# Benefiting Colleagues but not the City: Localized Effects from the Relocation of Superstar Inventors

Paolo Zacchia[*]

February 2018

### Abstract

In this paper I examine episodes in which superstar inventors relocate to a new city. In particular, in order to assess whether the beneficial effects of physical proximity to a superstar have a restricted network dimension or a wider spatial breadth (spillovers), I estimate changes in patterns of patenting activity following these events for two different groups of inventors: the superstar's close collaborators, and all the other inventors in a given urban area, for both the locality where the superstar moves to and for the one that is left behind. In the case of collaborators, I restrict the attention to patents realized independently from the superstar. The results from the event study register a large and persistent positive effect on the collaborators in the city of destination, as well as a simultaneous negative trend affecting those still residing in the previous location. In the long run, these effects translate into an increased difference between the two groups of about 0.16 patents per inventor. Conversely, no city-wide spillover effect can be attested, offering little support to place-based policies aimed at inducing a positive influx of top innovators in urban areas.

---

[*]IMT School of Advanced Studies, Lucca Italy

Theories of knowledge spillovers are predicated on a variety of social mechanisms, which all explain – across different settings – the occurrence of information exchange through individual interaction. Among these mechanisms, geographic proximity occupies an ambiguous position. On the one hand, spatial vicinity can be thought as a factor that is complementary to the process of knowledge creation, which occurs through formal professional collaborations within or between organizations. On the other hand, proximity may also favor the diffusion of new ideas among spatially close, but professionally unrelated workers. The latter scenario entails agglomeration externalities, a circumstance that can explain the endogenous co-location of individuals and firms across space, with far-reaching policy implications.

Despite the relevance of this issue, economic research does not offer a clear description and supporting evidence of how geographical proximity drives the emergence and diffusion of economically valuable new ideas. In order to isolate local knowledge spillovers, in this paper I analyze episodes in which "superstar" inventors – those in the top 5% of the patent distribution – relocate across North American or European cities. In particular, I examine the impact of these events over time on the "residual" patent output (patents not coauthored with the superstar) of four groups of inventors: that is, the individuals who belong to the network of patent collaborators of the superstar and those who do not, for both localities of departure and destination. In addition, I analyze how these effects vary between superstars' coworkers and other kinds of collaborators, with the relative position of the moving star in the patent distribution, across technological fields, and by the geographical extent of the move (whether it involved moving across the Atlantic, and in what direction).

The main findings from this event analysis can be summarized as follows. The relocation of superstar inventors appears to follow periods of time when both groups of collaborators – those in the locality of destination, and those left behind – experience similarly increasing trends of their residual patent production. Following the event, however, the trend relative to the colleagues left behind experiences a reversal, which manifests itself with a one year lag (this might reflect the lag of the R&D-to-patenting process). Conversely, the residual patent output of the newly neighboring colleagues keeps increasing for a few years to eventually remain stable. Thus, following the event the difference between the two groups increases – relative to the baseline year – up to about 0.16 extra yearly patents per inventor in favor of the benefiting network. The main qualitative result is robust to a variety of patent measures, and it appears stronger if coauthors from either locality have also been coworkers of the moving superstar, if the latter belongs to the top 1% of the patent distribution, and in the Electro-ICT and Mechanical sectors.

Nevertheless, there seems to be no strong evidence in favor of a similar effect on the patent output of all other inventors in both cities involved. Instead, the results indicate that superstars move towards cities that, relative to their city of departure, experience a slightly stronger

1

increase in their total patent output in the years before and after the event, but not an increase in the overall quality of such patents (as measured by patent citations) or in the average patent production of their inventors. Taken together, these findings indicate that knowledge spillovers, if they exist, appear largely confined within the circle of one inventor's close connections; by the terminology introduced in this paper, local knowledge spillovers do not seem to be very much *external* to the superstar's network. As a consequence, the evidence in favor of "big push" type of place-based innovation policies – realized on the expectation that attracting few highly creative individuals may result into a wider equilibrium shift in local innovativeness – is mixed.

This paper builds on the tradition of economic studies searching for geographically localized externalities associated with R&D activity. In their renown seminal work, Jaffe, Trajtenberg and Henderson (1993, JTH) show how patent citations – which are seen as the "paper trail" of knowledge spillovers[1] – tend to come from the same urban area as the cited patent. In subsequent research, however, JTH are notoriously criticized by Thompson and Fox-Kean (2005, TF). The latter show that, upon "zooming in" to finer technological classes, the result by JTH vanish, suggesting that the originally evidenced correlation reflects the spatial distribution of economic and technological activity. By contrast Breschi and Lissoni (2009, BL), in a contribution that anticipates some of the ideas explored in this paper, show that the results by JTH become smaller in magnitude – but are still economically significant – when controlling for the "social distance" (in the network sense) between the two inventor teams behind both citing and cited patents, even within narrow technological sector as per the analysis by TF.

In addition to BL, other studies about knowledge spillovers incorporate a social or network dimension. Agrawal, Kapur and McHale (2008) show that while reducing either the network or the geographical distance between inventors increases the probability of citing a patent, these two factors seem to be substitutes rather than complements in knowledge production, which is consistent with the results presented in this paper. In another study, Agrawal, Cockburn and Rosell (2010) attest that in "company towns" patent citations are concentrated within the same firm, suggesting an organizational dimension of spillovers. Breschi, Lissoni and Miguelez (2017) study both the "diaspora" and "brain gain" effects at work among expatriate inventors, that is respectively the role of ethnic and home country ties in determining patent citations. In this paper, social networks shaped by co-patenting relationship are interacted with geography, with the aim of disentangling the spatial and social dimensions of knowledge flows.

Studies of knowledge spillovers typically differ by the geographical scope of the knowledge flows they examine. Some analyses focus on cross-country spillovers (Branstetter, 2001; Keller, 2002; Singh, 2007; Mancusi, 2008), while others delve into smaller geographical units, like states

---

[1]This definition is based upon one famous critical statement by Krugman (1991), according to whom "knowledge flows […] are invisible; they leave no paper trail by which they may be measured and tracked."

and regions (Audretsch and Feldman, 1996; Almeida and Kogut, 1999; Singh and Marx, 2013; Bloom et al., 2013). In this paper, I consider the city or metropolitan area as the relevant spatial dimension of knowledge spillovers. In this respect, a notable precursor is the study by Moretti (2004), who shows that plant productivity is predicted by the share of college-educated workers from other industries in the city where the plant operates. Furthermore, the analysis of firm-to-firm R&D spillovers by Lychagin et al. (2016) provides an interesting take on the determination of spillovers' geographical scope: by estimating models featuring effects of external R&D that decay with distance, they argue that spillovers are largely confined within small areas.

The central identification problem faced by all these researchers is how to separate genuine geographic externalities from other common factors that are shared within the localities of interest. In fact, this methodological issue is not specific to the analysis of knowledge spillovers, as it is shared by all urban and trade economists searching for general agglomeration economies (of which knowledge spillovers is usually thought to be one main determinant). A celebrated approach, employed by Greenstone, Hornbeck and Moretti (2010) in their analysis of the effect of large plants on local productivity, is to exploit tail events affecting "winning," but not similar "losing" localities, and compare the outcomes of the two places against one another. This paper is based on a similar idea: specifically, I compare the patenting outcomes of selected groups of inventors between places that "receive" and places that "lose" superstar inventors. While the superstar's decision to move might be endogenous to the characteristics of the two groups, the machinery of the event study analysis allows to evaluate if their pre-event trends are sufficiently similar to one another and, if they are, how do they compare with the post-event dynamics.

This paper is also related to all those studies examining the economics of innovative "superstar" professionals, with a focus on very successful inventors and academics. Perhaps most famously among these studies, the work by Azoulay, Graff Zivin and Wang (2010) provides evidence on the role of superstars in stimulating the intellectual production of other scientists, by documenting the negative consequences of a superstar's "extinction" on his network of strict collaborators.[2] Oettl (2012) extends their analysis and methodology to what he calls *mavens*, that is, scientists of below-stellar productivity but of high "helpfulness" towards other scientists. In another related and influential study, Waldinger (2011) finds however more mixed evidence of star effects following the exile of Jewish top academics from Nazi Germany. There is also a growing interest in the analysis of the migration patterns of superstar inventors. In particular, recent work has assessed the role of marginal tax rates in determining the location choice of top inventors, both between U.S. cities and states (Moretti and Wilson, 2014, 2017) and countries (Akcigit, Baslandze and Stantcheva, 2016).

---

[2]The identification strategy employed in their paper, however, is not capable to provide an unambiguous test of the interaction between collaboration with a superstar and common spatial location (see the discussion at pg. 577 of the article).

The remainder of this paper is organized as follows. Section 1 provides a discussion of the conceptual framework underlying the paper. Section 2 introduces the data and describes the mobility pattern of highly skilled inventors. Section 3 outlines the empirical framework for the analysis of superstar relocation events. Section 4 presents and discusses the empirical results of the analysis. Finally, Section 5 is dedicated to some conclusive remarks.

# 1   Conceptual Framework

In his classical taxonomy of knowledge spillovers, Griliches (1979) distinguishes between *horizontal*, *vertical*, and *technological* spillovers, which respectively operate between firms that are related in product markets, in supply chains or in technological domains. Scholars occasionally amend this list by adding *geographic spillovers*, on the grounds that the easier communication due to spatial proximity expedites the exchange of ideas. However, listing spatial proximity as separate from other types of spillovers can be argued as misleading. Reduced communication costs, in fact, affect all forms of interactions that lead to the emergence of spillovers, across and within industries. The role of geography is perhaps best interpreted in terms of different *agglomeration* forces which, since Marshall (1890), are typically invoked in order to explain the emergence of productivity advantages due to co-location. In what follows, I interpret the classification of agglomeration forces by Duranton and Puga (2004), which builds upon the original Marshallian analysis, in a context of knowledge or – more specifically – patent production that involves both ordinary inventors and superstars.

The first agglomeration force is *sharing*, which is generally about the advantages from the use of common inputs by multiple firms or agents. In the context of knowledge production, different inventors might benefit from easy access to the same local superstar. The second force is *matching*, which results from the increased likelihood of generating appropriate matches between inputs and their buyers if both types are more numerous in one place. For example, inventors (and superstars) with specific characteristics might more easily pair up with colleagues with the right, complementary skills if they can search among a larger local pool of scientists. The last force is *learning*, which generally gives rise to productivity and technological enhancements that are facilitated by spatial proximity. For inventors and superstars this has a straightforward interpretation in terms of exchange of ideas. In this paper, I interpret the effect of superstars on the residual patent output of their colleagues as a combination of sharing and learning. Specifically, I hypothesize that the ingenuity of ordinary inventors benefits by learning from the non-rivalrous, shared knowledge of connected superstars, resulting in more independent patents. Since both in the thought experiment that motivates the empirical framework and in the actual analysis I take collaboration networks as given, I do not focus on the role of matching.

For both interpretation and policy purposes, it is important to characterize in what cases any measured effect of superstar location on inventor productivity consists of actual *spillovers* (externalities), between firms. In this paper, I distinguish between the effect on inventors who belong to the circle of collaborators and connections of the superstars, and that on other, unconnected local inventors. The latter effect is certainly an externality; in this paper, I refer to it as *external spillovers.* The interpretation is more ambiguous in the case of the superstar's own network. Clearly, the consequence of the superstar's spatial proximity on inventors belonging to the superstar's organization can be internalized by the latter; thus, one shall not speak of spillovers. As I document in the next sections, however, superstars maintain collaborations with many inventors whose patents are prevalently assigned to organization different than the superstar's; in their case, the effect may constitute an externality.[3] Due to this ambivalence, I refer to the consequences of the superstar's closeness on his colleagues with the more neutral expression *internal effects*; nevertheless, in the empirical analysis I also measure this effect separately for inventors classified as *coworkers* and as *non-coworkers,* in order to argue to what extent the internal effect can be given an interpretation in terms of spillovers. The internal effect on coworkers, while not an externality, is nonetheless an interesting quantity to measure *per se,* as it adds to studies that document how within-firm productivity depends on the quality of workers' interaction (Mas and Moretti, 2009; Bandiera, Barankay and Rasul, 2010).

It is useful to fix ideas and relate these concepts to the variables employed in the empirical analysis by means of some simple formalism. Consider a set of inventors $\mathscr{I}$ of which superstars are a subset $\mathscr{S} \subset \mathscr{I}$.[4] In addition, consider a set of cities $\mathscr{C}$. For every superstar $i \in \mathscr{S}$, the *network of local collaborators* in city $c \in \mathscr{C}$ is expressed as $\mathscr{N}_{ic} \subset \mathscr{I} \setminus \mathscr{S}$. Now, denote any inventor of such a network as $n \in \mathscr{N}_{ic}$, and express the circumstance where $n$ is also a coworker of the superstar $s$ through the binary indicator $C_n^i = 1$, with $C_n^i = 0$ if the two inventors are not coworkers. Similarly, write $G_n^i = 1$ if superstar $i$ lives in the same city $c$ as inventor $n$, and $G_n^i = 0$ otherwise. Finally, call $P_n^T > 0$ the total number of patents produced by inventor $n$, of which $P_n^i > 0$ are realized jointly with superstar $i$, while a number $P_n > 0$ are obtained independently ($P_n^T = P_n^i + P_n$). With this notation in mind, the expected production function of inventor $n$, conditional on $C_n^i$ and $G_n^i$, can be written as:

$$
\begin{aligned}
\mathbb{E}\left[ P_n^T \,\middle|\, C_n^i, G_n^i \right] &= \mathbb{E}\left[ P_m^i \,\middle|\, C_n^i, G_n^i \right] + \mathbb{E}\left[ P_m \,\middle|\, C_n^i, G_n^i \right] \\
&= \mathbb{E}\left[ P_m^i \,\middle|\, C_n^i, G_n^i \right] + \pi_0 + \pi_1 C_n^i + \rho_0 G_n^i + \rho_1 C_n^i G_n^i
\end{aligned}
\tag{1}
$$

---

[3]The effect of superstars' proximity on connected inventors in other firms may also be internalized. In the case of R&D joint ventures that involve the superstar's firm, for example, these productivity advantages might be taken into account by the partnering firms at the time of establishing the relationship.

[4]Here I use the term "superstars" loosely. In fact, this setup can also accommodate the so-called *mavens* from the conceptualization by Oettl (2012), that is inventors with non-stellar productivity but high "helpfulness."

where $\mathbb{E}\left[ P_m^i \middle| C_n^i, G_n^i \right]$ is left unrestricted.[5] Consider instead the set $\mathcal{M}_{ic}$ of all inventors who live in city $c$, but are *not* connected with the superstar $i$ (for simplicity, I assume that they cannot be coworkers of $i$ either), with $\mathcal{N}_{ic} \cap \mathcal{M}_{ic} = \varnothing$. I write the expected patent production $P_m$ of an inventor $m \in \mathcal{M}_{ic}$ as follows.

$$\mathbb{E}\left[ P_m^T \middle| G_m^i \right] = \sigma_0 + \sigma_1 G_m^i \tag{2}$$

It is straightforward to attribute to the parameters of (1) and (2) an interpretation in terms of the concepts discussed above. Consider first equation (1), and specifically the expected production of *residual* patents $P_n$. There, parameters $\rho_0$ and $\rho_1$ represent the overall *internal effect* due to the superstar's spatial proximity, respectively for non-coworkers and for coworkers. Parameter $\pi_1$ instead measures the productivity advantages that inventor $n$ enjoys by simply being a coworker of superstar $i$; these may also be thought as the consequence of some sort of learning effect that is made easier by interacting in the same organization, but is independent of physical closeness. In equation (2) instead, $\sigma_1$ represents the *external spillovers* that a superstar exerts, on average, on all the other inventors in a city. Note that (1) and (2) are to be interpreted here as steady-state relationships deprived of a temporal dimension. By the event study methodology adopted in this paper, one can evaluate the parameters that depend on $G_n^i$ as the long-run differences between inventors that receive and those that lose proximity the superstar, while at the same time investigating the dynamics of the adjustment process.

This setup is stylized, but allows to make meaningful considerations about the effectiveness of place-based policies targeting R&D, patent production and innovation. The analysis of place-based policies is generally framed in a spatial equilibrium model *à la* Rosen (1979) and Roback (1982), see e.g. Glaeser and Gottlieb (2008, 2009); Moretti (2011); Kline and Moretti (2014) for more recent extensions. These models generally feature multiple cities populated by different firms; workers (possibly of different type) that choose their location on the basis of their place-specific wage and other factors that affect their utility; and a general equilibrium analysis of the local labor markets. The main insight delivered by these models is that place-based policies can be efficient, and enhance aggregate output, only if agglomeration externalities are sufficiently strong. The intuition is that only in this case can a place based policy re-allocate employment in such a way that the gain obtained in one place is larger than the loss experienced elsewhere. A symmetric conclusion can be extended to the setting under analysis.

---

[5]This expression for expected patent production nests an *O-Ring* (Kremer, 1993) production function. To see this, write the residual patent production as $\mathbb{E}\left[ P_n | C_n^i, G_n^i \right] = q_n$ and suppose that the patents produced jointly by $n$ and $i$ are a linear function of $n$'s independent productivity $q_n$: $\mathbb{E}\left[ P_n^i \middle| C_n^i, G_n^i \right] = (Q_i - 1) q_n$, where $Q_i > 1$ is a measure of superstar $i$'s overall "helpfulness." In this case, (1) assumes the multiplicative form $\mathbb{E}\left[ P_n^T \middle| C_n^i, G_n^i \right] = Q_i q_n$ that is typical of the O-Ring production function, and under appropriate normalizations both $q_n$ and $Q_i$ can be given the usual interpretation in terms of *quality* of the two inventors. This idea can be extended to firms' patent output realized by networks of inventors. An O-Ring production function is particularly suitable for modeling intellectual production, since it highlights the role of complementarities between individual inventors.

To see why, consider superstar $i$, two cities $c, c' \in \mathcal{C}$, and take the associated groups of connected and unconnected local inventors from both cities as given. In addition, let local collaboration networks coincide with firms (that is, $G_n^i = C_n^i$ for all $n \in \mathcal{N}_{ic}$, and similarly for city $c'$). Each network-firm attempts to attract the superstar to their locality by offering a wage; assuming for simplicity that patents translate to currency units one-to-one, the maximum wage $\overline{W}_{\mathcal{N}_{ic}}$ that a network-firm is willing to offer is the total productivity gain that the superstar would yield to the network-firm.

$$\overline{W}_{\mathcal{N}_{ic}} = \sum_{n \in \mathcal{N}_{ic}} \left\{ \mathbb{E}\left[ P_n^i \middle| G_n^i \right] + (\rho_0 + \rho_1) G_n^i \right\} \tag{3}$$

In equilibrium, the superstar chooses the network-firm that offers the highest wage, thus its choice internalizes the internal effects $\rho_0$ and $\rho_1$. Social welfare, on the other hand, equals the total patent production in both cities, and depends on the spatial allocation of the superstar.

$$SW = \sum_{n \in \mathcal{N}_{ic} \cup \mathcal{N}_{ic'}} \mathbb{E}\left[ P_n^T \middle| G_n^i \right] + \sum_{m \in \mathcal{M}_{ic} \cup \mathcal{M}_{ic'}} \mathbb{E}\left[ P_m^T \middle| G_m^i \right] \tag{4}$$

Clearly, the social optimum is guaranteed to coincide with the equilibrium outcome if external spillovers are zero, that is if $\sigma_1 = 0$. Intuitively, a place-based policy that attempts to re-allocate innovative activity can be motivated in aggregate terms only so long as the benefits extend to unrelated local innovators.[6] This result is easy to generalize to multiple locations, as well as multiple superstars and their associated networks. If firms and local networks do not coincide, superstars may not internalize the internal effect $\rho_0$ on connected non-coworkers, but the main qualitative conclusions would remain unchanged.

# 2 Data and Relocation Events

This section is divided in two parts. In the first part I introduce the data, commenting on the steps taken in order to reconstruct paths of inventors' mobility. In the second part I describe in more detail the relocation events that involve superstars.

## 2.1 Inventor Data

In this study I utilize the patent dataset elaborated by Morrison, Pammolli and Riccaboni (2017) – henceforth MPR. This is a comprehensive elaboration of about 8.5 millions patents granted by the US Patent and Trademark Office (USPTO), by the European Patent Office (EPO) or under the

---

[6]Even then, short of introducing centrifugal forces into the model, the main policy recommendation would be that of spatially segregating R&D and innovative activity, which may not be necessary since these activities are already considerably spatially concentrated in advanced economies.

framework of the Patent Cooperation Treaty (PCT) since 1974, and whose application dates precede 2015. MPR apply a geolocalization algorithm in order to assign specific geo-coordinates to all addresses associated with any inventor-patent combination (this is a novel feature for such a comprehensive patent database). The accuracy of the geolocation varies, but it is satisfactorily good in the case of US, Canadian and European inventor-patents. In addition, MPR apply a disambiguation algorithm similar to the one by Li et al. (2014) in order to distinguish individual inventor identities in the data. They further refine the disambiguation algorithm by incorporating their geolocation output as one of its inputs,[7] and they also extend it to patent assignees, to which MPR assign unique identifiers.

As a further step, I select 6,762,683 patents with at least one inventor out of 3,139,253 who are geolocated in either North American or Europe, and I employ the geo-coordinates provided by MPR in order to attach each inventor-patent to a city. In particular, I utilize definitions of statistical "metropolitan area" which are as homogeneous as possible across the US, Canada and Europe. They are: *i.* the 2013 definition of Core Based Statistical Area (CBSA) in the US; *ii.* the 2011 definition of Census Metropolitan Area (CMA) of the Canadian Census; *iii.* the 2011-2014 definition of "European City" (EC) provided by Eurostat. These definitions, while differing in their specific details, are based on a common principle: a larger metropolitan area is defined in terms of the possibility to commute within its borders; hence, the urban area in question constitutes a single local labor market. This choice implies that the external spillovers that I attempt to estimate in the empirical analysis are confined within commuting zones.[8]

**Table 1:** Relocation Example 1

| Inventor No. | Patent No. | Name on Patent | App. Year | City ID | Sequence |
|---|---|---|---|---|---|
| 17380242 | US1369841 | JOHN SMITH | 1987 | 114 | 1 |
| 17380242 | US6432207 | J. SMITH | 1989 | 114 | 1 |
| 17380242 | US5578017 | JOHN SMITH | 1990 | 114 | 1 |
| 17380242 | US6047949 | JOHN SMITH | 1990 | 114 | 1 |
| 17380242 | US6841760 | JOHN SMITH | 1991 | 114 | 1 |
| 17380242 | US1086679 | SMITH, J. F. | 1991 | 876 | 2 |
| 17380242 | US6184582 | JOHN SMITH | 1992 | 876 | 2 |
| 17380242 | US5219248 | JOHN SMITH | 1994 | 876 | 2 |

By attaching inventor-patents to cities I can track inventors' mobility. Unsurprisingly, the vast majority of inventors in the sample (91.35%) are associated to patents filed from a single urban area, and less than 2% realize patents from more than two different cities over the course

---

[7]In fact, the dataset is constructed by letting the disambiguation and the geolocation algorithm "speak" and feed back on each other following an iterative procedure. Consult the paper by MPR for the details.

[8]Singh and Marx (2013) show that even controlling for distance, patent citations are more likely within US state or CBSA borders. Thus, metropolitan areas are arguably a relevant dimension of knowledge diffusion.

of their life. The typical pattern of patent-level information for individuals who have been relocating across cities is sharply discontinuous, denoting well-defined moving episodes. An example of this, fabricated for illustratory purposes, is represented in Table 1. The residence address of this hypothetical individual (John Smith[9]), as indicated on patent applications, is associated to two different urban areas corresponding to two consecutive "sequences": one for patent applications received until 1991, and the other one for later patent applications. In such a case, the data would unambiguously inform that inventor John Smith has moved to a different city in 1991, which defines the *event year* of Mr. Smith's *relocation event*.

In a minority of cases, however, the data are less unequivocal, because multiple urban areas overlap within time periods of varying length. This might happen for a variety of reasons: for example, an inventor collaborates across multiple laboratories at the same time, he or she visits some institutions for a short period of time, or perhaps really has multiple residences (but can only report one at a time). To deal with such instances, I assign inventors to cities according to the following procedure: *i.* I first split the patent history of each inventor into temporal intervals, which are delimited by the years when a different city occurs for the first time; *ii.* to each of these intervals I assign a unique city by choosing the one that is observed most frequently; ties are resolved conservatively in favor of the city that occurred earlier in the inventor's patent history;[10] *iii.* final "sequences" and event years are defined.

**Table 2:** Relocation Example 2

| Inventor No. | Patent No. | Name on Patent | App. Year | City ID | Sequence |
|---|---|---|---|---|---|
| 04857953 | EP8414094 | PAUL VOGEL | 1989 | 417 | 1 |
| 04857953 | EP2110077 | PAUL VOGEL | 1992 | 417 | 1 |
| 04857953 | EP5644092 | PAUL VOGEL | 1992 | 417 | 1 |
| 04857953 | EP2648021 | PAUL VOGEL | 1993 | 559 | 2 |
| 04857953 | EP9477426 | PAUL VOGEL | 1994 | 559 | 2 |
| 04857953 | EP2769154 | P. VOGEL | 1994 | 417 | 2 |
| 04857953 | EP1180159 | P. VOGEL | 1995 | 559 | 2 |
| 04857953 | EP4079703 | P. W. VOGEL | 1996 | 271 | 2 |
| 04857953 | EP7219492 | P. W. VOGEL | 1996 | 559 | 2 |
| 04857953 | EP8716910 | P. W. VOGEL | 1997 | 559 | 2 |
| 04857953 | EP5948404 | P. W. VOGEL | 1998 | 271 | 2 |

To clarify the procedure, I provide another constructed example in Table 2. In the case of Paul Vogel, another imaginary inventor, three cities appear on record: one corresponding to his first registered patent in 1989, a second one appearing in 1993, and a third city which occurs

---

[9]Notice how the name reported on patents differs in this constructed example. This illustrates the importance of disambiguating individual identities in patent data by exploiting available information (such as uniformity in the residence address, patent class, names of collaborators and so on).

[10]This also helps dealing with potential issues of incorrect address geolocation; see MPR paper for more on these.

twice, in 1995 and 1998. I would easily assign the first city (ID 417) to the time interval 1989-1992. Between 1993 and 1995 the second city (ID 559) is the one that is observed most often. Finally, between 1996 and 1998 both the second and the third city (ID 271) occur with equal frequency; thus I assign the earlier city to the last time interval. As a result, if this piece of data were real, I would define two inventor-city sequences delimited by the *event year* 1993, when Mr. Vogel would have moved from city 417 to city 559.

## 2.2 Moving Superstars

I restrict the analysis of inventor relocation to "superstars," who are defined as the individuals in the top 5% of the distribution of patents per inventor. Individuals with more than 18 patents fall in this category. Superstar inventors are considerably more mobile than other individuals. A staggering percentage of 21.5% is characterized by one relocation event as defined in the previous part of this section, and about 5% by more than one event. More prolific superstars seem to be, as expected, slightly more mobile. Throughout the rest of this study I subdivide superstar inventors in three subgroups by their position in the patents distribution. I adopt an astronomical terminology for this classification, and I call "Dwarves" the inventors between the 5% and the 1% percentiles; "Giants" those between the 1% and the 0.5% quantiles; and "Bright" Giants the individuals in the top 0.5% of the distribution (the 1% and 0.5% cutoffs are respectively 51 and 72 patents). The number of individuals in each of these three groups, broken down by number of observed moves, is reported in Table 3. I find it worthwhile to distinguish between ordinary and bright giants, so to ascertain if the size of any estimated effect varies monotonically as a function of a star's position in the patent distribution.[11]

**Table 3:** Number of Moves by Superstar Group (1975-2012)

| Group | Dwarves | (%) | Giants | (%) | Bright | (%) |
|---|---|---|---|---|---|---|
| 0 Moves | 90,625 | (74.9%) | 10,182 | (68.5%) | 10,118 | (66.5%) |
| 1 Move | 25,021 | (20.7%) | 3,644 | (24.5%) | 3,866 | (25.4%) |
| 2 Moves | 4,560 | (3.7%) | 832 | (5.6%) | 943 | (6.2%) |
| 3+ Moves | 826 | (0.7%) | 213 | (1.4%) | 297 | (1.9%) |
| Total | 121,032 | (100%) | 14,871 | (100%) | 15,224 | (100%) |

The event study analysis reported in this paper examines the patent consequences of superstars' relocation up to six years before and after the event, for a total of thirteen years. As it is typical in event studies, this demands a careful selection of the events sample. First, I exclude

---

[11]Choosing a cutoff at 0.5% results in a symmetrical split of the top 1% of the distribution, which contrasts with the fact that the dwarves' group is one order of magnitude larger than the other two. While the 1% cutoff is clearly symbolic, opting for a more restricted definition of "bright giants" results in too small a subsample.

multiple moves of the same superstar inventor that are not at least 13 years apart (as Table 3 suggests, this implies a negligible loss of data). In addition, I restrict the analysis to relocation episodes spanning from 1980 to 2000: by selecting this temporal window, I am able to utilize observations as early as 1974 for 1980 events and as late as 2006 for 2000 events. I remove later events on purpose, since the patent grant lag disproportionately affects patent outcomes for later years, with consequent issues of interpretation that are particularly severe for citation-based measures. While this choice implies a loss of about 8,000 moves, the results are not substantially affected by it.[12] The final sample is based on 12,687 events; it is a panel in event year that is balanced for the first six leads and lags around each event. The number of selected events in each year is displayed for each group of superstars in Figure 1: notably, it increases over time for dwarves and giants alike.
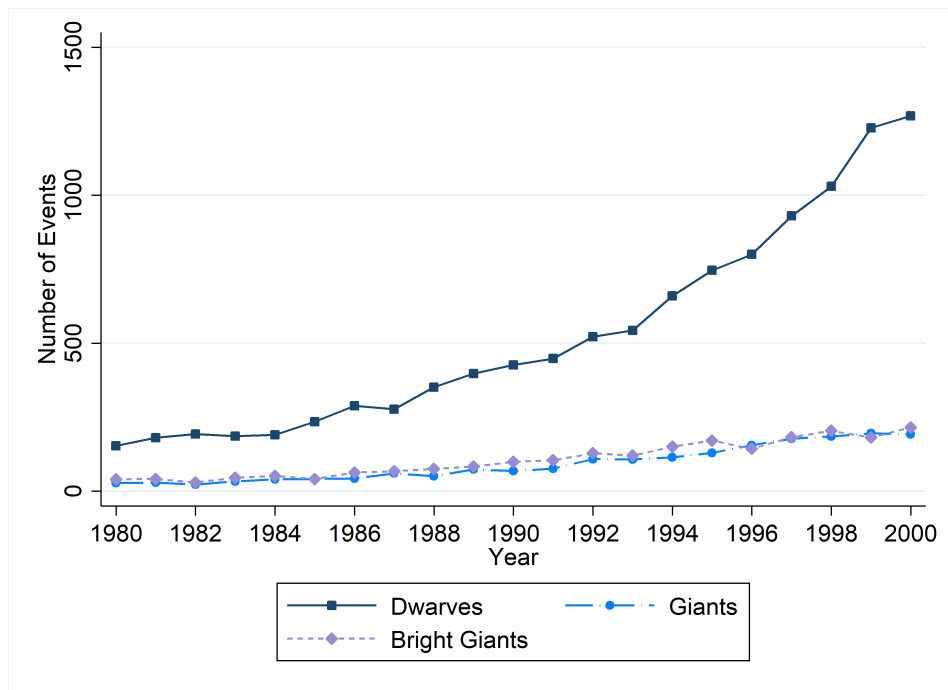


**Figure 1:** Number of Relocation Events in Each Year

In addition, I classify relocation events by geographical scope of the move and by technological field. In terms of geography, I distinguish whether the move happened between two urban areas of North America (US and Canada), two European Cities, or rather it implicated crossing the Atlantic (in either direction). I also assign macro-technological areas to the events, on the basis of the prevalent field of the moving star's patents. In particular: *i.* I select the first World

---

[12]In a previous versions of this work the final sample included 2001-2006 events. The associated results, available upon request, were very similar to those in the present article. Arguably, the use of a balanced panel in event date combined with the inclusion of year-specific effects allowed to eliminate most if not all the bias due to the patent lag truncation problem.

Intellectual Property Organization (WIPO) technological class which is reported on each patent, *ii.* I match it to one of six WIPO macro-classes (Chemistry, Biotech-Pharmaceutical, Electrical Engineering and ICT, Instruments, Mechanical Engineering, Others); and *iii.* finally, I associate one of these macro-classes to a superstar-event, so long as the former appears in at least 75% of all the patents of the superstar in question. This stringent criterion implies that about 40% of all events, arguably those related to the most eclectic superstars, are not classified within any of the six WIPO macro-areas (which also include the residual category "Others").

**Table 4:** Breakdown of Moves by Geographical Scope and Technological Field (1980-2000)

| Technological Field | Within North America | Europe to North America | North America to Europe | Within Europe | Total |
|---|---|---|---|---|---|
| Chemical | 1,618 | 108 | 87 | 1,336 | 3,149 |
| | (12.75%) | (0.85%) | (0.69%) | (10.53%) | (24.82%) |
| Biotech-Pharma | 397 | 13 | 22 | 124 | 556 |
| | (3.13%) | (0.10%) | (0.17%) | (0.98%) | (4.38%) |
| Electrical eng. & ICT | 1,589 | 102 | 102 | 430 | 2,223 |
| | (12.52%) | (0.80%) | (0.80%) | (3.39%) | (17.52%) |
| Instruments | 539 | 22 | 18 | 108 | 687 |
| | (4.25%) | (0.17%) | (0.14%) | (0.85%) | (5.41%) |
| Mechanical engin. | 414 | 25 | 21 | 475 | 935 |
| | (3.26%) | (0.20%) | (0.17%) | (3.74%) | (7.37%) |
| Others | 48 | 5 | 1 | 16 | 70 |
| | (0.38%) | (0.04%) | (0.01%) | (0.13%) | (0.55%) |
| Unclassified | 3,546 | 171 | 103 | 1,247 | 5,067 |
| | (27.95%) | (1.35%) | (0.81%) | (9.83%) | (39.94%) |
| Total | 8,151 | 446 | 354 | 3,736 | 12,687 |
| | (64.25%) | (3.52%) | (2.79%) | (29.45%) | (100.00%) |

The relative frequency of all events characterized by the same combination of geographical scope and technological field, along with marginal frequencies, is reported in Table 4. The vast majority of all the events (about 64%) involves a move within North America, while about 30% is internal to Europe. Less than 7% of all the moves is trans-Atlantic, with a prevalence of those directed from east to west. The geographical scope of the events at the level of metropolitan areas is graphically displayed by the two maps A.1 (North America) and A.2 (Europe) in the Appendix. Unsurprisingly, most of the action revolves around the major US urban centers, the Rhine basin, London and Paris. In terms of WIPO technological macro-classes, the largest share of classified superstar-events are in the chemical field (25% of all the events) and in electrical engineering and ICT (17.5%). The others are split between mechanical engineering (7.4% of the grand total), instruments (5.4%) and the biotech-pharmaceutical sector (4.4%).

# 3 Empirical Methodology

In the first part of this section I discuss the thought experiment that informs the empirical analysis and the associated regression framework. In the second part I introduce the different outcome measures employed in the estimates, along with some descriptive statistics.

## 3.1 Studying Relocation Events

It is useful illustrate the thought experiment on which the empirical analysis is based through a stylized example. Figure 2 represents the relocation event of a superstar inventor, along with his or her collaboration network. The superstar in question (the larger node) moves from City $O$ ('Out') to City $I$ ('In'), and has collaborators in both places – as well as some that do not reside in either city. The event study featured in this paper evaluates the impact of the relocation event on the residual patent outcomes of two groups: the "blue" nodes in the graph – connected inventors in the locality of destination – and the "red" nodes – those in the city left behind. If the superstar's spatial proximity generates local productivity effects, one should expect a positive effect on the former and a negative effect on the latter. While the graph does not distinguish between collaborators who are also coworkers against non-coworkers, in the analysis I attempt to measure the effect separately for the two groups in question by exploiting information about patent assignees. In addition, the impact of the event is evaluated against the patent output of all other non-connected inventors in either City $O$ or City $I$, who are identified residually.
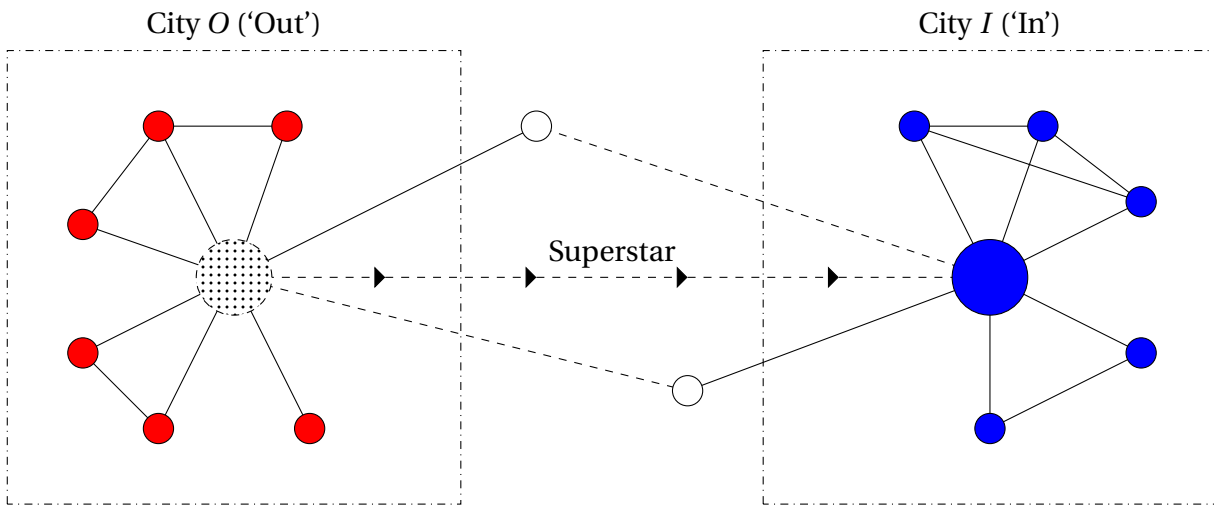


**Figure 2:** A Moving Superstar and Local Coauthorship Networks

In the actual analysis, a superstar's "connection" is defined as one inventor who has signed a patent application jointly with the superstar within an interval of six years before or after the relocation event. Therefore, networks are static, implying that the results are not driven by the

varying composition of the network, but by changes in the propensity to independently produce new knowledge. To avoid further issues of confoundedness, other *mobile* superstar inventors are excluded from the construction of local networks.[13] Finally, I classify individual connections as "coworkers" or "non-coworkers" whether the prevalent assignee associated with the patents of both the superstar and his or her linked collaborator coincides; this assignment is performed twice for each connected inventor, both before and after the event.[14] Summary statistics about the average size of superstars' local networks in both types of cities over the years are reported in Figure 3; team size appear to increase twofold over time, from an average of about 4 connections in 1980 to an average network size around 8, for both types of local network.[15] In addition, Figure 3 reports the analogous statistics restricted to connected inventors who are also coworkers of the superstar. Since coworkers appear to be a robustly constant fraction of total connections over time – about two thirds – the corresponding trends parallel closely the overall averages.
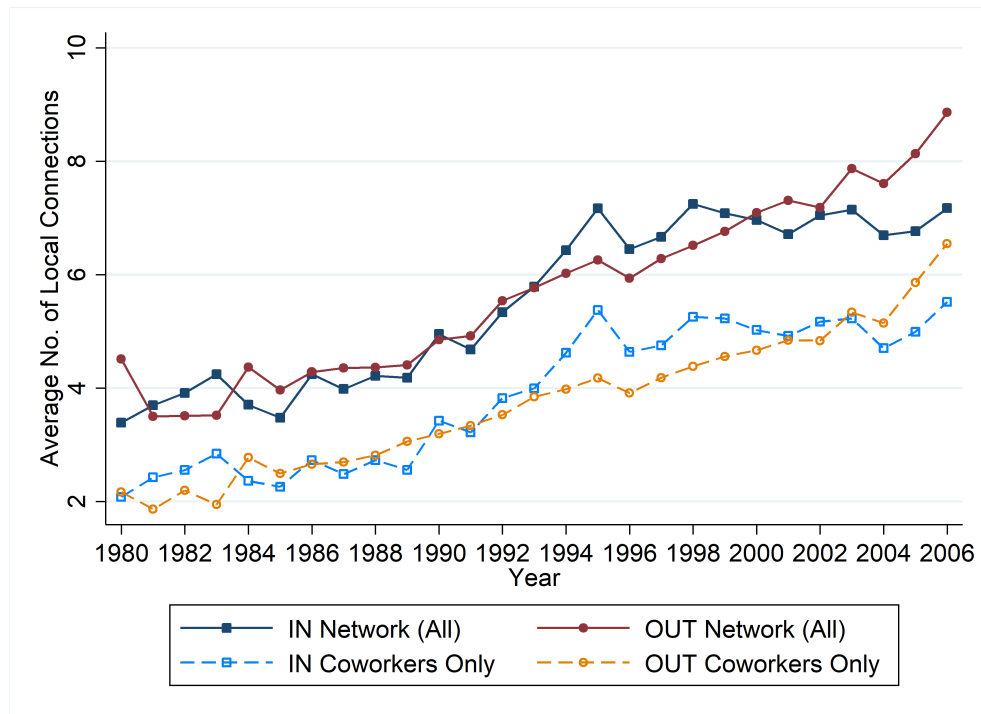


**Figure 3:** Average Network Size by group

---

[13]However, the few *mobile non-*superstar inventors are included. They enter a superstar's local network for a given statistical metropolitan area if they are assigned to that particular city on the superstar's event year.

[14]That is, an inventor connected with a superstar can be classified as a coworker before the event and as a non-coworker after the event, and vice versa. This should account for cases of joint mobility within the network.

[15]Interestingly, breaking down this analysis by superstar group does not reveal that to a more prolific superstar usually corresponds a higher number of local connections. In fact, typically the collaboration networks of non-bright "Giants" are, on average, larger than those of "Dwarves," but smaller than those of "Bright" Giants. This can be explained by the fact that superstars in higher quantiles have a disproportionate number of connections dispersed across places where they have never been residing.

In order to estimate the internal effects and external spillovers associated with this thought experiment, I employ a standard event study framework, which is common to all the estimates featured in this paper. To characterize the econometric model, let again $i \in \mathscr{S}$ be a superstar, and denote: *i.* the event date as $s(i) \in \{1980, 1981, \ldots, 2000\}$; *ii.* a set of $2\bar{K}+1$ dummy variables as $\mathscr{D} = \left\{ D^{-\bar{K}}, \cdot, D^k, \cdot, D^{\bar{K}} \right\}$, where $D_{ict}^k = \mathbb{1}\left[t = s(i) + k\right]$ for every integer $k$ with $-\bar{K} \le k \le \bar{K}$; *iii.* two additional "lower" and "upper" bound dummies as $D_{ict}^{\ell}$ and $D_{ict}^u$, which are respectively defined as $D_{ict}^{\ell} = \mathbb{1}\left[t < s(i) - \bar{K}\right]$ and $D_{ict}^u = \mathbb{1}\left[t > s(i) + \bar{K}\right]$. The model reads as:

$$Y_{ict} = \alpha_{ic} + \sum_{k=-\bar{K}}^{\bar{K}} \beta_k D_{ict}^k + \gamma_\ell D_{ict}^\ell + \gamma_u D_{ict}^u + \delta_{ct} + \varepsilon_{ict} \tag{5}$$

$$\beta_{-1} = 0$$

where $c \in \{I, O\}$ denotes either the city of destination or that of departure, while $t = 1975, \ldots, 2006$ indexes the year when the outcome variable $Y_{ict}$ – some measure of patent counts – is observed.

In this model, the estimates of parameters $\left\{\beta_{-\bar{K}}, \ldots, \beta_0, \ldots, \beta_{\bar{K}}\right\}$ track the effect of the event on the outcome of interest $Y_{ict}$ over time. The advantage of the event study framework relative to a simple regression model with a binary treatment – such as those expressed in equations (1) and (2) – is that it allows to visually examine the extent of potential endogeneity through the estimation of the "pre-trend" parameters $\beta_{k'}$ with $k' < -1$. In the case in question, if the *timing* of the event correlates with changes in the patterns of productivity of different groups, this should be visualized in terms of non-parallel pre-trends for the two groups. If, for example, the location choice of superstars is often at least in part driven by the observation of increasing productivity of their actual or potential local collaborators in the 'In' city, one should expect a steeper pre-trend for the 'In' group relative to the 'Out' group. In order to facilitate this assessment, it is useful to re-estimate the model – for each outcome $Y_{ict}$ subsequently detailed – by employing the following dependent variable.

$$D_{it} = Y_{iIt} - Y_{iOt}$$

I call $D_{it}$ the *In-Out difference* measure. Note that all the coefficients $\left\{\beta_{-\bar{K}}, \ldots, \beta_0, \ldots, \beta_{\bar{K}}\right\}$ of the In-Out difference model are, by construction, the difference between the corresponding coefficients for the outcomes $Y_{iIt}$ and $Y_{iOt}$; therefore, any non-flat estimated pre-trend for the In-Out differenced model should be interpreted as a warning sign of endogeneity.

In this study I set $\bar{K} = 6$, so that the model can be estimated on the 1975-2006 sample for all events occurred between 1980 and 2000. All regressions include event-city fixed effects $\alpha_{ic}$, which control for local time-invariant characteristics (like network size), and time effects $\delta_{ct}$. The model is estimated via OLS and the standard errors used to construct confidence intervals are all clustered at the superstar-event level.

## 3.2 Patent Outcomes

In order to evaluate internal productivity effects at the local network level, I estimate the effect of superstar relocation on eight alternative patent outcome variables $Y_{ict}$. I start with defining the simplest measure, the pure patent count. For the superstar-event $i$, city $c \in \{I, O\}$, network $\mathcal{N}_{ic}$ and year $t$ the "pure patent count" $Y_{ict} = P_{ict}$ of the corresponding network is given by

$$P_{ict} = \sum_{n \in \mathcal{N}_{ic}} \hat{p}_{nit}$$

where $\hat{p}_{nit}$ is the *residual* number of patent *applications* of inventor $n$ in year $t$ (that is, patents not realized jointly with superstar $i$, a property denoted by the hat in $\hat{p}$). While all empirical estimates control for network size, the pure patent count might not be a meaningful measure given the large variation in the dimension of the local networks of superstars. To account for this one may alternatively look at *average* patent count, dividing $P_{ict}$ by network size $|\mathcal{N}_{ic}|$.

$$\bar{P}_{ict} = \frac{1}{|\mathcal{N}_{ic}|} \sum_{n \in \mathcal{N}_{ic}} \hat{p}_{nit}$$

Another issue is that some technological fields are characterized by larger team sizes, which often result in patents under the name of many inventors. This fact may bias the analysis in either way, should the superstar relocate to an area closer to a group of collaborators with relatively more researchers per project, or vice versa. In order to address this, I also employ residual patent *shares* $\hat{s}_{nit}$. A patent share is the inverse of the number of authors of a patent; in the case of a patent with three authors, for example, the share of each author is 1/3. Shares $s_{nit}$ of patents realized jointly with the superstar are once again discarded. The share count $\tilde{P}_{ict}$ is the sum of all patent shares of inventor $i$. In analogy with average patent count, I also consider the *average share* measure $\tilde{\bar{P}}_{ict}$ which just like $\bar{P}_{ict}$ accounts for heterogeneous network size.

$$\tilde{P}_{ict} = \sum_{n \in \mathcal{N}_{ic}} \hat{s}_{nit}$$
$$\tilde{\bar{P}}_{ict} = \frac{1}{|\mathcal{N}_{ic}|} \sum_{n \in \mathcal{N}_{ic}} \hat{s}_{nit}$$

Figure 4 displays the averages of the patent count $P_{ict}$ and the patent shares $\tilde{P}_{ict}$ measures at the event-year level from 1976 to 2006, for networks networks located in both "In" and "Out" cities. These descriptive statistics show that until 2000, the networks left behind by the moving superstar are typically more innovative than those that are joined, although the early 2010's suggest a reversal of this trend. Identical patterns are observed by looking at the average measures $\bar{P}_{ict}$ and $\tilde{\bar{P}}_{ict}$ (not reported in Figure 4 for clarity).
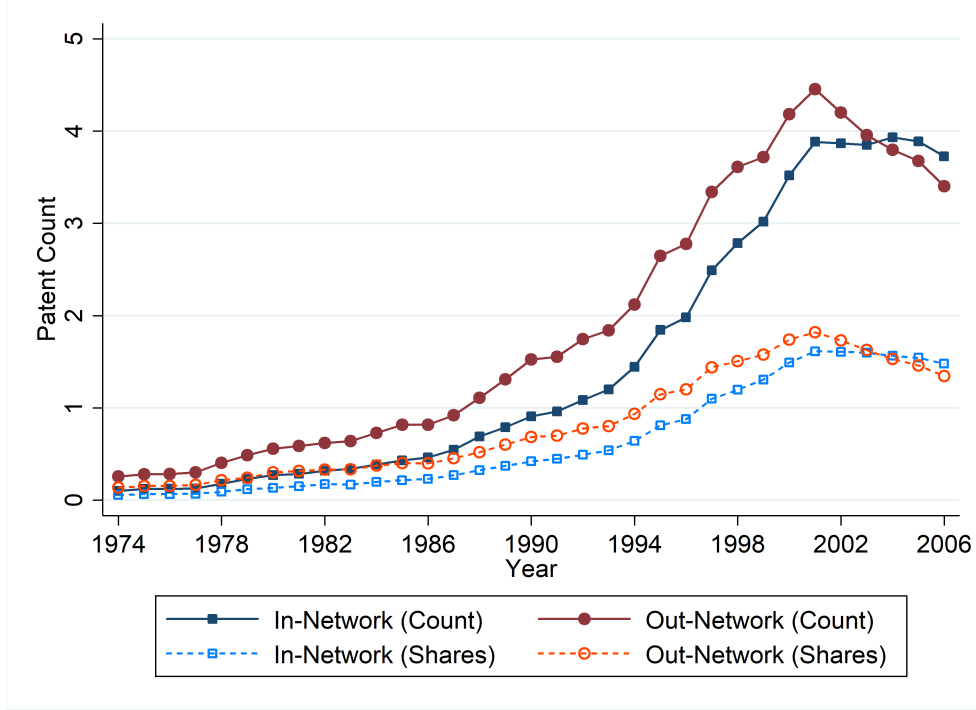
16

**Figure 4:** Network-level Patent Count averages, by year

None of these alternative measures account for patent heterogeneity – while certainly some patents are more valuable than others. A frequently used proxy of patent quality is the number of citations; thus I also estimate the effect of the events on appropriate citations-weighted patent outcomes. The citation-weighted analogues of $P_{ict}$, $\bar{P}_{ict}$, $\tilde{P}_{ict}$ and $\tilde{\bar{P}}_{ict}$ are, respectively:

$$B_{ict} = \sum_{n \in \mathcal{N}_{ic}} \hat{b}_{nit}$$

$$\bar{B}_{ict} = \frac{1}{|\mathcal{N}_{ic}|} \sum_{n \in \mathcal{N}_{ic}} \hat{b}_{nit}$$

$$\tilde{B}_{ict} = \sum_{n \in \mathcal{N}_{ic}} \hat{z}_{nit}$$

$$\tilde{\bar{B}}_{ict} = \frac{1}{|\mathcal{N}_{ic}|} \sum_{n \in \mathcal{N}_{ic}} \hat{z}_{nit}$$

where $\hat{b}_{nit}$ is the *citation-weighted, residual* patent count of inventor $n$ and $\hat{z}_{nit}$ is the citation-weighted analog of $\hat{s}_{nit}$.[16] In order to account for the so-called "truncation problem" of patent citations – the fact that more recent patents have yet to receive citations at the time they are

---

[16]I provide here more formal definitions of $\hat{p}_{nit}$, $\hat{b}_{nit}$, $\hat{s}_{nit}$ and $\hat{z}_{nit}$, not given in the text for the sake of easier legibility. Define $\mathbb{K}$ as the set of all patents, $|\mathbb{K}| = K$, and index patents by $k = 1, \ldots, K$. Notation $\mathscr{P}_{nt}$ denotes the set of patents for which inventor $n$ filed an application in year $t$. Define $N_k$ as the number of authors of patent $k$, while

observed – I apply the correction method devised by Jaffe and Trajtenberg (1996) and employed, among the others, by Hall et al. (2001, 2005).[17] Note that all network-based measures described so far are also separately calculated for the coworkers and non-coworkers subgroups that are associated with each superstar network.

For the estimation of external spillovers I use an analogous set of measures, one which is based on the residual set of non-connected local inventors $\mathcal{M}_{ic}$. The city-wide patent count measure $C_{ict}$ – the analogue of $P_{ict}$ – and the corresponding average count $\bar{C}_{ict}$ (which is normalized by the number $|\mathcal{M}_{ic}|$ of non-connected inventors in a city) are defined respectively as:

$$C_{ict} = \sum_{m \in \mathcal{M}_{ic}} \hat{p}_{mit}$$

$$\bar{C}_{ict} = \frac{1}{|\mathcal{M}_{ic}|} \sum_{m \in \mathcal{M}_{ic}} \hat{p}_{mit}$$

and $\tilde{C}_{ict}$, $\bar{\tilde{C}}_{ict}$, $G_{ict}$, $\bar{G}_{ict}$, $\tilde{G}_{ict}$ and $\bar{\tilde{G}}_{ict}$ are defined similarly, where here $G$ denotes patent or shares measures that are weighted by citations (like $B$ in the case of network-level measures). Figure 5 reports summary statistics about average city patent counts $\bar{C}_{int}$ and average patent shares $\bar{\tilde{C}}_{int}$ that are analogous to those from Figure 4 about local networks.[18] It appears that "Out" cities are slightly more innovative than "In" cities throughout the time interval under consideration. Remarkably, the time trends of all these averages looks noticeably similar for both "In" and "Out" localities, showing little systematic association of superstars' moving choices with changes in local innovativeness.

$C_k$ is the associated truncation-adjusted citation weight. The definitions are as follows.

$$\hat{p}_{nit} = \sum_{k=1}^{K} \mathbb{1}\,[k \in \mathcal{P}_{nt} \wedge k \notin \mathcal{P}_{it}]$$

$$\hat{s}_{nit} = \sum_{k=1}^{K} \mathbb{1}\,[k \in \mathcal{P}_{nt} \wedge k \notin \mathcal{P}_{it}] \cdot \frac{1}{N_k}$$

$$\hat{b}_{nit} = \sum_{k=1}^{K} \mathbb{1}\,[k \in \mathcal{P}_{nt} \wedge k \notin \mathcal{P}_{it}] \cdot C_k$$

$$\hat{z}_{nit} = \sum_{k=1}^{K} \mathbb{1}\,[k \in \mathcal{P}_{nt} \wedge k \notin \mathcal{P}_{it}] \cdot \frac{C_k}{N_k}$$

[17]This method is based on two stages. The first stage is about estimating the empirical distribution of the forward citation lags, usually through a flexible specification of a double exponential distribution with parameters differing by technological field. In the second stage, one divides the number of actually observed citations for a patent by the estimated fraction of patents in the corresponding lag year for similar patents (this correction can be very imprecise for the more recently observed patents with few observed lag years; note that while 2007-2014 patents are used for estimating the citation model, they are excluded from the actual analysis). I apply an extension of this method to the MPR data; the estimation details, its results and the calculated weights are available upon request.

[18]The statistics from Figure 4 can be thought as yearly weighted averages patent counts across metropolitan areas, with weights corresponding to the share of "In"-"Out" events for each city in every year.
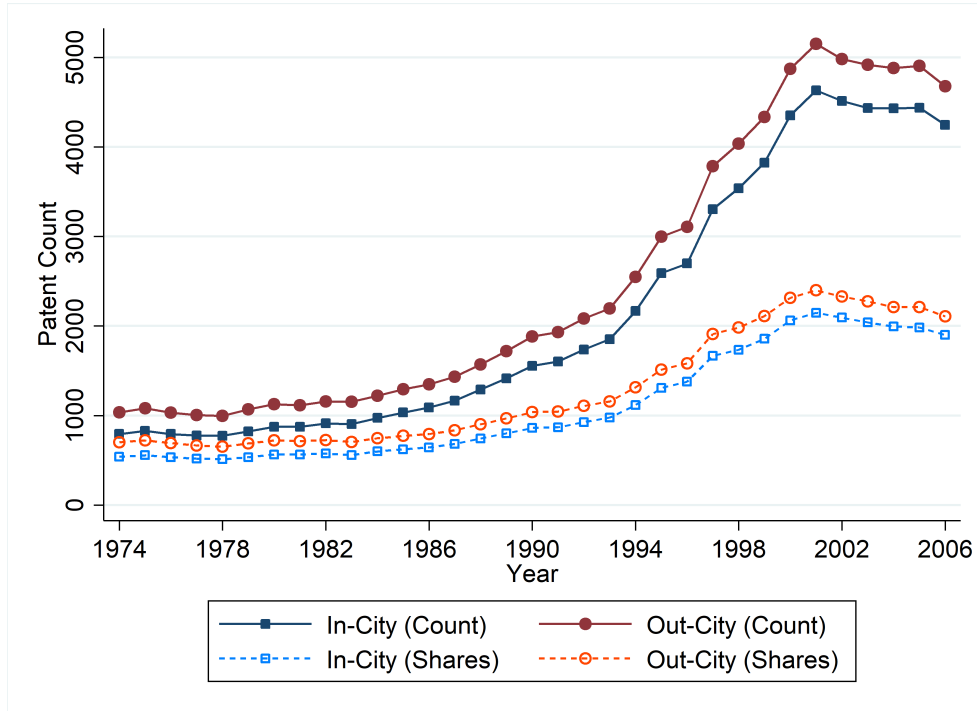
**Figure 5:** City-level Patent Count averages, by year

# 4 Event Study: Empirical Results

In this section I discuss the empirical results from the event study analysis. Along with presenting the baseline results at the network and city level, I break down the estimates by coworker status, superstar group, geographical scope of the event, and technological macro-field. All estimates of the relevant parameters $\{\beta_{-6},\dots,\beta_0,\dots,\beta_6\}$ from model (5) are reported in graphical form, as it is customary in event studies.

## 4.1 Baseline Results: Network-level Measures

Figure 6 reports the estimates of coefficients $\{\beta_{-6},\dots,\beta_0,\dots,\beta_6\}$ from three regressions based on model (5): one where the dependent variable is pure patent count measures $P_{ict}$ for the "In" network, one for the "Out" network, and a third one for the In-Out difference measure $D_{it}$. The estimates show that before the event both networks enjoyed rising, almost identical and statistically indistinguishable trends in the number of yearly patent applications (increasing by about 1 more patent per year over six years). After the event, the patent output of the "In" network rises more markedly to eventually stabilize at around 0.8-1 extra patents relative to the baseline year. Conversely, the productivity of the "Out" network undergoes a constant decline since the second year following the event, resulting in about 0.5 *less* patents per year relative to the base-

19

line after six years. The observed lag in the trend reversal experienced by these inventors might be tentatively attributed to the nature of R&D activity: following the event, the collaborators left behind by the superstar might keep elaborating for a while on research projects and ideas whose seeds were sown at the time when the superstar used to be physically close.
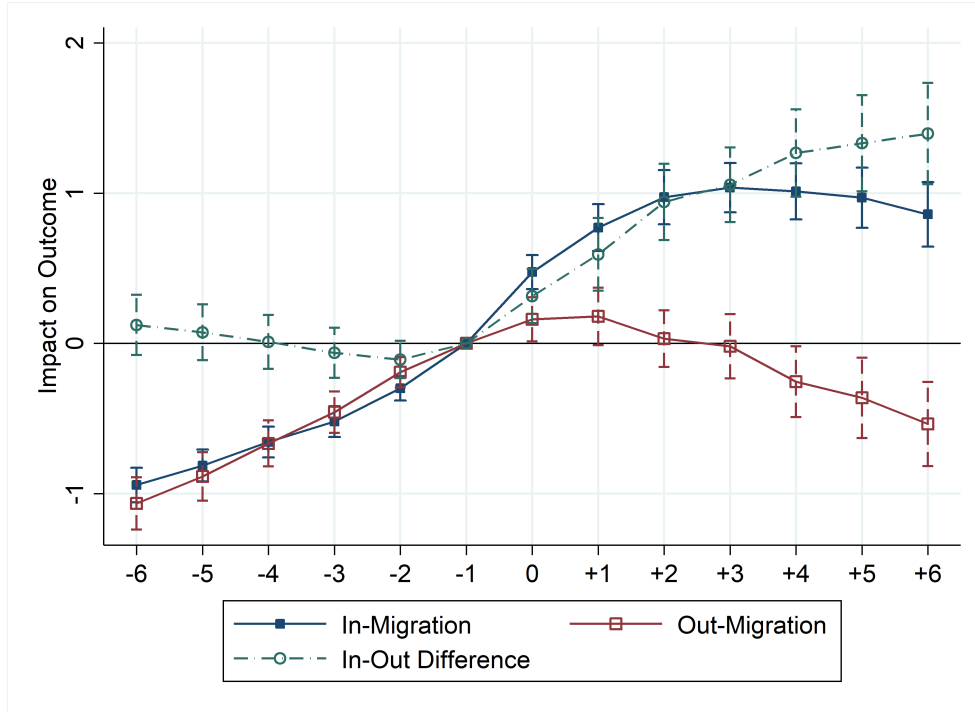


**Figure 6:** Network-level Estimates, Patent Counts $P_{ict}$

The same pattern is registered for all other network-level measures that are not weighted by citations. Figure 7 displays the results corresponding to the average patent counts measure $\bar{P}_{ict}$. The temporal trends about the estimated coefficients are identical to those from Figure 6. This indicates that the results are not driven by differences in *network* size (the superstars might, say, move from "small" to "large" networks) that are possibly not adequately captured by the fixed effect $\alpha_{ic}$. Similarly, the results do not seem to be caused by differences in the size of research *teams* responsible for the "In" and "Out" patents: in fact, the estimates relative to the share count measure $\tilde{P}_{ict}$ (not shown for brevity) and those for the average share count measure $\bar{\tilde{P}}_{ict}$ (shown in Figure 8) display identical patterns. Estimates about average measures are more easily related to individual productivity. In particular, they indicate how relative to the baseline year, in the long run inventors from the "In" network improve their performance versus those from the "Out" network by about 0.16 extra residual patents $\hat{p}$, or almost 0.08 extra residual fractional patents $\hat{s}$. Moreover, the losses experienced by the losing group erase on average about one half of the productivity gains that are experienced in the years leading to the event.
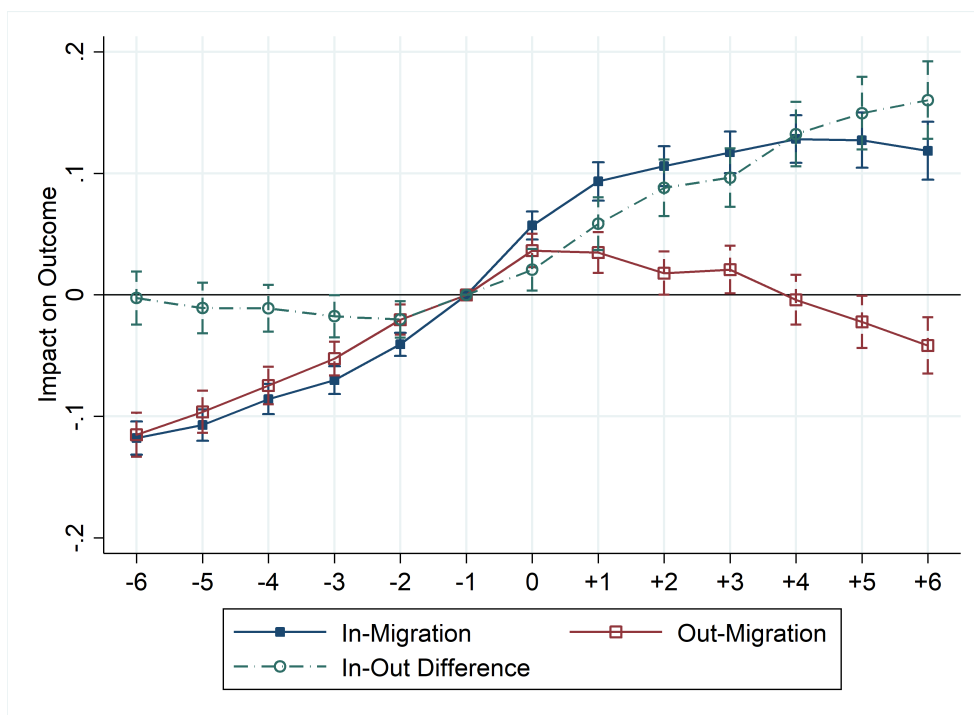
20

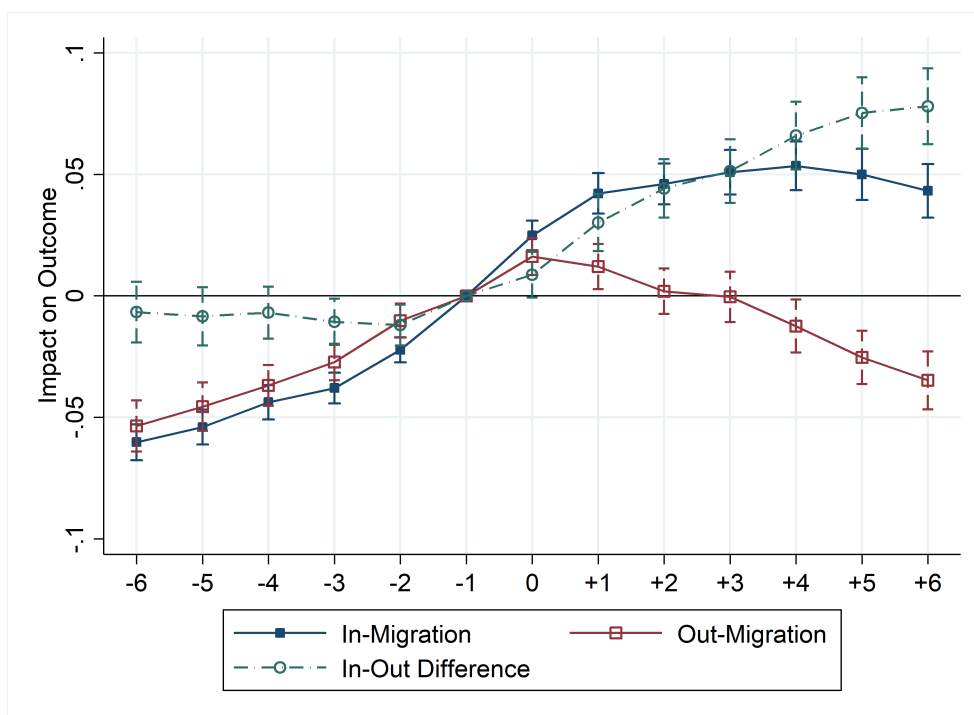**Figure 7:** Network-level Estimates, Average Patent Counts $\bar{P}_{ict}$



**Figure 8:** Network-level Estimates, Average Patent Shares $\tilde{\bar{P}}_{ict}$

In Figure 9 I report the results relative to the citation weighted average patent count measure $\bar{B}_{ict}$. Unlike in the case of unweighted measures, the estimated quality of the residual patents due to the "In" Network also registers a relative decline which starts 3-4 years after the event. A possible interpretation of these results is that the superstar's internal effect is indeed limited in time, but is also associated with an advantage of reputational kind – that is, individuals associated with a local superstar may become better able to achieve a higher number of low-quality patents, potentially crowding out the good ones.[19] If this hypothesis were true, it would represent a less unambiguously positive social effect associated with connecting to superstars.[20] Connected inventors in the "Out" network, instead, experience even a more marked decline of their independent production relative to the previous results, and in this case without a very visible lag. As a result, the In-Out difference measure for to these estimates switches to about 0.3 more citation-weighted patents for the "Out" group six years before the event (relative to the baseline) to about 1.6 more counts in favor of the "In" group. The results for the other three citation-weighted measures ($B_{ict}$, $\tilde{B}_{ict}$ and $\tilde{\bar{B}}_{ict}$) are not shown here, but are analogous.
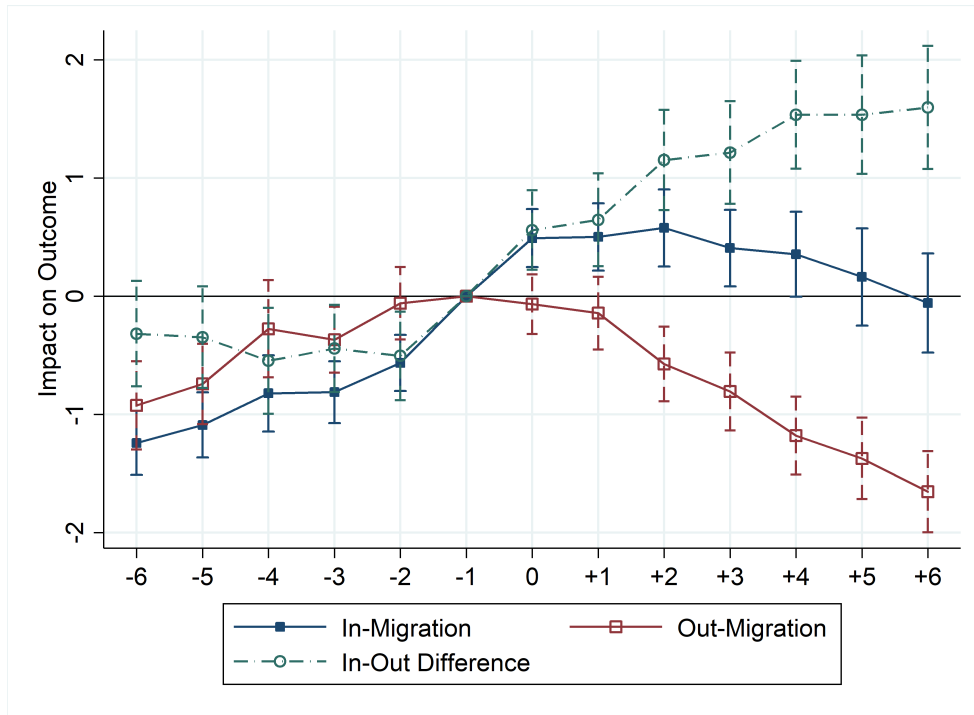


**Figure 9:** Network-level Estimates, Citation Weighted Average Patent Counts $\bar{B}_{ict}$

---

[19]One could argue that this result is due to measurement error, which is amplified by the truncation correction procedure, which disproportionately affects later dates – with increased error. However, I obtain similar results by weighting only for forward citations received after five years from the patent application; with this procedure, measurement error should not disproportionately affect patents from later years.

[20]It is difficult to provide a precise judgment about this hypothesis with the present data, but the issue certainly deserves further investigation.

Given the striking similarity between the pre-trends of the two groups across all these estimates, it is unlikely that the results are driven by the endogenous choice of superstars (due to, say, an inclination of superstars to move from relatively "worsening" to relatively "improving" networks). In fact, the clean shape of the curves traced out by all In-Out difference estimates – about flat before the event, approximately flat in the years following the adjustment phase – suggests that endogeneity issues are unlikely to play a role in explaining the results. Arguably, the location choices of successful inventors are perhaps only marginally affected by the relative quality of their local collaborators. By walking one one step ahead, one can interpret the long-run In-Out differences observed six years after the event as *twice* the structural effects of the superstars' proximity on the independent productivity of their collaborators, as it is expressed in equation (1) of the conceptual framework, and where the double count records both one network's gain and the other's loss. To this end, however, it is necessary to split the results between the two subgroups characterized in that framework: coworkers and non-coworkers.

## 4.2   Breaking Down the Results: Coworkers vs. Non-Coworkers

In Figure 10 I report the estimates of model (5) for the average patents $\bar{P}_{ic}$ outcome measure, split between subgroups of each local network: coworkers and non-coworkers ("others"). Because collaborators classified as coworkers are about two thirds of the total, normalized measures of patent output suit this exercise better. In order not to overcrowd the graph, the corresponding In-Out differences are not displayed. Overall, the results appear qualitatively similar to those from Figure 7, although some differences between coworkers and non-coworkers are apparent. First, the former enjoy faster pre-trends before the event, regardless of which local network they belong to. Second, the post-event differences are much wider for coworkers than for other collaborators; in fact, the confidence intervals of the $\beta_k$ estimates for "In" and "Out" non-coworkers do not overlap only for $k = 5, 6$. By relating the long-run differences observed six years after the event to the structural effects of superstar proximity on average patent output from equation (1), these results reveal parameter estimates of about $\widehat{\rho}_0 \simeq 0.35$ for non-coworkers, and $\widehat{\rho}_1 \simeq 0.75$ for coworkers – the latter is more than twice the former.

It is useful to reconnect these figures to the earlier discussion about the interpretation of internal effects. While productivity gains enjoyed by coworkers are certainly interesting from the point of view of the internal organization of R&D and possibly relate to learning mechanisms, they can hardly be considered spillovers. Internal effects on non-coworkers, on the other hand, *might* be interpreted as spillovers in certain circumstances, but not always (for example, this interpretation should not apply to the case of R&D joint ventures). The empirical results indicate that the estimated internal effects overwhelmingly affect coworkers; therefore, the evidence in favor of spillovers *mediated by geography* in collaboration networks is limited. In addition, since

23

the classification of superstars' connections as coworkers can switch following the event, part of the effects reported in Figures 6-9 is likely attributed to cross-firm superstar mobility.[21]
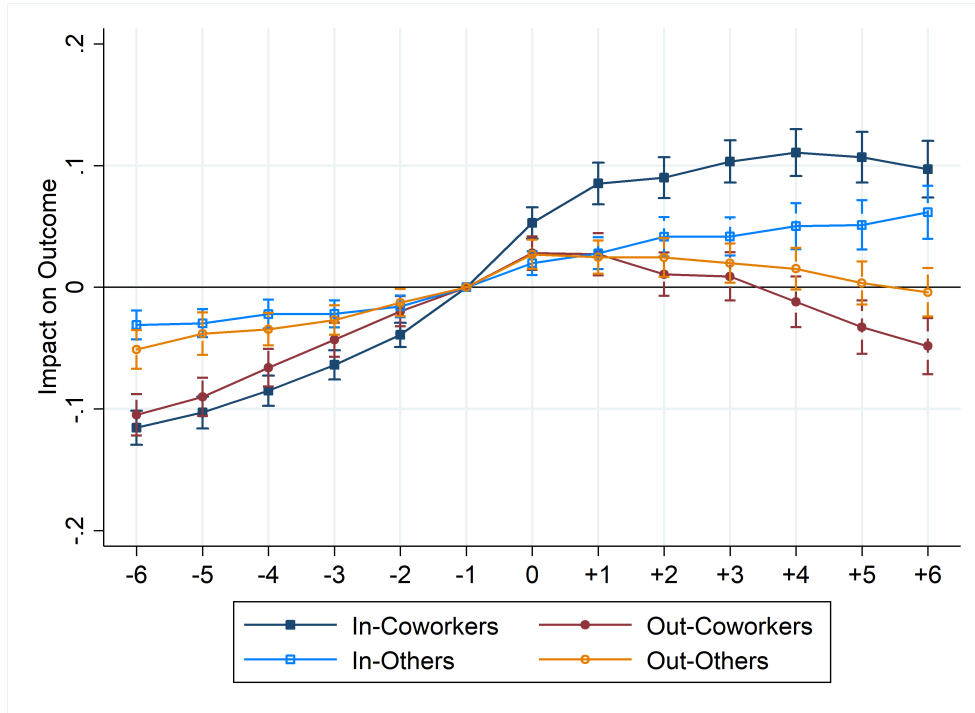


**Figure 10:** Network-level Estimates, Average Patent Counts $\bar{P}_{ict}$ by coworker status

## 4.3 Baseline Results: City-level Measures

In what follows, I analyze the consequences of superstars' relocation events on city-level innovation measures, in order to ideally uncover external spillovers. Figure 11 displays the estimates of model (5) relative to the pure patent count measure $C_{ict}$. The results are less clear-cut than in the case of network-level outcomes. Following the event, cities of destination innovate more in relative terms; in the sixth year after the event, the In-Out measure registers an increased difference of about 85 more patents between the two localities. However, one immediately notices a difference in pre-trends that is also clearly favorable to "In" cities, which indicates that superstars typically move towards places that are already on a track of faster innovation (in relative terms). Overall, the estimates from Figure 11 cannot be seen as conclusive evidence of external spillovers at work, and suggest to examine additional city-level outcomes.

In fact, the analysis of other measures casts serious doubts on the external spillovers hypothesis. For example, the estimates relative to the average city-level count $\bar{C}_{ict}$ are very noisy,

---

[21] In order to accurately quantify to what extent the results depend on cross-firm mobility, more precise data about inventors' and superstars' affiliation to firms are necessary.
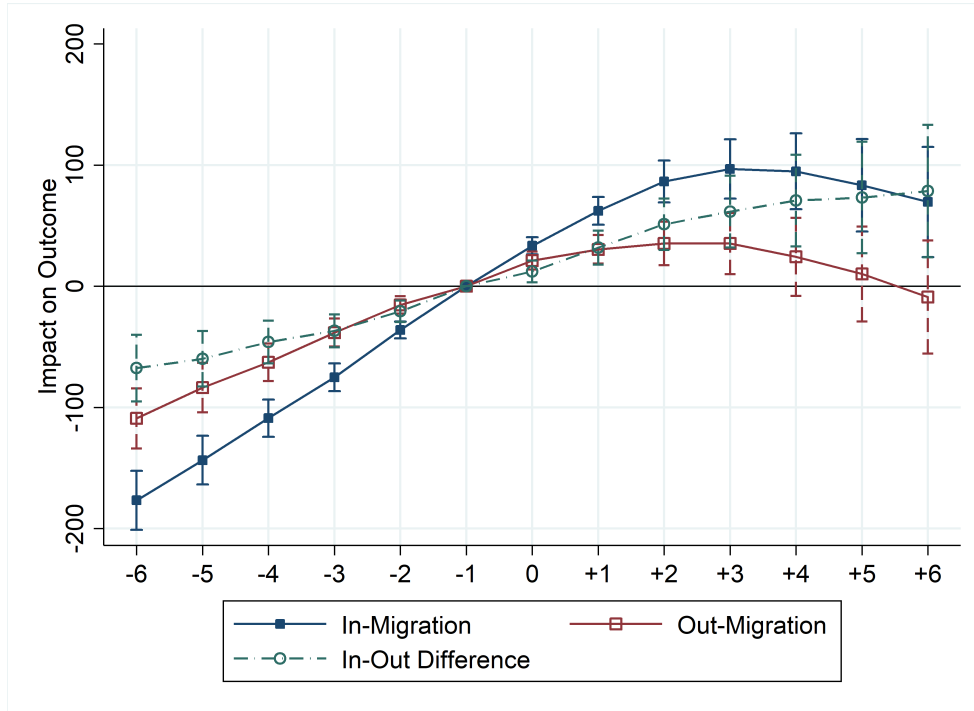
**Figure 11:** City-level Estimates, Patent Counts $C_{ict}$

never statistically different from zero and they display no clear pattern (hence, they are not reported for space reasons); in any event, they do not suggest the presence of any effect on the average productivity of local unconnected inventors. A clearer picture is offered by the analysis of citation-weighted average counts $\bar{G}_{ict}$, whose event-study estimates are reported in Figure 12. By controlling for patent quality, it appears that both cities involved in the events experience similarly decreasing trends in average inventor productivity, with no statistically discernible differences. These pieces of evidence can be reconciled with that from figure 11 through a simple explanation: superstars tend to move to places with faster inventor population growth; the patent output obtained by the new inventors appearing in these places, however, is typically not higher while it seems to be of lower quality (as measured by the adjusted number of citations). In this last respect, "In" cities do not seem to be much different from "Out" locations.

In summary, the city-level estimates do not register relevant external spillovers on inventor productivity at work, in the sense expressed by parameter $\sigma_1$ in equation (2) of the conceptual framework. Consequently, this evidence withdraws support and motivation for "big push" type of place-based innovation policies, especially those that are aimed at attracting highly skilled inventors in the hope that they would boost the growth of local high-tech industries. However, these results still leave open questions about what mechanisms drive the spatial mobility of superstars. An hypothesis that is worth exploring has to do with the Marshall-DP agglomeration

force left aside in this empirical analysis, that is *matching*. Specifically, if superstars pair up with collaborators of the right skills more easily in cities with a larger inventor population, both superstars themselves and inventors of various skill levels may endogenously sort across places, in a way that is consistent with the results from Figures 11 and 12.
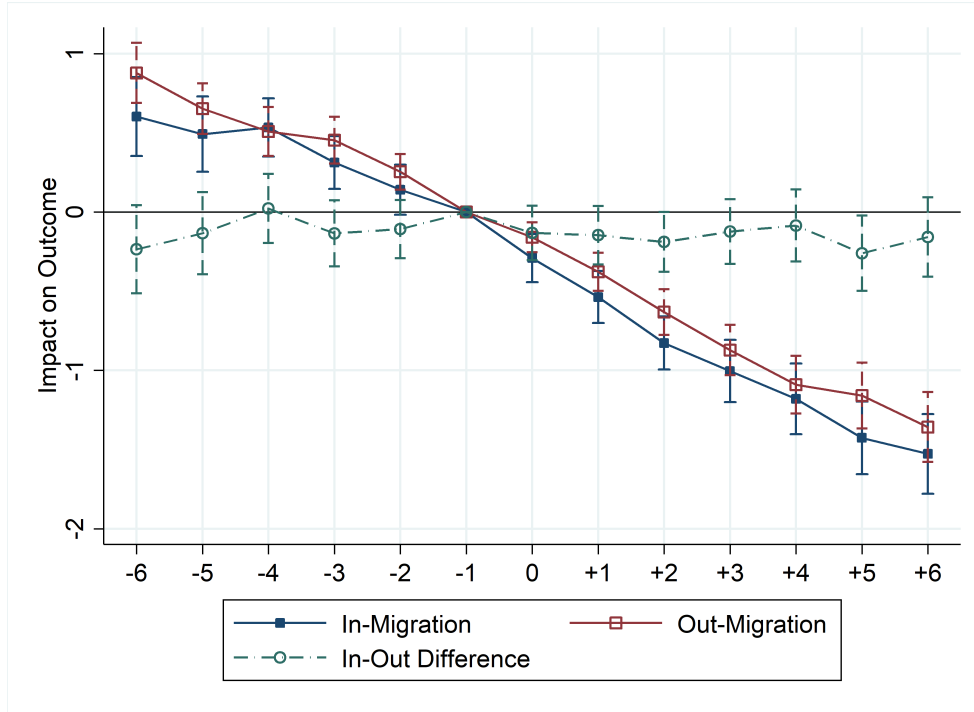


**Figure 12:** City-level Estimates, Average Citation-Weighted Patent Counts $\bar{G}_{ict}$

## 4.4 Breaking Down the Results: Superstar Heterogeneity

It may be argued that the effects of superstars' mobility may not be identical if the superstars in question are responsible for 20 patents, 100 patents, or even 1,000. Thus, it is interesting to re-examine the results by looking at separate groups of superstars, distinguished by their relative position in the patent distribution. Figure 13 displays network-level estimates of the overall internal effect based on average patent counts $\bar{P}_{ict}$, and split by the three groups of inventors defined in Section 2: Dwarves (top 5-1%), Giants (top 1-0.5%) and Bright Giants (top 0.5%). In the graph, I report both event coefficients relative to the "In" network (in shades of red-orange) and to the "Out" network (in shades of blue), but for readability I omit confidence intervals as well as the "In-Out" estimates. It appears that for both "In" and "Out" localities, the post-event outcomes related to non-bright Giants are *higher* than those relative to the Dwarves. In the long run, geographical closeness with a superstar inventor in the lower top 0.1% of the distribution is associated with about 0.2 more yearly patents per coauthor. Unexpectedly, the converse is

true for Bright Giants: the post-event outcomes associated with the "top of the top" inventors are actually *lower* than for those related to Dwarves, for either local network. Therefore, there seems to be no monotonic relationship between the position of superstar inventors in the patent distribution and the internal effects associated with their moves.[22] A potential explanation of this finding is that Bright Giants have collaborators dispersed across wider areas, and thus they have relatively less opportunities to interact with those residing in their own location.
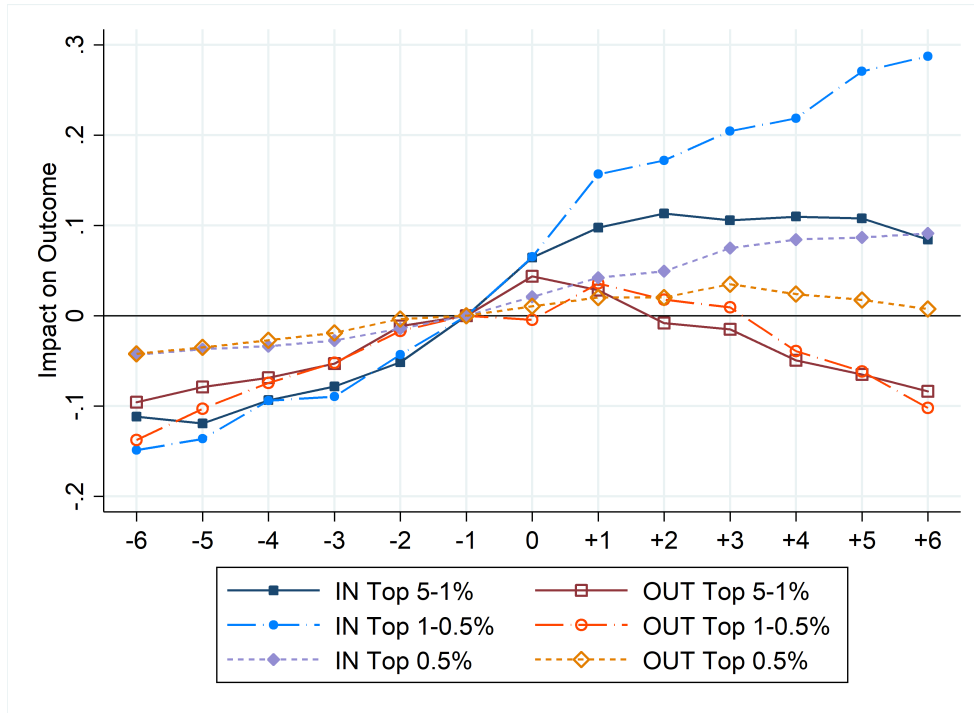


**Figure 13:** Network-level Estimates by Group, Average Patent Counts $\bar{P}_{ict}$

Figure 14 displays the analogous estimates relative to the city-level patent count measure $C_{ict}$, distinguished by superstar group.[23] Once again, the results are not favorable to the existence of sizable external spillovers, unless one restricts the analysis to very top superstars. In fact, the estimates for Dwarves appear similar to those from Figure 11, featuring divergent pre-trends and small, marginally significant post-trend differences. In the case of the lower group of top 1% superstars (non-bright Giants) the pre-trends appear coincident; however, while the "In" cities typically register more patents following the event, the difference with the "Out" localities is not statistically different from zero. Regarding Bright Giants instead, pre-trends are

---

[22]It must be remarked that the differences across these groups are not statistically different from zero. Hence, the estimates reported in Figure 13 as well as the following Figure 14 must be interpreted as a statistical decomposition of the overall effects from Figures 7 and 11, respectively.

[23]While confidence intervals are once again omitted for the sake of clarity, in this discussion I make references to inference results; more detailed results featuring the estimated confidence intervals for Figures 13, 14 and those that follow are available upon request.

not exactly parallel but they are similar enough, while the post event dynamics for the two types of cities is statistically significant and amounts to an amplified difference of about 100 patents in favor of the "In" locality six years after the event. This suggests that if external spillovers are to be found, they may lie in the extreme right tail of the patent per inventor distribution. Note that this result is consistent with that from Table 13: if very bright stars spread their knowledge among a larger population, the relative benefit on their direct collaborators is probably reduced.
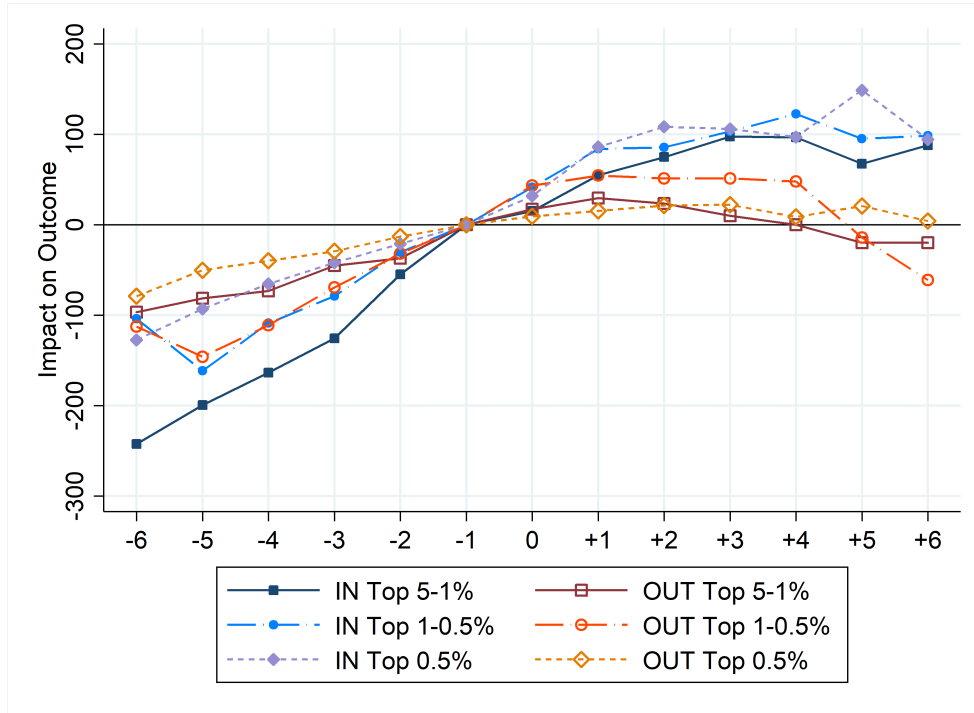


**Figure 14:** City-level Estimates by Group, Patent Counts $C_{ict}$

## 4.5   Breaking Down the Results: Geographical Scope of the Move

In Figure 15 I break down the results about internal effects on local collaborators by the geographic scope of the superstars' moves. Specifically, I consider the same geographic classification of events summarized in Table 4 (within North America, from N.A. to Europe, from Europe to N.A., within Europe). The graph reports, for each group, estimates relative to the "In-Out" differences in average patent counts, $D_{it} = \bar{P}_{iIt} - \bar{P}_{iOt}$. As a further step to make the figure easier to interpret, I once again omit confidence intervals. For moves directed to a North American city, the estimated post-event differentials are typically statistically significant; notably, in the case of moves from Europe to N.A. the effect seems to widen twofold in later years, although the estimated coefficients are not statistically different from those corresponding to the events

that involve only North American cities. In the case of events directed to Europe, instead, the estimates generally do not register any statistically significant effect.[24] It would be intriguing to explore further the reasons behind these continental differences, and in particular whether they are due to institutional differences at play.
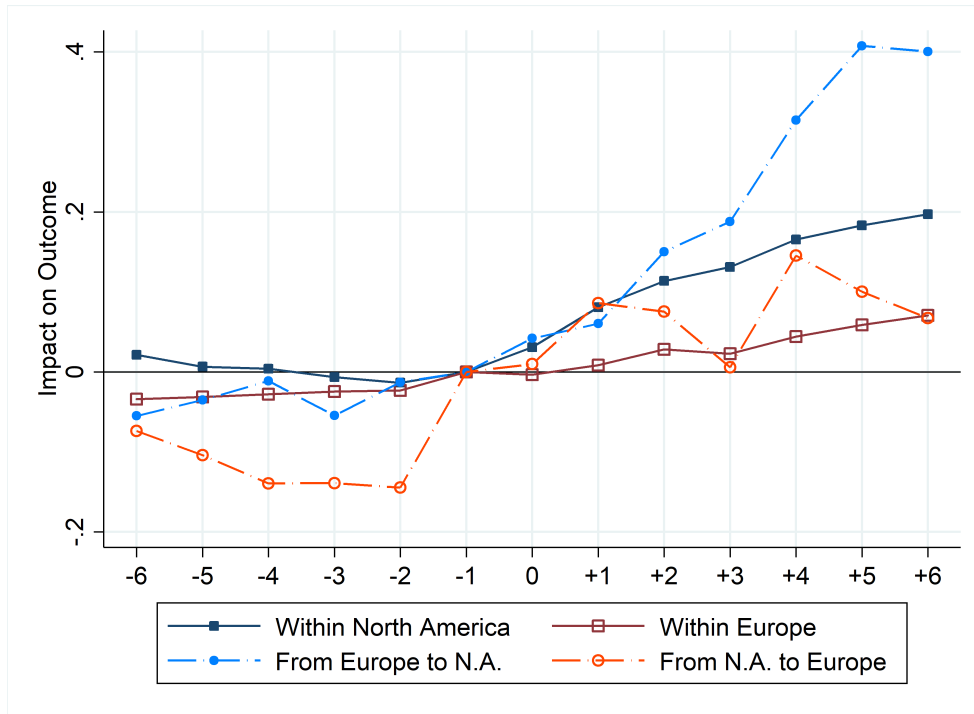


**Figure 15:** Network-level Estimates by Geographical Scope, Average Patent Counts $\bar{P}_{ict}$

## 4.6 Breaking Down the Results: By Technological Macro-Field

Finally, I also look at separate network-level estimates of the internal effect decomposed by wider technological area. Recall that I assign an event to a macro technological group if more than 75% of the superstar's patents fall in a specific category as per the first patent class listed on each patent, and that this conservative choice implies the loss of about 40% of the event sample. In Figure 15 I report estimates of the "In-Out" difference in average patent counts $\bar{P}_{ict}$ for the five non-residual macro-areas: Chemical (excluding Biopharma); Electrical, Electronic & ICT; Biotechnology and Pharmaceutical; Mechanical; and Instruments. I omit confidence intervals again. The post-effect estimates appear particularly strong in the case of the Electro-ICT and Mechanical areas; in these sectors, the post-event differenced coefficients are from two to five times larger than in the other areas, and are statistically significant. In the other sectors

---

[24]In the case of events from North America to Europe, it appears that if an effect exists it reveals itself one year in advance; however, this may also be an artifact of statistical noise in this particular (small) subsample.

instead, the estimates appear noisy – especially so in the highly innovative and R&D-intensive Biopharma sector – and are not statistically significant; however, this may also be due to the smaller size of the corresponding subsamples.
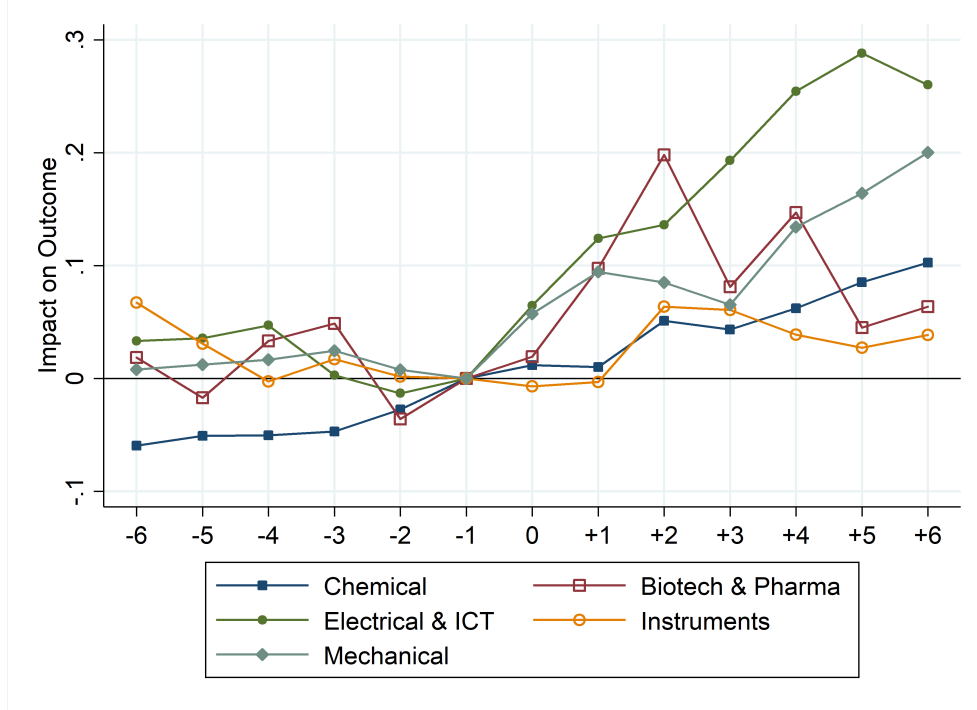


**Figure 16:** Network-level Estimates by Technological Field, Average Patent Counts $\bar{P}_{ict}$

# 5   Conclusions

This study analyzes patterns in patent production that precede and follow the relocation of a superstar inventor from one city to another. The objective of the analysis is to separate "internal" learning effects that benefit the direct collaborators of a superstar, from "external" spillovers that may affect other local inventors as well. The empirical results show that the location choices of superstars have very strong effects on the residual productivity of their proximate collaborators: on average, following the superstar's move the gain enjoyed by the inventors in the new city of residence of the superstar and the loss experienced by those in the old city combine into a difference of about 0.16 more yearly patents per inventor in favor of the former group. By contrast, there seems to be little to no evidence for external spillover effects on the larger population of local inventors. Consequently, this study offers little empirical support to policies aimed at attracting leading technological figures to a local area, on the expectation that this would generate wider agglomeration effects through the action of externalities.

This study raises more questions than the answers which it attempts to provide. In particular, why the effect on citation-weighted patents relative to the "In" network declines over time? To what extent are internal effects explained by superstar mobility between firms? What causes the co-location of superstars and inventors of lower productivity which is observed by means of the city-level analysis? What is exactly the role of "very bright" giants, and do at least they cause external spillovers upon larger populations of inventors? Finally, why no internal effects due to superstars' proximity are observed at play in European networks? Each of these questions is relevant enough to motivate a separate analysis, to which deservedly devote some future work.

# References

Agrawal, Ajay, Devesh Kapur, and John McHale (2008) "How do spatial and social proximity influence knowledge flows? Evidence from patent data," *Journal of Urban Economics*, Vol. 64, No. 2, pp. 258–269.

Agrawal, Ajay, Iain Cockburn, and Carlos Rosell (2010) "Not invented here? Innovation in company towns," *Journal of Urban Economics*, Vol. 67, No. 1, pp. 78–89.

Akcigit, Ufuk, Salomé Baslandze, and Stefanie Stantcheva (2016) "Taxation and the International Migration of Inventors," *American Economic Review*, Vol. 106, No. 10, pp. 2930–2981.

Almeida, Paul and Bruce Kogut (1999) "Localization of Knowledge and the Mobility of Engineers in Regional Networks," *Management Science*, Vol. 45, No. 7, pp. 905–917.

Audretsch, David B. and Maryann P. Feldman (1996) "R&D Spillovers and the Geography of Innovation and Production," *American Economic Review*, Vol. 86, No. 3, pp. 630–640.

Azoulay, Pierre, Joshua S. Graff Zivin, and Jialan Wang (2010) "Superstar Extinction," *The Quarterly Journal of Economics*, Vol. 125, No. 2, pp. 549–589.

Bandiera, Oriana, Iwan Barankay, and Imran Rasul (2010) "Social Incentives in the Workplace," *The Review of Economic Studies*, Vol. 77, No. 2, pp. 417–458.

Bloom, Nicholas, Mark Schankerman, and John Van Reenen (2013) "Identifying Technology Spillovers and Product Market Rivalry," *Econometrica*, Vol. 81, No. 4, pp. 1347–1393.

Branstetter, Lee G. (2001) "Are knowledge spillovers international or intranational in scope? Microeconometric evidence from the U.S. and Japan," *Journal of International Economics*, Vol. 53, No. 1, pp. 53–79.

Breschi, Stefano, Francesco Lissoni, and Ernest Miguelez (2017) "Foreign-origin inventors in the USA: testing for diaspora and brain gain effects," *Journal of Economic Geography*, Vol. 17, No. 5, pp. 1009–1038.

Breschi, Stefano and Francesco Lissoni (2009) "Mobility of skilled workers and co-invention networks: an anatomy of localized knowledge flows," *Journal of Economic Geography*, Vol. 9, No. 4, pp. 439–468.

Duranton, Gilles and Diego Puga (2004) "Micro-foundations of Urban Agglomeration Economies," in Duranton, Henderson, and Thisse eds. *Handbook of Regional and Urban Economics*, Vol. 4, pp. 2063–2117: Amsterdam: North-Holland.

Glaeser, Edward L. and Joshua D. Gottlieb (2008) "The Economics of Place-Making Policies," *Brookings Papers on Economic Activity*, pp. 155–239.

——— (2009) "The Wealth of Cities: Agglomeration Economies and Spatial Equilibrium in the United States," *Journal of Economic Literature*, Vol. 74, No. 4, pp. 983–1028.

Greenstone, Michael, Richard Hornbeck, and Enrico Moretti (2010) "Identifying Agglomeration Spillovers: Evidence from Winners and Losers of Large Plant Openings," *Journal of Political Economy*, Vol. 118, No. 3, pp. 536–598.

Griliches, Zvi (1979) "Issues in assessing the contribution of research and development to productivity growth," *The Bell Journal of Economics*, Vol. 10, No. 1, pp. 92–116.

Hall, Bronwyn H., Adam B. Jaffe, and Manuel Trajtenberg (2001) "The NBER patent citation data file: Lessons, insights and methodological tools." NBER Working Paper No. 8498.

——— (2005) "Market Value and Patent Citations," *The RAND Journal of Economics*, Vol. 36, No. 1, pp. 16–38.

Jaffe, Adam B., Manuel Trajtenberg, and Rebecca Henderson (1993) "Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations," *The Quarterly Journal of Economics*, Vol. 108, No. 3, pp. 577–598.

Jaffe, Adam B. and Manuel Trajtenberg (1996) "Flows of Knowledge from Universities and Federal Labs: Modeling the Flow of Patent Citations Over Time and Across Institutional and Geographic Boundaries." NBER Working Paper No. 5712.

Keller, Wolfgang (2002) "Geographic Localization of International Technology Diffusion," *American Economic Review*, Vol. 92, No. 1, pp. 120–142.

Kline, Patrick M. and Enrico Moretti (2014) "People, Places, and Public Policy: Some Simple Welfare Economics of Local Economic Development Programs," *Annual Review of Economics*, Vol. 6, pp. 629–662.

Kremer, Michael (1993) "The O-Ring Theory of Economic Development," *The Quarterly Journal of Economics*, Vol. 108, No. 3, pp. 551–575.

Krugman, Paul (1991) *Geography and Trade. Gaston Eyskens Lecture Series.*: Cambridge, MA: MIT Press.

Li, Guan-Cheng, Ronald Lai, Alexander D'Amour, David M. Doolin, Ye Sun, Vetle I. Torvik, Amy Z. Yu, and Lee Fleming (2014) "Disambiguation and co-authorship networks of the US patent inventor database (1975–2010)," *Research Policy*, Vol. 43, No. 6, pp. 941–955.

Lychagin, Sergey, Joris Pinkse, Margaret E. Slade, and John Van Reenen (2016) "Spillovers in Space: Does Geography Matter?" *Journal of Industrial Economics*, Vol. 64, No. 2, pp. 295–335.

Mancusi, Maria Luisa (2008) "International spillovers and absorptive capacity: A cross-country cross-sector analysis based on patents and citations," *Journal of International Economics*, Vol. 76, No. 2, pp. 155–165.

Marshall, Alfred (1890) *Principles of Economics*: London: Macmillan. (First Edition).

Mas, Alexandre and Enrico Moretti (2009) "Peers at Work," *American Economic Review*, Vol. 99, No. 1, pp. 112–145.

Moretti, Enrico (2004) "Workers' Education, Spillovers, and Productivity: Evidence from Plant-level Production Functions," *American Economic Review*, Vol. 94, No. 1, pp. 656–690.

———— (2011) "Local Labor Markets," in O. C. Ashenfelter and D. Card eds. *Handbook of Labor Economics*, Vol. 4, pp. 1237–1313: Elsevier.

Moretti, Enrico and Daniel J. Wilson (2014) "State incentives for innovation, star scientists and jobs: Evidence from biotech," *Journal of Urban Economics*, Vol. 79, pp. 20–38.

———— (2017) "The Effect of State Taxes on the Geographical Location of Top Earners: Evidence from Star Scientists," *American Economic Review*, Vol. 107, No. 7, pp. 1858–1903.

Morrison, Greg, Fabio Pammolli, and Massimo Riccaboni (2017) "Disambiguation of Patent Inventors and Assignees Using High-Resolution Geolocation Data," *Scientific Data*, Vol. 4. No. 170064.

Oettl, Alexander (2012) "Reconceptualizing Stars: Scientist Helpfulness and Peer Performance," *Management Science*, Vol. 56, No. 6, pp. 1122–1140.

Roback, Jennifer (1982) "Wages, Rents and the Quality of Life," *Journal of Political Economy*, Vol. 90, No. 6, pp. 1257–1278.

Rosen, Sherwin (1979) "Wage-based indexes of urban quality of life," in Miezkowski and Straszheim eds. *Current Issues in Urban Economics*, pp. 74–104: Johns Hopkins University Press.

Singh, Jasjit (2007) "Asymmetry of knowledge spillovers between MNCs and host country firms," *Journal of International Business Studies*, Vol. 38, No. 5, pp. 764–786.

Singh, Jasjit and Matt Marx (2013) "Geographic Constraints on Knowledge Spillovers: Political Borders vs. Spatial Proximity," *Management Science*, Vol. 59, No. 9, pp. 2056–2078.

Thompson, Peter and Melanie Fox-Kean (2005) "Patent Citations and the Geography of Knowledge Spillovers: A Reassessment," *The American Economic Review*, Vol. 95, No. 1, pp. 450–460.

Waldinger, Fabian (2011) "Peer Effects in Science: Evidence from the Dismissal of Scientists in Nazi Germany," *Review of Economics Studies*, Vol. 72, No. 2, pp. 838–861.

# Appendix: Maps of Relocation Events

Maps A.1 and A.2 display the geographical scope of superstars' moves. Events internal to North America are shown by dark blue lines, moves from Europe to North America by light blue lines, those from North America to Europe by orange lines, and intra-European events by red lines.
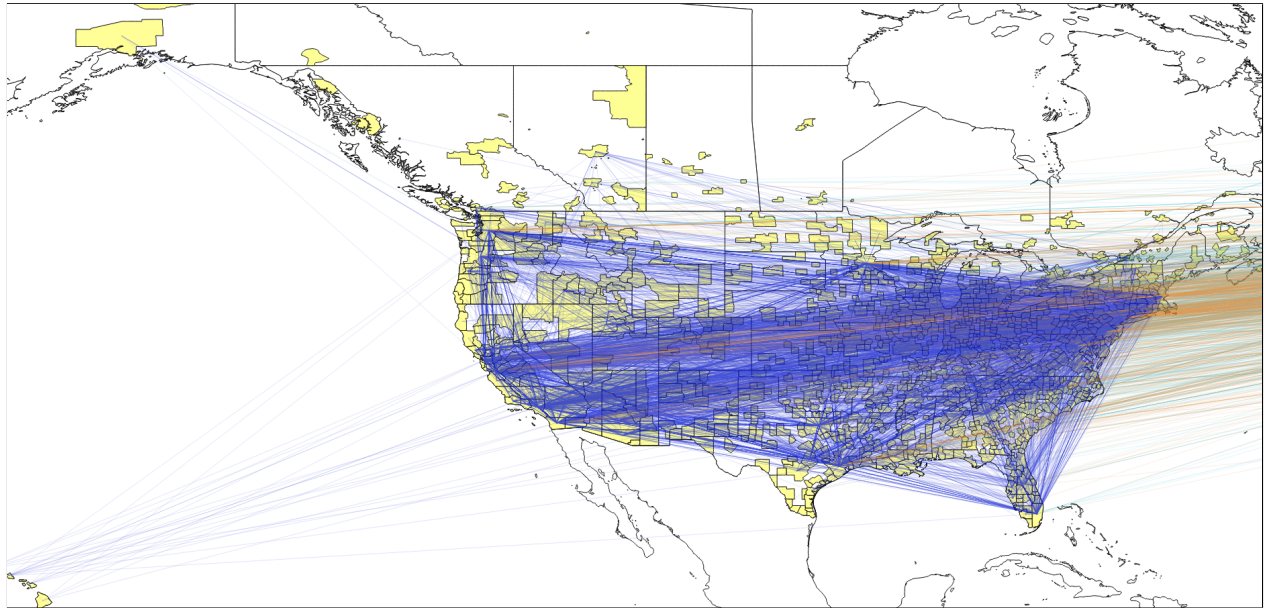


**Map A.1:** Relocation Events in North America. Yellow polygons represent US CBSAs and Canadian CMAs.
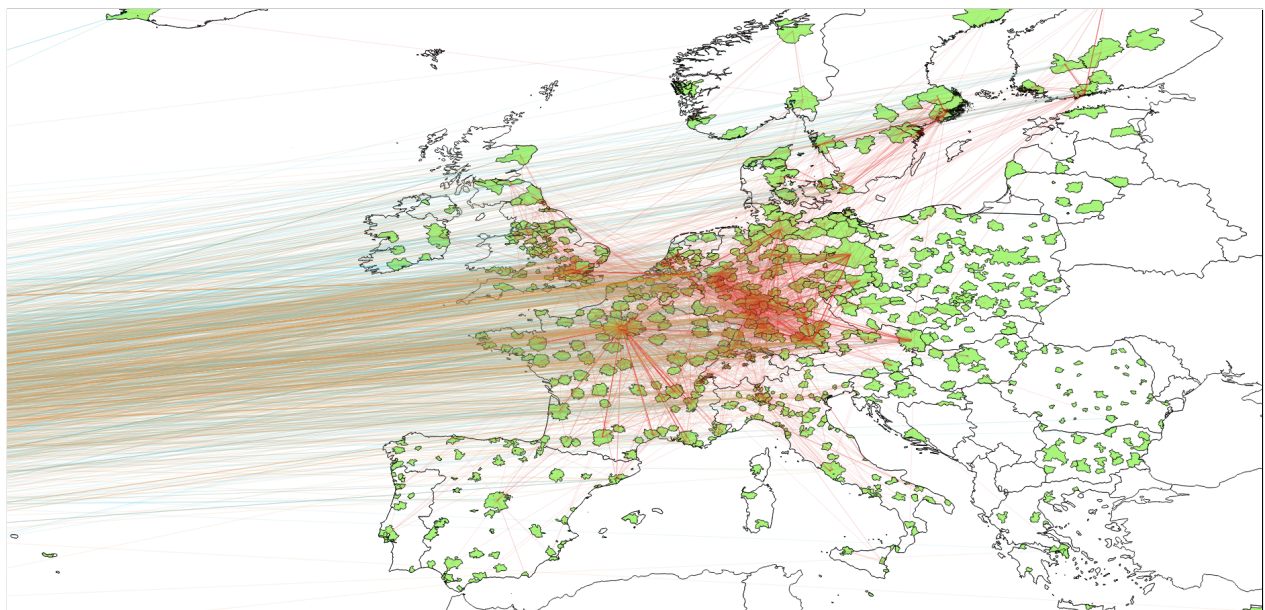


**Map A.2:** Relocation Events in Europe. Green polygons represent "European Cities" (Eurostat definition).