The Unintended Impact of International R&D Networks on Innovation.*

Justin Hicks Ph.D.
jhicks@pacific.edu
University of the Pacific

November 12, 2013

*I would like to thank Alex Whalley for his advice and continuous support. I would also like to thank Shawn Kantor, Robert Innes, Todd Neumann and Katie Winder of the University of California, Merced for their invaluable comments and suggestions given during the development of the project. In addition, I would like to thank Bruce Weinberg, Trevon Logan, Joel Blit, Adam Storeygard, Grid Thoma, Megan MacGarvie, and Ajay Agrawal for their comments and suggestions, and advice on data usage. I would like to thank conference participants at the Western Economics Association annual meetings especially Mohammad Ashraf for his comments and suggestions. I would like to thank Trinidad Beleche, Deniz Baglan and Daniel Farhat for their comments and contributions and support. Finally, I want to thank seminar participants from the School of Social Sciences, Arts and Humanities at UC Merced including Tom Hansford for their comments and suggestions. All errors are my own.

Abstract

Networking and collaboration has increased dramatically in recent years; especially in R&D. R&D networks take many forms based upon their intended purpose, but their importance in spurring further innovation is unclear. Networks may enable knowledge flows, but worries over leakage of proprietary knowledge can mitigate them. In order to gain insight, I consider the impact of natural disasters on international R&D networks. This particular setting allows me to observe an exogenous collapse of the R&D network that only impacts one of the collaborators directly. I then compare the innovative output of the unaffected network partners before and after the shock, to infer how important the knowledge flows across the network are in the production of new innovation within the firm. I use European Patent Organization data at the firm-level for OECD countries from 1989-2008 as a measure of innovative output. In sum, my analysis reveals no broad based effects of network partner access on innovation. I do find that disruptions to networks with international subsidiaries, where any competition effects are likely to be small, do reduce patenting by the home firm. A disaster impacting a subsidiary based network implies a 10-20% negative impact on home-firm innovation output over the following three years.

Keywords: Innovation, Networks, International
JEL Classifications: O3, L2, F2

1 Introduction

Today, in the wake of the wake of the Great Recession, firms are finding it tougher than ever to remain profitable. Given a monopolistically-competitive setting, unless a firm innovates, the firm will cease to earn profits. Without innovation by firms, macroeconomic growth will not occur and standards of living will not increase or at least regain what has been lost (Romer, 1990). Generally, innovation occurs at the culmination of human capital and knowledge accumulation. Therefore, only once a relatively complete comprehension of the existing body of knowledge in a firm's industry is acquired can innovation occur. This is referred to as the burden of knowledge, and the burden is increasing over time (Jones, 2009). These combined facts have compelled firms to push the boundaries when sourcing knowledge.

One way that firms have reacted to the increases in the burden of knowledge and the recession is through the increasing use of teams in research (Jones, 2009)(Wuchty et al., 2007). By hiring individuals who have highly specialized knowledge and bundling complementary knowledge sets together, firms assemble teams that are able to innovate. However, this leads to increased competition between firms over these specialists and their knowledge. This, in combination globalization and the internet, has resulted in the market for ideas becoming a global one. Firms who intend to remain profitable are now looking abroad to find new ideas and individuals who can drive their firms forward with new innovation. This has lead to a dramatic increase in international R&D networks (Narula and Santangelo, 2009).[1]

One reason for the push towards international knowledge networks is the diversity of knowledge available abroad. This fact is substantiated by the work relating the impact of highly skilled immigrants on native innovation by Kerr and Lincoln (2010), Hunt and Gauthier-Loiselle (2008), and Maskus et al. (2010). All three studies show the high potential for knowledge spillovers and increased innovative productivity associated with knowledge flows from abroad with estimates showing that a 1% increase in highly skilled immigrants increases innovative output by approximately 15%.

Rather than attempting to attract these highly skilled individuals, firms are actively seeking them on their own turf. By creating networks, firms open their doors directly to the probable

[1] Griffith et al. (2011) show that distance and agglomeration economies are becoming less important except in a few industries such as pharmaceuticals where both laboratory and research locations (hospitals) must all be co-located. Agrawal and Goldfarb (2006a) shows that decreasing costs of communication has facilitated gains from trade through the specialization of research tasks.

sources of new knowledge. These networks generally take one of two forms. First, a firm may choose to network by creating a subsidiary abroad. Second, the firm may choose to collaborate with a foreign firm. These two settings have very different implications on knowledge flows within the network, yet the relative costs and benefits to firms innovative productivity has yet to be explored.[2] One reason for the lack of prior research is that the measurement of the causal impact of networking on productivity in any form is a difficult task. Because the choice to enter into a network is endogenous to the profit maximization problem, those that do choose to network may be better suited to either invest abroad or work with other firms. This will make a simple comparison of innovative productivity between those firms that network and those that do not, an inappropriate approach. A plausible source of exogenous variation in networks must be found.

In order to best measure the causal effects of international R&D networks on home-firm innovative productivity, I implement a difference-in-differences (DD) identification strategy. After selecting the top 1000 innovative firms (the *home-firms*) within the OECD, I identify existing networks by considering co-patenting behavior from 1989-1998. I then look at the impact of large-scale natural disasters that occur in the network-partner's country from 1999-2008 on home-firm patenting productivity. This shock effectively shuts down the existing international R&D network for a time. In this way, I emulate the studies of Azoulay et al. (2010) and Waldinger (2011). However, I am able to expand the analysis by considering directly how competition over ideas effects knowledge flows and the relative importance of networks as inputs.

My findings point to the fact that competition within networks is highly correlated with the long term importance of the network in home-firm innovative production. Specifically, I find that when a network is comprised of multiple independent firms, there are no negative effects on home-firm patent output when the network is potentially shut down by a disaster. However, when the network is comprised of a firm and and international subsidiary, then a disaster has significant negative effects on home-firm patenting. Over the following three years, home-firm patenting drops on the order of 10 to 20%, which reveals two things. First, subsidiaries are key inputs in the home-firm's patent production function, rather than being utilized for a single project or other purpose. Second, because there is no competition over potential R&D discoveries within a subsidiary based network, all potential knowledge is captured within the network, whereas in the between-firm networks competition mitigates this effect.

[2]Bloom et al. (2007) show that product market rivalry can lead to a secondary spillover that actually hurts firms.

My results can explain the contrary results that exist in recent literature measuring potential spillovers from networks. Azoulay et al. (2010) attempt to pin down the effects of networking on innovative productivity within academia. By considering the unexpected death of a research partner in an academic atmosphere, they find evidence of long-lasting knowledge spillovers. However, using a similar identification strategy, Waldinger (2011) finds little evidence to support this fact. He shows that the unexpected removal of high-quality peers in response to early laws passed by the Nazi regime have no significant effect on the productivity of those who remain in academic research positions. These studies consider the breakdown of an existing network with mixed results. So, Borjas and Doran (2011) consider what happens when a potential network partner enters an existing research environment from abroad. He finds that the unexpected insurgence into the US of Soviet mathematicians after the breakup of the USSR actually leads to significantly lower levels of innovative productivity amongst US researchers in similar fields. These mixed findings highlight the importance of the increasing burden of knowledge and competition within the market for ideas in the productivity levels of incumbent researchers.

As the burden of knowledge increases, firms will use knowledge sourcing networks to bolster innovative activity. However, as competition over ideas increases, returns to networking will decrease. It is the fact that these effects work in opposite directions that can explain why we see such conflicting results in the prior literature. It is possible that in a given setting that either of these two effects will dominate, or they may cancel each other out. In this study I look to shed light on this open question, by separating the two effects and then identifying the causal impact of international networks on home-firm innovative productivity.

At the macroeconomic level, the results have bearing on policy maker's incentive programs such as the National Science Foundation's Partnerships for International Research and Education program (PIRE) which funded over $55 million directly towards projects that were international in nature. However, a much larger endeavor is the European Commission's Framework Research Programs (FP7). The FP7 funds 32 billion Euros directly for the purpose of incentivizing collaboration amongst researchers from 2007-2013. If competition over ideas overshadows the positive impacts of overcoming the burden of knowledge, then policy makers will need to adjust their incentive programs accordingly. At the microeconomic level, for firms it is clear that innovation and patenting is key to profitability. In a recent publication by Siemens AG, it was stated that the patents they hold are their most important asset; and "To maintain an edge in innovation on the international stage, companies need to utilize their global knowledge networks as effectively as possible" (Webel, 2011).[3]

[3]Bloom et al. (2005) show that social spillovers are about twice as large from R&D than private, thus generating

I identify networks that exist between the home-firm and subsidiaries as well as those that exist with other firms. Subsidiaries are an extension of the home-firm itself, and thus there will not be any competition over new knowledge. However, networks built between firms will potentially have a coordination problem where both firms seek to exploit the network for its own interests. With my identification strategy I am able to look at the analogue to Azoulay et al. (2010) and Waldinger (2011) as well as stratify amongst these different types of networks to get an idea of how important competition over ideas is. If these networks are generally key inputs in patent production at home, then the DD estimate will accurately measure their importance. However, if competition has a significant impact, then additionally stratifying of the disasters between subsidiary networks and between-firm networks will allow for further understanding into the causal impact of networking on innovative productivity.[4] [5]

The rest of the paper will proceed as follows. In Section 2, I will describe the experimental setting, describe the data and explain the identification strategy. In Section 3, I will specify the econometric model, present results and conclude with checks of robustness and sample sensitivity. Finally in Section 4, I will conclude.

2 Setting, Data, and Identification Strategy

Before moving into the analysis, a brief discussion of the incentives faced by firms who source knowledge abroad is worth while. All firms who source knowledge abroad face costs. Firms who use wholly owned subsidiaries, pay all the sunk costs associated with the R&D process. This

further evidence that if incentives are well aligned, government support should be used.

[4]Networks that exist between firms is the similar to Borjas and Doran (2011) immigration effect on networking. However in Borjas and Doran (2011), the incumbent researchers have no choice in the entrance of competitors. In my study, these relationships do exist based upon mutual choices of both firms. However, even within this setting, competition and coordination is potentially a severe problem.

[5]Aghion et al. (2009) show similar effects to Borjas, in that new entry of potential partners/competitors has differential effects depending upon the level of competition over knowledge that exists between the players. Singh (2007, 2008), Song et.al. (2011), and Blit (2011) all explore what drives collaboration between firms and the flows of knowledge between firms, specifically citing knowledge sourcing as playing a key role. However, none explore causality between networks and innovative productivity (Blit, 2010)(Singh, 2008)(Singh, 2007)(Song et al., 2011)(Kim and Song, 2007). Abramovsky et al. (2005) show clearly that competition, appropriability and government incentives play a key role in whether or not firms take part in international R&D networks. For a nice review of the theoretical underpinnings for international collaboration see Veugelers (1998) or Audretsch et al. (1996). In a recent study Griffith et al. (2006) do find significant evidence of spillovers between firms that network between the US and UK.

type of network is built from the ground up, including all capital and labor costs. In addition, all cultural and market characteristics in the international setting must be accounted for by either learning over time, or paying additional workers with local specific knowledge.[6] The benefit to this tactic, is that all knowledge sourced by the subsidiary will flow freely back to the home-firm. This occurs because their are well-aligned profit incentives across the network. Therefore, any potential knowledge spillovers are captured by the home-firm.

On the other hand, firms that partner with other firms abroad are able to share the substantial sunk costs of R&D. Also, the home-firm will gain at least some access to the other firm's location specific knowledge as well as their technology and know-how. However, there are two general categories of coordination problems that occur in networks comprised of two firms. First, there is a question as to how potential profits will be split. This arises because there are differences in the relative inputs that each firm introduces into the R&D process, and assessing value of these inputs is difficult. Second, and key to this analysis, is the asymmetry of information problem.[7] From both the home-firm's and the international partner's perspective, limiting unnecessary information leakage while maximizing knowledge attainment is the goal. This becomes a game of semi-coordination and thus will have a less optimal equilibrium output of new knowledge when compared to the subsidiary case ceteris paribus. However, it is possible that even given the coordination problem, the potential for the total amount of knowledge flow and spillovers between firms is so great that it could lead to between firm networks generally having more importance in the home-firms innovation production function. Therefore, it is an empirical question as to which form of network is relatively more important to home-firm innovative productivity.[8]

No matter the network structure, monitoring costs and barriers to knowledge flows will exist. It is generally true that the further away the network partner or subsidiary is from the

[6]Standard problems with information security arise, but they are not unlike those faced at home.

[7]If two firms are direct competitors in the market for ideas, then there will be a tenuous relationship where each firm will only want to share what they have to while attempting to glean as much as they can from the other. Even if the two firms deal in markets that are generally complimentary there is still potentially a similar problem. If it were possible to identify the level of knowledge complementarity between the home-firm and international partner, I would be better able to identify even further the levels of competition over ideas.

[8]Griffith et al. (2004) show that there are indeed two very different incentives at play in between-firm networks where both innovation as well as extraction of knowledge from the partner is a goal. If firms can legally replicate proprietary technology, it gives them entry into markets where real profits are feasible. Lerner and Malmendier (2010) illustrates the complicated structures of contract design that are necessary to optimally protect profits and guarantee appropriability. Aghion et al. (2002) show how the level of competition directly impacts the incentives to innovate.

home-firm, the more difficult it will be to actively monitor operations. This carries with it two implications. First, efficiency may fall the further away the home-firm is. However, firms may know this before starting the network, and therefore only enter into the relationship if the probable returns are higher. Therefore it is an empirical question as to which factor is stronger, or if distance has a significant impact at all. In addition, if cultural or societal differences lead firms to produce R&D in different ways, then it may be difficult to utilize the network for knowledge sourcing. Therefore in addition to geographic barriers, social distance may be a factor that impacts potential returns. If the network exists between firms in countries where language or cultural differences are large, then the same two implications will hold as with geographic distance. Efficiency may fall, but if firms know this before entering into the network, they may only enter if the potential returns outweigh the perceived risks. It remains an empirical question as to how these potential costs, stemming from geographic and social distance, affect innovation outcomes.

Finally, because competition for new ideas is potentially different between industries, there could be differential effects to home-firms based upon the types of innovation they are involved in. Firms that find themselves in frontier industries will experience two effects. First, the level of competition over new ideas is greater, therefore frontier industries should have more international R&D networks. However, finding new information will potentially be more difficult, limiting the effectiveness of these networks. However, if firms accurately predict these costs beforehand, then they will only enter into networks where the probable returns are justified. So, again, it is an empirical question as to which effect is dominant.

I look to answer these questions and estimate the causal effects of networking using a difference in differences identification strategy. In order to do so, I first select home-firms which are the set of the top 1000 patenting firms within the OECD or European Patent Community (EPC) over the ten year period from 1999-2008.[9] The treatment period covers the ten year period from 1999 through 2008. The reason for these criteria are both conceptual as well as driven by the data.

[9]The set of countries are those that existed in the OECD and EPC as of 1989, the first year of the network building period. The only country added to those from the OECD through inclusion of the EPC is Lichtenstein.

2.1 Sample

The sample of home-firms are the 1000 most innovative firms in the OECD and EPC based upon average patent counts per year.[10] Specifically, I only consider firms within the OECD and EPC as of 1989. The choice is primarily based upon the fact that I do not want major changes in trade channels to drive networking choices amongst the home-firms. Second, because most firms patent very infrequently, it is highly unlikely that they will produce or absorb knowledge spillovers through networks. As a result, estimation based upon the 1000 most innovative firms, those firms that actively take part in the innovative process, is fundamental.

Firms that regularly innovate and are responsible for most of the patent applications filed are most likely to create, as well as be recipients of spillovers from collaboration. Also, just as the data on individual innovation is highly skewed, the data on firm-level innovation is also highly skewed (Azoulay et al., 2010).[11] In addition, inference would be murky at best if firms that only innovate on an extremely limited basis were compared with highly innovative firms such as Toyota, IBM or Siemens. It is important to note that even after limiting to the top 1000 innovative firms, there is still a wide range of innovative productivity, but inference throughout the analysis is made much clearer using these selection criteria from which the treatment and control groups will be identified.

In addition, because countries vary widely in infrastructure and property rights enforcement, I focus on a set of firms located within countries of similar characteristics. This is further motivation for including only firms whose primary location is in the OECD or EPC as of 1989. Incentives to patent and the expectations tied to those patents are similar and evolve similarly amongst these countries over the time period in the analysis. Also, clarity is added to the statistical inference by only comparing like-firms that exist in countries with similar infrastructure.

Finally, the time period is chosen due to both theoretical and data constraints. The treatment period is from 1999 through 2008. A critical technological change which dynamically influenced information and transaction costs around the world was the introduction of the internet and email (Agrawal and Goldfarb, 2006b). By 1999, adoption of these technologies, as well as high-

[10] I have utilized citation-weighted patents as an alternative metric and the sample is not highly sensitive to the change.

[11] There are a total of 66708 firms in the patent data from 1989-2008. Therefore, selecting the top 1000 is the same as considering the top 1.5% of patenting firms. To highlight the skewness of the distribution of patenting, consider that the top 1.5% of firms produce 48% of all patents from 1999-2008.

speed transmission was no longer new amongst the OECD, and thus should not dramatically influence the results of the analysis.[12] Thus, the primary reason this time period is selected is that it is primarily post-internet. In addition, the primary data source only extends with complete records through 2008, so this is implemented as the final year of the study.

2.2 Primary Data

The primary data was purchased from the European Patent Organization (EPO). The data source is the April, 2010 snapshot of the EPO Worldwide Patent Statistical Databases (PATSTAT). PATSTAT is a snapshot of the EPO master documentation database (DOCDB) with worldwide patent activity coverage. It has 20 tables, including bibliographic data, citations and family links.[13] It is specifically designed to be used for statistical research.

From PATSTAT I extract all patents and application data on firms located in the OECD and EPC. I then restrict the sample to patents filed by firms rather than individuals. Next, I drop all patents filed with authorities other than the EPO, including all patents filed with the USPTO. This is a key data choice within the framework of my analysis.[14] The theoretical reason for the use of only EPO patent filing data is as follows.

I intend to measure the effect of collaboration on innovation productivity directly, and accurately measuring the timing of the network effects is a key part of the contribution of this paper. Although there is no perfect measure of the moment that innovation takes place, patent application dates within the EPO are the best proxy available on a global scale. The reason that this is the case is that patents in the EPO are given rights based upon the filing date, not the date of discovery. In other words, it is a first-to-file incentive system. This is very different than the setting in the US, where a firm can capture existing patent rights of another party if it can prove that it had produced the idea first. In addition, if a firm publicly displays, publishes, or

[12] I utilize time fixed effects for all specifications which should control for any global trends due to technological change, economic trends or otherwise.

[13] A patent family is a unique identifier given to identical patents filed within multiple organizations such as the EPO and USPTO. Some firms will file with the EPO and subsequently file with the USPTO depending upon their marketing and competitive production strategies. The patent family identifier allows the researcher to limit duplicate counts of the same innovation.

[14] For details on data acquisition and computing requirements, see the PATSTAT website. As the data acquisition and compiling is not trivial, it is important for researchers new to the EPO data to consider resource limitations before attempting its use.

files for a patent elsewhere regarding a technology, then the EPO will not consider the application. Therefore, when patenting with the EPO, there is a very clear incentive to file the patent application as soon as possible after innovation occurs. This is quite different when comparing to the USPTO and most other patent organizations, ceteris paribus. Also, PATSTAT includes actual filing dates which are critical to identification of spillovers in a time based analysis, which are not available in the USPTO data.[15]

Finally after selecting the appropriate patents from which the sample of firms will be selected, consideration is given to cleaning the data to ensure that only one identifier per firm is included. First, it is possible that multiple unique firm-identifiers in PATSTAT are actually the same firm. This can occur because of small discrepancies in the text-based name field. Combining these multiple identifiers into a singleton is an obvious solution to this problem. However, it is also crucial to ensure that firms whose IDs are combined not be subsidiaries of each other. Because I am interested in identifying the spillovers of international R&D networks, if subsidiaries are not identified appropriately, the potential effects on the results are not trivial. Rather, as outlined previously, the incentives and flow of information between the home-firm and a subsidiary are potentially quite different than in between-firm networks.[16]

Following a technique in Melamed et al. (2006), I create an algorithm that codes probable duplicate unique identifiers based upon naming conventions and the reported firm locations. After running the name-based portion of the matching algorithm, firms that are coded as probably matches are then compared based upon location. If both the name and location match, the unique identifiers are combined. However, if the locations are different, I code the firm with the highest patent count as a potential home-firm while the other is coded as a subsidiary.[17] I then create the outcome variable; a patent count indicator and sum it at the annual level. I create two additional permutations of the outcome, one is high-impact patent counts where I measure high-impact has having greater than five forward citations.[18] Next, I simply weight each patent by its total forward citations.[19] After sorting the 66708 firms and keeping only firms whose location is in the OECD or EPC, I am able to identify the top 1000 innovative firms. This set

[15]For an excellent discussion of the incentives for patenting see Hall (2008).

[16]For a detailed discussion of the use of international collaboration as knowledge sourcing see Belderbos et al. (2005). For a detailed discussion on the use of international expansion through the use of subsidiaries see Blit (2010).

[17]While further development of the firm-matching/firm-filtering algorithm is possible, the use of the frequency of patent attribution as a proxy for primary status is reasonable. For a detailed discussion of how the algorithm was generated see Melamed et al. (2006).

[18]Hall et al. (2009) indicates that patents with lower citations may have lower market value.

[19]Waldinger (2011) uses this metric to look at the quality of patents, not just the proliferation.

of firms will be referred to as the *home-firms* throughout the analysis.

2.3 Network Dyads: Construction and Considerations

Critical to the study is the construction of the fixed international R&D networks through which the disaster effects will be estimated. I identify networks from patent filings. The source of the network dyads is the set of all patents filed from 1989 through 1998, which will be referred to throughout the analysis as the *network building period*.

The initial process of identifying a network is straight forward. From all patents filed during the network building period, I extract patents on which at least one of the home-firms is a contributor. This leaves approximately 48% of all patents filed. Then, I drop any patent for which there is a sole contributing innovator, or any patent on which all contributing firms are located in the same country. This leaves 2547 patents with 147 home-firms out of the top 1000 and 245 other firms from which collaboration networks are to be identified.

Coding network dyads from the potential network pairs is not a trivial undertaking. The appropriate network dyad identification algorithm identifies and properly codes each network pair as either a between-firm network dyad or a subsidiary dyad. This is a relatively simple task on patents that have only two firms included in the filing. As previously stated, it must be that at least one of the firms on the patent be a home-firm, and the two firms are located in different countries. Therefore on a patent with only two authors there are either one or two network dyads that will be identified. If only one of the firms is a home-firm, then one dyad is identified. If both firms are home-firms, then two unique network dyads will be identified, one going in each direction where the potential network partners alternate between being the home-firm and the international-partner. However, when three or more are included, properly identifying and coding network dyads is not a straightforward task and justifies further illustration. To show the permutations that are possible, and how I construct the unique network dyads, let us consider two different hypothetical patents.[20]

First, consider a patent application with three firms listed; firm one is a home-firm based in Germany (METO International GMBH), firm two is a home-firm based in the US (Sugen Inc) and firm three is not in the sample and is from Australia (Biosignal LTD). From this set of

[20] See Figure F2 in the Appendix of Additional Figures and Tables for a visual illustration of the dyads created from a similar patent to that presented in third example.

three firms, four unique dyads will be coded. First, firm one will be the *home-firm* while firm three is the *international-partner*. Second, firm two will be the home-firm while firm three is the international-partner. For the third and fourth dyads; firms one and two will alternate between being the home-firm as well as international-partner as outlined in the simplest case.

Next, consider a more complex example of a patent with four firms listed: 1. Firm one is a home-firm from France (Alcatel-Lucent), 2. Firm two is a home-firm firm from Germany (Standard Elektrik Lorenz AG), 3. Firm three is a subsidiary of firm one located in France (Alcatel Cable) and 4. Firm four is not a home-firm and is located in the US (Advanced Micro Devices Inc). From this set of four firms, six unique network dyads will be coded: 1. Firm one with firm four, 2. Firm two with firm four, 3. Firm one with firm two, 4. Firm two with firm one and finally, 5. Firm two with firm 3. The last dyad between firm one and firm three will be tagged as a subsidiary network is that between Alcatel Cable and Alcatel-Lucent.

After running the algorithm, I am left with 343 network dyads, 75 of which are tagged as being subsidiary networks. It is through these dyads that the identification of network spillovers is identified. At this point, I generate various network-dyad variables from the patent data. First, I generate a measure of the international R&D network's duration, which is simply the length of time between the earliest co-patent and last co-patent observed between the home-firm and international-partner between 1989 and 2008. I also generate a measure of network dyad productivity which is a count of the number of co-patents that a particular network dyad produce. The final dyad characteristics must come from external data.[21]

Finally, data is introduced to measure both geographic and social distance between home-firms and their collaborative partners. As outlined, theory would indicate that distance may impact the returns to networking on home-firm productivity. The French research center in international economics (CEPII) has generously made public a database which includes all bilateral country distances as well as a myriad of social distance measures (Mayer and Zignago, 2006). The social distance measures include border contiguity, language commonality and prior-historic commonalities such as colonial roots.[22] I merge these to the international network dyads by home-country and partner-country. At this point, the data is appropriately prepared to merge

[21] There is a question as to whether or not these networks exist in the treatment period. It is possible that a dyad that is indicated from a co-patent in 1988 does not necessarily exist as a network that coordinates innovative activity in 2008. In this case, the results would be attenuated towards zero, as I would be estimating a causal effect when there is no link. In further sensitivity analysis, I use a relatively more strict definition of a network as one where the network dyad co-patents at least two times, and find no significant difference in the results.

[22] Details on the data construction can be found on their website, as well as ready to use files for use in STATA.

to the disaster data and then the analysis can be completed.

2.4 Identification Strategy and Treatment Data

In the ideal scenario I could randomly assign firms with similar characteristics to internationally collaborate and not collaborate, then simply compare their levels of innovative productivity to reveal the impact of R&D networks. This is the experimental setting that I intend to emulate with my quasi-experimental identification strategy. In the network building period I have identified 343 international R&D network dyads. Mirroring Azoulay et al. (2010) and Waldinger (2011), I will look at the impact of exogenously shutting down these networks. It will then be revealed dynamically how important the different networks are in home-firm patent production. Specifically, in the treatment period, I will look at the impact of large scale natural disasters that occur in the international-partner's country on home-firm patent production. In order for the quasi-experiment to be valid, the effect of the natural disaster on home-firm innovative productivity should only occur because of the existing international R&D network.[23]

To illustrate the idea, consider the recent earthquake and tsunami that impacted Japan in March of 2011. Though day-to-day economic activity has mostly returned to normal, most of the infrastructure that was destroyed has not yet been rebuilt as recovery is a slow and lengthy process (Ford, 2011). If a firm based in the US had an existing R&D network with a firm in Japan, and the Japanese lab was impacted by the earthquake or tsunami, then the natural disaster serves as a quasi-experiment of the exact type I intend to emulate. The occurrence of a natural disaster exogenously shuts down the network, allowing me to infer the causal effect of international R&D networks on home-firm innovative productivity.

The coefficient estimated will be that of a difference-in-differences identification strategy. Specifically, the treatment group would be any home-firm whose international-collaborator is impacted by a large scale natural disaster. The control group are those home-firms with existing international R&D networks that are not impacted by any disaster. The causal effect of the international R&D networks will be obtained from taking the difference in patent rates among those firms that are treated before and after the disaster, and then comparing this with the difference in patent rates among those firms that are not treated. In order to implement the DD

[23] It is worth noting that any disaster that impacts multiple countries in close proximity are coded separately. This allows me to control for the fact that some disasters may not only impact the international-partner, but the home-firm directly.

framework, data on natural disasters must be merged to the international dyad data previously constructed.[24]

The international disaster database (EM-DAT) is obtained from the Center for Research on the Epidemiology of Disasters (CRED) and is considered to be the most comprehensive source for data on natural disasters available (de Louvain Brussels Belgium, 2011). The database includes all natural disasters across the globe from 1900 to the present that meet at least one of the following criteria: 1. 10 or more people are killed, 2. 100 or more people need immediate assistance, 3. There is a formal declaration of a state of emergency or 4. There is a call for international assistance. The database includes data on the disaster type, country, and severity. The severity measures include number of people killed, number of people impacted, and the number of dollars ($US) in damage incurred. Because the data is collapsed at the country level, the raw measures do not do a good job of scaling the disasters from one country to another. I therefore convert all the severity measures to be per-capita, based upon the annual country population. This will provide a standard relative measure of severity unlike the level data.[25][26]

Because the disaster data are highly skewed in the severity measures, I limit possible treatment to those disasters in the 99th percentile of devastation in at least one of the three measures. This restriction captures disasters that have the potential to destroy capital infrastructure (laboratories etc.) or interrupt day-to-day economic activity at a level of severity that the international-partner's activity will potentially cease.[27]

Finally, in order to appropriately utilize firm fixed effects I must collapse the data to the home-firm-annual level. The reason for collapsing the data is that running the quasi-experiment while leaving the data at the dyad level is inappropriate as the model will be over-identified. Because I am measuring the outcome at the home-firm level, any firm that has multiple inter-

[24] I have also considered using the entire set of the 1000 most innovative firms in the control group. However, this is not the appropriate setting as most of the firms in the top 1000 can never be "treated" as they are not involved with any firms internationally through a network. In any case, firms are randomly assigned to these groups by the exogenous occurrence of the natural disasters, fitting with the intended quasi-experimental design.

[25] Disaster data is readily available at www.emdat.be. Population data is obtained from the Penn World Tables (Heston et al., 2011).

[26] Cavallo and Noy (2009) shows explicitly that the EM-DAT database captures many small events that are generally correlated with the level of government infrastructure and reporting ability. By using only the top disasters in per-capita levels, I am able to identify truly large events that are uncorrelated with the database's data capturing guidelines. There should not be a bias on the countries that have disasters, rather they must be truly exogenous.

[27] Further justification of the use of these specific disasters is explained and analyzed in detail later in the paper.

national network dyads would be used multiple times. In addition, if a given home-firm only co-patents with another firm one time while co-patenting 50 times with another, then coding a disaster to be 1 if it occurs in either country is not theoretically sound. In other words, the disaster indicator should represent the percentage of the existing network which it is actually shutting down. Therefore, if a home-firm only collaborates with one other firm, then the disaster indicator would be 1 in all cases, as the home-firms entire network is shut down if a disaster occurs. However, for a highly networked home-firm such as Siemens AG, a disaster impacting Japan would only shut down a small percentage of its entire network. I thus weight the country level disaster indicators by the proportion of the total network-driven patents a particular network dyad is responsible for. This will mean that I will interpret the impact of a disaster on home-firm patent output as the effect of shutting down its average network-partner.

3 Descriptive Statistics and Trends

Before considering the econometric model and results, it is worth while to consider what the data looks like. There are three tables of descriptive statistics to help illuminate how the outcomes and various network dyad characteristics are distributed amongst the different sets of home-firms thus far considered in the setting. Table 1 presents in column means and proportions for all 1000 top innovative firms and then compares these measures between those home-firms who network and those who do not. Table 2 looks within only those home-firms that network and then compares those that are treated versus those that are not. Finally, Table 3 looks within only those home-firms that network and then compares those that have subsidiary based networks versus between-firm networks.

Specifically within Table 1, the first column presents descriptive statistics for the entire top 1000 innovative firms. In section (1), means for the three outcome measures are presented. We see that there are drastic differences in patent output between those firms that network and those that do not by looking at columns (2) and (3).[28] On all patents, we see that firms that are part of international R&D networks patent approximately 8 times more than those that do not. Even when only considering high-impact patents, those who network produce five times more.[29] When considering forward citation weighted patent counts, the same patter holds as with raw

[28]See Figure F1 in the Appendix of Additional Figures and Tables for a histogram of patent counts per year amongst the 1000 most innovative firms.

[29]High impact patents are those with greater than 5 forward citations.

patent count. The fourth column present the t-statistic with $H_0 : Mean(2) - Mean(3) = 0$, which shows that these differences are statistically significant. So, networking firms do produce between 4 and 8 times more successful patent applications per year. However, to place on this difference a causal interpretation would be incorrect. Again, this is because the networking decision is endogenous to the profit maximization problem. The fact that there are significant differences does provide justification that networking itself may be a key input into the patent production function at home.

The second section looks at the dispersion of patenting amongst selected International Patent Classifications (IPCs).[30] It appears that firms who network have a greater proportion of their innovation focus amongst those fields that include innovation frontier industries such as nano-technology, bio-tech and medical devices, and also information technology. As proposed, those firms that exist in industries with higher levels of knowledge competition are forced to expand their knowledge sourcing efforts in order to be competitive.

Finally, section (3) considers how patent counts may vary between firms located in different home-countries.[31] I present the distribution of patents amongst the three top patenting countries in the OECD. Though there is variation in the average patent production between firms in the different countries, the relative comparison beween networked and non-networking firms is clear. In all three cases, networking home-firms produce significantly more patents than the non-networked home-firms. Again, this is further justification that networking does have positive effects on home-firm patent production, but cannot be interpreted as a causal effect of international R&D networks.

In Table 2 the same format is presented with the broad category of all networked-firms in column (1).[32] In columns (2) and (3) are those whose networks are impacted by the natural disasters and those whose networks are not respectively. On average there are differences between the two groups. In section (1), I present the data on the outcome variables. We see that generally the treated firms tend to produce more patents. In section (2) we see that those firms that are impacted by disaster's tend to have larger networks, which does increase the probability that

[30]See the table notes for a detailed breakdown of the IPC classifications. Further, IPC patent classifications are not mutually exclusive, so these patterns may be indicative, but cannot be relied upon to be true proxies of industry.

[31]A between country comparison is outside the scope of this paper, as I am interested in general networking effects. However, further research might look at how patent outcomes, networking and home-country may be related.

[32]This is the same as column (2) in Table 1.

a network disaster will occur. These firms also tend to have stronger ties with their network partners. These differences could be indicative that the identification strategy is in jeopardy, as based on the averages it appears that those firms who's networks are impacted by disasters are somehow different that the others. However, because of the nature of innovation, patenting rates are extremely skewed, and having one firm that is far out on the tail of the distribution located in either group will pull the average for the sub-group dramatically. In the analysis, I present clear evidence on the dynamics that include firm fixed effects showing that pre-disaster trends are not statistically significant. Finally, in section (3) we see that there are no significant differences in the geographic dispersion or social distances between the two types of networks.

In Table 3, column (1) is the same as that of Table 2. Columns (2) and (3) stratify the networked firms into those that are subsidiary-based and those that are between firms. Here we see in section (1) that firms that have subsidiaries tend to produce about 4 times as many patents as those that partner with other firms. In section (2) it is revealed that subsidiary-networks tend to produce more patents. The home-firms that have subsidiaries on average have more network partners. Also, the length of time over which the subsidiary networks are active is longer on average. This again does increase the probability that the network is impacted by a disaster. The distance measures in section (3) are not significantly different. Based upon the evidence presented in the descriptive statistic tables, and from the previously outlined theoretical basis, consideration of the types of networks is of further interest when thinking about the causal effect of international R&D networking on home-firm patent productivity.

4 Econometric Specification and Results

I present the econometric results in five sections. First, I introduce the estimation technique utilized throughout the analysis and subsequently the model of interest. Second, I present an empirical justification of the identification strategy. Third, I present the results of model of interest, including different network and disaster specifications based upon the methodological considerations. Fourth, I look to disentangle some of the possible mechanisms through which the main effect may be explained. I do this by considering how different dyad characteristics impact the results including measures of distance and industry proxies. Finally, I look to test the sensitivity of the model to various specifications, as well as consider robustness of the model through introduction placebo treatments.

4.0.1 Estimation Technique

The outcome variables for innovative productivity are a count variables, non-negative and highly skewed. Specifically, I utilize either a firm-level count of successful patent applications in a given year, or the cumulative citations on patents occurring in a given year. I therefore utilize a Poisson pseudo-maximum likelihood model (PPML) as developed in Silva and Tenreyro (2006). Silva and Tenreyro (2006) show that the log-linear specification often used is not only less efficient, it can be inconsistent when the dependent variable exhibits heteroscedasticity or has a significant number of zeros, as does my data. The PPML estimator is shown to be both consistent and more efficient than the OLS framework for count models with the characteristics my data carry.

The PPML method allows convergence even when the outcome variable has a large number of zeros and/or the independent variables have high values. In addition, the negative-binomial and zero-inflated models are sensitive to the scale of the dependent variable, giving different point estimates. This is important as I use three different outcome measures,and comparability of the elasticity estimates is important. The PPML method does not have the same inherent problems. Following the prior literature, such as Waldinger (2011), I directly implement the PPML procedure as specified in Silva and Tenreyro's 2011 update (Silva and Tenreyro, 2011). It is reassuring to know that the model does not have to follow a Poisson distribution in order for coefficient estimates to be efficiently and consistently estimated using the technique. It only requires that the conditional mean of the dependent variable be appropriately specified. This is a pretty general assumption, which the data should satisfy. Finally, the estimates are able to be interpreted as elasticities.

4.0.2 The Econometric Model

The model of interest relates innovative output of home-firm i in time t to disasters that impact international-partner $j's$ country in time $t-1$ through $t-4$:

$$\begin{aligned}(1)\quad y_{i,t} = \beta_0 + \beta_1 InternationalDisaster_{j,t-1} + ... + \beta_4 InternationalDisaster_{j,t-4} + \\ \theta_1 HomeDisaster_{i,t-1} + ... + \theta_4 HomeDisaster_{i,t-4} + \delta_t + \gamma_i + \varepsilon_{i,t}\end{aligned}$$

where y is a measure of innovative output. $InternationalDisaster$ denotes an binary indicator of a large scale natural disaster impacting international-collaborator j. $HomeDisaster$ denotes an binary indicator of a large scale natural disaster impacting the home-firm i's country.[33] The $delta_t$ is time fixed effects, and $gamma_i$ are firm fixed effects. The use of both firm and time fixed effects are consistent with my approach to analyze changes in i's output following probable shut down of firm j. The time fixed effects control for general changes in the technological and economic landscape. Also, the firm fixed effects control for firm characteristics that could influence output including industry and firm size, location and country-specific infrastructure and incentives. Finally $varepsilon$ is the error term, and standard errors are clustered at the firm-level. [34]

4.1 Empirical Basis for the Identification Strategy

Before implementing the model, it is important to validate the use of the DD estimation technique in two ways. First, the disaster's must actually impact the international-partner's operations. Second, the treatment and control groups in the quasi-experiment should have similar pre-disaster trends.

In order to show the magnitude and impact on innovative productivity that these disaster's have, I present Table 4. The model is as follows:

$$(2) \quad y_{i,t} = \beta_0 + \beta_1 HomeCountryDisaster_{i,t-1} + ... \\ + \beta_4 HomeCountryDisaster_{i,t-4} + \delta_t + \gamma_i + \varepsilon_{i,t}$$

Here the outcome variables in country i are regressed on disasters that occur in country i. This exercise is done to show that these disasters are large enough in scale to disrupt economic activity within a country as a whole, as well as negatively impact or potentially stop firms innovative productivity within the country. If this is not true, then using these natural disasters as a source of variation would not be a valid identification mechanism. In column (1), I the outcome is NASA lights data that Henderson et al. (2011) have recently been shown to be an excellent indicator of economic activity. Specifically, it is a measure of the man-made light that is observed by NASA satellites, which is then population weighted and summed at the country-year level. The data

[33] This is simply to control for any home-country disasters, and eliminate the possibility that results be biased when a firm who's network is relative close is directly impacted by the disaster as well as the network.

[34] As described prior, the dyad data is weighted by it's relative importance based upon the ratio of a given dyad's patent productivity to the entirety of a home-firms network driven patent count. This allows me to accurately collapse the data to the firm level. Interpretations of the results are straight forward and are indicative of the effect of the home-firm's average network partner.

shows strong negative effects that persist for at least three years on average.[35] In column (2) I present clear evidence that these home-country disasters do impact innovative productivity of firms in that given country. Here we see the normal cycle of patenting and innovation impacted, where the negative impact is implied one year after, but is large and significantly negative two years after the disaster. It is well documented that it takes on average two years to take an idea from start to patent, and this is confirmed by a 25% decrease in patent production two years after the disaster, and rebounding thereafter. Clearly, these large scale disasters have the potential to shutdown economic activity, and thus will provide a compelling source of variation for international R&D networks.

Next, I consider the pre-disaster trends amongst those home-firms with networks that are impacted by disasters versus those that do not. The key difference between this test of comparability and that of simply comparing the treatment and control groups based on averages is that I can include firm fixed effects. When only comparing averages as was presented in Tables 2 and 3, becuase of the highly skewed nature of patent production, the random inclusion of a just a few highly productive firms will pull the average dramatically. By running a dynamic comparison after removing the effects of firm heterogeneity, we can see the actual pre and post disaster trends. The results are presented in Figure 3. Here I regress home-firm patent outcomes from three years prior to three years after the international-disaster occurs within their network partner's country. The model is as follows:

$$\begin{aligned} (3) \quad y_{i,t} = \beta_0 + \beta_1 InternationalDisaster_{j,t+3} + ... + \beta_7 InternationalDisaster_{j,t-3} + \\ \theta_1 HomeDisaster_{i,t+3} + ... + \theta_7 HomeDisaster_{i,t-3} + \delta_t + \gamma_i + \varepsilon_{i,t} \end{aligned}$$

The regression includes both firm and time fixed effects as well as controls for home-country disasters. The standard errors (in parenthesis under the point estimates) are clustered at the firm level. After including firm fixed effects, there are no statistically significant differences prior to the time of the disaster which is reassuring that the identification strategy is valid. In addition, this is the first look at they model of interest. There is a weakly significant drop in home-firm patenting two years after the disaster occurs. However, the estimate is only significant at the 10% level, and thus it is difficult to say whether these networks are or are not key inputs into the home-firm innovative production function.[36]

[35] Though I have attempted to reveal why the effect is not significant in the second lag, I believe it is simply a feature of the data. The specification included a full set of year-dummies, so this is not a feature of serial correlation driving the result.

[36] See Figure F2 in the Appendix of Additional Figures and Tables for a similar exercise using the entire set of top 1000 innovative firms based upon Table 1.

4.2 Main Effects

Next, I present the primary findings in order to answer the key question in the analysis; whether or not there is evidence that international R&D networks are significant inputs in home-firm innovation. The results are presented in Tables 5 through 8. In Table 5, the model is as follows:

$$(4) \quad y_{i,t} = \beta_0 + \beta_1 InternationalDisaster_{j,t-1} + ... + \beta_4 InternationalDisaster_{j,t-4} + \theta_1 HomeDisaster_{i,t-1} + ... + \theta_4 HomeDisaster_{i,t-4} + \delta_t + \gamma_i + \varepsilon_{i,t}$$

In Table 5, each column is indicative of a separate regression, and all models include time and firm fixed effects. In this model, the dependent variable is the given outcome for all home-firms that take part in international R&D networks, regardless of the type. This will therefore answer the question in the most general case.[37] In columns (1) through (3), the results are consistent. There does not appear to be a significant effect of shutting down the international R&D network. This would indicate that on average, these networks are not key inputs into the R&D production function of the home-firms. However, there is a pattern that emerges. In the second year after the disaster, all point estimates are negative, though imprecisely estimated. This leads us to consider more deeply the theoretical implications outlined prior.

As explained prior, the incentives and costs associated with subsidiary versus between-firm networks differ substantially. I therefore present Table 6, where I separate the disasters into those that occur through subsidiaries, and those that occur through between-firm networks. It is worth noting that many firms have both types, and therefore, I include the full set of disasters in the model as follows:

(5)

$$y_{i,t} = \beta_0 + \beta_1 InternationalSubsid.Disaster_{j,t-1} + ... + \beta_4 InternationalSubsid.Disaster_{j,t-4} + \tau_1 InternationalCollaboratorDisaster_{j,t-1} + ... + \tau_4 InternationalCollaboratorDisaster_{j,t-4} + \theta_1 HomeDisaster_{i,t-1} + ... + \theta_4 HomeDisaster_{i,t-4} + \delta_t + \gamma_i + \varepsilon_{i,t}$$

The results in Table 6 show a strong contrast between the effect of subsidiary-network disasters and between-firm network disasters. There is evidence that it is subsidiary-networks that are key inputs into home-firm patent production, whereas between firm disasters are insignificant in all specifications and all lags. Though the first column is imprecisely estimated, as we move

[37] As there may be unobservable time-trends that affect innovation, whether that be a market downturn or other, time fixed effects will eliminate general trend changes from biasing the causal estimates. In addition, firm fixed effects will control for firm heterogeneity including size, innovative potential, country and industry trends.

into columns (2) and (3) where the outcome is also theoretically weighted by a measure of quality, the effect of subsidiary-network disasters are significant with large and persistent effects through the third year post disaster. This is indicative of the fact that the competition over ideas that occurs between firms in the same network mitigates any possible spillovers to home-firms. Whereas within a subsidiary-network, all useful knowledge is able to be channelled home, and the home-firm relies upon the network to some degree for sources of new ideas.[38]

In Tables 7 and 8, I stratify the firms into those that have only between-firm networks, and those that have only subsidiary networks. In Table 7 there are generally insignificant results, however an interesting result occurs in the second column. Here it is inferred that home-firms actually increase their patenting in the year just after the disaster while they fall in the same magnitude in the second year. It seems that firms that have only between firm-networks are able to "cash-in" based upon their work with the other firms, but then they do falter in the year following. However, there is no persistence, and the net effect as of just three years after the disaster is zero.

In Table 8, a very different story emerges. The model is run on home-firms that have only subsidiary-networks.[39] The results clearly demonstrate the negative, persistent and large impacts of a breakdown in these networks. When a home-firm invests in an international subsidiary, it is clear that the network is in fact a key input in home-firm innovative productivity. In fact, the elasticities in this setting are extremely large, indicative of the fact that nearly all innovation within the firm is tied to information or research being done in conjunction with the subsidiary. Again, it is worth noting that this is a very special case, where the home-firm's only network connection abroad is through the subsidiary. In spite of this fact, it does clearly demonstrate that there are measurable differences between the two types of networks and the returns to innovative productivity.[40]

[38]The null finding on between-firm networks can be interpreted in at least two ways. First, it may be that home-firms are able to substitute easily away from the resources provided in the network. Second, it may be that these networks are instigated as knowledge sources, rather as market searches or for various other purposes. In either case, they are obviously not critical in home-firm patent production.

[39]Note that the number of firms in this is quite small, with only 11 out of the 146 networking firms included. Yet, the results are still precisely estimated.

[40]See Tables A1 and A2 in the Appendix of Additional Figures and Tables for detailed descriptive statistics of the relative sets of home-firms and network dyads used in Tables 7 and 8.

4.3 Possible Causal Mechanisms

Next, I consider the generalizable incentive and cost sources that occur across all network types. First, I will consider both geographic and social distance. Theoretically, they both may have an effect on the potential returns to networking abroad. Second, I consider how industry may play a roll in the incentives to network and the related returns to networking.

4.3.1 Geographic and Social Distance

As outlined, geographic and social distance may effect the potential returns to knowledge sourcing. However, whether these cost are anticipated prior to entering into the network is unknown. To investigate this further I present Table 9. Here in columns (1) and (2) I stratify the home-firms by above and below median geographic distance. In columns (3) and (4) I stratify the home-firms based upon whether or not their network partner is located in a country with a common language to its own. Interestingly, there are no significant effects present. It is not completely surprising to find no effect of geographic distance within this setting. Once the distance between the home-firm and network-partner is sufficiently large, further distance should not significantly effect the costs of monitoring. If the median distance between firms were something on the order of 10 or 20 miles, then the results might fall in line with the work of Henderson et al. (1993) where it is found that distance plays a key roll in knowledge spillovers.

Though the results in columns (3) and (4) are imprecisely estimated, in column (3) the point estimates are significantly different than the rest in the table, being negative and relatively large in magnitude. Unfortunately, the number of firms that fit the stratification is quite small. If the estimates were significant, it would imply that having a common tongue does significantly increase the potential spillovers from international R&D networking. Unfortunately, without higher precision, making further inference would be unjustified.[41]

[41] Unfortunately the sample size becomes too small to estimate the specification when stratifying amongst both between-firm and subsidiary networks using this specification. However, I do present the same specification amongst only between-firm networks, and find large and significant effects. The information can be found in Table A3 in the Appendix of Additional Figures and Tables.

4.3.2 Industry Sensitivity

Finally, it is well recognized that different industries have very different trajectories when it comes to the speed, proliferation and appropriability of new innovation (Audretsch et al., 1996). Industries such as IT or biotechnology are considered to be at the frontier, with very high quantities and proliferation of new innovation. However, appropriability of knowledge in these industries is often quite low. It is therefore an open question as to how networking within these highly competitive industries may effect home-firm innovative productivity. In Table 10, I stratify the model by considering effects within the three IPCs presented in Table 1: Human Necessities, Performing Operations and Chemistry. Though these three IPCs are quite broad, they do contain most utility patents filed within the frontier fields such as IT, bio-tech and nano-tech. The results do not point to significant differences in the importance of networks as inputs across the different IPCs. The results are imprecisely measured across all three IPCs; and though based on the point estimates there are differences between the different categories in the importance of international R&D networks, conjecture as to what may be driving them is all that is possible.[42]

4.4 Sensitivity Tests

As has been discussed, patenting and innovation is highly skewed. it is therefore reasonable to wonder if the results are being driven by "super-stars." In order to investigate this further, I consider a specification where I remove the top 5% of patenting home-firms as well as the top 5% of the network dyads from the sample and run the primary model. It may be that the top performers in R&D have such diversified knowledge sources that they are able to adapt easily to changes in their network structures. It could be that the null result in Table 4 is driven by this fact. The results are presented in Table 11. There is no evidence that the "super-stars" are driving the general null result. In columns (1) through (3) I remove the top 5% of home-firms

[42] In addition to the burden of knowledge motivation, a story that would support differences in network effects between industries is their relative input factor weights between labor and capital. In industries that rely heavily upon capital goods in research versus those that are labor intensive would have very different costs associated with expanding abroad through a subsidiary versus partnering with another firm. These differences in labor and capital intensities would also effect how firms react to a breakdown in the network. Labor intensive networks may easily substitute geographically by simply moving workers. However, capital intensive industries would not be able to accommodate the shock as easily, leading to much worse outcomes in the face of a disaster. In order to explore this further, firm-level industry information would be necessary and is beyond the scope of this paper.

from the sample. and in columns (4) through (6) I remove the top 5% based upon the network-specific patent counts. None of the estimates are statistically significant, which lends evidence to the fact that the estimation strategy is stable even with the biggest innovation players missing from the game.

5 Conclusion

I examine the role of international R&D networks in the innovation production function of firms. After identifying existing international R&D networks from 1989-1998 within the OECD, I use large scale natural disasters from 1999-2008 as a source of variation. Specifically, I look at how disasters that occur in the country of a network partner affect home-firm patent rates. I do not find generalizable effects. However, I do find that networks built through international subsidiaries are key inputs in the production of new ideas and patenting at home, whereas between-firm networks do not have a significant impact. A disaster impacting a subsidiary based network implies a 10-20% negative impact on home-firm R&D productivity over the following three years.

This highlights the fact that competition over ideas is a key factor in whether or not knowledge spillovers occur within networks. In the between-firm networks there is a coordination problem. Each firm attempts to extract as much information out of its international partner while at the same time trying to only release the information absolutely necessary to appease the partner. This conflict mitigates the flow of knowledge and therefore the home-firms cannot rely on these networks as reliable sources of new knowledge. On the other hand, networks based upon the expansion abroad through a subsidiary face no such problem. Incentives between the international partner and the home-firm are well aligned and information flows freely. This allows any party within the entire network to capitalize on new information and build upon it. My results shed light on the previously contradictory evidence of Azoulay et al. (2010), Waldinger (2011), and Borjas and Doran (2011) on the direction and magnitudes of potential spillovers by highlighting the importance of incentive alignment.

My results add to the growing literature on the underlying processes that drive innovation and growth; however they should be interpreted cautiously. The study is based specifically upon existing international R&D networks where both the home-firm and international partner have willingly worked together and successfully patented in the past. The effects of exogenously moving a firm out of a network may not be symmetric with the effect of a firm starting a new international R&D network. Therefore, to say that a firm will see 10-20% increases in patent productivity if it opens an international subsidiary is inaccurate. Rather, this study does specifically measure the losses in home-firm patent productivity associated with the exogenous destruction of an existing international R&D network.[43]

Though careful consideration was made in identifying an appropriate experimental setting and conservatively interpreting the results, many questions are raised and further investigation is justified. A better investigation into industry impacts would be invaluable.[44] Though I attempt to shed light on the potential differences using the IPC classifications, this still leaves much to be desired. Because IPC classification are not mutually exclusive nor do they map cleanly to industries, it is still unclear how introduction of firm-specific industry data would impact the results. This information would be especially valuable for firms.

It is clear, that internationally collaborative research efforts are increasing, and therefore identifying the most productive international partners is integral for firm profitability in today's global marketplace. The results of this paper reveal that firms entering into partnerships abroad, or funding international expansion must consider carefully the potential impacts of knowledge competition. If the goal is not to source knowledge, rather to gain market experience and regional specific social capital, then it is unclear as to which type of network should be sought. However, if knowledge sourcing is the role a firm intends to fill by networking internationally, then utilization of a subsidiary will allow complete capture of the potential knowledge spillovers. For policy makers such as the European Commission, it is key to recognize how competition over knowledge completely changes the way in which knowledge is shared amongst network partners. Incentives must be well aligned amongst all the network partners or the positive spillovers will potentially be mitigated. Somehow, these policy measures must incite all the firms involved in a given project to share knowledge freely, otherwise the output produced will not be socially optimal.

[43]Unfortunately, it is hard to think of any instance in which the reverse circumstance would happen exogenously, making direct measurement a difficult task indeed.

[44]Azoulay (2004) shows that firm organization, which is correlated with industry has a significant impact on whether or not particular processes are outsourced. A further look into how industries, networking and innovation are related would shed further light on his work. In addition, Peri (2005) shows that industry is tightly correlated to the amount of knowledge that flows across boarder, and how far it may travel.

References

Abramovsky, Laura, Elisabeth Kremp, Alberto Lpez, Tobias Schmidt, and Helen Simpson, "Understanding Co-Operative R&D Activity: Evidence From Four European Countries," IFS Working Papers W05/23, Institute for Fiscal Studies October 2005.

Aghion, Philippe, Nicholas Bloom, Richard Blundell, Rachel Griffith, and Peter Howitt, "Competition and Innovation: An Inverted U Relationship," Working Paper 9269, National Bureau of Economic Research October 2002.

_, Richard Blundell, Rachel Griffith, Peter Howitt, and Susanne Prantl, "The Effects of Entry on Incumbent Innovation and Productivity," *The Review of Economics and Statistics*, October 2009, *91* (1), 20–32.

Agrawal, Ajay and Avi Goldfarb, "Restructuring Research: Communication Costs and the Democratization of University Innovation," *SSRN eLibrary*, 2006.

Agrawal, Ajay K. and Avi Goldfarb, "Restructuring Research: Communication Costs and the Democratization of University Innovation," NBER Working Papers 12812, National Bureau of Economic Research, Inc December 2006.

Audretsch, David, Albert Menkveld, and Roy Thurik, "The Decision Between Internal and External R&D," Technical Report 1996.

Azoulay, Pierre, "Capturing Knowledge within and Across Firm Boundaries: Evidence from Clinical Development," *American Economic Review*, December 2004, *94* (5), 1591–1612.

_, Joshua S. Graff Zivin, and Jialan Wang, "Superstar Extinction," *The Quarterly Journal of Economics*, May 2010, *125* (2), 549–589.

Belderbos, Rene, Elissavet Lykogianni, and Reinhilde Veugelers, "Strategic R&D Location by Multinational Firms: Spillovers, Technology Sourcing and Competition," CEPR Discussion Papers 5060, C.E.P.R. Discussion Papers May 2005.

Blit, Joel, "Multi-Location Firms as a Medium for the Geographic Diffusion of Knowledge." PhD dissertation, University of Toronto (Canada) 2010.

Bloom, Nicholas, Mark Schankerman, and John Van Reenen, "Identifying Technology Spillovers and Product Market Rivalry," 2007.

Bloom, Nick, Mark Schankerman, and John Van Reenen, "Identifying Technology Spillovers and Product Market Rivalry," *Proceedings*, 2005.

Borjas, George and Kirk Doran, "The Collapse of the Soviet Union and the Productivity of American Mathematicians," Working Papers, Harvard University 2011.

Cavallo, Eduardo and Ilan Noy, "The Economics of Natural Disasters - A Survey," Working Papers 200919, University of Hawaii at Manoa, Department of Economics November 2009.

de Louvain Brussels Belgium, Universite Catholique, "EM-DAT: The OFDA/CRED International Disaster Database," 2011.

Ford, Peter, "Japan's Tsunami Recovery Stalls," *Christian Science Monitor*, July 2011.

Griffith, Rachel, Rupert Harrison, and John Van Reenen, "How Special Is the Special Relationship? Using the Impact of U.S. R&D Spillovers on U.K. Firms as a Test of Technology Sourcing," *The American Economic Review*, 2006, *96* (5), pp. 1859–1875.

_ , Sokbae Lee, and John Van Reenen, "Is Distance Dying at Last? Falling Home Bias in Fixed Effects Models of Patent Citations," *Quantitative Economics*, 07 2011, *2* (2), 211–249.

_ , Stephen Redding, and John Van Reenen, "Mapping the Two Faces of R&D: Productivity Growth in a Panel of OECD Industries," *The Review of Economics and Statistics*, December 2004, *86* (4), 883–895.

Hall, Bronwyn, "Strategic Use of Patents, Lecture at the European Summer School in Industrial Dynamics (ESSID)," September 2008.

_ , Jacques Mairesse, and Pierre Mohnen, "Measuring the Returns to R&D," Working Paper 15622, National Bureau of Economic Research December 2009.

Henderson, Rebecca, Adam Jaffe, and Manuel Trajtenberg, "Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations," *Quarterly Journal of Economics*, 1993, *108* (3), 577.

Henderson, Vernon, Adam Storeygard, and David N. Weil, "A Bright Idea for Measuring Economic Growth," *American Economic Review*, May 2011, *101* (3), 194–99.

Heston, Alan, Robert Summers, and Bettina Aten, "Penn World Tables 7.0," May 2011.

Hunt, Jennifer and Marjolaine Gauthier-Loiselle, "How Much Does Immigration Boost Innovation?," Working Paper 14312, National Bureau of Economic Research September 2008.

Jones, Benjamin F., "The Burden of Knowledge and the "Death of the Renaissance Man": Is Innovation Getting Harder?," *Review of Economic Studies*, 2009, *76* (1), 283–317.

Kerr, William and William Lincoln, "The Supply Side of Innovation: H-1B Visa Reforms and U.S. Ethnic Invention," *Journal of Labor Economics*, 07 2010, *28* (3), 473–508.

Kim, Changsu and Jaeyong Song, "Creating New Technology Through Alliances: An Empirical Investigation of Joint Patents," *Technovation*, 2007, *27* (8), 461 – 470.

Lerner, Josh and Ulrike Malmendier, "Contractibility and the Design of Research Agreements," *American Economic Review*, March 2010, *100* (1), 214–46.

Maskus, Keith, Ahmed Mobarak, and Eric Stuen, "Skilled Immigration and Innovation: Evidence from Enrollment Fluctuations in U.S. Doctoral Programs," *SSRN eLibrary*, 2010.

Mayer, Thierry and Soledad Zignago, "CEPII Distances," May 2006.

Melamed, Ran, Gil Shiff, and Manuel Trajtenberg, "The 'Names Game': Harnessing Inventors Patent Data for Economic Research," CEPR Discussion Papers 5833, C.E.P.R. Discussion Papers September 2006.

Narula, Rajneesh and Grazia Santangelo, "Location, Collocation and R&D Alliances in the European ICT Industry," *Research Policy*, March 2009, *38* (2), 393–403.

Peri, Giovanni, "Determinants of Knowledge Flows and Their Effect on Innovation," *The Review of Economics and Statistics*, 06 2005, *87* (2), 308–322.

Romer, Paul, "Endogenous Technological Change," *Journal of Political Economy*, 1990, *98* (5), pp. S71–S102.

Silva, J. M. C. Santos and Silvana Tenreyro, "The Log of Gravity," *The Review of Economics and Statistics*, 09 2006, *88* (4), 641–658.

_ and _ , "Poisson: Some Convergence Issues," Economics Discussion Papers 695, University of Essex, Department of Economics January 2011.

Singh, Jasjit, "Asymmetry Of Knowledge Spillovers Between MNCs And Host Country Firms," *Journal of International Business Studies*, 2007, *38* (5), 764–786.

_ , "Distributed R&D, Cross-Regional Knowledge Integration and Quality of Innovative Output," *Research Policy*, 2008, *37* (1), 77 – 96.

Song, Jaeyong, Kazuhiro Asakawa, and Youngeun Chu, "What Determines Knowledge Sourcing From Host Locations Of Overseas R&D Operations?: A Study Of Global R&D Activities Of Japanese Multinationals," *Research Policy*, 2011, *40* (3), 380–390.

Waldinger, Fabian, "Peer Effects in Science - Evidence from the Dismissal of Scientists in Nazi Germany," Technical Report 2011.

Webel, Sebastian, "Research Without Borders: Networking Knowledge," Spring 2011.

Wuchty, Stefan, Benjamin F. Jones, and Brian Uzzi, "The Increasing Dominance of Teams in Production of Knowledge," *Science*, 2007, *316* (5827), 1036–1039.

6 Appendix: Tables and Figures

Table 1: Descriptive Statistics for the Top 1000 Innovative Firms in the OECD

	Top 1000	Home-Firms with Networks	Non-Networked Home-Firms	(2)-(3) t-stat [p-value]
	(1)	(2)	(3)	(4)
(1) Outcomes:				
Successful Patent Applications per Year	12	44	5	5.20
	(41)	(90)	(8)	(0.00)
Successful High-Impact Patent Applications per Year	6	20	4	5.21
(>5forward citations per patent)	(19)	(42)	(9)	(0.00)
Forward Citations per Application per Year	45	162	20	5.20
	(150)	(328)	(30)	(0.00)
(2) Proportions of Patent Applications by Patent Type*:				
Human Necessities	32	52	29	7.0
	(46)	(50)	(45)	[0.00]
Performing Operations	43	56	41	4.72
	(50)	(49)	(49)	[0.00]
Chemistry, Metallurgy	34	63	29	11.1
	(47)	(49)	(45)	[0.00]
Observations	10000	1460	8540	10000
(3) Successful Patent Applications Stratified by the Top and Bottom 3 Innovative Countries:				
United States	7.9	28	4.7	2.72
N=2480	(22)	(50)	(9.6)	[0.01]
Germany	12	55	3.9	2.45
N=2140	(54)	(124)	(3.9)	[0.02]
Japan	23	73	6.6	3.47
N=1460	(63)	(114)	(8.4)	[0.00]

Notes: The sample contains one observation for each firm-year observation over ten years from 1999-2008. Outcomes are measured over the period. The main entries in columns (1) through (3) are the mean of the selected variable. The entries in parentheses in columns (1) through (3) are the standard deviation of the selected variables. In section 1, outcomes are measured during the pre-treatment period. In section 2, patent classifications are not mutually exclusive. In section 3, firms are stratified by country and then means on the outcome variable are compared. Reported t-statistics are obtained from a regression of the selected variable on an international-collaboration indicator variable. The only variables for which there are differences in sample sizes are those stratifying by firm location, for which sample sizes are indicated under the country name.

*Patent Types are not mutually exclusive. A further breakdown of these three patent classifications shows that they include most industries at the frontier including nano-technology, bio-tech and medical devices and most IT.

Detailed International Patent Classification (IPC) Breakdown:

Human Necessities includes: Agriculture, Foodstuffs, Tobacco, Personal or Domestic Articles, Health: Life-Saving and Amusement.

Performing Operations includes: Transporting, Separating: Mixing, Shaping, Printing, Micro-Structural, Technology; Nano-Technology.

Chemistry, Metallurgy includes all Combinatorial Technology.

Table 2: Descriptive Statistics for Home-Firms with International R&D Networks within the Top 1000				
	Home-Firms with Networks	Disaster Impacted Networks	Non Impacted Networks	(2)-(3) t-stat [p-value]
	(1)	(2)	(3)	(4)
(1) Outcome:				
Successful Patent Applications per Year	44	62	17	3.62
	(90)	(111)	(26)	[0.00]
Successful High-Impact Patent Applications per Year	20	29	8	3.97
(>5forward citations per patent)	(42)	(52)	(16)	[0.00]
Forward Citations per Application per Year	162	229	63	3.69
	(328)	(405)	(89)	[0.00]
(2) Dyad Characteristics and Controls:				
Per Dyad Patent Count	16	25	4	2.52
	(61)	(78)	(9)	[0.01]
Network Size (total number of co-patenters)	2	3	2	3.59
	(4)	(3)	(1)	[0.00]
Dyad Strength:				
% of Dyads: Close (>=10 co-patents)	16	22	7	2.9
	(34)	(39)	(25)	[0.00]
% of Dyads: Standard(10>co-patents>=3)	18	20	14	1.01
	(34)	(35)	(33)	[0.31]
% of Dyads: Casual(3> co-patents)	67	58	79	-2.99
	(43)	(45)	(39)	[0.00]
Duration and Recency:				
Duration of Collaboration (Years)	2	3	1	2.46
	(4)	(4)	(3)	[0.02]
% of Dyads: Active Relationships	10	13	6	2.16
	(29)	(32)	(23)	[0.03]
% of Dyads: Recent Relationships	14	15	12	0.99
	(32)	(33)	(32)	[0.32]
% of Dyads: Old Relationships	76	72	82	-2.24
	(40)	(42)	(37)	[0.03]
Measures of Distance: Geographic and Social				
Distance (Miles)	6277	6328	6201	0.19
	(3905)	(3892)	(3926)	[0.848]
% of Dyads: Contiguous Border	18	15	23	-1.27
	(36)	(32)	(41)	[0.21]
% of Dyads: Common Language	16	13	21	-1.24
	(34)	(31)	(39)	[0.22]
Observations	1460	870	590	1460

Notes: The sample contains one observation for each firm-year observation amongst internationally networked firms over ten years from 1999-2008. The main entries in columns (1) through (3) are the mean of the selected variable. The entries in parentheses in columns (1) through (3) are the standard deviation of the selected variables. Reported t-statistics are obtained from a regression of the selected variable on an international-collaboration-treatment indicator variable. Observed differences in means amongst outcomes are driven by fixed characteristics of the firms. Because of the highly skewed nature of innovative production, if the top producer of patents is randomly lumped into either the treated or untreated group, this single firm's patent counts will sway the average patent counts of its group to be larger than the other. For example, if the top 5% of patent producers are removed from the sample, the mean for collaborative firms that are treated and untreated respectively are: 14.2 and 9.3, and they are not significantly different from each other at the 5% level. The same holds for the rest of the characteristics that are reportedly significantly different from each other. The inclusion of firm fixed effects will remove this distortion. See Figure 1 for an illustration of the dynamics between treatment and control groups with firm fixed effects which accounts for this distortion.

Table 3: Descriptive Statistics for Home-Firms with Subsidiary and Between-Firm Networks

	All Collaborative Firms	Home-Firms with Subsidiary Networks	Home-Firms with Between-Firm Networks	(2)-(3) t-stat [p-value]
	(1)	(3)	(2)	(4)
(1) Outcome:				
Successful Patent Applications per Year	44	94	24	2.34
	(90)	(140)	(49)	[0.02]
Successful High-Impact Patent Applications per Year	20	36	14	1.98
(>5forward citations per patent)	(42)	(62)	(30)	[0.05]
Forward Citations per Application per Year	162	345	93	2.38
	(328)	(519)	(172)	[0.02]
(2) Dyad Characteristics and Controls:				
Per Dyad Patent Count	16	38	8	2.42
	(61)	(81)	(50)	[0.02]
Network Size (total number of co-patenters)	2	4	2	2.5
	(3)	(4)	(1)	[0.01]
Dyad Strength:				
% of Dyads: Close (>=10 co-patents)	16	42	6	5.6
	(34)	(44)	(23)	[0.00]
% of Dyads: Standard(10>co-patents>=3)	18	16	18	-0.82
	(34)	(28)	(36)	[0.41]
% of Dyads: Casual(3> co-patents)	67	42	76	-4.91
	(43)	(43)	(40)	[0.00]
Duration and Recency:				
Duration of Collaboration (Years)	2	5	1	4.85
	(4)	(5)	(2)	[0.00]
% of Dyads: Active Relationships	10	20	6	3.21
	(29)	(36)	(24)	[0.00]
% of Dyads: Recent Relationships	14	16	13	2.17
	(32)	(32)	(32)	[0.03]
% of Dyads: Old Relationships	76	64	81	-3.78
	(40)	(43)	(38)	[0.00]
Measures of Distance: Geographic and Social				
Distance (Miles)	6277	6506	6189	0.33
	(3905)	(3439)	(4066)	[0.74]
% of Dyads: Contiguous Border	18	14	20	-1.29
	(36)	(29)	(38)	[0.20]
% of Dyads: Common Language	16	17	16	-0.27
	(34)	(33)	(35)	[0.78]
Observations	1460	400	1060	1460

Notes: The sample contains one observation for each firm-year observation amongst internationally collaborative firms over ten years from 1999-2008. The main entries in columns (1) through (3) are the mean of the selected variable. The entries in parentheses in columns (1) through (3) are the standard deviation of the selected variables. Reported t-statistics are obtained from a regression of the selected variable on an international-collaboration-treatment indicator variable. .

TABLE 4: The Impact of Home-Country Disasters on Economics Activity and Innovative Productivity

Dependent Variable:	*ln*(NASA Lights) $_i$	Patent Count$_i$
Estimation Technique:	OLS	PPML
	(1)	(3)
Home Country Disaster$_{i,t-1}$	-0.106***	-0.205
	(0.019)	(0.133)
Home Country Disaster$_{i,t-2}$	0.058	-0.253***
	(0.068)	(0.094)
Home Country Disaster$_{i,t-3}$	-0.297***	0.001
	(0.077)	(0.113)
Home Country Disaster$_{i,t-4}$	0.076	-0.101
	(0.081)	(0.186)
Pseudo-Log Likelihood:	-1150	-42500
N:	1460	1460
Mean of Dependent Variable:	2.11	29

Notes: Source: Author's Calculations. Each column represents a separate regression. Standard errors are clustered at the country level and are in parenthesis. The dependent variables are the log of the light density as measured by NASA and released by Henderson et al. (2011) and patent-counts summed at the country-year level. Each time period represents 1 year. All regressions include time and country fixed effects and controls for contemporaneous disaster effects. Column (1) is estimated using OLS, and the dependent variable is the log-transform of the level lights data to give a point estimate of the elasticity. In column (2) I present the OLS estimation of the regression of patent counts summed at the country-year level on the disaster indicator. However, with count data that is highly skewed, a more appropriate structural model is that of the Poisson, and I use the Poisson-Pseudo-Maximum-Likelihood estimation technique outlined in the section *Estimation Technique*. Effects are negative and significant especially in the two years following the disaster. This shows the general economic downturn in economic activity tied with the large scale disasters used throughout the study.

* indicates 10% level of statistical significance
** indicates 5% level of statistical significance
*** indicates 1% level of statistical significance

TABLE 5: The Affect of International$_j$-Network Disasters on Home-Firm Innovative Productivity

Estimation Technique:	PPML		
Dependent Variable:	Patent Counts$_i$	High Impact Patent Counts$_i$	Citation Weighted Patent Counts$_i$
	(1)	(2)	(3)
International Disaster$_{j,t-1}$	0.001	0.025	-0.001
	(0.083)	(0.091)	(0.082)
International Disaster$_{j,t-2}$	-0.063	-0.148	-0.101
	(0.091)	(0.105)	(0.085)
International Disaster$_{j,t-3}$	0.005	0.063	0.041
	(0.186)	(0.156)	(0.171)
International Disaster$_{j,t-4}$	0.020	0.0306	-0.006
	(0.135)	(0.114)	(0.118)
Pseudo-Log Likelihood:	-9600	-5230	-37900
N:	1460	1460	1460
Mean of Dependent Variable:	44	20	162

Notes: Source: Author's Calculations. Each column represents a separate regression. Columns (1) through (3) are estimated using Poisson-Pseudo-Maximum-Likelihood. PPML estimates are inferable as elasticities rather than levels. Standard errors are in parenthesis under all point estimates and are clustered at the firm level. Dependent variables are collapsed to the sum at the firm-year level. All regressions include year and firm fixed effects as well as controls for home-country disasters and contemporaneous disaster effects. Similar to the findings presented in Figure 2, there is some, but limited evidence of short-term spillover effects. There is no statistically significant evidence that spillovers occur in the medium run.

* indicates significantly different from zero at the 10% level of significance
** indicates significantly different from zero at the 5% level of significance
*** indicates significantly different from zero at 1% level of significance

TABLE 6: Differential Effects of Subsidiary and Between-Firm Network Disasters

Estimation Technique:	PPML		
Dependent Variable:	Patent Counts$_i$	High Impact Patent Counts$_i$	Citation Weighted Patent Counts$_i$
	(1)	(2)	(3)
International Subsidiary Disaster$_{j,t-1}$	-0.137	-0.113	-0.120
	(0.148)	(0.228)	(0.166)
International Subsidiary Disaster$_{j,t-2}$	-0.136	-0.160	-0.196***
	(0.089)	(0.105)	(0.070)
International Subsidiary Disaster$_{j,t-3}$	-0.278	-0.444*	-0.313*
	(0.200)	(0.253)	(0.185)
International Subsidiary Disaster$_{j,t-4}$	0.030	-0.057	0.038
	(0.187)	(0.244)	(0.178)
International Between-Firm Network Disaster$_{j,t-1}$	0.054	0.105	0.061
	(0.094)	(0.087)	(0.092)
International Between-Firm Network Disaster$_{j,t-2}$	-0.072	-0.185	-0.110
	(0.111)	(0.118)	(0.109)
International Between-Firm Network Disaster$_{j,t-3}$	0.075	0.172	0.132
	(0.224)	(0.182)	(0.205)
International Between-Firm Network Disaster$_{j,t-4}$	0.078	0.054	0.031
	(0.145)	(0.144)	(0.148)
Pseudo-Log Likelihood:	-9560	-5190	-37600
N:	1460	1460	1460
Mean of Dependent Variable:	44	20	162

Notes: Source: Author's Calculations. Each column represents a separate regression. Columns (1) through (3) are estimated using PPML. PPML estimates are inferable as elasticities rather than levels. Standard errors are in parenthesis under the point estimates and are clustered at the firm level. Patent counts are collapsed to the sum at the firm-year level. All regressions include year and firm fixed effects as well as controls for home-country disasters and contemporaneous disaster effects. Columns (1) through (4) present a regression where the dependent is regressed upon only disasters that occur to countries in which an international subsidiary to the home firm is located. Columns (5) through (8) present a regression where the dependent is regressed upon only disasters that occur to countries in which a different firm is the collaborator. It is critical to note that if a given home firm has both within and between firm collaborations in the same country, then only the net effect is identified. However, if a firm has does both intra and inter-firm knowledge sourcing in different countries, I am able to use this variation to identify the differential effects. This will shed light as to what the effects of entering into a relationship with a competitor versus international expansion may have on knowledge sourcing. ***NOTES***
* indicates significantly different from zero at the 10% level of significance
** indicates significantly different from zero at the 5% level of significance
*** indicates significantly different from zero at 1% level of significance

TABLE 7: The Affect of International$_j$-Network Disasters on Home-Firm Innovative Productivity: Between-Firm Effects Only

Estimation Technique:	PPML		
Dependent Variable:	Patent Counts$_i$	High Impact Patent Counts$_i$	Citation Weighted Patent Counts$_i$
	(1)	(2)	(3)
International Disaster$_{j,t\text{-}1}$	0.09	0.215**	0.12
	(0.09)	(0.09)	(0.09)
International Disaster$_{j,t\text{-}2}$	-0.05	-0.211*	-0.12
	(0.11)	(0.12)	(0.11)
International Disaster$_{j,t\text{-}3}$	-0.03	0.14	0.06
	(0.22)	(0.18)	(0.21)
International Disaster$_{j,t\text{-}4}$	0.10	0.04	0.04
	(0.16)	(0.16)	(0.16)
Pseudo-Log Likelihood:	-4960	-2860	-18900
N:	805	805	805
Mean of Dependent Variable:	44	20	162

Notes: Source: Author's Calculations. Each column represents a separate regression. Columns (1) through (3) are estimated using Poisson-Pseudo-Maximum-Likelihood. PPML estimates are inferable as elasticities rather than levels. Standard errors are in parenthesis under all point estimates and are clustered at the firm level. Dependent variables are collapsed to the sum at the firm-year level. All regressions include year and firm fixed effects as well as controls for home-country disasters and contemporaneous disaster effects. After stratifying the sample to be home-firms that have only between-firm networks, we see relatively weak evidence of spillovers. In fact, even in column (2) where there appears to be some evidence of an effect, the evidence is mixed.

* indicates significantly different from zero at the 10% level of significance
** indicates significantly different from zero at the 5% level of significance
*** indicates significantly different from zero at 1% level of significance.

TABLE 8: The Affect of International$_j$-Network Disasters on Home-Firm Innovative Productivity: Subsidiary Effects Only

Estimation Technique:	PPML		
Dependent Variable:	Patent Counts$_i$	High Impact Patent Counts$_i$	Citation Weighted Patent Counts$_i$
	(1)	(2)	(3)
International Disaster$_{j,t\text{-}1}$	0.41	0.419	0.40
	(0.28)	(0.31)	(0.28)
International Disaster$_{j,t\text{-}2}$	-0.20	-0.331*	-0.401*
	(0.17)	(0.19)	(0.21)
International Disaster$_{j,t\text{-}3}$	-0.693***	-0.862***	-0.711***
	(0.26)	(0.24)	(0.26)
International Disaster$_{j,t\text{-}4}$	0.06	0.14	0.06
	(0.25)	(0.28)	(0.22)
Pseudo-Log Likelihood:	-414	-285	-1470
N:	110	110	110
Mean of Dependent Variable:	44	20	162

Notes: Source: Author's Calculations. Each column represents a separate regression. Columns (1) through (3) are estimated using Poisson-Pseudo-Maximum-Likelihood. PPML estimates are inferable as elasticities rather than levels. Standard errors are in parenthesis under all point estimates and are clustered at the firm level. Dependent variables are collapsed to the sum at the firm-year level. All regressions include year and firm fixed effects as well as controls for home-country disasters and contemporaneous disaster effects. After stratifying the sample to be home-firms that have only subsidiary networks, we see relatively consistent evidence of spillovers. Here we see that the home-firm does suffer a negative impact on patent production starting in the second year after their network incurs the disaster. Large negative effects do imply that these home-firms do rely upon their network subsidiaries in the production of new innovation.

* indicates significantly different from zero at the 10% level of significance
** indicates significantly different from zero at the 5% level of significance
*** indicates significantly different from zero at 1% level of significance.

TABLE 9: The Interaction of Collaborative Disasters and Geographic Distance and Social Distance

Estimation Technique:	PPML			
Dyad Characteristic:	Distance from Average Collaborator		Common Official Language of Collaborator	
Stratification Categories:	Less than Median Distance	Greater than Median Distance	Country Shares Common Official Language	Country Does Not Share Common Official Language
Dependent Variable:	Patent Counts$_i$			
	(1)	(2)	(3)	(4)
Disaster$_{j,t-1}$	0.065	0.029	-0.095	-0.009
	(0.142)	(0.136)	(0.194)	(0.088)
Disaster$_{j,t-2}$	-0.145	0.029	-0.191	-0.025
	(0.180)	(0.159)	(0.230)	(0.092)
Disaster$_{j,t-3}$	0.164	-0.260	-0.163	-0.064
	(0.308)	(0.281)	(0.281)	(0.203)
Disaster$_{j,t-4}$	0.059	-0.400	-0.323	0.079
	(0.252)	(0.252)	(0.367)	(0.154)
Pseudo-Log Likelihood:	-5240	-4080	-548	-6090
N:	719	459	136	856

Notes: Source: Author's Calculations. Each column represents a separate regression. Standard errors are in parenthesis and are clustered at the firm level. Patent counts are collapsed at the firm-year level. Each time period represents 1 year. All regressions include time and firm fixed effects and controls for home-country disasters and contemporaneous disaster effects. Surprisingly, there is no evidence that distance plays a role in determining the relative impact of international R&D networks in home-firm patent production.

* indicates significantly different from zero at the 10% level of significance
** indicates significantly different from zero at the 5% level of significance
*** indicates significantly different from zero at 1% level of significance

TABLE 10: The Effect of the Collaborator-Disasters on Home-Firm Productivity Stratified by Patent Classification

Estimation Technique:	Poisson Pseudo-Maximum Likelihood (PPML)		
Dependent Variable:	Patent Counts		
IPC Classification:	Human Necessities	Performing Operations	Chemistry, Metallurgy
	(1)	(2)	(3)
$\text{Disaster}_{j,t\text{-}1}$	-0.042	0.039	-0.061
	(0.123)	(0.095)	(0.137)
$\text{Disaster}_{j,t\text{-}2}$	-0.172	-0.135	0.019
	(0.118)	(0.097)	(0.149)
$\text{Disaster}_{j,t\text{-}3}$	-0.054	0.033	-0.099
	(0.145)	(0.172)	(0.258)
$\text{Disaster}_{j,t\text{-}4}$	0.193	0.084	0.221
	(0.197)	(0.189)	(0.261)
Frontier Measure: Proportion of Total Patents in Classification	32	43	34
Pseudo-Log Likelihood:	-5370	-6000	-6380
N:	602	668	728

Notes: Source: Author's Calculations. Each column represents a separate regression. Standard errors are in parenthesis and are clustered at the firm level. Patent counts are collapsed at the firm-year level. Each time period represents 1 year. All regressions include time and firm fixed effects and controls for home-country disasters and contemporaneous disaster effects. As previously noted, the different IPC classifications are not mutually exclusive as one patent may serve several purposes. There is no general significant difference amongst the effects based upon industry effects.

* indicates significantly different from zero at the 10% level of significance
** indicates significantly different from zero at the 5% level of significance
*** indicates significantly different from zero at 1% level of significance

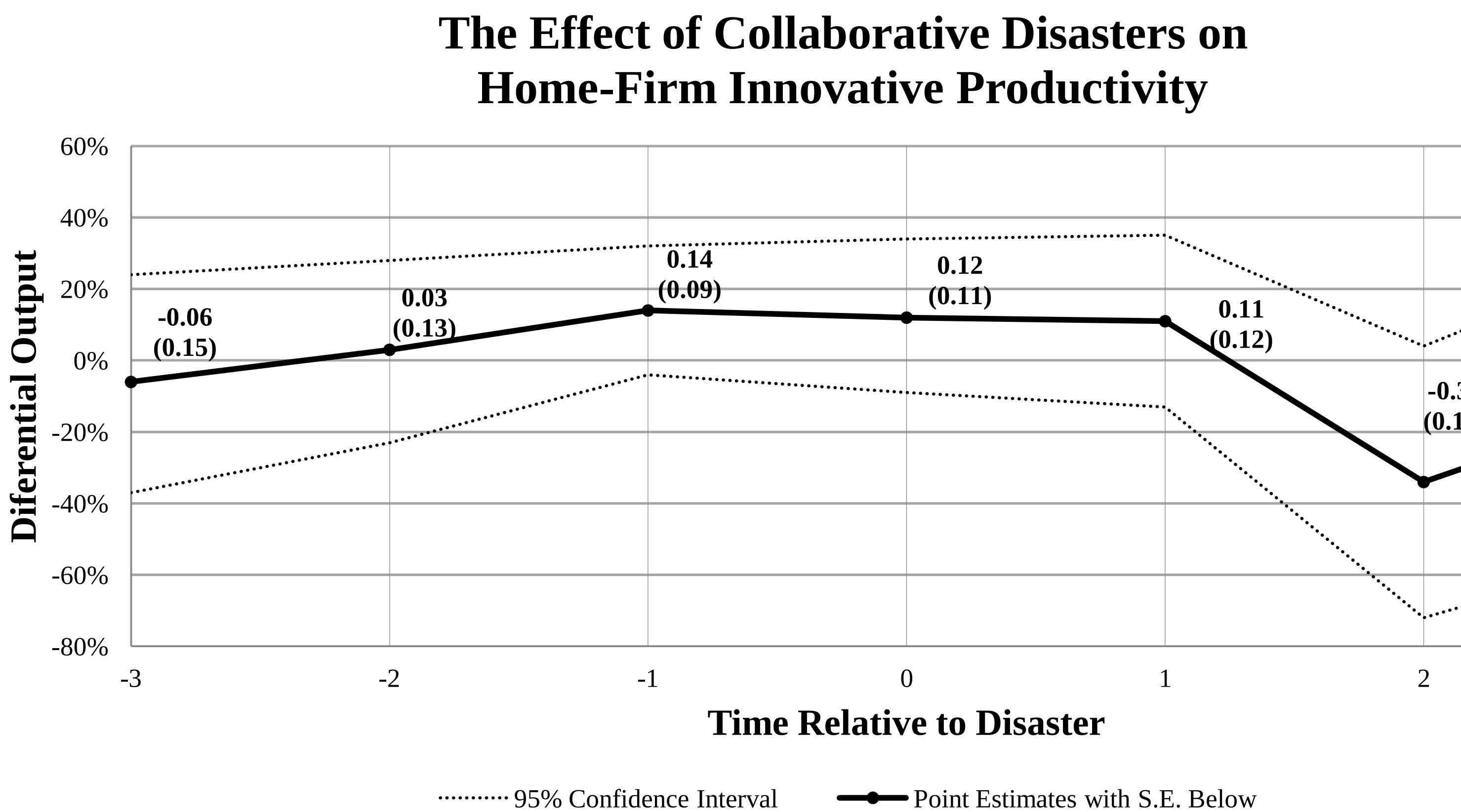
TABLE 11: Removal of the Super-Star Firms and the Effect on the Main Result

Estimation Technique:	PPML					
Sample Adjustment:	Bottom 95% Based on Home-Firm Patent Counts			Bottom 95% Based on Dyad Patent Counts		
Dependent Variable:	Patent Counts_i	High Impact Patent Counts_i	Citation Weighted Patent Counts_i	Patent Counts_i	High Impact Patent Counts_i	Citation Weighted Patent Counts_i
	(1)	(2)	(3)	(4)	(5)	(6)
$\text{Disaster}_{j,t\text{-}1}$	-0.048	-0.038	-0.046	-0.007	0.017	-0.008
	(0.073)	(0.084)	(0.081)	(0.083)	(0.091)	(0.083)
$\text{Disaster}_{j,t\text{-}2}$	-0.048	-0.134	-0.094	-0.061	-0.143	-0.096
	(0.087)	(0.095)	(0.078)	(0.092)	(0.107)	(0.086)
$\text{Disaster}_{j,t\text{-}3}$	-0.004	0.028	0.013	-0.002	0.056	0.034
	(0.133)	(0.130)	(0.133)	(0.184)	(0.155)	(0.169)
$\text{Disaster}_{j,t\text{-}4}$	0.050	0.041	-0.009	0.029	0.039	0.001
	(0.114)	(0.101)	(0.098)	(0.140)	(0.122)	(0.123)
Pseudo-Log Likelihood:	-6800	-3820	-26100	-9000	-4870	-35400
N:	1110	1110	1110	1100	1100	1100

Notes: Source: Author's Calculations. Each column represents a separate regression. Standard errors are in parenthesis. Patent counts are collapsed at the firm-year level. Each time period represents 1 year. All regressions include time and firm fixed effects and controls for home-country disasters and contemporaneous disaster effects. Here we see no evidence that the inclusion or exclusion of the "super-star" firms or networks impacts the null result found in Table 5.

* indicates significantly different from zero at the 10% level of significance
** indicates significantly different from zero at the 5% level of significance
*** indicates significantly different from zero at 1% level of significance

Figure 1: The Effect of Collaborative Disasters on Home-Firm Innovative Productivity



Notes: Source: Author's Calculations. The trend line represents the differential output that occurs between firms whose collaborator is shocked by a disaster versus those whose collaborator is not. Before the disaster, there is no statistically significant difference in trend between the two groups. There is limited evidence that spillovers occur, with only a short term statistically weakly significant drop in the second year after the shock. However, immediately thereafter, the trends between the two types of firms converge clearly both in terms of the point estimates and their statistical significance. The trend line is estimated from a single regression of the home-firm patent counts on large-scale natural disasters that impact their collaborative partners. The regression follows the dynamic framework outlined in the section *Setting, Data, and Identification Strategy* with firm-year patent counts regressed on the collaborator-disaster from contemporaneous to both three years in the past and future. The regression includes year and firm fixed effects, as well as home-disaster controls. The standard errors are clustered at the firm level.

7 Appendix: Additional Tables and Figures

Figure F1: The Distribution of Annual Patent Counts Among the 1000 Most Innovative Firms in the OECD

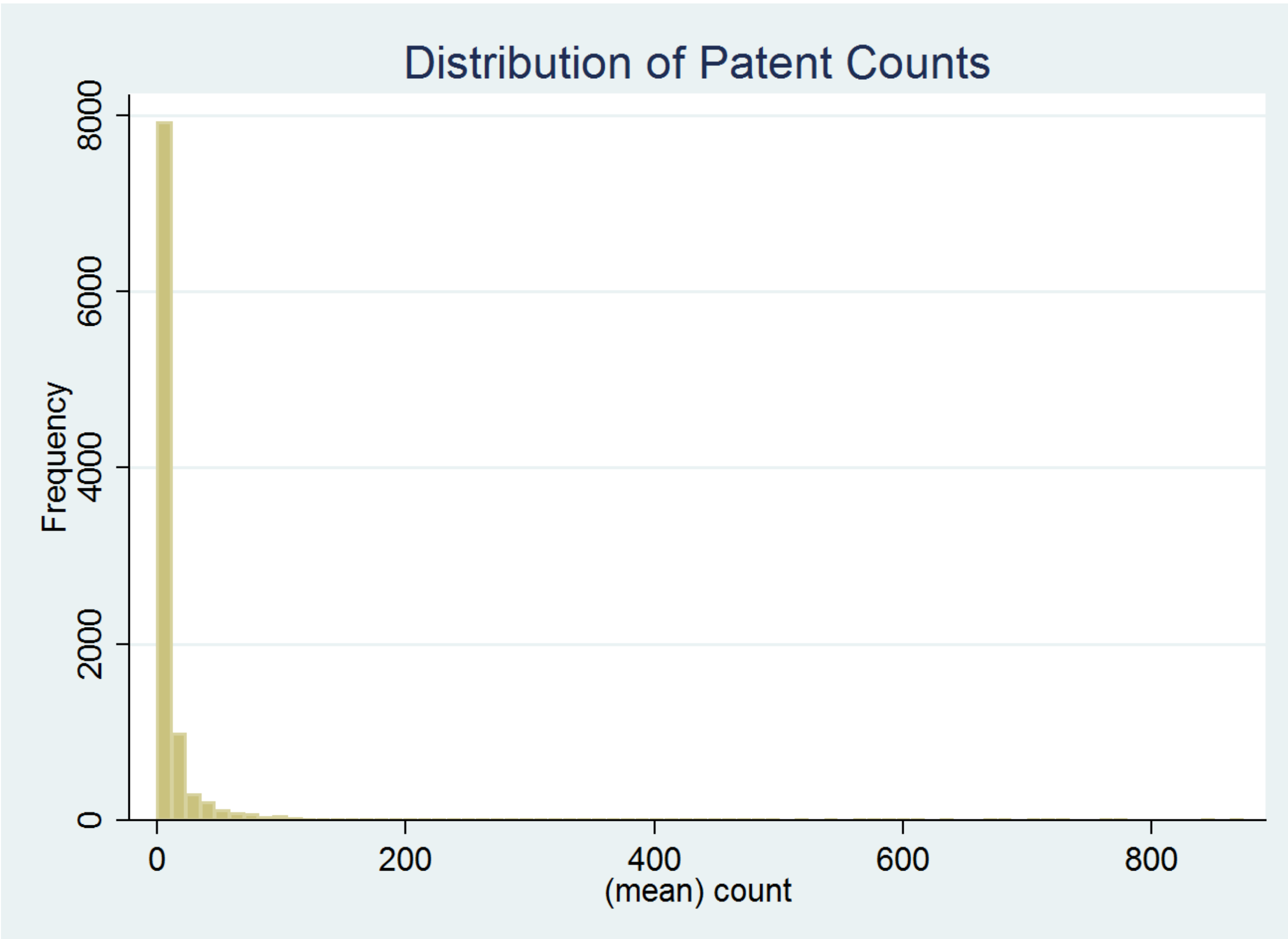


Figure F2: Dyad Creation Illustration and Disaster Weights

A=Top 1000 Innovative Firm 1

B=Top 1000 Innovative Firm 2

Ã=Subsidiary of Firm 1

C=Firm 3 (Not in the Top 1000)

Disaster Weighting

If a disaster occurs in A, it will receive a weight of 1/3 as only 1/3 of Firm B's potential knowledge comes from this source

Dyads Created:

Home-Firm $\rightarrow$ International Knowledge Source

Dyad 1:A $\rightarrow$ Ã and Copatents = 10 in building period

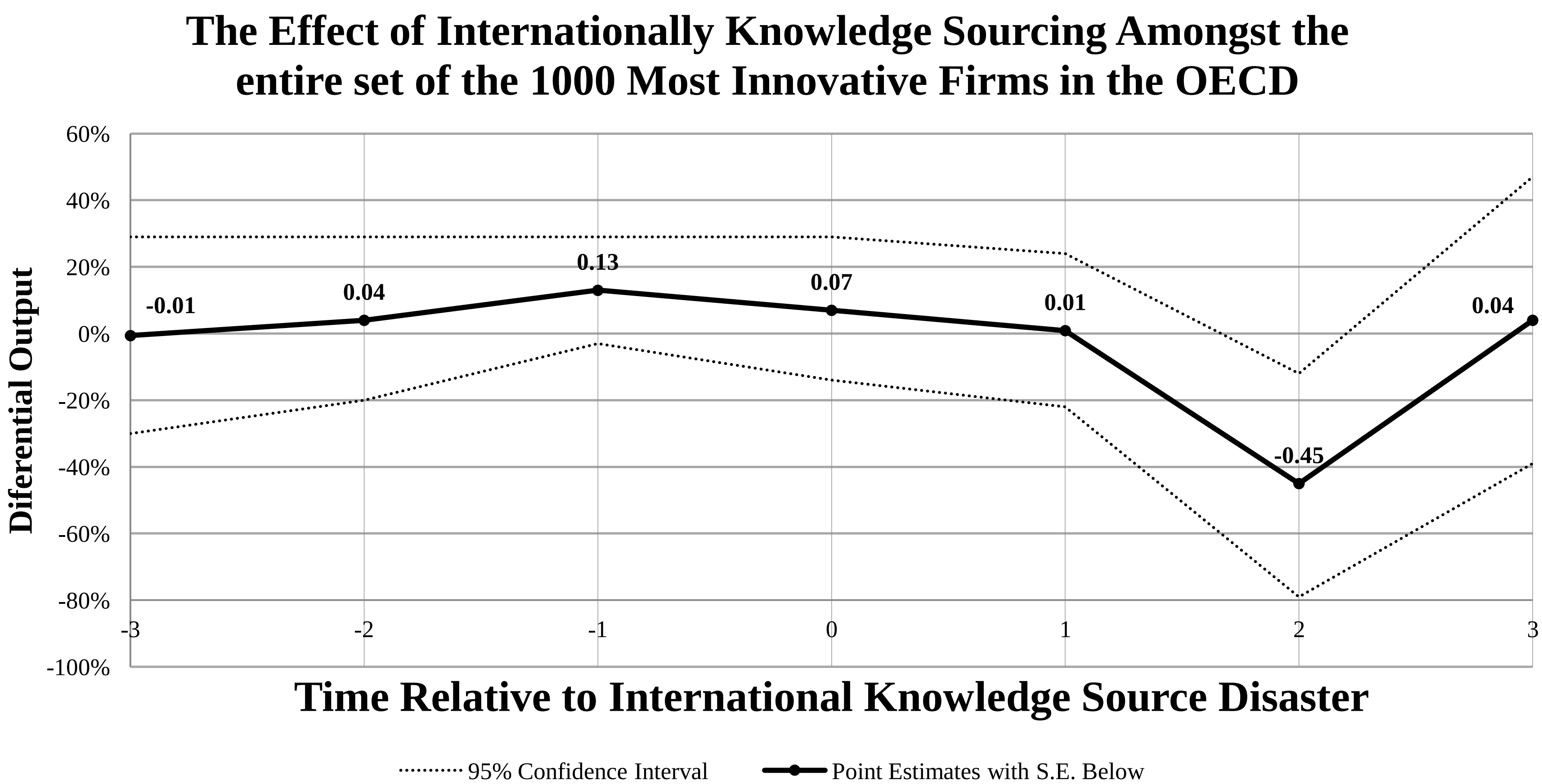
Dyad 2:A $\rightarrow$ B and Copatents = 10 in building period

Dyad 3:A $\rightarrow$ C and Copatents = 20 in building period

Dyad 4:B $\rightarrow$ A and Copatents = 10 in building period

Dyad 5:B $\rightarrow$ Ã and Copatents = 20 in building period

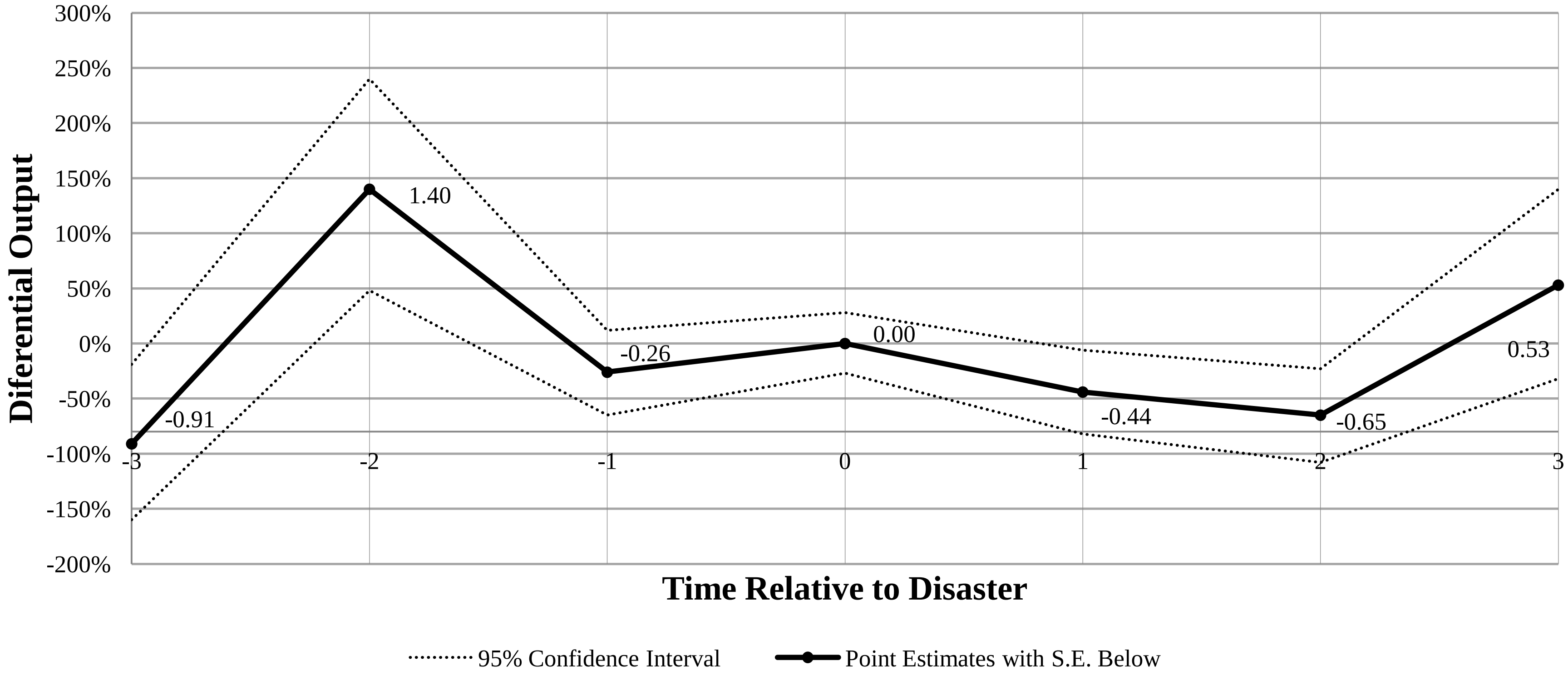
Figure F3: The Effect of Collaboration Shock to Collaborative firms Versus All Other Firms in the Top 1000



Notes: Source: Author's Calculations. The trend line represents the differential output that occurs between firms whose collaborator is shocked by a disaster versus those unshocked. Before the disaster, there is no statistically significant difference in trend between the two groups. There is significant evidence that spillovers occur in the second year. However, immediately thereafter, the trends between the two types of firms converge clearly both in terms of the point estimates and their statistical significance. The trend line is estimated from a single regression of the home-firm patent counts on large-scale natural disasters that impact their collaborative partners, whether subsidiary or not. The regression follows the dynamic framework outlined in the section *Setting, Data, and Identification Strategy* with firm-year patent counts regressed on the collaborator-disaster from contemporaneous to both three years in the past and future. The regression includes year and firm fixed effects, as well as home-disaster controls. The standard errors are clustered at the firm level.

Figure F4: The Effect of Disasters on Subsidiary Networks

The Effect of Collaborative Disasters on Subsidiaries



Notes: Source: Author's Calculations. The trend line represents the differential output that occurs between firms whose collaborator is shocked by a disaster versus those unshocked. Prior to the disaster the data is noisy with large movements around the horizontal axis, but with no pattern. In the year prior and disaster year, trends between home-firms whose subsidiaries are impacted and those who are not are statistically zero. There is significant evidence that spillovers occur in the second and third years after the disaster. However, immediately thereafter, the trends between the two types of firms converge clearly both in terms of the point estimates and their statistical significance. This is inline with the recovery period experienced after large scale disasters. The trend line is estimated from a single regression of the home-firm patent counts on large-scale natural disasters that impact their collaborative partners, whether subsidiary or not. The regression follows the dynamic framework outlined in the section *Setting, Data, and Identification Strategy* with firm-year patent counts regressed on the collaborator-disaster from contemporaneous to both three years in the past and future. The regression includes year and firm fixed effects, as well as home-disaster controls. The standard errors are clustered at the firm level.

Table A1: Descriptive Statistics for Home-Firms with Only Subsidiary Based Networks

	Home-Firms with Networks (1)	Disaster Impacted Networks (2)	Non Impacted Networks (3)	(2)-(3) t-stat [p-value] (4)
(1) Outcome:				
Successful Patent Applications per Year	20	23	13	1.3
	(18)	(20)	(10)	[0.22]
Successful High-Impact Patent Applications per Year (>5forward citations per patent)	14	16	9.8	0.76
	(21)	(25)	(10)	[0.46]
Forward Citations per Application per Year	89	108	55	1.36
	(87)	(98)	(45)	[0.20]
(2) Dyad Characteristics and Controls:				
Per Dyad Patent Count	15	14	18	-0.27
	(22)	(18)	(27)	[0.79]
Network Size (total number of co-patenters)	1.1	1.2	1	1.54
	(0.35)	(.41)	(0.00)	[0.148]
Dyad Strength:				
% of Dyads: Close (>=10 co-patents)	42	44	40	0.13
	(49)	(49)	(49)	[0.90]
% of Dyads: Standard(10>co-patents>=3)	14	11	20	-0.41
	(35)	(32)	(40)	[0.69]
% of Dyads: Casual(3> co-patents)	43	45	40	0.18
	(49)	(49)	(49)	[0.86]
Duration and Recency:				
Duration of Collaboration (Years)	3.9	3.4	4.9	-0.56
	(4.5)	(4.1)	(5.1)	[0.59]
% of Dyads: Active Relationships	15	11	22	-0.82
	(35)	(31)	(42)	[0.43]
% of Dyads: Recent Relationships	21	25	12	1.77
	(40)	(43)	(33)	[0.10]
% of Dyads: Old Relationships	64	63	66	-0.12
	(47)	(48)	(48)	[0.91]
Measures of Distance: Geographic and Social				
Distance (Miles)	6432	7722	4108	1.67
	(3676)	(2487)	(4304)	[0.12]
% of Dyads: Contiguous Border	14	0	40	-1.75
	(35)	(0.00)	(49)	[0.10]
% of Dyads: Common Language	14	0	40	-1.75
	(35)	(0.00)	(49)	[0.10]
Observations	140	90	50	140

Notes: The sample contains one observation for each firm-year observation amongst internationally networked firms over ten years from 1999-2008. The main entries in columns (1) through (3) are the mean of the selected variable. The entries in parentheses in columns (1) through (3) are the standard deviation of the selected variables. Reported t-statistics are obtained from a regression of the selected variable on an international-collaboration-treatment indicator variable.

Table A2: Descriptive Statistics for Home-Firms with Only Between-Firm Networks

	Home-Firms with Networks	Disaster Impacted Networks	Non Impacted Networks	(2)-(3) t-stat [p-value]
	(1)	(2)	(3)	(4)
(1) Outcome:				
Successful Patent Applications per Year	24	34	13	2.4
	(49)	(62)	(20)	[0.02]
Successful High-Impact Patent Applications per Year	14	21	6	3.2
(>5forward citations per patent)	(30)	(38)	(11)	[0.00]
Forward Citations per Application per Year	93	129	48	2.71
	(172)	(216)	(64)	[0.01]
(2) Dyad Characteristics and Controls:				
Per Dyad Patent Count	8	13	2	1.26
	(50)	(67)	(4)	[0.21]
Network Size (total number of co-patenters)	1.5	1.8	1.3	2.31
	(1.4)	(1.8)	(0.66)	[0.02]
Dyad Strength:				
% of Dyads: Close (>=10 co-patents)	6	8	2	1.49
	(23)	(27)	(1.5)	[0.14]
% of Dyads: Standard(10>co-patents>=3)	18	23	12	1.53
	(36)	(38)	(33)	[0.13]
% of Dyads: Casual(3> co-patents)	76	69	85	-2.21
	(40)	(43)	(35)	[0.03]
Duration and Recency:				
Duration of Collaboration (Years)	1.1	1.4	0.7	1.54
	(2.4)	(2.7)	(2.0)	[0.13]
% of Dyads: Active Relationships	6	8	4	1.05
	(24)	(27)	(20)	[0.30]
% of Dyads: Recent Relationships	13	13	13	0.24
	(32)	(32)	(33)	[0.81]
% of Dyads: Old Relationships	81	79	83	-0.98
	(38)	(39)	(37)	[0.33]
Measures of Distance: Geographic and Social				
Distance (Miles)	6189	6215	6156	0.07
	(4066)	(4231)	(3857)	[0.94]
% of Dyads: Contiguous Border	20	16	24	-1.00
	(38)	(35)	(42)	[0.32]
% of Dyads: Common Language	16	13	20	-0.88
	(35)	(33)	(38)	[0.38]
Observations	1060	590	470	1060

Notes: The sample contains one observation for each firm-year observation amongst internationally networked firms over ten years from 1999-2008. The main entries in columns (1) through (3) are the mean of the selected variable. The entries in parentheses in columns (1) through (3) are the standard deviation of the selected variables. Reported t-statistics are obtained from a regression of the selected variable on an international-collaboration-treatment indicator variable.

TABLE A3: The Affect of Between-Firm Network-Disasters on Home-Firm Innovative Productivity

Estimation Technique:	PPML	
Stratification Category:	Common Official Language of Collaborator	
	Country Shares Common Official Language	Country Does Not Share Common Official Language
Dependent Variable:	Patent Counts$_i$	
	(1)	(2)
Disaster$_{j,t\text{-}1}$	-0.903***	0.133
	(0.257)	(0.100)
Disaster$_{j,t\text{-}2}$	-1.100**	0.021
	(0.446)	(0.141)
Disaster$_{j,t\text{-}3}$	-1.307***	-0.113
	(0.415)	(0.265)
Disaster$_{j,t\text{-}4}$	-2.020***	0.236
	(0.406)	(0.159)
Pseudo-Log Likelihood:	-385	-3980
N:	116	626
Mean of Dependent Variable:	44	162

Notes: Source: Author’s Calculations. Each column represents a separate regression. Standard errors are in parenthesis. Patent counts are collapsed at the firm-year level. Each time period represents 1 year. All regressions include time and firm fixed effects and controls for home-country disasters and contemporaneous disaster effects. There is evidence that social barriers are important. However, without further inquiry into what the CEPII measure of common language is capturing, making strong claims based on this evidence is questionable.

* indicates significantly different from zero at the 10% level of significance
** indicates significantly different from zero at the 5% level of significance
*** indicates significantly different from zero at 1% level of significance