University of Groningen

Geographical patterns in US inventive activity 1977-1998

Castaldi, Carolina; Los, Bart

Published in:
Research Policy

DOI:
[10.1016/j.respol.2017.04.005](https://doi.org/10.1016/j.respol.2017.04.005)

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2017

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Castaldi, C., & Los, B. (2017). Geographical patterns in US inventive activity 1977-1998: The "regional inversion" was underestimated. *Research Policy*, 46(7), 1187-1197.
<https://doi.org/10.1016/j.respol.2017.04.005>

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Geographical patterns in US inventive activity 1977–1998: The “regional inversion” was underestimated

Carolina Castaldi^{a,*}, Bart Los^b

^a School of Innovation Sciences, Eindhoven University of Technology, The Netherlands

^b Faculty of Economics and Business, University of Groningen, The Netherlands

ARTICLE INFO

Keywords:

Geography of invention
Regional inversion
Superstar inventions
United States of America

ABSTRACT

Towards the end of the previous century, the geography of US inventive activity changed drastically. The old hotbeds of invention (the Northeast and Midwest) lost much of their prominence, and rates of invention in Western states grew considerably. In this paper, we argue that this well-known “regional inversion” has been underestimated. We arrive at this conclusion by addressing an important concern regarding the use of raw patent counts in the previous literature: raw patent counts do not tell much about inventive performance, since the importance of patents in terms of their impact on future technological and economic developments varies much. We focus on “superstar” patents, which are disproportionately important. We identify these employing a statistical regularity in citation patterns. We find that the West did not only outpace the Northeast and Midwest in the numbers of patents produced, but also specialized much more in patents that “really matter”.

1. Introduction

Inventive activity is clustered in space. Contributors to the literature have identified several causes for such clustering, like economies of scale in R & D activities, the uneven presence of highly skilled researchers, the limited geographical distance that knowledge spillovers travel, etc.¹ As a consequence of this spatial clustering and the cumulativeness that characterizes the innovation process, some regions end up as “hotbeds” of invention, whereas many other regions are much less successful in this respect. With the advent of new technologies, the location of invention hotbeds is likely to change. Ceh (2001), Hicks et al. (2001), Johnson and Brown (2004) and Ó hUallacháin and Leslie (2005), for example, all found that US-based inventive success steadily relocated in the 70s, 80s and 90s, from clusters in the Northeast and Midwest to locations in Western states. This “regional inversion” is virtually undisputed.

In this paper, we argue that the magnitude of the regional inversion that characterized changes in the geography of US invention has been underestimated considerably. We arrive at this conclusion after addressing an empirical issue that is widely acknowledged in the literature. The methods applied in the literature on the regional inversion are based on “raw” patent counts, which implies that the inventions

underlying patents are assumed to be equally important. It is well-known, however, that patents are extremely heterogeneous with regard to the importance of the underlying invention, both in a technological and an economic sense.² The degree of heterogeneity has most probably increased considerably over time, for at least two reasons. First, the emergence of markets for technology has led to an increased use of patents for strategic reasons, with hosts of incremental inventions being patented (partly to enhance the protection of the truly important inventions). Second, the capabilities of the US Patent and Trademark Office (USPTO) to examine prior art appear to have deteriorated considerably, due to budget cuts and unfavorable financial incentives for patent examiners (Jaffe and Lerner, 2004).

In this paper, we reconsider trends in the location of true inventive success, by looking at really important inventions only. We start this analysis by summarizing a statistical method to identify so-called “superstar” patents. Next, we use information about the state where inventors associated with such superstar patents were located to produce quantitative results about the recent history of geographical invention patterns in the United States.

The paper is structured as follows. In Section 2, we discuss our methodology to identify superstar patents. Section 3 focuses on trends in the numbers of superstar patents over time and compares the extents

* Corresponding author.

E-mail address: c.castaldi@tue.nl (C. Castaldi).

¹ See, for instance, Audretsch and Feldman (1996), Anselin et al. (1997) and Breschi and Lissoni (2001).

² Trajtenberg (1990), for example, found that about 50% of US patents were never cited, while a very limited number of patents accounted for the vast majority of citations to his samples of patents. Harhoff et al. (1999) found a similar degree of heterogeneity when inventors were asked (ex-post) to rate the monetary value of their patented inventions.

of geographic concentration of superstar invention and invention measured using raw patent counts. We also provide quantifications of our claim that the regional inversion has been much more marked than thought already. Western states did not only experience much higher growth rates in the numbers of patents produced than the Northeast and Midwest, but also specialized much more in producing superstar patents (as opposed to producing “run-of-the-mill” patents). Section 4 concludes by discussing the limitations of our approach, and by sketching some avenues for future research.

2. “Superstar” patents vs. regular patents

Before analyzing trends in the distribution of superstar patents over US states, a procedure to tell such superstar patents apart from “regular” patents should be introduced.

By now, many researchers have acknowledged the fact that raw patents counts do not convey much information about the inventiveness of firms, regions or countries. Nowadays, many patents are granted to very minor inventions (see, e.g., Jaffe and Lerner, 2004). Applicants often do not seek protection against imitation of these inventions, but use patents as change in cross-licensing negotiations. Simultaneously, however, a much smaller fraction of patents are awarded to inventions that have a big impact, in economic terms and/or with respect to subsequent technological developments. Given that run-of-the-mill inventions are much more common than what we will label “superstar” inventions, just counting patents might well yield misguided conclusions about inventiveness.³ In the recent literature, several approaches to focus on the technologically most important subset of patents have been proposed.⁴

In their analysis into the inventiveness of US Metropolitan Statistical Areas, Clark et al. (2010) looked at the subset of ‘triadic patents’. Since application at various patent offices involves more costs and efforts, triadic patents can be viewed as representing inventions that were considered by applicants as relatively important. In a more or less similar vein, Usai’s (2011) study regarding invention in OECD regions focused on the subset of patents that were applied through the Patents Cooperation Treaty, which implies that applicants at least contemplated filing for a patent at multiple national or supranational patent offices.

Most methods follow Trajtenberg (1990) and suppose that numbers of citations received by patents are a useful indication of their importance. Schoenmakers and Duysters (2010), who consider patents issued by the European Patent Office (EPO), defined all patents that received at least 20 citations within the first five years after application as “radical”, irrespective of the year of application or the technology category concerned.

Phene et al. (2006), Akkermans et al. (2009) and Singh and Fleming (2010) defined inventions as belonging to the “breakthrough” and “radical” categories if they are part of the specific upper quantiles of citation distributions. Phene et al. (2006), who considered biotech patents only, selected the upper 2% of patents in terms of the numbers of citations and checked the robustness of their results by taking the upper 1%, 5% and 10% percentiles as well. Akkermans et al. (2009) considered several citation-based indicators of invention importance for an economy-wide set of industries, and considered the top-5% and top-10% of patents in terms of these indicators as radical inventions. These quantiles were constructed per cohort of patents (characterized by year of application) in an industry-of-origin context. Hall and Trajtenberg (2006) did not distinguish between industries-of-origin or technology categories, but roughly selected all patents that received at least 3 times

the number of citations received by the patent that demarcated the 99th percentile of the entire distribution of citations received. Ejermo (2009) based his criterion on several aspects of patents: numbers of forward and backward citations, family size and the number of oppositions against a patent. By means of factor analysis, he came up with a single variable indicating the importance of a patent. Finally, Dahlin and Behrens (2005), who studied a narrow set of patents related to tennis rackets, specified criteria based on the citation profiles of patents.

To study geographical trends in inventive activity in US states, we must adopt an identification method that is suitable for analyses of a wide range of technologies. The cohort-specific technology-specific computation of upper quantiles adopted by Akkermans et al. (2009) and Singh and Fleming (2010), among others, recognizes that patenting and citing behavior vary across technologies and over time (see Hall et al., 2002; for an early plea in this regard). Considering, let’s say, the 5% most cited patents in each cohort of patents associated with each technology category as breakthrough is not without problems, however. Such an approach disregards differences (over time and across technologies) in the degree to which strategic considerations affect patenting behavior. If strategic patenting is widespread in a technology category, the 5% most heavily cited patents will most probably contain relatively many incremental, run-of-the-mill patents, while 5% might well be an underestimation of the share of radical inventions if strategic patenting is virtually absent. Furthermore, the magnitude of the selected upper quantile is rather arbitrary. It is hard to tell whether 5% would be close to the true share of breakthrough patents in the total population of patents. 1% or 10%, for example, might well be better approximations, but we do not know. Hence, we will rely on data constructed by Castaldi and Los (2015), who employ a stylized fact in distributions of numbers of received citations to obtain cohort-specific and technology-specific sets of “superstar” patents.

Silverberg and Verspagen (2007) discovered that the numbers of citations that patents receive can be described by a mix of a lognormal distribution and a Pareto distribution. For the vast majority of patents, a lognormal distribution fits the empirical distribution very well. There are a few patents, however, that receive many more citations than one would expect on the basis of the lognormal distribution and this right tail of the distribution (which is “fat”) is generally Pareto-distributed.

The lognormal distribution emerges if the growth rates of numbers of received citations are independent of the number of citations received already. In such a situation, the fact that some of these patents got more citations than others is the consequence of the stochastic nature of the citing process and does not reflect intrinsic differences between these patents (Gibrat’s Law). The finding that some patents received many more citations than would be in line with the lognormal distribution, however, suggests that citations to a very limited subset of patents are governed by a different process. As Sanditov (2006, Ch. 6) argues, small initial differences between inventions and path dependencies in innovation processes together lead to situations in which some inventions act as “focusing devices” for future innovation processes within new technological paradigms (Dosi, 1982). Following the literature in studies on income distributions, we will make a distinction between “superstar” patents (those for which the numbers of citations received cannot be reconciled with the lognormal distribution for the majority of patents) and regular patents.

The procedure adopted by Castaldi and Los (2015) starts following Silverberg and Verspagen (2007), who used the Drees-Kaufman-Lux approach (based on the Hill-estimator, see Lux, 2001) to estimate the cut-off point between the lognormal and the Pareto parts of the distribution of received citations. This cut-off point gives the minimum number of citations a patent must have received to belong to the subset of patents of which the numbers of citations are Pareto-distributed. Next, two problems had to be overcome. First, the statistical properties of the cut-off point estimators cannot be derived analytically, but it is well-known that it can be hard to tell a lognormal and a Pareto distribution apart in empirical cases. Castaldi and Los followed Clause

³ Several labels have been introduced to stress that some patents reflect inventions that have more impact than others. Examples are “important”, “radical” and “breakthrough”. For reasons that will become apparent later, we prefer to introduce the term “superstar” patents.

⁴ The list of procedures is by no means exhaustive, but gives an impression of the most common approaches.

et al. (2009), and constructed confidence intervals by means of bootstrapping methods. These confidence intervals appeared to be very wide for some technology categories, but much narrower for many others. Furthermore, the point estimates obtained by the raw Drees-Kaufman-Lux procedure appeared to be relatively far from the mean taken over 1000 bootstraps. Castaldi and Los decided to base their assignment of patents to the superstar and regular patents categories on their numbers of received citations in comparison to the mean cut-off number of citations over the bootstrapped pseudo-samples. For old cohorts of patents, all patents with a number of citations at least as high as the number of citations associated to their technology-specific, cohort-specific cut-off point are labeled “superstar” patents.

The second problem was caused by truncation issues and relates to younger cohorts of patents. Since the citations data do not extend beyond 2006, patents applied for in relatively recent years did not have much time to gather citations included in subsequent patents. Analyses of older cohorts of patents showed that there are some differences in the citations-receiving behavior of superstar and regular patents at young age already, but that the most important difference between the two types of patents is the length of the period during which patents continue to receive citations: superstar patents stay “alive” for a much longer period.⁵ Castaldi and Los (2015) adopted a probabilistic approach for more recent patents. For such patents, it is uncertain whether they will be superstar patents, but the probability of this event can be estimated based on data on citation-related variables in the first years after application (the independent variables are (i) the number of cumulated citations obtained at the patent age for which the most recent citation data was available, (ii) an indicator of the share of these cumulated citations received in the most recent years for which citation data was available, and (iii) the generality measure of Hall et al., 2002; which quantifies the number of technological fields in which the patent has provoked citations). The second variable captures the effect that superstar patents tend to continue attracting many citations at relatively old age, as discussed above. The third indicator is incorporated to take into account that application of the knowledge contained in a patent in many fields enhances the chances that it will help creating a new technological paradigm in at least one field. This implies that Castaldi and Los have samples of clearly superstar patents and clearly regular patents for the years 1976–1985. For the years 1986–1999, they have the estimated probabilities for each and every patent being a superstar after 20 years. In this paper, we are not interested in individual patents and inventions, but focus on groups of patents. We opted for a procedure in which a patent with a 0.2 probability of becoming superstar counts for 0.2 superstar patent and for 0.8 regular patent and added these probabilities. Hence, two more recent patents with a 0.5 probability of becoming a superstar patent count as “heavily” as a single early patent that Castaldi and Los unconditionally attributed to the class of superstar patents.

In total, Castaldi and Los (2015) determined the superstarness of slightly more than 800,000 patents applied for in the period 1976–1999, of which the first inventor was located in the United States. The NBER Patent Citations Datafile they used has some characteristics that make it inappropriate to take patents into account that do not have a clear US-basis. As a consequence, superstarness in this paper should be interpreted as “superstarness relative to other patents originating in the U.S.”.

Hall et al. (2002) gave a number of health warnings about the use of

⁵ An example: regular patents in the 1980-cohort of the category Information Storage had received 50% of the citations it had accumulated in 2006 after on average 7 years already, whereas superstar patents attained this after as many as 12 years, on average. Rysman and Simcoe (2008) also find that patents by standard-setting organizations (patents related to standards are generally important and receiving above-average numbers of citations) generate citations over long periods of time. This observation is in line with Sanditov's (2006) above-mentioned theoretical model based on path dependency.

patent citation data, in particular with respect to studies in which such data for different technologies and different patent cohorts are used. Their discussion, however, did not address the question whether the shares of superstar patents in total patents vary across technology categories or not. If these shares would be more or less identical, the approach chosen in studies like Akkermans et al. (2009) and Singh and Fleming (2010) might be not too far off the mark, in particular if sensitivity analyses for somewhat smaller or larger common quantiles lead to similar conclusions. Fig. 1 shows clearly that the share of superstar patents in the sets of total patents has been very specific to technologies.

For Computers, Biotechnology, and Medical Instruments, we find that the share of superstar patents in total patents has roughly been 10–13%, with generally lower shares in the first years of the analysis. On the other side of the spectrum, we find technological categories with considerably lower shares. We find, for example, that the share of superstar patents in total patents is only 5–6% for Drugs. For Communications equipment and Semiconductors, we find intermediate proportions. The percentages that we find are generally higher than the 2% or 5% most cited patents considered in earlier studies about radical or breakthrough inventions. The extent to which application of such fixed proportions leads to distorted results might well vary across technology classes and, of course, across research questions.

Fig. 1 clearly shows that statistical regularities about received citations warrant the use of technology-specific (and also cohort-specific) rules to define sets of superstar patents. These will be used to investigate the performance of US states in producing inventions that really mattered, rather than in producing run-of-the-mill (or even junk) patents.

3. The production of superstar patents in US states

Patents can be attributed to regions according to several rules. Attributing a patent to the location of the assignee is probably the most appropriate procedure in contexts in which patents are linked to profitability. These profits (if any) are likely to be recorded in the accounts of the headquarters of the firm owning the patent, or in one or more of its subsidiaries. If one is mainly interested (as we are) in the location of processes that lead to patented inventions, however, it is more useful to consider the location of the inventors. Such information is contained in USPTO patent documents and included in the NBER Patent Citations Datafile. In line with other authors (see, e.g. Singh and Fleming, 2010), we chose to attribute a patent to the geographical entity (the US state in this study) to which the location of the first inventor belonged. Patents often have two or more inventors, which are not necessarily co-located. In principle, it would be possible to “split” such patents and attribute fractions to the locations of all inventors. Given the large number of patents on which our analysis is based, it is unlikely that such a somewhat cumbersome procedure would have substantial effects on the results. According to USPTO (2010), legal rights of joint inventors are unrelated to the order in which their names appear in the patent document. In order to get some idea about potential biases due to an approach based on the location of the first inventor only, we conducted some analyses based on the data underlying Li et al. (2014).⁶ Li et al. used a Bayesian learning algorithm to assign inventors other than the first one to locations, in order to study networks of inventors.⁷

Based on a sample of about 380,000 patents in the period 1976–1985 that we could match to the Li et al. (2014) data, we find that patents with inventors in multiple states have a higher propensity (4.9%) to be superstar patents than patents with a single inventor

⁶ See also Balland and Rigby (2017).

⁷ The NBER Patent-Citations Data File (Hall et al., 2002) only contains information on the state in which the first inventor is located.

Fig. 1. Shares of superstar patents in all patents, by technology category*. **Three-year moving averages; Selected technology categories.

or multiple inventors in a single state (3.6%).⁸ This result would imply biases in our analysis if the order of states in which inventors are located would not be random. We consider the by far largest subsample of “multiple state” superstar patents, those with two inventors in two states. If the order of inventors is random, we should find that a state should be the state of the first inventor in close to 50% of the superstar patents for which it hosted one of the inventors. At a five percent significance level, we must reject the hypothesis of a random order for only three states. These are Indiana, Michigan and Ohio, which together accounted for less than 11% of all superstar patents in 1976–1985.⁹ These results lead us to think that reliance on the location of the first inventor only does not cause biased results.

3.1. Geographical clustering of superstar patents

We first study how the geography of invention differs between superstar patents and regular patents. To this end, we first aggregated patents over technological categories and computed totals per state.¹⁰ Next, we computed Herfindahl indexes, which would yield a minimum value of 1/51 in case of a completely even distribution over states. Given that Texas is much larger than Vermont, it is unlikely that such a situation would occur. Hence, Fig. 2, which depicts the geographical spread of patenting, includes the Herfindahl index for the distribution of the U.S. population over the constituent states as a benchmark. In the time period considered, this index has always been slightly higher than 0.04 and has risen slowly (the population data have been taken from US Census Bureau, 1995, 1996, 1999).

Fig. 2 clearly confirms earlier findings in the literature, which state that inventive activity is clustered in space. For the Herfindahl index for regular patents we find averages of about 0.06, with clearly higher values in the latest years in our analysis. More importantly, however, we find that the clustering of superstar patents is even much more geographically focused, with a Herfindahl index of about 0.08.¹¹ Apparently, companies locate their search for breakthrough inventions in very specific places (most likely in order to benefit from high-skilled labor forces and from geographically bounded knowledge spillovers

originating from top universities and other firms).¹² Apparently, more regular inventions can also be produced in other locations. These effects appear to have become stronger in the early 1990s, the period for which Hicks et al. (2001) found their regional inversion of inventive activity in the U.S.

3.2. Distribution of superstar patent production over census regions and states

Table 1 presents trends in the numbers of superstar patents in the four U.S. Census regions. To make comparisons meaningful, the numbers of superstar patents have been scaled by the regional population (following, for example, Ó hUallacháin and Leslie, 2005). The numbers represent 3-year moving averages. These results confirm some of the main results obtained by means of raw patent counts, as in Ceh (2001), Hicks et al. (2001) and Johnson and Brown (2004), among others: from the mid-1980s onwards, the Western part of the U.S. has overtaken the Northeast as the most inventive part of the country and has steadily increased its lead (the regional “inversion”).

Table 2 provides results at a lower degree of geographical aggregation. It shows how US states fared in terms of the production of superstar patents, again scaled by the number of inhabitants. As can be seen, a large majority of states managed to produce more superstar patents per year towards the end of our period of analysis than before. This reflects the fact that the share of superstars in total patents has been more or less stable for most technologies (see Fig. 1 above), while the total number of patents has grown considerably. The degrees to which states managed to increase their production vary wildly, however. The leader in 1977, Delaware, shows such a low growth rate for 1977–1998 (from 11 to 16 superstars per million inhabitants) that it was overtaken by as many as 13 states.¹³ Only one state (Massachusetts) was among the top-5 in all three years for which we present results in Table 2, while Minnesota, California and to a lesser extent Connecticut have also been consistently present in the set of most inventive states.

Setting Wyoming and South Dakota aside (these states did not produce a single superstar patent in 1976–1978), the fastest grower was Idaho, which saw its per capita production increase about 75-fold and was more than twice as productive as the next most productive state in 1998 (Minnesota). Next on the list of fast growers is Vermont (nr. 3 in 1998) and a couple of states that produced hardly any superstars in 1977: Nebraska, Georgia and, particularly, Iowa. The Western states Washington, Oregon and Colorado are all among the states that belonged to the group of states with just above-average superstar patent productivity in the late 1970s, but managed to get into the top-10 of producers of superstar patents by the end of the 1990s. These states increased their output of superstar patents in per capita terms approximately ten times.

Not a single state saw its per capita production of superstar patents reduced. West Virginia recorded the lowest growth rate (31%). Other Northeast coast states are also among the slowest growers. Apart from the islands of Hawaii, the states with the lowest ranks for inventive productivity growth are West Virginia, Delaware, Pennsylvania and New Jersey. The latter three states all belonged to the top-10 of most inventive states in the late 1970s, but lost their positions.

Do the results based on superstar patents as presented in Table 1 also yield the same magnitude of this inversion? The previous results would not change if the shares of superstar patents in all patents would have remained stable over time (or would have changed uniformly).

⁸ Close to 7% of the matched patents for 1976–1985 had inventors in more than one state. An ANOVA analysis on almost a million patents in the period for which we adopted the probabilistic approach yielded significant and similar differences regarding the share of superstar patents in all patents between single-state and multiple-state patents.

⁹ Inventors in Indiana and Michigan appeared too frequently as first inventors, while those in Ohio were relatively often appearing as second inventors.

¹⁰ We decided to remove patents with a first inventor in Guam, Puerto Rico and the American Virgin Islands from the dataset, in view of the extremely few patents that these unincorporated organized territories generated and their “non-state” status.

¹¹ Using his indicator for patent quality, Ejermo (2009) presents a similar diagram for the distribution of patents and breakthrough patents (“innovations” in his terminology, identified in a different way) over Swedish regions in the period 1982–1999.

¹² Verspagen and Schoenmakers (2004), for example, find that multinational firms cluster their technological activities in a very limited number of European regions. They attribute this to “asset-seeking” behavior (Dunning and Narula, 1995) by these firms, in combination with limited mobility of researchers with specific skills.

¹³ Results at the level of states are most likely affected by small number problems. This is why the main research question is addressed at the level of US Census regions, at which such problems do not prevail.

Fig. 2. Geographical clustering of patents (Herfindahl indexes)^a: 3-Year moving averages

Table 1
Numbers of superstar patents per million inhabitants, US Census Regions.^a

	1977	1988	1998
West	4.8	10.3	28.5
Northeast	5.4	9.1	17.5
Midwest	3.3	5.8	11.8
South	1.8	4.3	8.3

^a Aggregated over technology categories, 3-year moving averages (with indicated years as midpoints); Ordered by number of superstar patents per million inhabitants in 1998.

Then, a focus on superstar patents would not lead to new insights.

3.3. Superstar vs. regular patents: regional specialization patterns

We computed indicators of specialization of Census regions into the generation of superstar patents (as opposed to generating regular patents), by simply taking the shares of superstar patents in all patents, aggregated over technology categories. The results are presented in Table 3.

Part of the strong performance of Western states in terms of numbers of superstar patents per million inhabitants has been due to an increasing specialization towards the production of such very important inventions, which is a factor that has remained hidden in studies focusing on raw patent counts. In the West, the share of superstar patents in total patents grew by 4.4 percentage points, which is considerably more than the increases of 3.0–3.2 percentage points for the other Census regions.

Across states within regions, the shares of superstar patents appear to vary substantially, both in levels and in changes over (see Table 4).

In 1977, Utah had the highest proportion of superstar patents, followed by Hawaii, Nevada, California and Minnesota. For these states, 3.5–5.5% of all patents were superstar patents. Of all patents granted to inventors in states like Georgia and Iowa, however, only about 1% of all patents fell into the category of superstars.

Towards the end of the period that we analyze, the specialization patterns are very different. Some of the states with high proportions of superstar patents in the late 1970s dropped down the rankings (Utah from position 1 to position 8, Hawaii from 2 to 23 and Wisconsin from 6 to 33). Minnesota, however, had already been specialized in superstar patent production (position 5 in 1977), but had become the most

specialized of all states in the last years of the century. Nevada and California maintained their high positions. The states that showed the most remarkable tendency towards specialization in breakthrough invention are all located in the West: Washington moved from 37 to 2, Oregon from 16 to 5 and Idaho from 39 to 7. Iowa is another remarkable case, it moved from 48 to 20. Although inventive activity is still clustered in the big states that have had a strong record in this respect for a long time (e.g., California, Texas, New York and Massachusetts), specialization into those patents that really matter from a technological perspective has become much more even over states, with some “newcomers” taking the lead.

The results in Tables 3 and 4 suggest that the shift of innovative activity from the Northeast to the West as identified by Hicks et al. (2001) has been even more marked than thought before.

3.4. Accounting for growth in the production of superstar patents

To quantify the extent to which the intensity of the regional inversion of inventive activity in the U.S. has been underestimated in previous studies, we conduct a shift-share analysis. It starts from the perspective that is central to this paper: we should focus on a limited subset of all patented inventions to get a meaningful indication of the truly inventive character of regions. Hence, the left-hand side of the shift-share equation indicates the growth of the number of superstar patents generated in state i , between 1977 and 1998.

If P stands for a number of patents, irrespective of whether these are superstars or not (raw patent counts), while SS indicates a number of superstar patents, the number of superstar patents produced in state i in technology j in a given year t can be expressed as the multiplication of

- the number of patents produced in all N states together, in all J technology categories together ($\sum_{n=1}^N \sum_{j=1}^J P_{nj,t}$, indicated by $USInv$ in Eqs. (1) and (2) below);
- the share of state i in the patents produced in all N states together, in all J technology categories together ($\frac{\sum_{j=1}^J P_{ij,t}}{\sum_{n=1}^N \sum_{j=1}^J P_{nj,t}}$, $StateShare$);
- the share of patents in technology category j in all patents produced in state i (the state’s technological specialization, $\frac{P_{ij,t}}{\sum_{j=1}^J P_{ij,t}}$, $TechSpec$);
- the share of superstar patents in the set of all patents in technology category j produced in state i ($\frac{SS_{ij,t}}{P_{ij,t}}$, $SSSpec$).

Table 2
Numbers of superstar patents per million inhabitants, US States.^a

	1977	rank	1988	rank	1998	rank	Δ77–98	rank
ID	1.1	34	8.7	11	80.6	1	334.4	3
MN	5.5	6	14.8	3	38.7	2	28.7	12
VT	0.7	42	13.6	5	35.7	3	246.2	4
CA	6.7	5	13.1	6	35.3	4	20.2	20
MA	7.1	4	19.2	2	34.9	5	18.7	22
CO	2.8	18	11.1	9	25.4	6	37.8	10
NH	4.2	12	11.6	8	23.8	7	22.1	18
WA	1.8	27	5.7	22	23.6	8	59.0	7
OR	2.6	21	9.1	10	23.1	9	37.8	9
CT	7.2	3	14.1	4	22.3	10	9.9	37
NJ	9.5	2	12.2	7	22.2	11	6.4	47
UT	4.8	7	8.4	12	17.9	12	13.0	29
TX	2.7	20	7.5	13	16.4	13	23.8	17
DE	11.3	1	19.8	1	15.9	14	2.0	50
AZ	4.4	8	7.0	15	15.9	15	12.4	30
NY	4.2	11	6.8	16	13.7	16	10.7	35
MD	4.0	13	6.3	19	12.8	17	10.4	36
MI	3.6	16	6.7	17	12.6	18	12.0	31
WI	4.3	10	5.9	21	11.7	19	8.3	42
IL	3.9	15	7.3	14	11.6	20	9.5	38
NC	1.9	24	4.4	28	11.3	21	23.9	16
NV	3.4	17	4.9	25	11.2	22	10.8	33
IA	0.7	41	2.6	35	10.8	23	70.3	6
PA	4.4	9	4.8	26	10.2	24	6.3	48
OH	3.9	14	5.1	23	9.8	25	7.2	44
RI	1.8	28	6.2	20	9.6	26	21.3	19
NM	1.4	31	4.5	27	8.5	27	24.8	14
IN	2.8	19	3.9	29	7.6	28	8.4	41
GA	0.6	45	3.7	31	6.8	29	51.7	8
FL	2.1	23	5.0	24	6.4	30	9.4	39
VA	1.8	26	3.6	32	6.2	31	11.6	32
OK	2.4	22	3.7	30	5.9	32	6.8	45
KS	1.3	32	2.3	36	5.4	33	15.0	25
DC	1.0	37	6.6	18	5.1	34	19.9	21
TN	1.3	33	1.4	42	5.1	35	14.1	28
NE	0.2	49	1.5	41	4.7	36	100.8	5
MO	1.9	25	2.9	33	4.4	37	6.7	46
WY	0.0	50	0.9	48	4.2	38	∞	1
KY	0.7	43	1.4	43	4.1	39	24.8	15
ME	0.6	44	1.6	39	4.1	40	27.1	13
SC	1.4	30	2.7	34	3.7	41	7.5	43
MT	0.4	47	0.5	50	3.7	42	36.2	11
SD	0.0	51	0.9	47	3.7	43	∞	2
ND	1.0	36	1.2	45	3.4	44	10.8	34
AL	0.8	39	1.6	40	3.3	45	15.0	26
LA	0.7	40	1.8	38	3.0	46	14.6	27
AK	0.8	38	0.7	49	2.4	47	9.3	40
HI	1.1	35	2.0	37	2.4	48	5.6	49
WV	1.8	29	1.4	44	2.3	49	1.5	51
AR	0.5	46	0.3	51	2.0	50	16.0	23
MS	0.4	48	1.2	46	1.8	51	15.9	24

^a Aggregated over technology categories, 3-year moving averages (with indicated years as midpoints); Ordered by number of superstar patents per million inhabitants in 1998. 1977–1998 growth expressed in average annual percentage growth rates. A list of the two-letter state codes is included in [Appendix B](#).

Table 3
Share of superstar patents in all patents, US Census Regions.^a

	1977	1988	1998
West	0.034	0.066	0.078
Northeast	0.028	0.049	0.058
South	0.023	0.041	0.055
Midwest	0.024	0.047	0.056

^a Aggregated over technology categories, 3-year moving averages.

Table 4
Share of superstar patents in all patents, US States.^a

	1977	1988	1998	Δ77–98	R
MN	0.036	0.073	0.088	0.052	5
WA	0.019	0.042	0.087	0.069	1
NV	0.041	0.069	0.086	0.045	11
CA	0.038	0.072	0.080	0.042	14
OR	0.028	0.064	0.079	0.051	6
MA	0.031	0.078	0.076	0.045	12
ID	0.018	0.077	0.076	0.059	2
UT	0.053	0.069	0.069	0.016	45
DC	0.021	0.105	0.067	0.046	9
CO	0.020	0.070	0.067	0.047	8
NH	0.028	0.061	0.063	0.034	21
VT	0.009	0.088	0.063	0.053	4
TX	0.026	0.057	0.062	0.037	17
MD	0.030	0.051	0.061	0.031	24
AZ	0.032	0.047	0.060	0.028	29
NJ	0.030	0.049	0.059	0.029	27
NM	0.024	0.050	0.059	0.035	20
NC	0.030	0.049	0.059	0.029	28
ME	0.019	0.028	0.059	0.039	16
IA	0.008	0.027	0.058	0.050	7
WY	0.000	0.027	0.057	0.057	3
NY	0.030	0.044	0.056	0.026	34
HI	0.043	0.053	0.056	0.013	48
GA	0.012	0.057	0.054	0.042	15
VA	0.023	0.045	0.054	0.031	25
CT	0.026	0.048	0.053	0.028	32
OK	0.018	0.035	0.053	0.035	19
IL	0.021	0.045	0.052	0.031	26
KY	0.014	0.029	0.051	0.037	18
AL	0.023	0.031	0.051	0.028	31
FL	0.027	0.054	0.051	0.023	39
NE	0.005	0.028	0.051	0.046	10
WI	0.035	0.039	0.050	0.015	47
OH	0.025	0.034	0.050	0.025	36
KS	0.015	0.032	0.050	0.034	22
TN	0.024	0.021	0.048	0.024	38
RI	0.020	0.058	0.047	0.027	33
SD	0.000	0.026	0.045	0.045	13
MI	0.019	0.035	0.045	0.025	35
PA	0.026	0.031	0.044	0.018	43
MO	0.021	0.035	0.044	0.022	41
IN	0.021	0.032	0.044	0.022	42
AR	0.015	0.009	0.043	0.028	30
MT	0.008	0.011	0.042	0.034	23
LA	0.017	0.027	0.040	0.024	37
DE	0.030	0.045	0.040	0.010	50
MS	0.024	0.037	0.040	0.016	46
AK	0.017	0.019	0.039	0.023	40
SC	0.023	0.036	0.039	0.016	44
WV	0.025	0.020	0.037	0.012	49
ND	0.034	0.021	0.034	0.000	51

^a Aggregated over technology categories, 3-year moving averages; Ordered by share of superstar patents in 1998. Δ77–98: Change of share of superstar patents in total patents between 1977 and 1998, in percentage points. R: Rank of state in terms of Δ77–98. A list of the two-letter state codes is included in [Appendix B](#).

The number of superstar patents in technology class j as produced in state i in year t can now be written as¹⁴

$$SS_{ij,t} = \left(\sum_{n=1}^N \sum_{j=1}^J P_{nj,t} \right) \left(\frac{\sum_{j=1}^J P_{ij,t}}{\sum_{n=1}^N \sum_{j=1}^J P_{nj,t}} \right) \left(\frac{P_{ij,t}}{\sum_{j=1}^J P_{ij,t}} \right) \left(\frac{SS_{ij,t}}{P_{ij,t}} \right)$$

If we sum over technology classes, the multiplication introduced above leads to the aggregate number of superstar patents produced in state i . This can be expressed in mathematical terms as:

¹⁴ The numerator of the rightmost factor is the only remaining expression, as the numerator of any of the other factors cancels out against the denominator of the next factor.

$$\sum_{j=1}^J SS_{ij,t} = \sum_{j=1}^J \left[\left(\sum_{n=1}^N \sum_{j=1}^J P_{nj,t} \right) \left(\frac{\sum_{j=1}^J P_{ij,t}}{\sum_{j=1}^N \sum_{j=1}^J P_{nj,t}} \right) \left(\frac{P_{ij,t}}{\sum_{j=1}^J P_{ij,t}} \right) \left(\frac{SS_{ij,t}}{P_{ij,t}} \right) \right] = \sum_j [USInv_i \cdot StateShare_{i,t} \cdot TechSpec_{ij,t} \cdot SSSpec_{ij,t}] \quad (1)$$

Each of the four factors in the bottom part of Eq. (1) is equal to the corresponding expression in parentheses in the right-hand side of the upper part of the equation. Evaluating Eq. (1) at time $t + 1$ and dividing the result by Eq. (1) itself, allows us to write for the aggregate growth of production of superstar patents in state i between year t and year $t + 1$ (in ratio form). Note that the elements related to *TechSpec* and *SSSpec* in the first and second factors and those related to *USInv* and *StateShare* in the third and fourth factor are equal to one. They have been included in the present form to emphasize the structure of the shift-share decomposition framework:

$$\begin{aligned}
\frac{\sum_j SS_{ij,t+1}}{\sum_j SS_{ij,t}} &= \sum_j \frac{[USInv_{t+1} \cdot StateShare_{i,t+1} \cdot TechSpec_{ij,t+1} \cdot SS_{Specij,t+1}]}{[USInv_t \cdot StateShare_{i,t} \cdot TechSpec_{ij,t} \cdot SS_{Specij,t}]} \\
&= \left[\frac{USInv_{t+1}}{USInv_t} \cdot \frac{StateShare_{i,t+1}}{StateShare_{i,t}} \cdot \frac{\sum_j (TechSpec_{ij,t+1} \cdot SS_{Specij,t+1})}{\sum_j (TechSpec_{ij,t} \cdot SS_{Specij,t})} \right] \\
&\quad \cdot \left[\frac{USInv_t}{USInv_t} \cdot \frac{StateShare_{i,t+1}}{StateShare_{i,t}} \cdot \frac{\sum_j (TechSpec_{ij,t} \cdot SS_{Specij,t})}{\sum_j (TechSpec_{ij,t} \cdot SS_{Specij,t})} \right] \\
&\quad \cdot \left[\frac{USInv_t}{USInv_t} \cdot \frac{StateShare_{i,t}}{StateShare_{i,t}} \cdot \frac{\sum_j (TechSpec_{ij,t+1} \cdot SS_{Specij,t})}{\sum_j (TechSpec_{ij,t} \cdot SS_{Specij,t})} \right] \\
&\quad \cdot \left[\frac{USInv_t}{USInv_t} \cdot \frac{StateShare_{i,t}}{StateShare_{i,t}} \cdot \frac{\sum_j (TechSpec_{ij,t} \cdot SS_{Specij,t+1})}{\sum_j (TechSpec_{ij,t} \cdot SS_{Specij,t})} \right]
\end{aligned}$$

The first factor (which we label $\Delta USInv$) indicates the extent to which the number of superstar patents in state i would have increased if only the total number of patents in the U.S. would have grown as it actually did between t and $t+1$, and everything else would have remained as in t . The second factor (henceforth $\Delta StateShare$) gives the part of the growth (again as a ratio) in superstar patents production in state i that can be attributed to the change in its share in all patents granted to US inventors. The third factor ($\Delta TechSpec$) quantifies the part of the total change that can be attributed to shifts in inventive activity towards or from technological categories in which state i had high superstar to all patents ratios in t already. Fig. 1, for example, showed that these ratios are (nationwide) higher for Medical instruments than for Drugs, which implies that an increase in the share of regional patents in Medical instruments at the expense of the share in regional patents in Drugs will (ceteris paribus) lead to an increase in the region's number of superstar patents. $\Delta SSSpec$ (the fourth factor in Eq. (2)) isolates the part of the total change in i 's production of superstar patents that can be attributed to changes in the superstar patents to all patents ratios (like presented at the aggregate level in Table 4) for all technology categories. Finally, a factor that incorporates all interactions of changes in $USInv$, $StateShare$, $TechSpec$ and $SSSpec$ is included. Such an interaction emerges if, for example, a state starts patenting more in a technology category in which its specialization in the production of superstar patents is increasing. In this case, both changes in $TechSpec$ and $SSSpec$ play a role, while either of the two is held constant at the levels of period t in the third and fourth factor in Eq. (2).¹⁵ Given the presence of changes in $SSSpec$, this dynamic interaction factor also contributes to a difference in the estimated intensity of the regional inversion between our analysis and the previous literature as discussed

¹⁵ We could have proposed a shift-share equation without an interaction factor, by choosing the weights in the other factors differently (see [Dietzenbacher and Los, 1998](#); for a discussion of exhaustive vs. non-exhaustive decompositions). In the present context, the interaction factor provides useful information regarding the positive or negative effects of changing specialization patterns.

in the introduction.

The shift-share results obtained by Eq. (2) are presented in Table 5. The column titled *Ratio7798* shows the ratios between the numbers of superstar patents obtained in 1998 and in 1977 by a state (or a group of states), or, in other words, the left-hand side of Eq. (2). We applied the equation to the growth in the production of superstar patents in states, but also at a geographically more aggregated level. The results in the column $\Delta StateShare$ reflect the traditional wisdom regarding the regional inversion of inventive activity in the US. If only the shares of states in the total number of US patents would have changed and everything else had remained unchanged, the number of superstar patents in the Western states taken together would have increased by 66%, while those in the Northeast and the Midwest would have decreased by 28% and 24%, respectively. From the perspective of our study, however, the results in the columns $\Delta TechSpec$, $\Delta SSSSpec$ and $\Delta Interact$ are the most important ones. These indicate the consequences of changes in technological specialization, changes in the focus on producing superstar patents and dynamic interactions regarding these effects on the inventive performance of states if only really important inventions are considered. The values in the rightmost column ($\Delta RawCountsBias$) are obtained by multiplying the values in the preceding three columns and indicate the extent to which states and regions have changed in their relative inventive performance above and beyond what is found when disregarding the heterogeneity in the importance of patents.

The ratios in *ΔRawCountsBias* show that changes regarding the difference between superstar patents and regular patents have strongly contributed to the “true” regional inversion of inventive activity. For all Western states taken together, changes related to the production of superstar patents in specific taken in isolation (i.e., all other things equal) would have increased the number of superstar patents by 174%, while the corresponding growth rates for all Northeastern states taken together would only have been 103% (for the Midwest, an intermediate value is found). If just these effects are considered, the West would almost have surpassed the Northeast regarding the number of superstar patents in 1998 (534 for the West, 541 for the Northeast).

Focusing on the three constituent parts of $\Delta RawCountsBias$, we find that the dynamic interaction effects did not contribute to these findings. Actually, these effects had a more positive effect on Midwest superstar invention than on superstar invention in the West. $\Delta TechSpec$ and $\Delta SSSpec$ both contributed substantially to the inversion. Shifts towards activity regarding technological categories with a high share of superstar patenting in 1977 alone would have led to an increase in the number of Western superstar patents of 65%, while this growth rate would have been only 34% and 25% in the Northeast and Midwest, respectively. We find somewhat smaller, but still sizable, differences regarding the effects of specializing in superstar patenting within categories. In the West, such changes account for an increase of 56%, whereas the corresponding figures for the Northeast and Midwest are 40% and 36%, respectively.

Zooming in to the level of states, the results show that the remarkable performance of Western states cannot be described by a common pattern. This can be illustrated by focusing on Idaho and Washington, the two Western states the inventive output of which grew fastest over the period considered. Idaho (which started to patent much more anyway, reflected in a value exceeding 10 for $\Delta StateShare$) is characterized by a very low level of $\Delta TechSpec$, 0.07. This is caused by the fact that Idaho specialized strongly in the production of very important semiconductor patents, a technology category in which it did

Table 5
Sources of growth in the production of superstar patents, selected US States^a and Census regions.

	SS_{77}	$Ratio_{9877}$	$\Delta USInv$	$\Delta StateShare$	$\Delta TechSpec$	$\Delta SSSpec$	$\Delta Interact$	$\Delta RawCountsBias$
West	195.0	8.82	2.26	1.66	1.65	1.56	1.06	2.74
ID	1.0	99.21	2.26	10.57	0.07	1.68	35.99	4.14
WA	6.7	20.10	2.26	1.87	1.99	2.23	1.07	4.75
CO	7.7	13.15	2.26	1.67	1.56	1.32	1.68	3.47
OR	6.3	11.98	2.26	1.91	2.38	1.87	0.62	2.76
MT	0.3	9.83	2.26	0.93	1.67	1.54	1.82	4.69
NM	1.7	8.86	2.26	1.56	1.89	1.34	1.00	2.51
NV	2.3	8.37	2.26	1.76	2.31	1.96	0.47	2.10
CA	150.7	7.66	2.26	1.58	1.65	1.56	0.83	2.14
AZ	10.7	6.94	2.26	1.61	1.31	1.28	1.13	1.90
UT	6.3	5.93	2.26	1.90	1.26	1.45	0.76	1.38
AK	0.3	4.52	2.26	1.00	0.53	1.49	2.52	1.99
HI	1.0	2.81	2.26	0.95	0.72	1.13	1.60	1.31
WY	0.0	*	2.26	0.81	0	0	0	*
Northeast	266.7	3.38	2.26	0.72	1.34	1.40	1.08	2.03
VT	0.3	63.36	2.26	3.53	1.34	1.42	4.19	7.94
NH	3.7	7.70	2.26	1.53	1.48	1.54	0.97	2.22
ME	0.7	7.59	2.26	0.97	0.36	1.84	5.25	3.45
RI	1.7	5.69	2.26	1.04	1.09	1.07	2.08	2.42
MA	40.7	5.27	2.26	0.95	1.59	1.44	1.07	2.46
NY	75.3	3.31	2.26	0.77	1.40	1.36	1.00	1.89
CT	22.3	3.27	2.26	0.68	1.08	1.39	1.41	2.12
NJ	70.0	2.57	2.26	0.57	1.25	1.44	1.10	2.00
PA	52.0	2.36	2.26	0.62	1.29	1.37	0.96	1.69
Midwest	190.4	3.90	2.26	0.76	1.25	1.36	1.36	2.32
NE	0.3	23.64	2.26	0.97	0.39	1.18	23.31	10.73
IA	2.0	15.50	2.26	0.95	0.79	1.60	5.67	7.17
MN	22.0	8.32	2.26	1.50	1.13	1.62	1.33	2.45
KS	3.0	4.73	2.26	0.68	1.59	1.49	1.30	3.09
MI	32.7	3.78	2.26	0.72	1.35	1.25	1.36	2.31
ND	0.7	3.21	2.26	1.06	1.11	1.18	1.02	1.34
IL	44.0	3.17	2.26	0.57	1.28	1.31	1.47	2.47
WI	19.7	3.10	2.26	0.98	1.24	1.34	0.84	1.40
IN	15.0	3.01	2.26	0.64	1.28	1.33	1.20	2.06
MO	9.0	2.69	2.26	0.58	0.97	1.24	1.71	2.06
OH	42.0	2.62	2.26	0.58	1.24	1.41	1.14	2.00
SD	0.0	*	2.26	1.15	0	0	0	*

* The numbers in the column $Ratio_{9877}$ represent the left-hand side of Eq. (2), the number of superstar patents in 1998 divided by the number of superstar patents in 1977 (3-year moving averages). If SS_{77} equals zero, this ratio is not defined and indicated by an asterisk. The values in cells in the column $\Delta RawCountsBias$ are obtained by multiplying the results in the corresponding cells in the columns $\Delta TechSpec$, $\Delta SSSpec$ and $\Delta Interact$. A list of the two-letter state codes is included in Appendix B.

not have a high share of superstar patents back in 1977.¹⁶ Washington, in contrast, scores among the top-states with regard to $\Delta TechSpec$ (1.99). In the end of the 1990s, this state was specialized in breakthrough innovations in computer hardware, software and peripherals, categories in which it generated superstar patents in the late 1970s already: while Idaho's inventive performance is mainly due to starting new lines of inventive activity, Washington's success seems to be largely the result of fruitfully exploiting existing technological capabilities. Washington did not only specialize in producing inventions in which its share of superstar patents were already high. As the results for $\Delta SSSpec$ reveal, Washington also managed to increase the shares of superstar patents in its sets of all patents to a degree not equaled by any other state. Idaho's performance in this respect is much less remarkable. Idaho (like, in other parts of the U.S., Nebraska, and Iowa and Maine to a lesser extent) scores very high on $\Delta Interact$, which indicates that it managed to raise its shares of superstar patents in all patents in those technology categories in which it also specialized over time. Apparently, it managed to enter a virtual cycle, in which extraordinary performance led to an increasing specialization, which caused further improvements in the extent to which its inventions belonged to the superstar category. Washington's value for $\Delta Interact$, on the other hand, is very close to the average for the Western region.

¹⁶ Agrawal et al. (2010) explicitly label Idaho's capital city (Boise) as a "company town", dominated by one of the world's leading firms in semiconductors (Micron).

4. Conclusions

Ceh (2001), Hicks et al. (2001), Johnson and Brown (2004) and Ó hUallacháin and Leslie (2005) found that the location of inventive activity in the U.S. has markedly moved over the 1970s, 1980s and 1990s. The Northeast and Midwest lost out (in relative terms), while the importance of Western states grew. These authors used simple patent counts as indicators of inventive success. Recently, the large heterogeneity regarding the importance of patents has led to doubts regarding the usefulness of such "raw" patent counts as inventiveness indicators. In this paper, we proposed an alternative, reconsidering the extent to which U.S. inventive activity underwent the above-mentioned regional inversion indeed. We focused on subsets of very highly-cited patents, which we label superstar patents. The well-established fact that citations frequencies to these patents are governed by statistical laws that are different from those applicable to citations to all patents strongly suggests that they play a key role in the dynamics of technology.

Our findings suggest that the existing literature has even underestimated the regional inversion. The production of superstar patents moved to Western states at a faster pace than suggested when considering all patents as a homogeneous set. To quantify the underestimation, we developed an accounting framework along the lines of shift-share analysis. This shift-share analysis also shows that Western states did not improve their superstar patent production performance following a common pattern. A comparison between Washington and

Idaho illustrates this. While Washington improved its performance on the basis of existing capabilities in computer-related technologies, the marked improvement in performance of Idaho was due to successfully engaging in types of research (mainly related to semiconductors) for which it was initially not recognized.

The analysis in this study has an accounting nature. It quantifies important trends, but does not attempt to explain these. In view of the fact that the regional inversion of U.S. inventive activity has been even stronger than thought before, such explanations are needed. Some clues have been developed. [Kerr \(2010\)](#), for example, found that the generation of a limited number of breakthrough inventions in a region often leads to a subsequent shift of inventive activity in the corresponding technologies to such a region. This effect is stronger for technologies that rely heavily on the inventive input of immigrants, who are generally more mobile than other inventors. A different approach starts from the idea that inventions are the products of processes in which

existing pieces of knowledge are successfully recombined. [Castaldi et al. \(2015\)](#) found evidence that states with high shares of superstar patents in all patents generally had patent portfolios characterized by a high degree of unrelated variety (i.e., knowledge contained in patents from rather distant technology categories), while raw numbers of patents produced are positively affected by a high degree of related variety (i.e., knowledge contained in patents in closely related field of technology). It would be interesting to investigate in future research whether the increased specialization in the generation of superstar patents in Western states as observed in this paper can be explained by theories building on the recombinant nature of knowledge indeed.

Acknowledgment

We would like to thank Koen Frenken for useful comments on a previous draft of this paper.

Appendix A. Technology Categorization

Nr.	Description	Sub-category code in Hall et al. (2002)
1.	Agriculture, food, textiles	11
2.	Coating	12
3.	Gas	13
4.	Organic compounds	14
5.	Resins	15
6.	Communications	21
7.	Computer hardware and software	22
8.	Computer peripherals	23
9.	Information storage	24
10.	Drugs	31
11.	Surgery and medical instruments	32
12.	Biotechnology	33
13.	Electrical devices	41
14.	Electrical lighting	42
15.	Measuring and testing	43
16.	Nuclear and X-rays	44
17.	Power systems	45
18.	Semiconductor devices	46
19.	Materials processing and handling	51
20.	Metal working	52
21.	Motors, engines and parts	53
22.	Optics	54
23.	Transportation	55
24.	Agriculture, husbandry and food	61
25.	Amusement devices	62
26.	Apparel and textile	63
27.	Earth working and wells	64
28.	Furniture, house fixtures	65
29.	Heating	66
30.	Pipes and joints	67
31.	Receptacles	68

Appendix B. US States

Code	State
AK	Alaska
AL	Alabama
AR	Arkansas
AZ	Arizona
CA	California
CO	Colorado
CT	Connecticut
DC	Washington DC
DE	Delaware
FL	Florida
GA	Georgia
HI	Hawaii
IA	Iowa
ID	Idaho
IL	Illinois
IN	Indiana
KS	Kansas
KY	Kentucky
LA	Louisiana
MA	Massachusetts
MD	Maryland
ME	Maine
MI	Michigan
MN	Minnesota
MO	Missouri
MS	Mississippi
MT	Montana
NC	North Carolina
ND	North Dakota
NE	Nebraska
NH	New Hampshire
NJ	New Jersey
NM	New Mexico
NV	Nevada
NY	New York
OH	Ohio
OK	Oklahoma
OR	Oregon
PA	Pennsylvania
RI	Rhode Island
SC	South Carolina
SD	South Dakota
TN	Tennessee
TX	Texas
UT	Utah
VA	Virginia
VT	Vermont
WA	Washington
WI	Wisconsin
WV	West Virginia
WY	Wyoming

References

- Agrawal, A., Cockburn, I., Rosell, C., 2010. Not invented here? Innovation in company towns. *J. Urban Econ.* 67, 78–89.
- Akkermans, D.H.M., Castaldi, C., Los, B., 2009. Do 'liberal market economies' really innovate more radically than 'coordinated market economies'? *Res. Policy* 38, 181–191.
- Anselin, L., Varga, A., Acs, Z., 1997. Local geographic spillovers between university research and high technology innovations. *J. Urban Econ.* 42, 422–448.
- Audretsch, D.B., Feldman, M.P., 1996. R & D spillovers and the geography of innovation and production. *Am. Econ. Rev.* 86, 630–640.
- Balland, P.A., Rigby, D., 2017. The geography of complex knowledge. *Econ. Geogr.* 93 (1), 21–23.
- Breschi, S., Lissoni, F., 2001. Knowledge spillovers and local innovation systems. *Ind. Corp. Change* 10, 975–1005.
- Castaldi, C., Los, B., 2015. The Identification of Breakthrough Inventions Using Tail Estimators; Counting Superstar Patents, ECIS Working Paper 2015–5. Eindhoven

- University of Technology. <http://ecis.ieis.tue.nl/node/436>.
- Castaldi, C., Frenken, K., Los, B., 2015. Related variety, unrelated variety and technological breakthroughs: an analysis of US state-level patenting. *Reg. Stud.* 49, 767–781.
- Ceh, B., 2001. Regional innovation potential in the United States: evidence of spatial transformation. *Papers Reg. Sci.* 80, 297–316.
- Clark, J., Huang, H.-I., Walsh, J.P., 2010. A typology of ‘innovation districts’: what it means for regional resilience. *Cambridge J. Reg. Econ. Soc.* 3, 121–137.
- Clauset, A., Shalizi, C.R., Newman, M.E.J., 2009. Power-law distributions in empirical data. *SIAM Rev.* 51, 661–703.
- Dahlin, K.B., Behrens, D.M., 2005. When is an invention really radical? Defining and measuring technological radicalness. *Res. Policy* 34, 717–737.
- Dietzenbacher, E., Los, B., 1998. Structural decomposition analysis: sense and sensitivity. *Econ. Syst. Res.* 10, 307–324.
- Dosi, G., 1982. Technological paradigms and technological trajectories. *Res. Policy* 11, 147–162.
- Dunning, J.H., Narula, R., 1995. The R & D activities of foreign firms in the United States. *Int. Stud. Manage. Organ.* 25, 39–73.
- Ejermo, O., 2009. Regional innovation measured by patent data—does quality matter? *Ind. Innov.* 16, 141–165.
- Hall, B.H., Trajtenberg, M., 2006. Uncovering general purpose technologies with patent data. In: Antonelli, C., Foray, D., Hall, B.H., Steinmueller, W.E. (Eds.), *Frontiers in the Economics of Innovation; Essays in Honor of Paul David*. Edward Elgar, Cheltenham, UK, pp. 389–426.
- Hall, B.H., Jaffe, A.B., Trajtenberg, M., 2002. The NBER patent-citations data file: lessons, insights, and methodological tools. In: Jaffe, A.B., Trajtenberg, M. (Eds.), *Patents, Citations & Innovations*. MIT Press, Cambridge, MA, pp. 403–459.
- Harhoff, D., Narin, F., Scherer, F.M., Vopel, K., 1999. Citation frequency and the value of the value of patented inventions. *Rev. Econ. Stat.* 81, 511–515.
- Hicks, D., Breitman, T., Olivastro, D., Hamilton, K., 2001. The changing composition of innovative activity in the US—a portrait based on patent analysis. *Res. Policy* 30, 681–703.
- Jaffe, A.B., Lerner, J., 2004. *Innovation and Its Discontents*. Princeton University Press, Princeton, NJ.
- Johnson, D.K.N., Brown, A., 2004. How the west has won: regional and industrial inversion in U.S. patent activity. *Econ. Geogr.* 80, 241–260.
- Kerr, W.R., 2010. Breakthrough inventions and migrating clusters of innovation. *J. Urban Econ.* 67, 46–60.
- Li, G.-C., Lai, R., D’Amour, A., Doolin, D.M., Sun, Y., Torvik, V.I., Yu, A.Z., Fleming, L., 2014. Disambiguation and co-authorship networks of the U.S. patent inventor database, 1975–2010. *Res. Policy* 43, 941–955.
- Lux, T., 2001. The limiting extremal behaviour of speculative returns: an analysis of intraday data from the frankfurt stock exchange. *Appl. Financ. Econ.* 11, 299–315.
- Ó hUallacháin, B., Leslie, T.F., 2005. Spatial convergence and spillovers in American invention. *Ann. Assoc. Am. Geogr.* 95, 866–886.
- Phene, A., Fladmoe-Linquist, K., Marsh, L., 2006. Breakthrough innovations in the U.S. biotechnology industry: the effects of technological space and geographic origin. *Strateg. Manage. J.* 27, 369–388.
- Rysman, M., Simcoe, T., 2008. Patents and the performance of voluntary standard-setting organizations. *Manage. Sci.* 54, 1920–1934.
- Sanditov, B., 2006. *Essays on Social Learning and Imitation*, Unpublished PhD Thesis. Maastricht University, The Netherlands.
- Schoenmakers, W., Duysters, G., 2010. The technological origins of radical inventions. *Res. Policy* 39, 1051–1059.
- Silverberg, G., Verspagen, B., 2007. The size distribution of innovations revisited: an application of extreme value statistics to citation and value measures of patent significance. *J. Econom.* 139, 318–339.
- Singh, J., Fleming, L., 2010. Lone inventors as sources of breakthroughs: myth or reality. *Manage. Sci.* 56, 41–56.
- Trajtenberg, M., 1990. A penny for your quotes: patent citations and the value of innovations. *RAND J. Econ.* 21, 172–187.
- US Census Bureau (1995), *Population Estimates Program: Intercensal Estimates of the Total Resident Population of States, 1970–1980*, Washington, DC.
- US Census Bureau (1996), *Population Estimates Program: Intercensal Estimates of the Total Resident Population of States, 1980–1990*, Washington, DC.
- US Census Bureau (1999), *Population Estimates Program: State Population Estimates, Annual Time Series, July 1, 1990 to July 1, 1999*, Washington, DC.
- USPTO (2010), *Manual of Patent Examining Procedure 605-04(f)*, United States Patent and Trademark Office, Alexandria VI (website accessed on August 18, 2011).
- Usai, S., 2011. The geography of inventive activity in OECD regions. *Reg. Stud.* 45, 711–731.
- Verspagen, B., Schoenmakers, W., 2004. The spatial dimension of patenting by multinational firms in Europe. *J. Econ. Geogr.* 4, 23–42.