer data transferring, subclass 225 for controlling which of plural computers may transfer data via a communications medium.
710,	Electrical computers and digital data processing systems: Input/Output, subclasses 36 through 51 for regulating access of peripherals to computers or vice-versa; subclasses 107–125 for regulating access of processors or memories to a bus; and subclasses 200–240 for general purpose access regulating and arbitration.
711,	Electrical computers and digital processing systems: Memory, subclass 150 for regulating access to shared memories, subclasses 163–164 for preventing unauthorized memory access requests.
713,	Electrical computers and digital processing systems: Support, subclasses 150 through 181 for multiple computer communication using cryptography; subclasses 182–186 for system access control based on user identification by cryptography; subclass 187 for computer program modification detection by cryptography; subclass 188 for computer virus detection by cryptography; and subclasses 189–194 for data processing protection using cryptography.
714,	Error detection/Correction and fault detection/recovery, subclasses 1 through 57 for recovering from, locating, or detecting a system fault caused by malicious or unauthorized access (e.g. by virus, etc.).
726	Protection of data processing systems, apparatus, and methods as well as protection of information and services.

References

Acemoglu, D. 2009. *Introduction to Modern Economic Growth*. Princeton, NJ: Princeton University Press.

Agrawal, A., I. Cockburn, and J. McHale. 2006. “Gone but Not Forgotten: Labor Flows, Knowledge Spillovers and Enduring Social Capital.” *Journal of Economic Geography* 6 (5): 571–91.

Ahuja, G. 2000. “Collaboration Networks, Structural Holes, and Innovation: A Longitudinal Study.” *Administrative Science Quarterly* 45 (3): 425–55.

Akcigit, U., E. Miguelez, V. Sterzi, S. Stantcheva, and S. Caicedo. 2018. *Dancing with the Stars: Interactions and Human Capital Accumulation*. Cambridge, MA: NBER Working Paper no. 24466.

Alcacer, J., and M. Zhao. 2012. “Local R&D Strategies and MultiLocation Firms: The Role of Internal Linkages.” *Management Science* 58 (4): 739–53.

Almeida, P., J. Song, and G. Wu. 2001. “Mobility of Engineers and Cross-Border Knowledge Building: The Technological Catching-Up Case of Korean and Taiwanese Semiconductor Firms.” In *Research in Technology and Innovation Management*, Vol. 7, edited by H. Chesbrough, and R. Burgelman, 57–84. Somerville, MA: Emerald Publishing Company.

Arora, A., and A. Nandkumar. 2012. “Insecure Advantage? Markets for Technology and the Value of Resources for Entrepreneurial Ventures.” *Strategic Management Journal* 33 (3): 231–51.

Ballester, C., A. Calvó-Armengol, and Y. Zenou. 2006. “Who’s Who in Networks. Wanted: The Key Player.” *Econometrica* 74 (5): 1403–17.

Bernstein, J., and M. I. Nadiri. 1988. “Interindustry R&D Spillovers, Rates of Return, and Production in High-Tech Industries.” *American Economic Review Papers and Proceedings* 78 (2): 429–34.

Berry, H. 2014. “Global Integration and Innovation: Multi-Country Knowledge Generation within MNCs.” *Strategic Management Journal* 35 (6): 869–90.

Branstetter, L. 2001. “Are Knowledge Spillovers International or Intranational in Scope? Microeconometric Evidence from Japan and the United States.” *Journal of International Economics* 53 (1): 53–79.

Branstetter, L. 2006. “Is Foreign Direct Investment a Channel of Knowledge Spillovers: Evidence from Japan’s FDI in the United States.” *Journal of International Economics* 68 (2): 325–44.

Branstetter, L., and K. Saggi. 2011. “Intellectual Property Rights, Foreign Direct Investment, and Industrial Development.” *Economic Journal* 121 (555): 1161–91.

Branstetter, L., and M. Sakakibara. 2002. “When Do Research Consortia Work Well and Why? Evidence from Japanese Panel Data.” *American Economic Review* 92 (1): 143–59.

Breschi, Stefano, and Francesco Lissoni. 2009. “Mobiilty of Skilled Workers and Co-invention Networks: An Anatomy of Localized Knowledge Flows.” *Journal of Economic Geography* 9 (4): 439–68.

Calvó-Armengol, A., and M. O. Jackson. 2004. “The Effects of Social Networks on Employment and Inequality.” *American Economic Review* 94 (3): 426–54.

Calvó-Armengol, A., E. Patacchini, and Y. Zenou. 2009. “Peer Effects and Social Networks in Education.” *Review of Economic Studies* 76 (4): 1239–67.

Fershtman, C., and N. Gandal. 2011. “Direct and Indirect Knowledge Spillovers: the “Social Network” of Open-Source Projects.” *The RAND Journal of Economics* 42 (1): 70–91.

Fleming, L., C. King III, and A. I. Juda. 2007. “Small Worlds and Regional Innovation.” *Organization Science* 18 (6): 938–54.

Freeman, L. 1979. “Centrality in Social Networks: Conceptual Clarification.” *Social Networks* 1 (3): 215–39.
Gandal, N., and U. Stettner. 2016. “Network Dynamics and Knowledge Transfer in Virtual Organizations.” *International Journal of Industrial Organization* 48 (C): 270–90.
Gomes-Casseres, B., J. Hagedoorn, and A. Jaffe. 2006. “Do Alliances Promote Knowledge Flows?” *Journal of Financial Economics* 80 (1): 5–33.
Goyal, S., M. J. Van Der Leij, and J. L. Moraga-González. 2006. “Economics: An Emerging Small World.” *Journal of Political Economy* 114 (2): 403–12.
Griliches, Z. 1979. “Issues in Assessing the Contribution of R&D; to Productivity Growth.” *The Bell Journal of Economics* 10 (1): 92–116. Also reprinted in *The Economics of Productivity*, vol. I, edited by E. Wolff, 256–80. Cheltenham: Elgar, 1997.
Griliches, Z. 1992. “The Search for R&D Spillovers.” *The Scandinavian Journal of Economics* 94: S29–47.
Grossman, G., and E. Helpman. 1991. *Innovation and Growth in the Global Economy*. Cambridge: MIT Press.
Irwin, D., and P. Klenow. 1994. “Learning-by-Doing Spillovers in the Semiconductor Industry.” *Journal of Political Economy* 102 (6): 1200–27.
Jackson, M., and L. Yariv. 2007. “Diffusion of Behavior and Equilibrium Properties in Network Games.” *American Economic Review* 97 (2): 92–8.
Jaffe, A. 1986. “Technological Opportunity and Spillovers of R&D: Evidence from Firms’ Patents, Profits and Market Value.” *American Economic Review* 76 (5): 984–1001.
Jaffe, A., M. Trajtenberg, and R. Henderson. 1993. “Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations.” *Quarterly Journal of Economics* 108 (3): 577–98.
Jaffe, A., and M. Trajtenberg. 2002. *Patents, Citations, and Innovations: A Window on the Knowledge Economy*. Cambridge: MIT Press.
Jones, B. F. 2014. *The Knowledge Trap: Human Capital and Development Reconsidered (WP No. 14138)*. Cambridge, MA: National Bureau of Economic Research.
Karlan, D., M. Mobius, T. Rosenblat, and A. Szeidl. 2009. “Trust and Social Collateral.” *Quarterly Journal Economy* 124 (3): 1307–61.
Keller, W. 1998. “Are International R&D Spillovers Trade-Related? Analyzing Spillovers Among Randomly Matched Trade Partners.” *European Economic Review* 42 (8): 1469–81.
Keller, W. 2004. “International Technology Diffusion.” *Journal of Economic Literature* 42 (3): 742–82.
Keller, W., and S. Yeaple. 2013. “The Gravity of Knowledge.” *American Economic Review* 103 (4): 1414–44.
Kretschmer, T. 2008. “Splintering and Inertia in Network Industries.” *Journal of Industrial Economics* 56 (4): 685–706.
Marx, M., J. Singh, and L. Fleming. 2015. “Regional Disadvantage? Employee Non-Compete Agreements and Brain Drain.” *Research Policy* 44: 941–55.
Nerkar, Atul, and Srikanth Paruchuri. 2005. “Evolution of R&D Capabilities: The Role of Knowledge Networks Within a Firm.” *Management Science* 51 (5): 771–85.
Paruchuri, S. 2010. “Intraorganizational Networks, Interorganizational Networks, and the Impact of Central Inventors: A Longitudinal Study of Pharmaceutical Firms.” *Organization Science* 21 (1): 63–80.
Romer, P. 1986. “Increasing Returns and Long-Run Growth.” *Journal of Political Economy* 94 (5): 1002–37.

Romer, P. 1990. "Endogenous Technological Change." *Journal of Political Economy* 98 (5): S71–102.
Scherer, F. M. 1982. "Inter-industry Technology Flows and Productivity Growth." *The Review of Economics and Statistics* 64 (4): 627–34.
Schilling, M. A., and C. C. Phelps. 2007. "Interfirm Collaboration Networks: The Impact of Large-Scale Network Structure on Firm Innovation." *Management Science* 53 (7): 1113–26.
Senor, D., and S. Singer. 2009. *Start-Up Nation: The Story of Israel's Economic Miracle*. New York, NY: McClelland and Stewart.
Singh, Jasjit. 2005. "Collaborative Networks as Determinants of Knowledge Diffusion Patterns." *Management Science* 51 (5): 756–70.
Singh, J. 2008. "Distributed R&D, Cross-Regional Knowledge Integration and Quality of Innovative Output." *Research Policy* 37 (1): 77–96.
Trajtenberg, M. 1990. *Economic Analysis of Product Innovation – The Case of CT Scanners*. Cambridge, MA: Harvard University Press.
Trajtenberg, M., G. Shif, and R. Melamed. 2009. "The Names Game: Harnessing Inventors, Patent Data for Economic Research." *Annals of Economics and Statistics* 93/94: 79–108.
Ventura, S., R. Nugent, and E. Fuchs. 2015. "Seeing the Non-Stars: (Some) Sources of Bias in Past Disambiguation Approaches and a New Public Tools Leveraging Labeled Records." *Research Policy* 44 (9): 1672–701.