NBER WORKING PAPER SERIES

NETWORK-MEDIATED KNOWLEDGE SPILLOVERS:
A CROSS-COUNTRY COMPARATIVE ANALYSIS OF INFORMATION SECURITY INNOVATIONS

Lee Branstetter
Neil Gandal
Nadav Kuniesky

Working Paper 23808
http://www.nber.org/papers/w23808

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
September 2017

Branstetter, Gandal, and Kunievsky thank the Maurice Falk Institute for Economic Research in Israel, the National Science Foundation, and Start-Up Nation Central for financial support of this research. We thank Britta Glennon for excellent research assistance, and Avi Ben-Bassat, Reuben Gronau, Eugene Kandel, and Niron HaShai for helpful comments and suggestions. Research results and conclusions expressed are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Network-Mediated Knowledge Spillovers: A Cross-Country Comparative Analysis of Information Security Innovations
Lee Branstetter, Neil Gandal, and Nadav Kuniesky
NBER Working Paper No. 23808
September 2017
JEL No. O31,O33,O57

ABSTRACT

A large and growing literature has used patent and patent citation data to measure knowledge spillovers across inventions and organizations, but relatively few papers in this literature have explicitly considered the collaboration networks formed by inventors as a mechanism for shaping and transmitting these knowledge flows. This paper utilizes an approach developed by Fershtman and Gandal (2011) to examine the incidence and nature of knowledge flows mediated by the collaboration networks of inventors active in the information security industry. This is an industry in which a number of nations outside the United States, including Israel, have emerged as important centers of innovation. Using data from U.S. PTO patent grants in information security, we find that the quality of Israeli information security inventions is systematically linked to the structure of the collaborative network generated by Israeli inventors in this sector. Using the Fershtman and Gandal (2011) model, this suggests that there are knowledge spillovers from the network. In some other nations, invention quality is less closely linked to the collaboration networks of inventors. This research highlights the importance of direct interaction among inventors as a conduit for flows of frontier scientific knowledge.

Lee Branstetter
Heinz College
School of Public Policy and Management
Department of Social and Decision Sciences
Carnegie Mellon University
Pittsburgh, PA 15213
and the Peterson Institute for International Economics
and also NBER
branstet@andrew.cmu.edu

Neil Gandal
Eitan Berglas School of Economics
Tel Aviv University
Tel Aviv 69978, Israel
and CEPR
gandal@post.tau.ac.il

Nadav Kuniesky
Eitan Berglas School of Economics
Tel-Aviv University
Tel Aviv 69978, Israel
nadavkun@gmail.com

1. Introduction

Knowledge spillovers lie at the heart of modern theories of endogenous growth (Romer, 1986, 1990; Acemoglu, 2009), international trade (Grossman and Helpman, 1991; Branstetter and Saggi, 2011); international investment (Keller and Yeaple, 2011), and economic development (Jones, 2014). The late Zvi Griliches and several generations of his students, including Adam Jaffe and Manuel 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 the end products - remain imperfectly understood. Some early research (Griliches, 1979, 1992; Keller, 1996) 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 who work together and exchange ideas and information.

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 (Almeida and Song, 2000; Agrawal et al., 2006). The networks traced out by collaborations can become a key mechanism through which knowledge flows. Interestingly, though a great deal of the research has focused on measuring

[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.

knowledge spillovers in patents, over time and space, to the best of our knowledge, no previous research tried to link network (knowledge) spillovers to the quality of patents.

We apply a model developed by Fershtman and Gandal (FG 2011) (and applied to Open Source Software) to examine the existence and importance of collaborator network-mediated knowledge spillovers in the information security industry. This is an industry in which a number of nations outside the United States, including Israel, have emerged as important centers of innovation. Israeli prominence in this sector is often attributed, in part, to dense networks of personal collections and collaborations that has their genesis in elite intelligence units in the Israeli Defense Forces. Through service in these units, many Israeli information security inventors and entrepreneurs receive their first exposure to this domain.

Using data from U.S. PTO patent grants in information security, we find that the quality of Israeli information security inventions is systematically linked to the structure of the collaborative network generated by Israeli inventors in this sector. Using the FG (2011) model, this suggests that there are knowledge spillovers in the network which improve the quality of patents, as measured by the number of citations. In other nations, patent quality is less closely linked to the collaboration networks of inventors. This research highlights the importance of direct interaction among inventors as a conduit for flows of frontier scientific knowledge.

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 (1992) 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 connections between inventors that might mediate knowledge spillovers between them. Branstetter (2001, 2006), Singh (2009), Berry (2012), and Alcacer and Zhao (2013), 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, Jaffe, and Hagedoorn (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 and Song (2000) and Agrawal, Cockburn, and McHale (2010), among many others, have sought to measure the impact of the movement of specific individual inventors across organizational boundaries on knowledge spillovers between them. Interestingly, however, virtually no previous studies in the economics literature have examined the impact of the collaboration network traced out by coinventions (that is, inventors appearing together previously on the same patent document) on knowledge flows and invention quality.[2]

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, & Zenou, 2006; Calvo-Armengol & Jackson, 2004; Goyal, van der Leij and Moraga-Gonzalez, 2006; Jackson & Yariv, 2007; Karlan, Mobius, Rosenblat, & Szeidl, 2009) and the relationship between network structure and performance (Ahuja, 2000; Calvó-Armengol, Patacchini, & Zenou, 2009, Fershtman and Gandal, 2011, and Gandal and Stettner, 2016). 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 engineers named in patent documents, shape the pattern of knowledge spillovers and influence patent quality.

1.2 Our Analysis and Results

Our paper employs a methodology inspired by Fershtman and Gandal (2011), which utilizes the pattern of collaborations across programmers traced out in the Open Source Software community. It is believed that there are significant knowledge spillovers across R&D projects. In the case of Open Source Software (OSS), if such spillovers exist, it is likely that programmers take code, know-how, and experience gained from one OSS project they have worked on in the past and apply this learning to another OSS project. Using cross-sectional data, Fershtman and Gandal (2011) find that the structure of the product network is associated

[2] Breschi and Lissoni (2009) provide an exception. Their question and approach differs 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, but rather compare treatment patents to control patents.

with the project's success, which under the assumptions of the model, provides support for knowledge spillovers. Using panel data, Gandal and Stettner (2016) find additional evidence for the presence of knowledge spillovers across OSS projects.

In this paper, we use data on the inventors which 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 the patents and two patents are linked if there are inventors who work in both.[3] 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 jointly hold a patent.[4] (In section 3 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 all patents in patent classes that the USPTO defines as information security related classes; these are listed in detail in Appendix B, and discussed later in the paper. For each information security 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 patent network, in a manner defined below. Similarly, we calculate the centrality of inventors within the inventor network.

Following Fershtman and Gandal (2011), we then regress patent invention quality, measured by the total number of forward citations, on network centrality measures within the patent network and inventor networks, controlling for characteristics of the patent. We find that the network centrality measures are significantly associated with the variation in patent quality. In the context of the FG (2011) model, this result (which is very robust) provides evidence of both direct and indirect knowledge spillovers. We find that these spillovers are especially strong for Israeli patents/inventors.

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

While this paper will present a cross-country comparative analysis of innovation in the information security domain, it will place special emphasis on activity in Israel, which is increasingly recognized as one of the most innovative countries in the world. Widely cited

[3] Each link also has a value, which reflects the number of common inventors between two patents.
[4] Again, each link also has a value, which reflects the number of common patents between two inventors.

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 5 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. In 2015 alone, there were 95 mergers and acquisitions of Israeli high-tech companies.[7] In 2017, Intel's acquisition of Israeli computer vision firm Mobileye for $15.3 billion set a financial record; Mobileye will become the global center of Intel's efforts to develop new technology for the worldwide automobile industry. Israeli companies today 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. The Israeli National Cyber Bureau (established in 2011) estimates that the number of firms in cyber security doubled from 150 in 2010 to 300 in 2015. One-quarter of the world's venture capital-funded cyber-security startups are Israeli. Israel's success in this area has led to a few relatively large successful Israeli cybersecurity firms like Checkpoint and CyberArc and to many acquisitions of Israeli cybersecurity startups. Additionally, virtually all of the leading international information security firms (i.e., McAfee) have set up R&D centers in Israel.

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.[8] Unit 8200, and similar units,

[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)

[7] IVC Research Centre (2016) IVC high-tech yearbook. Tel Aviv: IVC Research Centre.

[8] 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/

effectively nudge a fraction of their most gifted alumni into high-tech entrepreneurship in information security 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.[9] 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 work has been conducted on this issue.

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 is correlated with 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 as containing information security patents.[10] This limits the scope of our analysis, but focuses it on a domain in which Israeli firms have emerged as important global leaders. To provide an international basis for comparison of Israeli networks, we construct similar networks for other countries that have generated a significant number of information security innovations in recent years.

We will use instances of "coinvention" – 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 "paper trail" left by coinvention.[11] To the extent that this omission introduces measurement error into our econometric specification, it may lead to a downward bias in the true incidence of spillovers channeled through inventor networks.

---

[9] "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 athttp://www.israel21c.org/high-tech-elites-to-nurture-arab-israeli-startups/

[10] These classes are reasonably broad, and contain within them patents that are not strictly information security inventions, *per se*. It was important for us to include all of information security classes as defined by the USPTO. Additionally, very narrowly defined fields have limited numbers of patents and make econometric work infeasible.

[11] While acknowledging this point, we argue that unmeasured communication and interaction is likely to be highly correlated in space and time with the coinvention episodes that we do observe in the patent data record.

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 occur whenever an inventor learns something from participating in 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 (FG 2011) show that whenever there are spillovers across open source software projects, there should be a positive correlation between project success and the *degree* of the project (the number of projects with which the focal project has a common developer). When there are both direct and indirect project spillovers, 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 thus measures how far each project is from all the other projects in the network. We formally define the relationship between the network centrality measures (*degree* and *closeness centrality*) and spillovers below.

2.1 An Example Constructing the Patent and Inventor Networks

Before we proceed, Figure 1 below provides a simple example in how to construct the patent and inventor networks in order to make the concepts more concrete. Suppose that there are six inventors and five patents with the following patent-inventor data:

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

The first sub-figure in figure 1 shows the two-mode network with both patents and innovators. The second sub-figure shows the "Inventor Network," where two inventors are connected if they work on a patent together. The third sub-figure is the "Patent Network." Two patents are connected if they have an inventor in common.[12]

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.

[12] Since we have very few inventors and patents, the Inventor and Patent networks look quite similar.

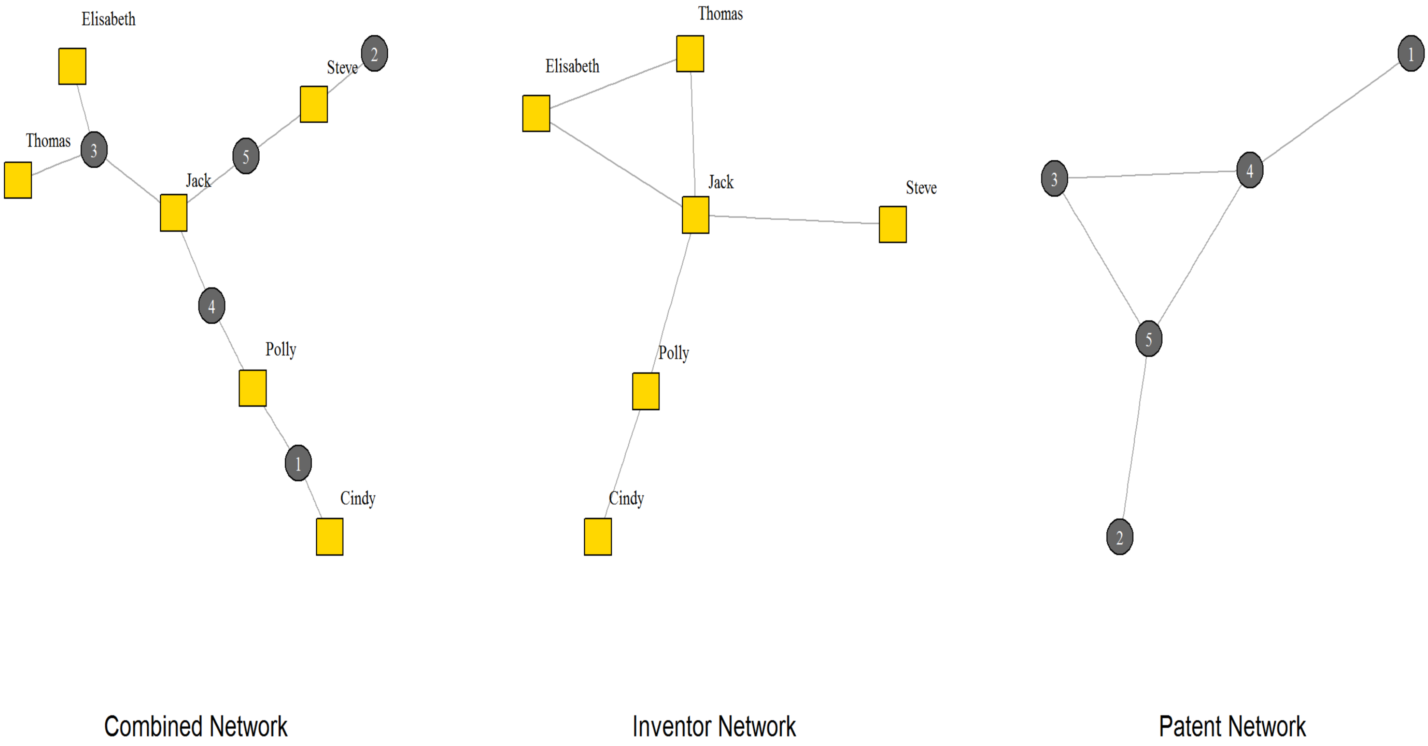

Figure 1: A Two-mode Network and Corresponding Inventor and Patent Networks

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 success level or impact (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 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 inventor(s), characteristics of the patent (including the number of backward citations), and characteristics of the firm holding the patent. We write:

$$(1)\ S_i = X_i\,\omega + \varepsilon_i$$

where ε_i is an error term.

The FG (2011) model shows how to measure the network ties that could become channels of knowledge spillovers. The model focuses on two network centrality measures: degree and closeness. We define two Israeli patents to be linked if they have an inventor in common. For each "Israeli" patent (denoted "i"), we calculate the (i) cited patent's "Israeli network" degree, which is the number of Israeli patents with which the focal patent has a direct link (i.e., an inventor in common) and (ii) the cited patent's "Israeli network" closeness centrality. A patent is an "Israeli" patent if it has at least one Israeli inventor, and an inventor is defined as Israeli based on the address associated with that inventor in a given patent document.[13] This means that an ethnically Irish inventor resident in Israel at the time of a patent application is counted as Israeli for our purposes. Conversely, an Israeli citizen temporarily assigned to the Silicon Valley lab of her multinational employer would be considered "American" if she lists her California address on the patent application.

Inspection of the data reveals that the vast majority of information security patents possess inventor teams who were all resident in the same country at the time of the patent application. Nevertheless, a nontrivial number of "Israeli" patents include foreign inventors (that is, inventors who possessed a non-Israeli address at the time of the patent application), and nearly half of our "pure" Israeli patents are "assigned to" (that is, owned by) a U.S.-based multinational. Our baseline regressions are conducted using the subsample of Israeli patents that have all-Israeli inventor teams regardless of whether the assignee is Israeli or not.

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

[13] We reiterate that our inference is based on U.S. patent grants to Israeli inventors, not patents held in the State of Israel. For the Israeli information security industry, the U.S. is the most important market for products and services, and there are strong incentives to patent new innovations in the U.S. The home market of Israel accounts for a small fraction of total revenues for most established Israeli firms in this sector.

centrality measures how far each patent is from all the other patents in a network and is calculated as:

$$(2) \qquad C_i \equiv \frac{(N-1)}{\sum_{j \in N} d(i,j)},$$

where N is the number of patents and $d(i,j)$ is the distance between Israeli patents i and j, as measured by the network of coinventions traced out in patent documents. For two Israeli 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, $d(i,j) = 2$. Patents that indirectly link other patents have a higher closeness centrality measure than patents near or at the edge of a network. (See Freeman (1979), pp. 225-226.)

The model assumes that (a) each Israeli patent "i" may receive a positive spillover denoted β from all "connected" Israeli patents, and (b) that a patent may enjoy positive spillovers from patents that are indirectly connected, but (c) that these spillovers are subject to decay that increases linearly as the distance between the patents in the patent network increases. When the distance between patent i and j is $d(i,j)$, this spillover is $\gamma / \sum_j d(i,j)$. Under these assumptions, the success level of each Israeli patent i can be written

$$(3) \qquad S_i = X_i\omega + \beta D_i + \gamma / \sum_j d(i,j) + \varepsilon_i.$$

D_i is the *degree* of patent i in the Israeli network, and β and γ are greater than or equal to zero. As noted, S_i is the number of forward citations received by the patents. Using (2), the expression for closeness centrality, patent k's success can be rewritten as

$$(4) \qquad S_i = X_i\omega + \beta D_i + \gamma C_i / (N-1) + \varepsilon_i.$$

This spillover specification is simple but quite general. When β and γ equal zero, there are no spillovers at all. When $\beta>0$ and $\gamma=0$, there are only direct spillovers. When $\beta=0$ and $\gamma>0$, there are both direct and indirect spillovers which are exclusively measured by the patents' closeness centrality. When $\beta>0$ and $\gamma>0$, there are additional spillovers from directly connected patents above and beyond those captured by its *closeness* measure: the spillovers have a "hyperbolic" structure. Hence, the theoretical model shows that spillovers depend on the network structure and that they can be measured by constructing the network linking the patents.

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. This abstracts from the possibility that some Israeli inventors might seek a (temporary) foreign posting for the explicit purpose of building a transnational network of collaborators, and we know that foreign postings are occasionally sought for exactly that reason. However, the well-documented challenges that arise in disambiguating similar names in patent documents become even greater when we attempt to track inventors' movements across national boundaries. Our focus on purely "intranational" networks in this study is also motivated, in part, by the importance assigned to intranational Israeli networks in conventional explanations of Israeli high-tech success.[14]

We also note that we construct patent networks independent of the calendar time in which patents are created. In other words, the degree and closeness of patent *i* reflect both the coinventions that occurred before patent *i* was created and the coinventions that occur after it is granted. This is intentional – the number of citations a given patent receives may reflect, in part, the evolution of the inventor network that takes place after it appears. On the other hand, while our language in this paper often implies the existence of a causal relationship between the strength of the network upon which inventors could rely when creating a given patent and the quality of that patent, we acknowledge here a more complicated reality: our measure of quality will be impacted by the evolution of the network after a given patent appears. Since a successful patent may lead to more partnerships for the inventors *ex post*, the relationship between network strength and patent quality probably reflects causal mechanisms running in both directions. The goal of this paper is not to demonstrate one-way causality, but rather to statistically validate the strength and robustness of an association between network centrality and patent quality in Israel.

The discussion in this section describes the definition of "Israeli" networks for Israeli information security patents. In order to enable explicit international comparisons of the relative importance of network-mediated knowledge spillovers across different countries, we

[14] To the extent that the international components of these networks, which we consciously omit, are important determinants of Israeli success, we are only capturing the "local" impact of network-mediated knowledge spillovers with error. In order to include international components, we will separately examine Israeli patents with US Assignees and Israeli assignees separately.

construct similar networks for information security patents of countries with large numbers of cyber security patents (Korea, Japan, Taiwan, Finland, Canada and Germany.[15] In all cases, we consider only "local" networks – a topic to which we return below.

3. Data and Empirical Work

3.1 Defining and Delimiting Our Patent Populations

We now turn to our empirical work. In order to begin, we need to define the relevant information-security patent classes. From detailed examination of United States Patent and Trademark Office (USPTO) patent class descriptions, we were able to determine the patent classes relevant for information security innovations, broadly defined. These patent classes are shown in Appendix B.[16]

We then collected data from the USPTO on all patents granted in the relevant information security patent classes. In this data set, we know the number of forward citations, backward citations (citations made to previously granted patents), grant year, application year, location of inventor (hence we know whether the inventor(s) are Israeli), patent class and subclass, patent title and abstract, number of inventors, and the assignee (owner) of the patent. The number of U.S. information security patents by country for the years 1985-2014 is given in Table 1. Since there were very few information security patents in general (and virtually no Israeli patents) in these patent classes before 1985, we start with that grant year. In the 1985-2014 period, the USPTO issued approximately 340,000 "information security" patents in which all inventors are from the same country, and the patent documents contain information on all the variables we use in our empirical analysis. We employ this subset of our data for the empirical analyses described below.[17]

In the case of Israel, complete data exist for 4,431 USPTO patents with Israeli inventors in this period. That is, for these patents, all inventors had an address in Israel. There are 4,582 Israeli

[15] By large, we mean at least 1500 patents in the largest connected component. We exclude the US.
[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 research by Arora and Nandakumar (2013), 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. In the case of Israel, there are missing data for 3% of the patents. Further, there is no selection issue.

inventors listed on these patents and 11,085 inventor-patent pairs (since many patents have more than one innovator.) There are 1,578 patents with both Israeli inventors and inventors from other countries - primarily the US. We exclude these patents from the main analysis, since we want to focus on the local network.[18]

In a similar way, we create samples of information security patents generated in all nations, other than the United States, that create at least 1,500 information security patents over our sample period. In each case, we restrict ourselves to patents where all inventors have addresses that place them as resident in the country of interest. No other country has as high a fraction of foreign co-inventors as Israel.

The number of Israeli "information security" patents is small relative to the total number of such patents. Table 2 shows that Israeli patents as a proportion of all information security patents granted by the USPTO increased steadily over the 1985-2014 period, but remained a small percentage of the total. The conventional wisdom regarding Israeli information security patents is that they stand out in terms of quality rather than quantity.[19]

3.2 Construction of the Israeli Information Security 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 (Trajtenberg et al., 2009; Ventura, Nugent, and Fuchs, 2015; Fleming et al., 2016).[20] For the purposes of our study, we think of the use of recorded inventor names in USPTO data as raising two main issues, which we refer to as "false positives" and "false negatives."

[18] Not surprisingly, when we include these patents, the measured correlation between network density and invention quality attenuates. One interpretation of this outcome is that it is really the local networks that matter, in terms of propagating knowledge spillovers.

[19] 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.

[20] We thank Manuel Trajtenberg for generously providing us with the data and code he used in his prior work on Israeli inventor name disambiguation. Future versions of this paper will utilize and update these data and methods.

A **false positive** means that we identify a connection between two patents in the coinvention network, where this connection does not actually exist. A false positive occurs if two (or more) separate inventors have the same name, and we therefore infer more coinventions than actually take place. In order to reduce the potential for false positives, we drop inventors with 100 patents or more patents.[21] 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 measurement. In the case of the Israeli network, we individually examined the names of all information security patent holders with more than between 20 and 100 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 probability of false negatives, we standardize all inventor names in the following ways:

1. We use only lower case letters for the names
2. We remove leading and following spaces.
3. We replace all "-" symbols with spaces between names.
4. We remove 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 of coinventions omits important connects, we are underestimating the extent of the network and therefore the knowledge spillovers that may flow through them.

Like many empirical networks, the network of Israeli information security patents includes one large connected component with 1,903 patents and many, much smaller components. We refer to the large component as the "giant component." It is indeed very large when compared to the second largest component, which has 80 patents in it. Recall that degree is the number of patents that are connected to the relevant patent. Tables 3a and 3b show the distribution of degrees for Israel during the 1985-2014 period.

[21] 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.

We follow a similar procedure in constructing patent networks for the other countries (other than the U.S.) that have generated large amounts of information security patents. Since we will also explicitly compare Israeli patent networks to those of other countries, we also include data for these other countries with significant numbers of information security patents in the tables below. The distribution of degree is similar for all of these countries. (Recall that a patent of degree one has an inventor in common with one other patent, while a patent with degree 4 has an inventor in common with four other patents.)

The variables used in the analysis are:

- The Number of Forward Citations
- Number of Forward Citations "no self-citations" (excluding forward citations from the same inventor and same assignee)
- Grant Year
- Number of Backward Citations received by the Patent
- Number of Inventors on the Patent
- Degree
- Closeness (Giant Component Only)

Descriptive Statistics for the Israeli network appear in Tables 4a and 4b.[22]

Israel is unique among countries in that many of its patents have US assignees. Fully 45% of all Israeli patents have US assignees, while 53% of the patents in the giant component have US assignees. For comparison, 19% of Canadian patents in the Cyber Security patent classes have US assignee, whereas in the Canadian giant component, only 5% have US assignees. Additionally, 13% of German patents in have US assignees, while the percentage of US assignees for other countries with relatively large numbers of information security patents in the giant component is less than 5%. In the case of Israel, two large U.S.-based technology firms, Intel and IBM, are assignees for approximately 18% of Israeli patents in the relevant patent classes. No other firm holds more than 2% of these Israeli patents. In these patent classes, virtually all non-Israeli assignees are US assignees. Hence, we refer to either Israeli assignees or US assignees.[23] We use these data as well in the analysis.

3.3 Measuring Spillovers via Connected Networks in Israeli Information Security Sector

[22] Correlations among the variables in the giant component appear in Appendix A.
[23] Since the data are from the USPTO, we know whether the assignees are US or foreign entities.

In this section, we estimate equation the FG (2011) model by estimating equation (4) which we repeat below:

$$(4)\ \ S_i = \omega X_i + \beta D_i + \gamma \frac{C_i}{N-1} + \epsilon_i$$

In this specification, we postulate that patent networks play a dual role in expanding the number of citations received by a given patent, which is our measure of quality. First, patent networks, as measured by degree and closeness, 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 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. Given the way we have constructed patent networks, we cannot disentangle these separate effects, nor can we fully disentangle the extent to which especially high quality inventions are a product of the network that produced them versus an instigator of later network connections that enhance the number of forward citations they receive. In all of these cases, though, networks are a mechanism through which spillovers are propagated, and a regression of the citations received by a patent on the inventor network linkages that pertain to it will allow us to make inference about the aggregate incidence of these various spillover flows.

Citations are highly skewed; additionally, some of the independent variables (like degree and number of inventors) are also highly skewed. Hence, it makes sense to use logarithms and employ the log/log specification.[24] The term "ln" before the variable means natural log. In Table 5, we provide the results of regressing (4) on data for all Israeli patents, while in Table 6, we use only the subsample of patents contained within the giant component. In the analysis summarized in Table 5, we cannot include closeness, since the patents are in different (i.e., unconnected) components. Hence, in Table 5, we can only test for direct network spillovers from connections in the Israeli patent network. The dependent variable used in the regressions reported in Tables 5 and 6 is the natural log of forward citations excluding citations from the same inventor and assignee. The independent variables are the number of inventors on each patent, the number of backward citations, and the degree of the patent, where degree is the

[24] We use ln(Forward Citations + 1), since some of the patents do not have any forward citations. Similarly, we use ln(degree+1) and ln(Backward Citations +1) since degree and backward citations can also take on the value zero. An alternative negative binomial specification gives quite similar results.

number of patents with which the relevant patent has an inventor in common.[25] We include dummy variables for patent classes and grant year. In the case of the giant component, we include closeness as well. In all regressions, we include dummy variables for patent classes and grant year, but do not report these coefficient estimates, for reasons of space.

In column 1 in Table 5, we include all patents with purely Israeli inventor teams. We find that the estimated coefficient on "Degree" is positive and statistically significant (0.054, t=4.36.) Thus suggests that there are (direct) spillovers from connections in the Israeli patent network. In columns 2 and 3 we run the analysis separately for patents with (and without) US assignees. In column 2 (US assignees) the estimated coefficient on "Degree" is equal to zero (0.0041, t=0.23), while in column 3 (Israeli assignees), the coefficient is positive and statistically significant (0.083, t=4.90***). This result is interesting and suggests that there is a larger spillover for locally assigned patents. One possible interpretation is that US assignees have more stringent intellectual property policies that restrict the flow of knowledge outside of the firm, while in the tightly connected Israeli network (of innovators and assignees,) knowledge flows more freely. Another possible interpretation is that patents held by U.S. assignees reflect the input of a broader, international network that is not fully reflected in the network ties across Israeli inventors, and when we regress patent quality on only a subset of the relevant network, we get weaker results.

Since we only have forward citation data through 2014, in column 4, patents issued after (say) 2011 had a very short time horizon to receive citations. It is true that we control for grant year, and thus, partly control for the truncation of citations experienced by more recent patent cohorts, but it is nevertheless interesting to examine the case when we restrict attention to patents granted through 2011. In such a case, the estimated coefficient on "Degree" is twice as large as in column 1 and is statistically significant as well. (0.10, t=4.93.) In all of the specifications, the estimated coefficient on backward citations is positive and significant. The estimated coefficient on the number of innovators is not statistically significant. Hence, controlling for degree, the number of innovators listed on the patent does not affect the quality or success of the patent.

[25] The results are robust to the log/linear specification as well. Both the log/log and log/linear specifications have a similar adjusted R-squared (approximately 0.56). A linear/linear specification on the other hand has a very low adjusted R-squared (0.14.)

3.5 Testing for Direct and Indirect Network-Mediated Knowledge Spillovers

Closeness is only defined for patents within a component, so we can include this variable only when we restrict attention to patents in the giant component. Hence, we now focus on the giant component. This enables us to examine whether there are both direct and indirect spillovers. Column 1 in Table 6 shows the results when we include all patents in the giant component. The estimated coefficient on closeness (γ) is positive and significant (0.46, t=4.20***) in column 1, suggesting that there are both direct and indirect knowledge spillovers from "connections" in the giant component. Recall that the coefficient on closeness (γ) captures both direct and indirect spillovers. Since the estimate for β is not statistically significant from zero (0.0082, t=0.32), there are not hyperbolic spillovers.

In columns 2 and 3, we again divide the patents into US assignees (Column 2) and Israeli assignees (Column 3.) Table 6 shows that our results are similar to those in Table 5. For US assignees, the estimated coefficient on closeness (γ) is positive and significant at the 10% level (0.23, t=1.69*). In column 3, however, the estimated coefficient on closeness is much larger and more statistically significant (0.63, t=3.81***). This again suggests that there are stronger spillovers when both assignees and inventors are Israeli. The estimated coefficient on the number of innovators is not statistically significant. Hence, controlling for degree and closeness, the number of innovators listed on the patent does not affect the quality or success of the patent.[26]

When we restrict attention to patents issued through 2011, the estimate of γ is 1.00 (t=4.95***, N=1,058). When we restrict attention to patents issued by 2011, and run the regressions separately for US and Israeli assignees, the estimate of γ is 0.40 (t=1.52, N=537) for US assignees and the estimate of γ is 1.46 (t=4.42, N=521) for Israeli assignees. When we restrict attention to patents with at least one forward citation, the estimate of γ is 0.74 (t=4.38***, N=1007). Hence, the results on knowledge network spillovers are extremely robust.

3.6 Adding Information from the Innovator Network

[26] We again include dummy variables for patent classes and grant year in all regressions, but do not report the coefficient estimates for these variables.

As discussed earlier, in addition to the patent network generated by connections among inventors, there is also the related inventor network (see Figure 1). Here we add information from the inventor network to the analysis. We can do this in several ways:

Star Inventors

The most intuitive is to include a dummy variable 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 takes on the value one if the patent has a top one-percent innovator on the patent and zero otherwise. Using the top one percent is ideal because in the giant component, roughly half (about 45 percent) of the patents have such an inventor. Outside of the giant component, only eight percent of the patents have such an inventor. In our Israeli patent data, 77% of the inventors have one or two patents, while 10% have more than five patents.[27]

These measures are quite similar to the open source software data employed by Fershtman and Gandal (2011) in the case of open source software. Two percent of contributors in open source projects worked on five or more projects. In the giant component in the open source data, 50 percent of the projects had a contributor who worked on five or more patents, while outside of the giant component, only eight percent had a contributor who worked on five or more projects. Overall, 90% of the contributors in open source software worked on one or two projects.[28]

We now examine whether - controlling for the patent network structure induced by the inventors - such "stars" affect the success of the patent. We find that in the case of Israel, the presence of such stars does not affect the success of the patent beyond the effect from the network structure induced by the "stars." The estimated coefficient on a dummy variable for the presence of "star inventors" on a patent inventor team is not statistically significant. More importantly, the estimate of (γ) is unaffected by including this variable. The estimated coefficient on γ remains positive and statistically significant in as shown in column 4 in table

[27] The distribution of patents per inventor for those who have more than 10 patents are shown in the appendix.
[28] Similarly, Goyal et. al (2006) shows that in the 1990s, the giant component of connected economists - where a connection exists between two economists if they have written a paper together – consisted of 40% of all economists, while the second largest component was tiny. The average degree in the network was 1.68. Fully 1% of the authors have more than 10 links and some of them have 40 to 50 links. The 20% most-linked authors account for about 60% of all the links.

6. (0.45, t=4.41***).[29] Hence, this suggests that, controlling for "stars," we again find that there are both direct and indirect knowledge spillovers.

Robustness: Degree and Closeness from the Inventor Network

Up until this point in the analysis, we have included the network centrality measures from the patent network. We will now instead include network centrality measures from the inventor network. We first use the inventor network to calculate network characteristics for each inventor. We then look at the group of inventors who participate in each patent and define (for each patent) measures that capture the network characteristics of these inventors. For each patent, we construct a list of inventors and construct the following variables:

(i) Average degree of the inventors on a patent.

(ii) The average closeness centrality of the inventors on a patent.

The above variables differ respectively from the degree of a patent and the closeness centrality of a patent. For example, consider patent A with two inventors (denoted I and II), each of whom works on one other patent. This means that patent A has a (patent) degree equal to two. Further, suppose that inventor "I" also works on patent B, and that there are three other distinct inventors on patent B. Similarly, suppose that inventor II also works on patent C, and that there are again three additional distinct inventors on patent C. The "inventor" degree of inventor I equals four (since he/she participates with four other inventors on two different patents). Similarly, the inventor degree of "II" is four as well. Hence, the average inventor degree of patent A is four.

In the tightly connected Israeli network, there is a high degree correlation between (i) the patent degree & the average degree of the inventors on a patent and (ii) between the patent closeness & the average closeness centrality of the inventors on a patent. Hence, when we replace patent degree and closeness with the average degree of the inventors on a patent and the average closeness centrality of the inventors on a patent, the results are virtually unchanged: the estimate of γ is 0.48 (t=4.08) when we use "inventor" centrality measures versus 0.46 (t=4.20) when we use "patent" centrality measures. In both cases, the estimate of β is essentially zero:

[29] When we run the analysis for all patents, the addition of this variable again has no effect and the estimated coefficient on "Degree" is again positive and statistically significant (0.046, t=2.72.)

(It is 0.010, t=0.24, when we use "inventor" centrality measures versus 0.082, t=0.32, when we use "patent" centrality measures. Hence, our results are robust to using degree and closeness from the inventor networks.

3.7 International Comparisons of Network-Mediated Knowledge Spillovers

Israel is not the only country with expertise in information security. In this section, we compare Israel with other countries with information security expertise and large numbers of information security patents (at least 1500) in their respective giant components.[30] Hence, we estimate the log/log specification of (4) for the following countries: Korea, Japan, Taiwan, Canada, Finland and Germany. For each of these countries, we are estimating the same regression model applied in the first column of Table 6 to Israeli patents.[31] The key estimated coefficients for all seven countries are summarized and compared in Table 7.

The results in Table 7 suggest that there are very strong network spillovers for Israel, relative to other countries with large numbers of information security patents in the giant component. The network "spillover" functions are smaller for Korea, Taiwan, Japan, and Canada. We find virtually no spillover effects for Finland and Germany.[32] When we restrict attention to patents issued by 2011, the differences in "network knowledge" spillovers between Israel and the other countries is even larger.

Where are these differences coming from, and what do they mean? It seems that we can draw a meaningful distinction between geographically compact economies like Israel, Taiwan, and South Korea, and larger nations like Canada, Germany, and Finland, all of which possess a national territory many times larger than that of Israel. In the former three economies, national territory is limited and innovative activity is highly concentrated, even within that limited national territory. Prior literature confirms a high concentration of patenting activity in Tel Aviv, the Seoul Metropolitan Area, and Taipei/Hsinchu respectively. This tight concentration of inventive activity would seem to facilitate the formation of tight networks and intense interaction between inventors in the information security domain. In contrast, many geographers have noted that Canada is a nation closer to the United States than to itself.

[30] We exclude the US from this analysis.
[31] The results are robust to including the dummy variable for star programmers.
[32] For Germany, the estimate of γ is negative. We set it equal to zero in Table 8.

German economic (and patenting) activity is divided across a number of cities and regions, and Finland possesses a much larger territory than most non-Finns recognize. Among the larger economies, Japan seems to be the outlier – a relatively large country with relatively strong estimated network effects. These two facts can be partly reconciled by the fact that inventive activity in Japan tends to be highly concentrated geographically (in a handful of reasonably proximate major cities), organizationally (in a handful of large companies), and technologically (for Japanese firms, information security patents are highly concentrated in a small number of patent classes).

Even within the set of smaller economies, Israel stands out in terms of the magnitude of its estimated spillover effects. Although the differences in the network coefficients are not statistically distinguishable between Israel and the other smaller countries included in the analysis at conventional levels of statistical significance, the magnitude of its direct and indirect spillover effects is still larger than those of any other economy. This provides some empirical validation for the notion that the quality of Israeli invention is linked to the unusual strength of its inventor networks.

In the final column of Table 7, we compare the average quality of Israeli information security patents to those generated by other economies. To do so, we again regress the log of forward citations on the log of the number of innovators on the patent, the log of the number of backward citations, the log of degree, and dummy variables for grant year and patent class. Here we include all patents in the information security patent classes from all countries, as in Table 1. We employ dummy variables for all seven countries in Table 7 (the ones with at least 1500 patents in the giant component). The results are shown in the final column of Table 7. The results show that after controlling for the characteristics noted above, Israeli patents have more forward citations than patents from any other sample country.[33] Hence, in addition to stronger network spillovers in the Israeli patent network, Israeli patents are of higher quality in the information security realm.[34]

[33] The differences are statistically significant for all countries except Canada and Finland.

[34] In fact, Israeli information security patents are nearly equal in quality to those generated in America's most innovative region, California. We again run the same regression with dummy variables for Israel and California. Israeli patents have on average 19% higher quality, while California patents have on average 22% higher quality.

4. Conclusions and Next Steps

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 function as mechanisms for the transmission of knowledge spillovers.

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 – the information security technology. 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 tight networks that characterize Israeli inventors operating in this domain. These networks allegedly 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 cybersecurity inventions is systematically related to the location of these patents within the Israeli invention network. Furthermore, when we compare Israeli information security inventions to those generated in other countries, we find a stronger relationship between invention quality and network structure in Israel than in other sample countries.[35] These networks may help (in part) to account for the fact that Israeli cybersecurity invention quality is essentially as high as that created within Silicon Valley itself.

These initial results suggest a number of potentially useful directions for further research. While network ties among inventors appear to be strongly correlated with cybersecurity invention quality in Israel, we still know little about the genesis of these ties. Conventional

[35] We acknowledge, though, that these differences are not always statistically significant at conventional levels.

wisdom points to the importance of military service within elite groups like Unit 8200, but no large-sample statistical study has formally tested this popular belief. However, it is possible, in principle, to measure the importance of veterans of Unit 8200, and other elite Israeli Defense Force units, as central nodes within these networks. Increasingly, veterans openly acknowledge their prior ties to these once secret units, and even list their service as a professional credential on social networks like LinkedIn. In future work, we will seek to use these data to probe the importance of the Israeli military as a source of network ties and a driver of cybersecurity invention quality.

Rapid development of machine learning and text mining techniques, applied to patent data, provide another interesting path forward. Gandal, Naftaliev, and Stettner (2017) were able to the movement of specific bits of software code across open source projects, and could therefore separately measure the network connections between inventors (and projects) as well as the movement of specific ideas and techniques across these projects. In principle, text mining and machine learning techniques could recognize particular cybersecurity techniques and technologies, as revealed by the text of patent documents, allowing us to track the movement and evolution of these ideas across patents, in both space and time. This would provide a measure of knowledge flows that is independent of the network, but plausibly influenced by it, allowing for a richer and more direct test of the idea that denser networks really do enhance the diffusion and evolution of useful knowledge.

Finally, our measures of network density are deliberately designed to be time invariant in this paper, but the reality is that the inventor and patent networks evolve over time in ways that we can track in our data. Allowing the networks to evolve temporally may enable us to better distinguish between the idea that denser networks create better ideas from the notion that better ideas create a denser network. As is usually the case in economics, much remains to be done.

References

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

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 (2): 571-591.

Ahuja, G. 2000. Collaboration Networks, Structural Holes, and Innovation: A Longitudinal Study. Adm. Sci. Q. 45(3) 425–455.

Alcacer, J. and M. Zhao, 2012. Local R&D Strategies and MultiLocation Firms: The Role of Internal Linkages, Management Science 58(4): 739-753.

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 H. Chesbrough and R. Burgelman (eds.) *Research in Technology and Innovation Management, 7,:57-84.*

Arora, A. and A. Nandkumar, 2012, “Insecure Advantage? Markets for Technology and the Value of Resources for Entrepreneurial Ventures,” *Strategic Management Journal*, 33(3) 231–251.

Ballester, C., A. Calvó-Armengol, Y. Zenou. 2006. Who’s who in networks. Wanted: the key player. *Econometrica* 74(5) 1403–1417.

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), pp. 429-434.

Berry, H., 2014. "Global Integration and Innovation: Multi-Country Knowledge Generation within MNCs," *Strategic Management Journal* 35(6): 869-890.

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*, vol. 68, pp. 325-344.

Branstetter, L., and Saggi, K., 2011. "Intellectual Property Rights, Foreign Direct Investment, and Industrial Development." *Economic Journal*, vol. 121, no. 555, pp. 1161-1191.

Branstetter, L., and Sakakibara, M., 2002, “When Do Research Consortia Work Well and Why? Evidence from Japanese Panel Data.” *American Economic Review*, 92, (1), 143-159.

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

Calvó-Armengol, A., M. O. Jackson. 2004. The effects of social networks on employment and inequality. Am. Econ. Rev. 94(3) 426–454.

Calvó-Armengol, A., E. Patacchini, Y. Zenou. 2009. Peer effects and social networks in education. Rev. Econ. Stud. 76(4) 1239–1267.

Fershtman, C., & Gandal, N. 2011. Direct and indirect knowledge spillovers: the “social network” of open-source projects. The RAND Journal of Economics, 42(1): 70–91.
(16) Freeman, L. 1979. Centrality in social networks: Conceptual clarification. Soc. Netw. 1(3) 215–239.

Freeman, L. 1979. Centrality in social networks: Conceptual clarification. Social. Networks 1(3) 215–239.

Gandal, N., and U. Stettner, 2016, “Network Dynamics and Knowledge Transfer in Virtual Organizations, International Journal of Industrial Organization, 48: 270-290.

Gandal, N., Naftaliev, P., and U. Stettner, 2017, “Following the Code: Spillovers and Knowledge Transfer,” CEPR Discussion Paper, DP11851.

Grossman, G. and E. Helpman, 1991. *Innovation and Growth in the Global Economy*. MIT Press, Cambridge.

Irwin, D. and P. Klenow, 1994. "Learning-by-Doing Spillovers in the Semiconductor Industry," *Journal of Political Economy* 102: 1200-1227.

Jaffe, Adam B., B. Gomes-Casseres and John Hagedoorn, 2006. "Do alliances promote knowledge flows?." Journal of Financial Economics 80. 1: 5-33.

Goyal, S., M. J. Van Der Leij, J. L. Moraga-González. 2006. Economics: An emerging small world. J. Polit. Econ. 114(2) 403–412.

Grewal, R., G. L. Lilien, G. Mallapragada. 2006. Location, location, location: How network embeddedness affects project success in open source systems. Manag. Sci. 52(7) 1043–1056.

Griliches, Zvi, 1979. "Issues in Assessing the Contribution of R&D; to Productivity Growth," The Bell Journal of Economics, vol. l0(l), 92-116. Also reprinted in E. Wolff, ed., The Economics of Productivity, vol. I, Cheltenham: Elgar, 1997, 256-80.

Griliches, Zvi, 1992. "The Search for R&D Spillovers," The Scandinavian Journal of Economics, vol. 94, 1992, Supplement 29-47.

Jackson, M. O., L. Yariv. 2007. Diffusion of behavior and equilibrium properties in network games. Am. Econ. Rev. 97(2) 92–98.

Adam B. Jaffe, 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., Trajtenberg, M., and R. Henderson. 1993. “Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations,” Quarterly Journal of Economics 108 (3): 577-598.

Jaffe, A. and M. Trajtenberg, 2002, “Patents, Citations, and Innovations: A Window on the Knowledge Economy, MIT Press, Cambridge.

Jones, B. F, 2014, "The knowledge trap: human capital and development reconsidered" (No. w14138). National Bureau of Economic Research.

Karlan, D., M. Mobius, T. Rosenblat, A. Szeidl. 2009. Trust and social collateral. Q. J. Econ. 124(3) 1307–1361.

Keller, W. and S. Yeaple, 2013. "The Gravity of Knowledge," American Economic Review, June 2013.

Keller, W., 2004. "International Technology Diffusion," Journal of Economic Perspectives.

Keller, W., 1998. "Are International R&D Spillovers Trade-Related? Analyzing Spillovers Among Randomly Matched Trade Partners," European Economic Review.

Romer, P., 1986. "Increasing Returns and Long-Run Growth," Journal of Political Economy, 94 (5), pp. 1002-1037.

Romer, P., 1990. "Endogenous Technological Change," Journal of Political Economy, 98 (5), pp. S71-S102.

Scherer, F. M., 1982. "Inter-industry Technology Flows and Productivity Growth," *The Review of Economics and Statistics*, 64 (4), pp. 627-634.

Senor, D., Singer, S., 2009, "Start-up nation: The story of Israel's economic miracle". McClelland & Stewart.

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", Harvard University Press, Cambridge, MA.

Trajtenberg, M., G. Shif, and R. Melamed., 2009. "The Names Game: Harnessing Inventors, Patent Data for Economic Research," Annals of Economics and Statistics, No. 93/94 (2009), pp. 79-108.

Ventura, S., Nugent, R., and Fuchs, E. 2015. Seeing the Non-Stars: (Some) Sources of Bias in Past Disambiguation Approaches and a New Public Tools Leveraging Labeled Records. Research Policy. Special Issue on Data. 44(9): 1672-1701.

Table 1
US Information Security Patents by Country for 1985-2014

	Country/State	% of Patents
Israel	4,431	1%
Korea	17,799	5%
Taiwan	8,200	2%
Japan	64,618	19%
Canada	8,057	2%
Finland	3,497	1%
Germany	10,472	3%
USA	190,392	56%
Other Countries	32,062	9%
Total	339,528	100%

The table presents the number of information security 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.

Table 2
Israeli Information Security patents 1985-2014

	(1)	(2)	(3)
	# of Israeli Patents	**# of All Patents**	**% of Israeli Patents**
1985-1989	32	11,253	0.28%
1990-1994	71	16,417	0.43%
1994-1999	256	36,492	0.70%
2000-2004	554	54,745	1.01%
2005-2009	980	82,732	1.18%
2010-2014	2,538	137,889	1.84%
Total	4,431	339,528	1.31%

Column 1 presents the number of Israeli information security patents that were granted at each five-year period, between 1985 and 2014. Column 2 presents the number of information security patents issued by all countries, 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 its inventors home addresses were listed under that country, according the USPTO data. Patents for which we cannot associate a specific country are not included.

Table 3a
Distribution of Degree – All patents granted between 1985-2014

Degree	Israel		Taiwan		South Korea		Japan		Canada		Finland		Germany	
	Number	%	Number	%	Number	%	Number	%	Number	%	Number	%	Number	%
0	680	15%	2,163	26%	2,686	15%	7,108	11%	1725	21%	513	15%	2589	25%
1	471	11%	1,144	14%	1,540	9%	5,312	8%	955	12%	358	10%	1434	14%
2-3	677	15%	1,285	16%	2,134	12%	7,522	12%	1202	15%	497	14%	1670	16%
4-5	457	10%	788	10%	1,378	8%	5,324	8%	724	9%	374	11%	1087	10%
6-9	668	15%	909	11%	2,291	13%	9,175	14%	947	12%	674	19%	1428	14%
10+	1,478	33%	1,911	23%	7,770	44%	30,177	47%	2504	31%	1081	31%	2264	22%
Total	4,431	100%	8,200	100%	17,799	100%	64,618	100%	8057	100%	3497	100%	10472	100%

The table presents the distribution of degree in the patent network, in the relevant country.

Table 3b

Distribution of Degree - Giant Component - patents granted between 1985-2014

	Israel		**Taiwan**		**South Korea**		**Japan**		**Canada**		**Finland**		**Germany**	
Degrees	Number of Patents	Percentage	Number of Patents	Percentage	Number of Patents	Percentage	Number of Patents	Percentage	Number of Patents	Percentage	Number of Patents	Percentage	Number of Patents	Percentage
1	54	3%	113	4%	397	3%	1480	3%	35	2%	110	5%	197	5%
2-3	160	8%	337	11%	1002	8%	4017	9%	116	5%	273	12%	552	13%
4-5	185	10%	320	11%	1041	9%	3957	8%	129	6%	256	11%	601	14%
6-9	309	16%	549	19%	2081	17%	8026	17%	300	13%	569	26%	948	22%
10>	1195	63%	1620	55%	7616	63%	29693	63%	1711	75%	1022	46%	1959	46%
Grand Total	1903	100%	2939	100%	12137	100%	47173	100%	2291	100%	2230	100%	4257	100%

The table presents the distribution of degree for patents in the giant component of each country.

Table 4a

Descriptive Statistics Israel – All Patents

	# of Obs.	Mean	Std. Dev	Minimum	Maximum
Forward Citations	4431	12.19	38.26	0	1017
Forward Citations – “No self cites”	4431	11.21	36.46	0	944
Grant Year	4431	2008.58	5.53	1985	2014
# of inventors	4431	2.42	1.49	1	19
Degree	4431	14.76	24.12	0	168
Backward Citations	4431	20.88	53.51	0	547

The table presents descriptive statistics for the all the Israeli information security patents, granted between 1985 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 a patent was approved by the USPTO. Number of inventors are the number of inventors listed as the patent inventors. Degree is the number of patents that are directly connected to the patent in the patents' network. Backward citations are the number of patents that were cited by the patent. 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.

Table 4b
Descriptive Statistics Israel – Giant Component

	# of Obs.	Mean	Std. Dev	Min	Max
Forward Citations	1903	10.25	39.96	0	1017
Forward Citations – “No self cites”	1903	8.99	37.21	0	944
Grant Year	1903	2009.36	4.70	1986	2014
# of inventors	1903	2.70	1.51	1	11
Degree	1903	28.90	31.05	1	168
Backward Citations	1903	29.81	76.42	0	547
Closeness (Giant Comp Only)	1903	0.0000646	0.0000147	0.0000311	0.0000927

The table presents descriptive statistics for the all the Israeli patents that are in the Giant component and were granted between 1985 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. Degree is the number of patents that are directly connected to the patent in the inventors' network. Backward citations are the number of patents that were cited by the patent. 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.

Table 5
Regression analysis for all patents

	(1)	(2)	(3)	(4)
Dependent Variable	ln(Forward_Citations)	ln(Forward_Citations)	ln(Forward_Citations)	ln(Forward_Citations)
	All Patents	All Patents	All Patents	Patents issued through 2011
		US assignees	Israeli assignees	
Independent Variables				
ln(# of Inventors)	0.0058 (-0.25)	0.030 (0.88)	0.041 (1.22)	-0.12 (-0.32)
ln(Backward Cites)	0.082 (6.42)***	0.083 (4.62)***	0.062 (3.53)***	0.11 (5.14)***
ln(Degree)	0.054 (4.36)***	0.004 (0.23)	0.083 (4.90)***	0.10 (4.93)***
Dummy variables for Patent Class	YES	YES	YES	YES
Dummy variables for Grant Year	YES	YES	YES	YES
Adjusted R squared	0.61	0.62	0.62	0.42
Observations	4,431	1,976	2,455	2,604

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. Backward Cites is one plus the number of citations made by the patent. Degree is the number of patents which are connected to the patent in the Israeli patent network. In column (1) we regress on all Israeli patents. In column (2) we restrict our sample to Israeli patents with US assignees. In column (3) we run the same regression only for Israeli patents with Israeli assignees. In column (4) we restrict our sample to patents that were granted until 2011. All specifications include patent class fixed effects, as well as grant year fixed effects. T-statistics appear in the parentheses.

*=significant at 10% level, **=significant at 5% level, ***=significant at 1% level.

Table 6
Regression Analysis for patents in the Giant Component

Dependent Variable	ln(Forward_Citations)	ln(Forward_Citations)	ln(Forward_Citations)	ln(Forward_Citations)
	Giant Component	Giant Component	Giant Component	Giant Component
		US Assignees	Israeli Assignees	
Independent Variables				
ln(# of Inventors)	-0.017 (-0.47)	0.059 (1.25)	0.013 (0.23)	-0.024 (-0.66)
ln(Backward Cites)	0.064 (3.55)***	0.067 (2.82)***	0.060 (2.12)**	0.065 (3.61)***
ln(Degree)	0.0082 (0.32)	-0.045 (-1.43)	-0.011 (-0.29)	0.037 (0.96)
ln(Closeness)	0.46 (4.57)***	0.23 (1.69)*	0.63 (3.81)***	0.45 (4.41)***
Star Innovator				-0.072 (-0.99)
Dummy variables for Patent Class	YES	YES	YES	YES
Dummy variables for Grant Year	YES	YES	YES	YES
Adjusted R^2	0.57	0.59	0.58	0.57
Observations	1,903	1,006	897	1,903

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. Backward Cites is one plus the number of citations made by the patent. Degree is the number of patents which are connected to the patent in the Israeli patent network. In column (1) we regress on all Israeli patents in the Israeli giant component. In column (2) we restrict our sample to Israeli patents in the giant component with US assignees. In column (3) we run the same regression only for Israeli patents in the giant component with Israeli assignees. In column (4) we restrict our sample to Israeli patents in the giant component that were granted until 2011. All specifications include patent class fixed effects, as well as grant year fixed effects. T-statistics appear in the parentheses.

*=significant at 10% level, **=significant at 5% level, ***=significant at 1% level.

Table 7
International Comparison

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Country	Information security Patents in Giant	Estimate of β	Estimate of γ	**Direct** (Neighbor) Spillover (β+ γ)	**Indirect** (Two-step) Spillover (γ/2)	Estimate of γ patents issued by 2011	% fewer citations than Israel
Israel	1,903	0.01	0.46***	0.47	0.23	1.00***	
Korea	12,137	0.02***	0.25***	0.27	0.125	0.39***	-25%
Taiwan	2,939	-0.02	0.23**	0.21	0115	0.31***	-34%
Japan	47,173	0.04***	0.15***	0.19	0.075	0.17***	-37%
Canada	2,291	-0.02	0.18*	0.16	0.09	0.28	-6%
Finland	2,230	0.04	0.04	0.08	0.02	0.00	-4%
Germany	4,257	0.00	0.00	0.00	0.00	0.00	-48%

The table compares the direct and indirect knowledge spillover effects between different countries. We estimate the specification in Table 7, column (1), for all patents in the giant component, in countries that have more than 1500 patents in their giant component. In column (1) of Table 8, we show the number of patents each country has in its giant component. In column (2) we show the estimate of β, the direct spillover premium, in column (3) we show the estimate of γ, the knowledge spillover parameter. In column (4) and (5) we show the estimated direct and indirect effects. In column (6) we show the estimate for γ using patents that were issued by 2011. In column (7,) we estimate the specification in Table 6, column 1, using all patents (from Table 2) and dummy variables for countries in Table 8.

* = significant at 10% level, ** = significant at 5% level, *** = significant at 1% level.

Appendix A:

Correlation among Variables – Giant Component (Israel) N=1903

	Fwd_cites	back_cites	inventors	Degree	Closeness
Fwd_cites	1				
back_cites	-0.07	1			
Inventors	-0.05	0.04	1		
Degree	-0.004	0.32	0.16	1	
Closeness	0.03	0.19	0.01	0.62	1

Appendix B: Relevant Patent Classes for Information Security:[36]

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.

[36] See https://www.uspto.gov/web/patents/classification/uspc726/defs726.htm, accessed 25 June 2017.

Appendix C: Distribution of Inventors who hold more than ten Patents