Paper to be presented at

DRUID15, Rome, June 15-17, 2015

(Coorganized with LUISS)

# Impact of patent-paper pairs on the quality of biotechnology and nanotechnology patents in Quebec


**Catherine Beaudry**
Polytechnique Montreal
Mathematics and industrial engineering
catherine.beaudry@polymtl.ca

**Narjes Naserbakht**
Polytechnique Montreal
Mathematics and industrial engineering
n.naserbakht@gmail.com



## Abstract

This article aims to assess whether patents issued from a patent-paper pair are of a higher quality than other patents. This research analyzes the impact of university-industry linkages on the quality of biotechnology and nanotechnology patents. This article measures quality first as the number of forward citations, second as the number of claims and third as an originality index combining forward citations, number of claims and a Herfindahl index of backward citations. Our results show that the impact of patent-paper pairs on the quality of patents is negative, regardless of the indicator chosen.



Jelcodes:O31,O34

# Impact of patent-paper pairs on the quality of biotechnology and nanotechnology patents in Quebec

**Abstract**

This article aims to assess whether patents issued from a patent-paper pair are of a higher quality than other patents. This research analyzes the impact of university-industry linkages on the quality of biotechnology and nanotechnology patents. This article measures quality first as the number of forward citations, second as the number of claims and third as an originality index combining forward citations, number of claims and a Herfindahl index of backward citations. Our results show that the impact of patent-paper pairs on the quality of patents is negative, regardless of the indicator chosen.

**Keywords**: Patent quality; citations; claims; research funding; biotechnology and nanotechnology

## 1 Introduction

In science-based industries, the product development process is based on the collaborative network of university scientists and other public research institutes with industrial scholars (Hsu and Bernstein, 1997; George et al., 2002; Oliver, 2004). Biotechnology and nanotechnology are science-based domains that drive other industries forward, hence, the critical role of collaboration between universities and biotechnology and nanotechnology organizations (Niosi and Reid, 2007; Oliver, 2004; Baba et al., 2009). In the biotechnology sector, collaborations between universities and other players, including firms and research institutes, are deemed essential for actors to compete and survive in this competitive area (Bowie, 1994; Peters et al., 1998; Oliver, 2004). Other researchers highlight the strategic role of university-industry alliances (Bowie, 1994; Liebeskind et al., 1996; Oliver and Liebeskind, 1998) and support the fact that this linkage significantly increases the number of patents as an essential source of market value (Zahra, 1996). Patents are indeed an important requirement of the product commercialization process (Almeida, 1996; Grant and Baden-Fuller, 1995). Biotechnology is considered as one of the few domains where the ideas and knowledge generated in universities and research labs can be transferred to firms quickly (Baba et al., 2009; Cohen et al., 2002). Many start-up biotechnology firms use intellectual property as one of their key assets to protect the rights over their idea generation (Gans et al., 2002; Arora and Merges, 2004; Giuri et al., 2007) and only one third of the patents generated by individual inventors have no collaborative ties with other partners, hence highlighting the prevalence of collaboration in this domain and in the scientific world in general (Wagner-Doebler, 2001).

The impacts of university-industry collaboration on scientific and technological production have been measured by several factors. Some researchers examined the role of university-industry collaboration on patent licensing, article publication, production performance, and productivity of R&D products (Hausman et al., 1984; Branstetter and Nakamura, 2003; Motohashi, 2005). Motohashi (2005) showed that small start-up firms achieve greater productivity from this collaborative network. Other studies assumed that the two networks of scientists and inventors have distinctive social structures, although in some aspects, their activities overlap (Dasgupta and David, 1994; Murray, 2002).

Traditional bibliometric methods are often used to assess university-industry linkages (Henderson and Cockburn, 1994; Jaffe, 1998; Podolny and Stuart, 1995; Zucker et al., 1998) and to measure how university-industry ties affect overall performance, specifically, when the technology is science-based (Henderson and Cockburn, 1994; Zucker et al., 1998). Murray (2002) identified three traditional university-industry collaborative networks: (i) the citation of papers in patents; (ii) the publication of papers by firms and industrial scientists; (iii) the co-publication of papers by academic scientists and industrial inventors. She proposed a novel concept called "patent-paper pairs" to try to understand which aspects of science and technology are linked together, and simultaneously to identify which firms or scientists have a significant impact on science and technology. The patent-paper pairs concept is based on the fact that both scientists and inventors contribute to idea generation through publications and patenting. This methodology tries to identify which patents and papers are paired and link science and technology (Ducor, 2000; Murray, 2002; Leopold et al., 2004; Bassecoulard and Zitt, 2004). A number of authors have used accurate content analysis to measure the similarity between the patents and papers in order to identify the pairs (Jaffe, 1998; Lissoni and Montobbio, 2006; Lubango and Pouris, 2010; Murray, 2002; Podolny and Stuart, 1995).

Inspired by the scientometric literature and initiated by Trajtenberg (1990), a number of scholars have explored various metrics to find good proxy measures of innovation quality (Jaffe et al., 1993; Narin and Olivastro, 1992; Narin et al., 1997; Lanjouw and Schankerman, 2001). There is ample evidence in the science-technology linkages literature to the effect that amongst appropriate patent quality indicators, patent citations are considered a very good proxy for innovation performance (Hall et al., 2005; Harhoff et al., 1999; Jaffe and Trajtenberg, 2002). Bonaccorsi and Thoma (2007) for example selected multiple indicators used by Hall and Trajtenberg (2004), Henderson et al. (1998), and Lanjouw and Schankerman (2004) to measure innovation productivity. The authors categorized the patent inventors in three groups: the first group is composed of the inventors of the patents for which all inventors are also named author(s) of scientific publication(s) in Nano Science and Technology (NST); the second group consists of inventors who have no scientific publication in NST; and the third group is composed of inventors of the patents for which at least one of inventors has published article(s) in this field. They analyzed the impact of each category of inventors on the quality of patents and found that the quality of patents developed by the inventor community alone is lower than the patent quality of the author-inventor network (Bonaccorsi and Thoma, 2007).

By surveying the literature on patent-paper pairs and on its impact on the performance of the innovation process, we found that less attention is given to trying to assess which related factors in addition to patent-paper pairs influence the quality of patents. This paper therefore aims to answer the following question: Are patents that are linked to papers of greater quality compared to other patents that are not linked to such publications? Less central to our analysis but nonetheless related, we also aim to address whether patents assigned to universities of greater quality, or whether better-positioned and more centralized scientists have a better patenting performance, or if better-funded scientists generate better quality patents, or if better-cited scientists produce higher quality patents. Our results show that patents issued from patent-paper pairs are of a lesser quality compared to other patents to which academic inventors have contributed. In addition, the type of assignee does not seem to have an influence on patent quality.

The remainder of this paper is structured as follows: Section 2 describes the conceptual framework of this study including a literature review on patent-paper pairs that justifies our proposed hypotheses; Section 3 describes the data and methodology; in Section 4, the descriptive statistics of the data are presented; Section 5 summarizes the results and finally Section 6 concludes.

## 2 Conceptual framework

Analyzing the impact of patent-paper pairs on the quality of patents requires finding the contribution of the interaction between authorship and inventorship on innovation performance, measured for instance by the number of citations and other measures (Hagedoorn and Cloodt, 2003). Haeussler and Sauermann (2013) present evidence to highlight the social impacts of authorship and inventorship on innovation performance. On an individual basis, a number of incitation to patent has been identified. Scientists share their ideas through publications and can receive the recognition from other researchers through citation of their scientific articles but also through benefits such as increased salaries and consulting opportunities (Cole and Cole, 1967; Haeussler and Sauermann, 2013; Haeussler et al., 2011; Merton, 1973; Stephan, 2012). Moreover, besides the revenues they can generate through licensing their innovations and ideas, academic inventors can receive peer recognition in the professional network from patenting (Dasgupta and David, 1987).

Patent-paper pairs are the patents and papers that are issued from the same project (Murray, 2002). Lubango and Pouris (2010) studied 70 patents from USPTO, EPO, and

WIPO and found 58 patents (82%) developed by scientists from South African universities that are linked to articles. They showed that authors have a propensity to develop patents and publish articles at the same time. Magerman et al. (2011) studied the impact of patent-paper pairs on the citation flows and demonstrated that there is no significant difference between the forward citations of patents belonging to patent-paper pairs and the patents which are not associated to these pairs. In contrast, their results show that the publication part of these pairs received significantly more citations compared with the publications without a patent counterpart.

Murray and Stern (2007) studied on the impact of intellectual property on knowledge diffusion putting patent-paper pairs at the heart of their research strategy: inventors tend to both publish articles and patent, in particularl, half of the publications in Nature Biotechnology are linked to patents within five years of their publications. They categorised patent-paper pairs in two groups: pre-granted period with no formal Intellectual Property Right (IPR) and post-granted period including IPR associated to the time period during which articles are published. Their results showed that the forward citation rate of publications associated to patent-paper pairs declines after the patents being granted (see also Heisey and Adelman, 2011; Kang et al., 2009).

Some scholars suggested that IPR offer financial and social benefits for innovative activities (see for instance Hellman, 2007; Kitch, 1977), while others abide to the “anti-common” perspective and argue that IPR has a negative impact on innovations. The debate between these two approaches is related to the question of how IPR affect a researcher’s tendency to generate more knowledge in future scientific activities. Some researchers state that IPR are more akin to “privatizing” knowledge and thus prevent knowledge flows between researchers’ ideas and their exploration (Heller and Eisenberg, 1998; Argyres and Liebeskind, 1998; David, 2001). Murray and Stern (2007) explored whether there is a difference in the citation rate of publications that are patented. According to Heisey and Adelman (2011) the Murray and Stern (2007) approach highlights that intellectual property acquisition has a negative impact on knowledge application by subsequent scientists and the number of publication citations decreased after the patents are granted. With this debate in mind, we therefore propose the following hypothesis that examines a somewhat reverse argument:

***H1***: *Patents invented by at least one inventor who has also authored a scientific article in similar field within a short time frame are of higher quality compared with the patents that have been developed without close links to publications.*

Other factors, have also been found to have an impact on patent performance. For instance, Sterzi (2013) examined the ownership of patents to show that academic patents which are owned by companies receive higher citations in their first years, diminish in the next few years to disappear completely in longer periods (6 years for instance). As a result, academic patents owned by industries but originally assigned to universities are of a remarkably higher quality compared with patents assigned directly to industrial assignees. This is due to the fact that patents assigned to companies mostly target direct commercial benefits in the short term, while patents owned by universities and other public institutions tend to answer scientific questions that have an impact for longer periods (Czarnitzki et al., 2012; Sterzi, 2013). Lissoni et al. (2010) showed that patents owned by industries are of greater quality compared to patents owned by universities. Crespi et al. (2010) examined the quality of patents owned by universities in a few European countries (France, the Netherlands, Spain, Italy, Germany and UK) but could not find any evidence of a relationship between ownership and patent quality. In light of this evidence, our second hypothesis therefore goes as follows:

***H2***: *Patents owned by universities are of a lesser quality compared to those owned by industry.*

## 3 Data and methodology

### *3.1 Data*

Finding potential patent-paper pairs requires integrating data from two sources: the United States Patent and Trademark Office (USPTO) for patents[1] and Elsevier's Scopus

[1] Canadian biotechnology and nanotechnology inventors generally patent in the US in addition, or in lieu of, patenting in Canada (Beaudry and Kananian, 2013). Furthermore, the Canadian Intellectual Property

for papers[2]. The Scopus database includes names of authors, their affiliations, publication date, title and abstract for each scientific article. The USPTO provides information on the names and addresses of inventors, the names and addresses of assignees, application and granting dates, the number of claims, etc. We first extracted all the papers and patents for which at least one author or one inventor has an address or affiliation in Canada using a keyword search (see Beaudry and Kananian, 2013 for details). These two databases were then merged using a roughly unique ID for each individual, i.e. identifying the individuals who have common names in both patents and papers[3]. In our biotechnology and nanotechnology database, there are 2,517 scientists and inventors in Quebec who were involved in patenting and in the publication of articles during the period 1985-2005 in biotechnology or nanotechnology. These individuals were involved in filing 1,222 patents over the period.

Because we are interested in academic patenting, funding may play an important role in the quality of patents. The third source of information necessary for this study is the Quebec University Research Information System (*Système d'Information sur la Recherche Universitaire* – SIRU) provided by the Quebec Ministry of Education. This database provides information about the yearly amounts of contracts and grants obtained by Quebec academics. Because no such database exists for Canada, we thus restrict our sample to the Quebec patent-paper pairs. We calculated yearly public and private funding, grants and contracts, for each individual Quebec university scientist.

### *3.2 Dependent variables*

The number of backward citations to a patent, the number of forward citations, the number of claims, the number of IPC-subclasses, renewal times of patents, and the number of applicants have been used by many scholars as dependent variables (Carpenter et al., 1980; Goetze, 2010; Guellec and van Pottelsberghe de la Potterie, 2000; Hirschey and Richardson, 2004; Narin et al. 1987; Trajtenberg, 1990). For example, the number of

---

Office (CIPO) does not provide consistent addresses for inventors, which adds to the difficulty of disambiguating inventor's names.

[2] Scopus generally links authors with their affiliations, which greatly facilitates matching with the USPTO database and disambiguation of names. Because of the large number of individuals to match for this research, this database was therefore favoured.

[3] Better precision is not necessary prior to the data mining similarity analysis.

forward citations counts the number of times that patents have been cited in subsequent patents during the 5-years after the grant date of patents (Trajtenberg, 1990; Burke and Reitzig, 2007). The number of backward citations count the number of patents referenced in the patent document as cited patents (Narin et al., 1987; Burke and Reitzig, 2007). Henderson et al. (1998) used the Herfindahl index as a measure of concentration of patents. Bonaccorsi and Thoma (2007) constructed a quality index built from different quality factors such as the number of forward citations, the number of backward citations, family size and the number of claims. They integrate 1 minus the Herfindahl index of backward citations as a component of their originality index. A higher value of originality generality shows that patents are less concentrated, and hence more diversified (Henderson et al., 1998).

Our dependent variables are amongst the commonly used proxies for patent quality. Considering the database and information to which we have access, the patent forward citations over five years after the grant year (*NbFCit5*) and number of claims (*NbClaims*) have been selected as two of the proxies for patent quality. Following Bonaccorsi and Thoma (2007), we also add an originality index (*InOriginality*) calculated using the principal components[4] of the number of forward citations, the number of claims, and 1 minus the Herfindahl index of the number of backward citations. Table 10 in appendix describe the principal component analysis (PCA) performed to build our originality index. The analysis clearly points towards the use of a single component, Comp1 (eigenvalue > 1). A complete variable description, including dependent, independent, endogenous and instrumental variables is provided in Table 8 in the appendix.

### *3.3 Patent-paper pairs*

Several researchers employ different yet slightly similar methodologies to construct patent-paper pairs. Murray and Stern (2007) tried to match the articles and papers published in *Nature Biotechnology Journal* by asking experts to find the connection between the matched articles and papers. Thompson et al. (2011) used an "Inventor-based matching" algorithm to extract the patent-paper pairs. Their algorithm is structured around the name of inventors that participate in both patenting and publishing. Two

[4] Principal component analysis with varimax rotation.

assumptions are necessary for this methodology: first, inventors who contribute to publications are considered as the link between science and technology; second, the patent application year is close to the publication date, i.e. within 2 years either side of the patent application date (Thompson et al., 2011). While, Murray and Stern (2007) limited their patent-paper pairs to one patent linked to one publication, Thompson et al. (2011) matched a number of common publications to one patent.

A number of researchers used text-mining tools to find inventor(s) who are also named author(s) in a similar domain (Lissoni and Montobbio, 2008). Lissoni and Montobbio (2006) selected the potential patent-paper pairs as those for which at least one inventor published scientific articles during the period [*t*-2, *t*+2], where *t* corresponds to the application date of the patents. In one of their methods (they compared five), they calculated the cosine similarity between the patent and paper documents to measure the similarity of the content. They identified as patent-paper pairs the top 10% of the potential patent-paper pairs, which corresponds to similarity measures ranging from 0.145 to 0.75. Ducor (2000) however showed that authors are not always matched with the inventors that publish (Haeussler and Sauermann, 2013).

In this study, we use a methodology very similar to that of Magerman et al. (2011) who used content similarity measurement to analyze the similarity of abstracts and titles of both patents and papers where at least one inventor is listed as an author of the publications. In their methodology, all the words of these documents are first indexed, removing evident stop words. Then the vector space is created based on a term matrix that is generated from patent and paper documents. In this matrix, the patent and article documents occupy the rows, and some terms of the abstracts and titles of patents and papers are added as the columns of this term matrix. Then a classic data mining technique is used, Term Frequency and Inverse Document Frequency (TF-IDF), to find the term frequencies of the words in the documents, hence measuring the similarities between the patent and paper documents (Magerman et al., 2011).

In our particular case, we first identified the potential patent-paper pairs by selecting the patents developed by authors-inventors. With the simple constraint that at least one of the inventors published one or many articles in the period [*t*-2, *t*+2], where *t* corresponds to the patent application date, we could find 22,688 potential biotechnology patent-paper pairs and 20,003 potential nanotechnology patent-paper pairs in Canada. We then used the text mining software Rapidminer to calculate the cosine similarity between the paired patent and paper documents. The similarity measures range, in our sample, from 0 to 0.78 for biotechnology, and from 0 to 0.53 for nanotechnology (theoretically similarity measures range from 0 to 1). Similarly to Lissoni and Montobbio (2006), we selected the patent-paper pairs from the top 10 percentile of the similarity measures. This roughly corresponds to a similarity measure of 0.30 for both biotechnology and nanotechnology samples. From this first selection mechanism, we created a dummy variable (*dPPP*) that takes the value 1 if the patent has a paper counterpart (i.e. the patents and papers are similar enough according to the top 10 percentile measure) and 0 otherwise. We then tested whether this threshold yielded significant results in our regression models. Using this 0.30 threshold for the similarity measure, in Canada, we find 249 actual nanotechnology patent-paper pairs and 376 biotechnology patent-paper pairs[5,6].

### *3.4 Other independent variables*

In this study, it is important to stress that all the variables were initially measured at the academic-inventor-year level and then aggregated (averaged) at the patent level for all the academic-inventors. Hence, the variable *AveGrant3* measures the average amount of grants raised by each academic inventor over the three years prior to the patent application, averaged over the Quebec academic inventors named on the patent document[7]. Similarly, *AveCont3* measures the amount of contract funding raised over the past three years, once again averaged over the Quebec academic inventors named on the patent document. Payne and Siow (2003) found a small but positive impact of funding on the number of patents to which researchers have contributed. Separating grants from contracts, Beaudry and Kananian (2013) showed that grants have little or no effect on the

[5] These patent-paper pairs were all checked individually to ensure that no two individuals with the same name were mistakenly associated in patent-paper pairs.
[6] Note that the nanobiotechnology field overlaps both biotechnology and nanotechnology fields, hence the total number of patent-paper pairs (PPP) found is less than the sum of both numbers of PPP.
[7] We have no means by which to evaluate the amount of funding raised by out of Quebec inventors and invested by the assignees. This variable is a control for the capacity of the academic team to raise funds.

number of patent citations, but follow an inverted U-shaped relationship with the number claims, suggesting a substitution effect between grants and contracts. Contracts however have a positive impact on both the number of citations and the number of claims.

Because scientists and inventors do not generally work alone, the team itself may influence the quality of the resulting patent. Beaudry and Schiffauerova (2011) for instance measured the impact of the network characteristics of Canadian Nanotechnology inventors on patent quality measured by the number of claims and found that more central individuals, i.e. good intermediaries, produce higher quality patents. In contrast, Wang et al. (2010) showed that high brokerage (similar to the intermediary position and measured by betweenness centrality) has a negative impact on the patent renewal decision. Using the social network analysis software Pajek we mapped the co-authorship and co-invention networks of scientists and inventors respectively. In these networks, the vertices represent the scientists or inventors and the edges between the vertices correspond to the collaborative links between scientists or inventors that led to articles or patents (Carrington et al., 2005). In the literature, the time window used to study the history of collaboration differs from one study to another. For example, Schilling and Phelps (2007) used three-year windows to map the network of firm collaboration, while Gulati and Gargiulo (1999) considered five-year windows. In this study, we measured different combinations of three- and five-year subnetworks to build the network metrics and chose the three-year subnetworks as they provided the most consistent result.

To characterize the position of researchers within these networks we constructed two indicators: betweenness centrality (Ave*BtwP3*) measures the importance of an inventor as an intermediary in the co-invention[8] network (Freeman, 1979; Nooy et al., 2005) averaged over all academic inventors of a given patent; and cliquishness represents the likelihood that the direct neighbours of researchers are also connected to each other (Nooy et al., 2005) in the co-publication network (*AveCliqA3*) and in the co-invention network (*AveCliqP3*), once again averaged over all researchers of a given patent.

[8] We also calculated the average betweenness centrality of scientists in the co-invention network but as this variable never showed any significance in our results, it was simply eliminated from the models.

The literature generally finds that individuals that are highly productive in terms of technological outputs generally also produce a lot of papers (Balconi and Laboranti, 2006; Breschi et al., 2005; Calderini and Franzoni, 2004; Calderini et al., 2007; Meyer, 2006; Van Looy et al., 2006). The intrinsic quality of the individual is probably what drives this high production and constitutes a latent variable in our analysis. It is therefore important to try to measure the quality of the individual. A first measure in this regard is the average number of citations received by the articles published in the past three years by the academic inventors that contributed to a patent (*AveNbACit3*). Hence, quality on the science side my drive quality on the technology side. A second measure of individual quality is the type of research chair that an academic inventor has held in her career. We define the variable *MaxChair* as an ordinal variable taking the value 0 if the academic inventor never held a chair, the value 1 for an industrial chair, the value 0 for an NSERC or CIHR chair and the value 3 for a Canada research chair.

To take into account the impact of the age of scientists on patent quality, we include the "career" age (*AveAge*) of a scientist as a proxy for real age. Career age corresponds to the age from the period at which an inventor appears in our database either from raising funds, publishing articles or patenting. This control variable expresses the fact that older scientists maybe more creative (Cole and Cole, 1973; Kyvik and Olsen, 2008; Merton, 1973), or *a contrario* that scholars may in contrast do their most important discoveries before the age of 40 (Adams, 1946; Gieryn, 1981; Stern, 1978; Zuckerman, 1977).

Finally, to control for the diversity of inventorship, we introduce the proportion of academic inventors amongst all inventors named in the patent documents (*propAcInv*).

### *3.5 Model Specification*

Because the number of forward citations and the number of claims are count measures, we started with Poisson regression models to analyze the impact of different independent factors such as grants, contracts, number of article citations, age of scientists, etc. Our models however suffered from over dispersion, and thus the preferred method of analysis chosen for this part of the study is the negative binomial regression (nbreg) that relaxes the assumption of the mean variable being equal to the variance. For the originality index models, a simple ordinary least squares (OLS) regression model was used.

Our models potentially suffer from endogeneity problems. There are several possible causes for this endogeneity: first, unobserved heterogeneity may plague the analysis because of a lack of quality of the data. In this research, we tried to clean the data as much as possible to correctly identify individuals and their names. As previously mentioned, this was performed manually but is not immune from human error. Second, we suspect that the number of patent citations depends on the total average amount of contracts received by academic-inventors over three years (*AveCont3*) but that the contracts are also related to the amount of grants raised, which is one of our explanatory variables. We therefore treat the contracts as an endogenous variable. To correct for potential endogeneity in our model, we employ the Two-Stage Residual Inclusion (2SRI) method suggested by Terza et al. (2008). The average amount of contracts raised over 3 years (*AveCont3*) is highly correlated with the three following variables, which will be treated as instruments. The average amount of past contracts received in the same university (*AveCont3U*) accounts for the fact that universities that traditionally collaborate a great deal with industry are probably closer to the so-called third mission of universities. The average amount of grants for equipment and infrastructure received by inventors (*AveGrantEI3*) is related to the sharing of important biotechnology and nanotechnology infrastructure that is often encouraged, if not necessary, to ensure the survival of these laboratories. The number of innovation loops (*NbLoopPast*) as coined by Beaudry and Kananian (2013) measure the number of times that academic-inventors received funds from companies for research purposes and own the patents on which these academics are named inventors. Hence researchers that have closer links with industry are likely to attract more contracts.

Three conditions are needed to capture the endogeneity. First, in the first-stage regressions the instrumental variables must be significant. Second, the correlation between the number of forward citations, the number of claims and the originality index (the dependent variables of the second stage regressions) and the instrumental variables should be less than 0.30. Third, there should be a low correlation (<0.50) between the instrumental variables and the exogenous variables. All these conditions are respected in our models. During the course of the analysis, we have estimated various lag structures for the instrumental variables and selected a two-year lag for past contracts of universities ($AveCont3U_{t-2}$), while for equipment and infrastructure grants ($AveGrantEI3_{t-1}$), a one-year lag yields the most consistent results.

Taking into account the number of forward citations, the number of claims and our originality index as dependent variables and other independent and interactive variables, the models that will be estimated are therefore given by Eq. 1, 2 and 3. In addition to standard non-instrumental variable models (negative binomial and ordinary least-squares), Eq. 2 will be estimated using two-stage residual inclusion negative binomial models and Eq. 3 will be estimated using two-stage least squares models.

*First-stage*

$$\ln\left(AveCont3_t\right) = f\begin{pmatrix} \ln\left(AveLoopPast_t\right), AveCont3U_{t-2}, \ln\left(AveGrantEI3_{t-1}\right), \\ Variables\ from\ 2^{nd}\ stage \end{pmatrix} \quad \textbf{(1)}$$

*Second stage – Two-stage residual inclusion*

$$\begin{Bmatrix} NbFCitP5_t \\ NbClaims_t \end{Bmatrix} = f\begin{pmatrix} \ln\left(AveGrant3_t\right), \left[\ln\left(AveGrant3_t\right)\right]^2, MaxChair_t, \\ AveAge_t, \left[AveAge_t\right]^2, \ln\left(AveNbCitA3_t\right), dLoop_t, \\ \ln\left(AveCliqA3_t\right), \ln\left(AveCliqP3_t\right), \ln\left(AveBtw3_t\right), \\ propAcInv_t, dAcAssignee_t, dnanoex_t, dPPP_t, \\ \ln\left(AveCont3_t\right)\ and\ residual\ 1^{st}\ stage\left(2SRI\right) \end{pmatrix} \quad \textbf{(2)}$$

*Second stage – Two-stage least squares*

$$\begin{Bmatrix} \ln\left(NbFCitP5_t\right) \\ \ln\left(NbClaims_t\right) \\ InOriginality_t \end{Bmatrix} = f\begin{pmatrix} \ln\left(AveGrant3_t\right), \left[\ln\left(AveGrant3_t\right)\right]^2, MaxChair_t, \\ AveAge_t, \left[AveAge_t\right]^2, \ln\left(AveNbCitA3_t\right), \\ \ln\left(AveCliqA3_t\right), \ln\left(AveCliqP3_t\right), \ln\left(AveBtw3_t\right), \\ propAcInv_t, dAcAssignee_t, dPPP_t, dnanoex_t \end{pmatrix} \quad \textbf{(3)}$$

## 4 Descriptive statistics

Before presenting the results, let us first briefly present the descriptive statistics of the variables (Table 1). The correlation matrix is provided in Table 9 in appendix.

**Table 1 – Descriptive statistics**

| Variable | Nb Obs | Mean | Std. Dev. | Min | Max |
|---|---|---|---|---|---|
| **Dependent variables** | | | | | |
| $NbFCitP5_t$ | 1064 | 1.4342 | 2.6667 | 0 | 33 |
| $NbClaims_t$ | 1064 | 17.5376 | 15.6906 | 1 | 151 |
| $HerfBCit5t$ [d] | 1064 | 0.3337 | 0.3055 | 0 | 0.9050 |
| $InOriginality_t$ | 1064 | -0.0287 | 1.0983 | -1.5893 | 6.6288 |
| **Independent variables** | | | | | |
| $dPPP_t$ | 1064 | 0.1767 | 0.3816 | 0 | 1 |
| $AveGrant3_t$ [a] | 1064 | 281,271.4 | 845,497.4 | 0 | 8,170,096 |
| $AveAge_t$ [a] | 1064 | 11.4735 | 4.0960 | 1 | 21 |
| $MaxChair_t$ [b] | 1064 | 0.9662 | 1.3188 | 0 | 3 |
| $AveNbCitA3_t$ [a] | 1064 | 38.6979 | 85.4982 | 0 | 806 |
| $AveCliqA3_t \times 10^3$ [a,c] | 1064 | 25.1008 | 26.2079 | 0 | 212.500 |
| $AveCliqP3_t \times 10^3$ [a,c] | 1064 | 527.3702 | 381.1754 | 0 | 1000 |
| $AveBtw3_t$ [a,c] | 1064 | 591.2744 | 1,649.0470 | 0 | 8,473.183 |
| $propAcInv_t$ | 1064 | 0.4197 | 0.2925 | 0 | 1 |
| $dAcAssignee_t$ | 1064 | 0.2782 | 0.4483 | 0 | 1 |
| $dNanoEx_t$ | 1064 | 0.1344 | 0.3412 | 0 | 1 |
| **Endogenous variable** | | | | | |
| $AveCont3_t$ [a] | 1064 | 76,643.2 | 499,209.5 | 0 | 8,690,699 |
| **Instrumental variables** | | | | | |
| $AveNbLoopPast_t$ [a] | 1064 | 1.2363 | 4.5487 | 0 | 40.5 |
| $AveGrantEI3_{t-1}$ [a] | 1064 | 23,308.6 | 160,124.4 | 0 | 2,561,704.0 |
| $AveCont3U_{t-2}$ [a] | 1064 | 26,275.6 | 24,303.8 | 0 | 124,651.7 |

Notes: (a) All the variables have been averaged over all academic inventors that contributed to a given patent; (b) Only this variable uses the maximum value of all academic inventors (*MaxChair*); (c) We added 1 to all these variables to insure that their minimum value is greater than 1; (d) This variable is used in the calculation of the Originality Index (*InOrginality*).

Figure 1 shows the number of patents and the number of those that are part of a patent-paper pairs for biotechnology and nanotechnology over the period used in the regression analysis. The graph shows a highly volatile evolution of the number of patents over the years. Of these patents, about 20% are linked through patent-paper-pairs, and about 30 % are assigned to academic institutions (Figure 2).

Figure 3 then shows that patents that are part of patent-paper-pairs are generally less cited five years after their official grant year and have a smaller number of claims, suggesting that their quality is lesser than those that do not have such links. The distinction is not so clear-cut if we compare patents owned by academic institutions versus other types of organizations (in Figure 4). While non-academic patents have a higher number of claims in general, they are not more or less cited than academic patents. We should therefore expect a negative sign of the patent-paper-pairs on the number of citations and on the number of claims in our regression analysis, and a negative sign of academic assignee on the number of claims. These results are also reflected in Table 2: only the number of forward citations between academic assignees and non-academic assignees yields similar results.

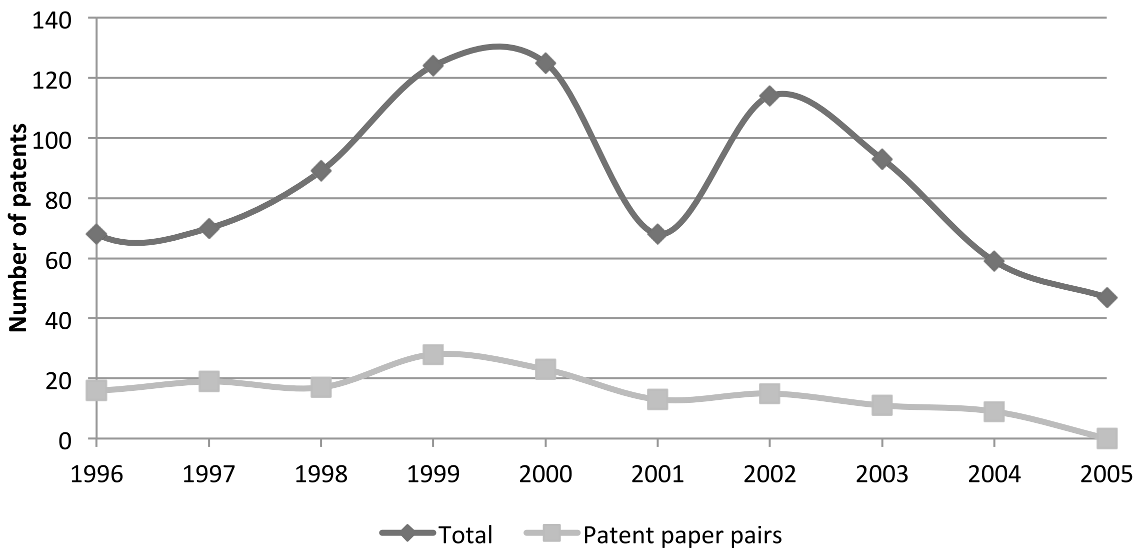



**Figure 1 – Total number of patents and number of patent-paper pairs in the combined fields of biotechnology and nanotechnology in Quebec**

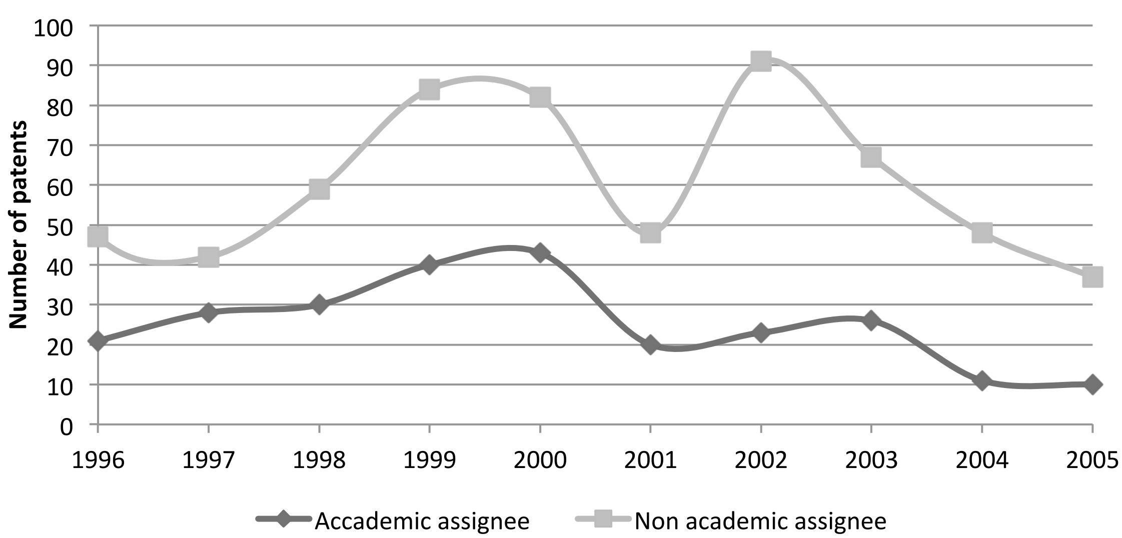
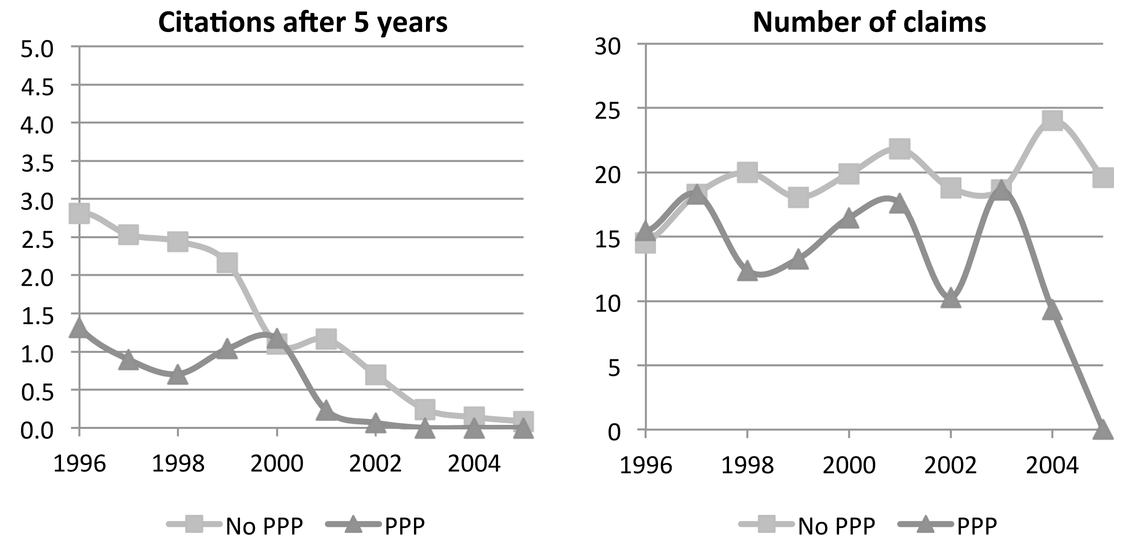
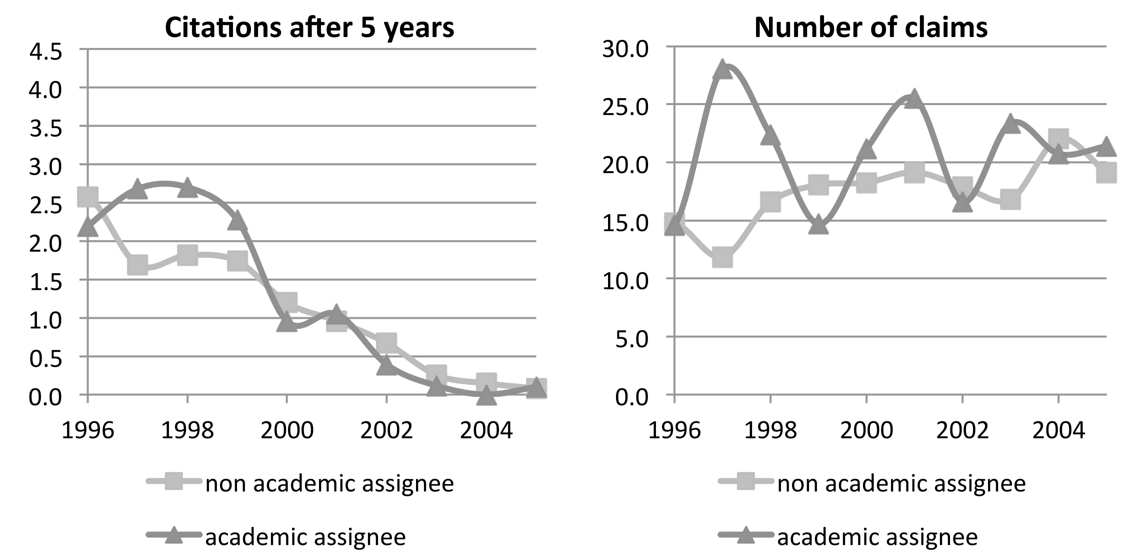



**Figure 2 – Total number of patents by type of assignee (academic and non academic) in the combined fields of biotechnology and nanotechnology in Quebec**




(a) (b)

**Figure 3 – Total number of (a) citations after 5-years and (b) total number of claims per patent for patents that are part of a patent-paper-pair or not in the combined fields of biotechnology and nanotechnology in Quebec**




(a) (b)

**Figure 4 – Total number of (a) citations after 5-years and (b) total number of claims per patent by type of assignee (academic and non academic) in the combined fields of biotechnology and nanotechnology in Quebec**

**Table 2 – Mean comparison between PPP-No PPP and Academic assignee-Non academic assignee groups for each of the dependent variables**

| Variable | No PPP | PPP | Pr(\|T\| < \|t\|) = | |
|---|---|---|---|---|
| Number of observations | 876 | 188 | | |
| $NbFCitP5_t$ | 1.5274 | 1.0000 | 0.0138 | ** |
| | (0.0951) | (0.1295) | | |
| $NbClaims_t$ | 18.2192 | 14.3617 | 0.0022 | *** |
| | (0.5356) | (1.0602) | | |
| $lnOriginality_t$ | 0.0466 | -0.3797 | 0.0000 | *** |
| | (0.0380) | (0.0642) | | |
| **Variable** | **Non academic assignee** | **Academic assignee** | **Pr(\|T\| < \|t\|) =** | |
| Number of observations | 768 | 296 | | |
| $NbFCitP5_t$ | 1.4076 | 1.5034 | 0.5996 | |
| | (0.0968) | (0.1528) | | |
| $NbClaims_t$ | 16.7122 | 19.6791 | 0.0057 | *** |
| | (0.5112) | (1.1013) | | |
| $lnOriginality_t$ | -0.0657 | 0.0671 | 0.0771 | * |
| | (0.0373) | (0.0726) | | |

Note: Standard deviation in parentheses.

Figure 5 shows the average amount of public (operating cost and infrastructure grants) and private funding (contracts) in constant Canadian dollars obtained per acedemic-inventor. In comparison, Beaudry and Allaoui (2012) found a slightly higher proportion of grants compared with contracts. The fact that contracts play a more important role stems from the fact that our sample eliminates scientists that do not patent, and that may not have as close a link with industry as their academic-inventors colleagues.

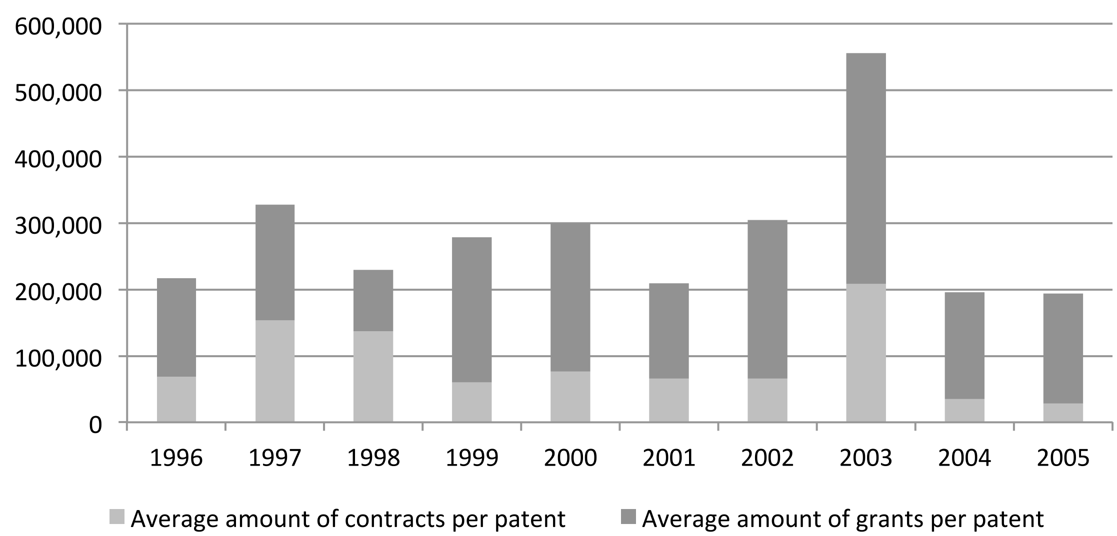



**Figure 5 – Average amount (for each patent) of contracts and grants which academic-inventors received in Quebec for the combined fields of biotechnology and nanotechnology**

## 5 Results

The first stage regression results (not shown here[9]) show that three instrumental variables ln(1+$AveCont3U_{t\text{-}2}$), ln(1+$AveGrantEI3_{t\text{-}1}$), and ln(1+$AveNbLoopPast_t$) are significant. The results for the two-Stage Residual Inclusion (2SRI) negative binomial regressions of the number of forward citations and the number of claims are presented in Table 3 and Table 4 respectively. Table 5 shows the ordinary least squares[10] results for the originality index.

Let us to examine the impact of each variable on patent quality measured by the number of forward citations, the number of claims or our originality index. The results clearly show the negative impact of patent-paper pairs on patent quality measured by the number of forward citations and by the number of claims. In other words, patents issued from patent-paper pairs receive less forward citations and have a lower number of claims. This outcome contrasts with previous studies that found no significant impact of patent-paper pairs on the number of forward citations received by patents (Magerman et al., 2011). Hence closeness to scientific work is far from being a guarantee of quality for patents to which academics have contributed, quite the opposite.

Turning now to the influence of patent ownership on patent quality, our results show that the patents owned by universities have a higher number of claims. However, this analysis finds no evidence that academic assignees receive more patent forward citations or produce patents with a higher originality index. Furthermore, taking the natural logarithm of the number of claims and estimating two-stage least squares regressions (see Table 6 for the citations and Table 7 for the claims), the positive impact of academic assignees completely disappears. The impact of the type of assignee is therefore not robust. Our results echo those of Crespi et al. (2010) who found no significant impact of the ownership structure on patent quality, but contrast with those of Lissoni et al. (2010) who demonstrated that academic patents are of lower quality compared those owned by corporations.

[9] The first stage regressions are available from the authors upon request.
[10] Durbin and Wu-Hausman tests could not reject the null hypothesis that the variables are exogenous after the Two-stage least-square (2SLS) regressions. The 2SLS results will therefore not be presented here.

To further investigate the impact of patent-paper-pairs and of academic assignees, we introduced a number of interactive variables to the regression, which for the sake brevity will not all be presented her. The fifth columns of Tables 3 to 7 present the resulting regressions including three interactive terms. Patents owned by universities where the named inventors are part of a very integrated co-publication clique will generally be less cited (see Table 3 and Table 6). If, however, the academic inventors are surrounded by a well-integrated clique of inventors, then the patents owned by universities will gain more citations and will also state a greater number of claims. In this regard, what matters most appears to be the integration of the inventors network close to the patent team's interests. Where the notoriety of the academic inventor team has an influence is in reducing the negative impact of the patent-paper-pair. The more cited are the papers published by the academic inventors, the more the proximity between science and technology will yield a greater number of citations and of claims. The overall coefficient (of *dPPP* and of its interactive term with *AveNbACit3*), remains nonetheless negative.

Having introduced one of the control variables for the latent quality of the academic inventors, the average number of citations received by their papers, let us now mention its individual positive impact on both the number of claims and the number of citations (it is weakly significant on the originality index). In addition to the notoriety granted by a research chair, having published highly cited papers will both positively influence the number of claims and citations of the patents produced by academic inventors. These positive and significant impacts however disappear when the number of claims and of citations are incorporated into our originality index.

What seems to matter across the board is the intermediary position of the academic inventors in the co-invention network. Our models show that better-positioned and more centralized inventors in the co-invention network produce higher patent quality measured by the number of forward citations and the number of claims, but only up to a point. Occupying a too central position in this network eventually contributes to the decline in quality of the resulting patents. Our results therefore only partially support the findings of Wang et al. (2010) regarding their observation that patents occupying higher brokerage positions obtain more forward citations and have longer renewal decision times.

**Table 3 – Regression results for the number of forward citations (*NbFCit5$_t$*) – Two-stage residual inclusion negative binomial (2SRI)**

| Variable | 2SRI(1) | 2SRI(2) | 2SRI(3) | 2SRI(4) | 2SRI(5) |
|---|---|---|---|---|---|
| ln($AveCont3_t$) | -0.0429 | -0.0429 | -0.0580 | -0.0742 | -0.0700 * |
| | (0.0357) | (0.0361) | (0.0373) | (0.0624) | (0.0373) |
| $[\text{n}(AveCont3_t)]^2$ | | | | 0.0012 | |
| | | | | (0.0044) | |
| ln($AveGrant3_t$) | -0.0243 | -0.0282 | -0.0165 | -0.0121 | -0.0098 |
| | (0.0173) | (0.0421) | (0.0179) | (0.0188) | (0.0179) |
| $[\ln(AveGrant3_t)]^2$ | | 0.0003 | | | |
| | | (0.0034) | | | |
| $AveAge_t$ | -0.1529 *** | -0.1529 *** | -0.1451 *** | -0.1467 *** | -0.1414 *** |
| | (0.0162) | (0.0164) | (0.0161) | (0.0162) | (0.0158) |
| $MaxChair_t$ | 0.1498 *** | 0.1486 *** | 0.1460 *** | 0.1386 *** | 0.1678 *** |
| | (0.0469) | (0.0467) | (0.0462) | (0.0458) | (0.0466) |
| ln($AveNbACit3_t$) | 0.0990 *** | 0.0977 *** | 0.1025 *** | 0.1097 *** | 0.1143 *** |
| | (0.0325) | (0.0324) | (0.0326) | (0.0335) | (0.0384) |
| $1+\ln(AveCliqA3_t \times 10^3)$ | 0.1718 ** | 0.1714 ** | 0.1988 ** | 0.2216 *** | 0.1946 ** |
| | (0.0798) | (0.0809) | (0.0789) | (0.0790) | (0.0783) |
| $1+\ln(AveCliqP3_t \times 10^3)$ | 0.0417 * | 0.0417 * | 0.0270 | -0.2825 | -0.0238 |
| | (0.0226) | (0.0226) | (0.0226) | (0.2339) | (0.0266) |
| $[1+\ln(AveCliqP3_t \times 10^3)]^2$ | | | | 0.0349 | |
| | | | | (0.0263) | |
| $1+\ln(AveBtwP3_t)$ | -0.0134 | -0.0129 | 0.2585 *** | 0.3094 *** | 0.2911 *** |
| | (0.0216) | (0.0217) | (0.0855) | (0.0979) | (0.0856) |
| $[1+\ln(AveBtw3_t)]^2$ | | | -0.0306 *** | -0.0326 *** | -0.0354 *** |
| | | | (0.0096) | (0.0099) | (0.0096) |
| $propAcInv_t$ | 0.1616 | 0.1613 | -0.0580 | -0.0756 | -0.2765 |
| | (0.2034) | (0.2242) | (0.2191) | (0.2205) | (0.2252) |
| $dAcAssignee_t$ | -0.0629 | -0.0645 | -0.0316 | -0.0360 | -0.9178 ** |
| | (0.1229) | (0.1250) | (0.1232) | (0.1232) | (0.3580) |
| *dNanoEx* | 0.5296 *** | 0.5310 *** | 0.5219 *** | 0.5097 *** | 0.5430 *** |
| | (0.1415) | (0.1415) | (0.1407) | (0.1406) | (0.1393) |
| $dPPP_t$ | -0.4402 *** | -0.4399 *** | -0.3812 *** | -0.3769 *** | -0.8854 *** |
| | (0.1406) | (0.1417) | (0.1421) | (0.1423) | (0.2507) |
| $dAcAssignee_t$ x ln($AveNbACit3_t$) | | | | | -0.1543 ** |
| | | | | | (0.0624) |
| $dAcAssignee_t$ x $[1+\ln(AveCliqP3t \times 10^3)]$ | | | | | 0.1902 *** |
| | | | | | (0.0496) |
| $dPPP_t$ x ln($AveNbACit3_t$) | | | | | 0.2062 *** |
| | | | | | (0.0782) |
| Residual of $AveCont3_t$ | 0.0855 ** | 0.0855 ** | 0.0982 ** | 0.1003 ** | 0.1078 *** |
| | (0.0369) | (0.0373) | (0.0386) | (0.0391) | (0.0386) |
| Constant | 0.9890 ** | 0.9950 ** | 0.6229 | 0.7745 * | 0.9167 ** |
| | (0.4354) | (0.4578) | (0.4306) | (0.4535) | (0.4451) |
| ln(alpha) | 0.4817 *** | 0.4816 *** | 0.4693 *** | 0.4675 *** | 0.4279 *** |
| | (0.0833) | (0.0833) | (0.0837) | (0.0838) | (0.0848) |
| Nb observations | 1063 | 1063 | 1063 | 1063 | 1063 |
| Log Likelihood | -1586.71 | -1586.68 | -1583.51 | -1583.25 | -1573.99 |
| $\chi^2$ | 194.058 *** | 194.112 *** | 200.451 *** | 200.978 *** | 219.485 *** |
| Pseudo $R^2$ | 0.0576 | 0.0576 | 0.0595 | 0.0597 | 0.0652 |

Notes: ***, **, * show significance at the 1%, 5%, and 10% levels. Standard errors are presented in parentheses.

**Table 4 – Regression results for the number of claims (*NbClaims*$_t$) – Two-stage residual inclusion negative binomial (2SRI)**

| Variable | 2SRI(6) | 2SRI(7) | 2SRI(8) | 2SRI(9) | 2SRI(10) |
|---|---|---|---|---|---|
| ln($AveCont3_t$) | -0.0426 *** | -0.0439 *** | -0.0508 *** | -0.1352 *** | -0.0493 *** |
| | (0.0156) | (0.0157) | (0.0162) | (0.0275) | (0.0164) |
| [n($AveCont3_t$)]$^2$ | | | | 0.0070 *** | |
| | | | | (0.0019) | |
| ln($AveGrant3_t$) | 0.0079 | -0.0243 | 0.0113 | 0.0167 ** | 0.0104 |
| | (0.0075) | (0.0196) | (0.0077) | (0.0079) | (0.0077) |
| [ln($AveGrant3_t$)]$^2$ | | 0.0027 * | | | |
| | | (0.0016) | | | |
| $AveAge_t$ | -0.0071 | -0.0083 | -0.0025 | -0.0060 | -0.0024 |
| | (0.0063) | (0.0064) | (0.0062) | (0.0063) | (0.0063) |
| $MaxChair_t$ | 0.0821 *** | 0.0804 *** | 0.0722 *** | 0.0680 *** | 0.0733 *** |
| | (0.0216) | (0.0215) | (0.0212) | (0.0210) | (0.0214) |
| ln($AveNbACit3_t$) | 0.0404 *** | 0.0350 ** | 0.0432 *** | 0.0498 *** | 0.0314 * |
| | (0.0136) | (0.0139) | (0.0137) | (0.0139) | (0.0164) |
| 1+ln($AveCliqA3_t \times 10^3$) | -0.1974 *** | -0.2026 *** | -0.1807 *** | -0.1667 *** | -0.1815 *** |
| | (0.0339) | (0.0343) | (0.0334) | (0.0332) | (0.0334) |
| 1+ln($AveCliqP3_t \times 10^3$) | 0.0328 *** | 0.0318 *** | 0.0219 ** | -0.2695 ** | 0.0055 |
| | (0.0101) | (0.0101) | (0.0101) | (0.1059) | (0.0122) |
| [1+ln($AveCliqP3_t \times 10^3$)]$^2$ | | | | 0.0329 *** | |
| | | | | (0.0119) | |
| 1+ln($AveBtwP3_t$) | -0.0029 | -0.0002 | 0.1869 *** | 0.2346 *** | 0.1876 *** |
| | (0.0096) | (0.0096) | (0.0365) | (0.0415) | (0.0372) |
| [1+ln($AveBtw3_t$)]$^2$ | | | -0.0208 *** | -0.0227 *** | -0.0209 *** |
| | | | (0.0039) | (0.0040) | (0.0041) |
| $propAcInv_t$ | -0.2995 *** | -0.3516 *** | -0.4159 *** | -0.4534 *** | -0.4375 *** |
| | (0.0894) | (0.0952) | (0.0931) | (0.0933) | (0.0958) |
| $dAcAssignee_t$ | 0.1402 ** | 0.1489 *** | 0.1544 *** | 0.1553 *** | -0.1331 |
| | (0.0563) | (0.0567) | (0.0562) | (0.0557) | (0.1533) |
| *dNanoEx* | 0.3011 *** | 0.3030 *** | 0.2960 *** | 0.2876 *** | 0.2963 *** |
| | (0.0669) | (0.0668) | (0.0665) | (0.0661) | (0.0663) |
| $dPPP_t$ | -0.2024 *** | -0.1905 *** | -0.1658 *** | -0.1479 ** | -0.3750 *** |
| | (0.0626) | (0.0629) | (0.0623) | (0.0620) | (0.1046) |
| $dAcAssignee_t$ x ln($AveNbACit3_t$) | | | | | -0.0080 |
| | | | | | (0.0273) |
| $dAcAssignee_t$ x [1+ln($AveCliqP3t \times 10^3$)] | | | | | 0.0467 ** |
| | | | | | (0.0212) |
| $dPPP_t$ x ln($AveNbACit3_t$) | | | | | 0.0814 ** |
| | | | | | (0.0331) |
| Residual of $AveCont3_t$ | 0.0549 *** | 0.0559 *** | 0.0607 *** | 0.0644 *** | 0.0590 *** |
| | (0.0163) | (0.0164) | (0.0169) | (0.0170) | (0.0171) |
| Constant | 3.5028 *** | 3.5838 *** | 3.2610 *** | 3.4672 *** | 3.4009 *** |
| | (0.1840) | (0.1940) | (0.1807) | (0.1926) | (0.1914) |
| ln(alpha) | -0.7455 *** | -0.7468 *** | -0.7637 *** | -0.7794 *** | -0.7703 *** |
| | (0.0471) | (0.0471) | (0.0473) | (0.0474) | (0.0474) |
| Nb observations | 1063 | 1063 | 1063 | 1063 | 1063 |
| Log Likelihood | -3988.79 | -3988.2 | -3979.72 | -3971.66 | -3976.41 |
| $\chi^2$ | 128.886 *** | 130.077 *** | 147.022 *** | 163.148 *** | 153.655 *** |
| Pseudo $R^2$ | 0.0159 | 0.0160 | 0.0181 | 0.0201 | 0.0190 |

Notes: ***, **, * show significance at the 1%, 5%, and 10% levels. Standard errors are presented in parentheses.

**Table 5 – Regression results for the originality index (*InOriginality*$_t$) – Ordinary least squares (OLS)[11]**

| Variable | OLS(11) | OLS(12) | OLS(13) | OLS(14) | OLS(15) |
|---|---|---|---|---|---|
| ln(*AveCont3*$_t$) | 0.0278 *** | 0.0273 *** | 0.0245 *** | -0.0590 * | 0.0225 *** |
| | (0.0068) | (0.0068) | (0.0068) | (0.0316) | (0.0068) |
| [n(*AveCont3*$_t$)]$^2$ | | | | 0.0074 *** | |
| | | | | (0.0027) | |
| ln(*AveGrant3*$_t$) | -0.0212 *** | -0.0456 * | -0.0196 ** | -0.0183 ** | -0.0192 ** |
| | (0.0079) | (0.0271) | (0.0079) | (0.0080) | (0.0079) |
| [ln(*AveGrant3*$_t$)]$^2$ | | 0.0020 | | | |
| | | (0.0021) | | | |
| *AveAge*$_t$ | -0.0288 *** | -0.0297 *** | -0.0247 *** | -0.0267 *** | -0.0241 *** |
| | (0.0086) | (0.0087) | (0.0087) | (0.0087) | (0.0087) |
| *MaxChair*$_t$ | 0.0073 | 0.0057 | -0.0009 | 0.0061 | 0.0057 |
| | (0.0260) | (0.0261) | (0.0261) | (0.0264) | (0.0260) |
| ln(*AveNbACit3*$_t$) | 0.0329 * | 0.0292 | 0.0352 * | 0.0353 * | 0.0273 |
| | (0.0189) | (0.0193) | (0.0189) | (0.0191) | (0.0223) |
| 1+ln(*AveCliqA3*$_t$ × 10$^3$) | -0.0914 ** | -0.0946 ** | -0.0741 * | -0.0821 * | -0.0781 * |
| | (0.0414) | (0.0416) | (0.0416) | (0.0432) | (0.0414) |
| 1+ln(*AveCliqP3*$_t$ × 10$^3$) | 0.0577 *** | 0.0572 *** | 0.0488 *** | 0.0795 | 0.0247 |
| | (0.0138) | (0.0138) | (0.0140) | (0.1466) | (0.0171) |
| [1+ln(*AveCliqP3*$_t$ × 10$^3$)]$^2$ | | | | -0.0033 | |
| | | | | (0.0164) | |
| 1+ln(*AveBtwP3*$_t$) | 0.0031 | 0.0052 | 0.1518 *** | 0.1409 *** | 0.1661 *** |
| | (0.0127) | (0.0129) | (0.0487) | (0.0542) | (0.0490) |
| [1+ln(*AveBtw3*$_t$)]$^2$ | | | -0.0162 *** | -0.0155 *** | -0.0180 *** |
| | | | (0.0051) | (0.0052) | (0.0052) |
| *propAcInv*$_t$ | -0.0393 | -0.0809 | -0.1365 | -0.1521 | -0.1975 |
| | (0.1238) | (0.1314) | (0.1271) | (0.1269) | (0.1281) |
| *dAcAssignee*$_t$ | 0.0375 | 0.0438 | 0.0505 | 0.0456 | -0.2953 |
| | (0.0766) | (0.0769) | (0.0764) | (0.0762) | (0.2105) |
| *dNanoEx* | 0.7216 *** | 0.7238 *** | 0.7140 *** | 0.7173 *** | 0.7156 *** |
| | (0.0945) | (0.0945) | (0.0941) | (0.0941) | (0.0937) |
| *dPPP*$_t$ | -0.3571 *** | -0.3489 *** | -0.3271 *** | -0.3126 *** | -0.6526 *** |
| | (0.0854) | (0.0859) | (0.0856) | (0.0856) | (0.1450) |
| *dAcAssignee*$_t$ x ln(*AveNbACit3*$_t$) | | | | | -0.0510 |
| | | | | | (0.0384) |
| *dAcAssignee*$_t$ x [1+ln(*AveCliqP3t* × 10$^3$)] | | | | | 0.0717 ** |
| | | | | | (0.0285) |
| *dPPP*$_t$ x ln(*AveNbACit3*$_t$) | | | | | 0.1351 *** |
| | | | | | (0.0465) |
| Constant | 0.2481 | 0.3049 | 0.0403 | 0.0767 | 0.2272 |
| | (0.2321) | (0.2398) | (0.2403) | (0.2500) | (0.2518) |
| Nb observations | 1064 | 1064 | 1064 | 1064 | 1064 |
| Log Likelihood | -1528.92 | -1528.47 | -1523.89 | -1519.98 | -1516.17 |
| F | 14.2271 *** | 13.1998 *** | 14.0131 *** | 12.7254 *** | 12.476 *** |
| $R^2$ | 0.1397 | 0.1405 | 0.1478 | 0.154075 | 0.1601 |
| Adjusted $R^2$ | 0.1299 | 0.1298 | 0.1373 | 0.141967 | 0.1473 |

Notes: ***, **, * show significance at the 1%, 5%, and 10% levels. Standard errors are presented in parentheses.

[11] Two-stage least squares regressions were performed taking the average amount of contracts raised as the endogenous variable. In contrast with the number of citations and the number of claims, exogeneity cannot be rejected and the instrumental variables results will therefore not be presented here.

**Table 6 – Regression results for the number of forward citations [ln($NbFCit5_t$)] – Two-stage least squares (IV)**

| Variable | IV(1) | IV(2) | IV(3) | IV(4) | IV(5) |
|---|---|---|---|---|---|
| ln($AveCont3_t$) | -0.0270 * | -0.0267 * | -0.0328 ** | -0.0327 ** | -0.0354 ** |
| | (0.0142) | (0.0144) | (0.0150) | (0.0152) | (0.0151) |
| ln($AveGrant3_t$) | -0.0077 | -0.0092 | -0.0054 | -0.0047 | -0.0041 |
| | (0.0069) | (0.0180) | (0.0071) | (0.0074) | (0.0072) |
| [ln($AveGrant3_t$)]$^2$ | | 0.0001 | | | |
| | | (0.0015) | | | |
| $AveAge_t$ | -0.0568 *** | -0.0568 *** | -0.0538 *** | -0.0542 *** | -0.0526 *** |
| | (0.0058) | (0.0059) | (0.0058) | (0.0059) | (0.0058) |
| $MaxChair_t$ | 0.0724 *** | 0.0720 *** | 0.0686 *** | 0.0658 *** | 0.0740 *** |
| | (0.0198) | (0.0197) | (0.0197) | (0.0196) | (0.0198) |
| ln($AveNbACit3_t$) | 0.0393 *** | 0.0390 *** | 0.0416 *** | 0.0435 *** | 0.0504 *** |
| | (0.0128) | (0.0129) | (0.0129) | (0.0132) | (0.0153) |
| 1+ln($AveCliqA3_t \times 10^3$) | 0.1059 *** | 0.1061 *** | 0.1154 *** | 0.1228 *** | 0.1121 *** |
| | (0.0314) | (0.0317) | (0.0311) | (0.0314) | (0.0311) |
| 1+ln($AveCliqP3_t \times 10^3$) | 0.0164 * | 0.0163 * | 0.0100 | -0.0743 | -0.0142 |
| | (0.0093) | (0.0093) | (0.0094) | (0.1002) | (0.0115) |
| [1+ln($AveCliqP3_t \times 10^3$)]$^2$ | | | | 0.0095 | |
| | | | | (0.0112) | |
| 1+ln($AveBtwP3_t$) | -0.0136 | -0.0134 | 0.0992 *** | 0.1127 *** | 0.1166 *** |
| | (0.0087) | (0.0088) | (0.0351) | (0.0401) | (0.0355) |
| [1+ln($AveBtw3_t$)]$^2$ | | | -0.0124 *** | -0.0128 *** | -0.0146 *** |
| | | | (0.0038) | (0.0039) | (0.0039) |
| $propAcInv_t$ | 0.1116 | 0.1095 | 0.0347 | 0.0329 | -0.0294 |
| | (0.0829) | (0.0889) | (0.0879) | (0.0882) | (0.0896) |
| $dAcAssignee_t$ | 0.0055 | 0.0056 | 0.0177 | 0.0166 | -0.2910 ** |
| | (0.0518) | (0.0521) | (0.0524) | (0.0524) | (0.1415) |
| *dNanoEx* | 0.2197 *** | 0.2198 *** | 0.2139 *** | 0.2103 *** | 0.2209 *** |
| | (0.0626) | (0.0626) | (0.0630) | (0.0631) | (0.0627) |
| $dPPP_t$ | -0.1723 *** | -0.1717 *** | -0.1500 *** | -0.1495 *** | -0.2667 *** |
| | (0.0567) | (0.0569) | (0.0573) | (0.0573) | (0.0972) |
| $dAcAssignee_t$ x ln($AveNbACit3_t$) | | | | | -0.0657 ** |
| | | | | | (0.0259) |
| $dAcAssignee_t$ x [1+ln($AveCliqP3_t \times 10^3$)] | | | | | 0.0727 *** |
| | | | | | (0.0197) |
| $dPPP_t$ x ln($AveNbACit3_t$) | | | | | 0.0550 * |
| | | | | | (0.0312) |
| Constant | 0.7419 *** | 0.7433 *** | 0.6009 *** | 0.6378 *** | 0.7339 *** |
| | (0.1695) | (0.1778) | (0.1692) | (0.1785) | (0.1796) |
| Nb observations | 1063 | 1063 | 1063 | 1063 | 1063 |
| $\chi^2$ | 207.252 *** | 207.471 *** | 212.554 *** | 212.715 *** | 232.061 *** |
| $R^2$ | 0.1090 | 0.1102 | 0.0992 | 0.1001 | 0.1083 |
| Adjusted $R^2$ | 0.0988 | 0.0992 | 0.0880 | 0.0880 | 0.0947 |

Notes: ***, **, * show significance at the 1%, 5%, and 10% levels. Standard errors are presented in parentheses.

**Table 7 – Regression results for the number of claims [ln($NbClaims_t$)] – Two-stage least squares (IV)**

| Variable | IV(6) | IV(7) | IV(8) | IV(9) | IV(10) |
|---|---|---|---|---|---|
| ln($AveCont3_t$) | -0.0410 ** | -0.0411 ** | -0.0481 ** | -0.0500 *** | -0.0479 ** |
| | (0.0179) | (0.0181) | (0.0189) | (0.0192) | (0.0190) |
| ln($AveGrant3_t$) | 0.0061 | -0.0028 | 0.0089 | 0.0116 | 0.0083 |
| | (0.0087) | (0.0227) | (0.0090) | (0.0093) | (0.0090) |
| [ln($AveGrant3_t$)]$^2$ | | 0.0007 | | | |
| | | (0.0018) | | | |
| $AveAge_t$ | -0.0135 * | -0.0138 * | -0.0096 | -0.0108 | -0.0096 |
| | (0.0073) | (0.0074) | (0.0074) | (0.0074) | (0.0073) |
| $MaxChair_t$ | 0.0638 ** | 0.0632 ** | 0.0585 ** | 0.0517 ** | 0.0618 ** |
| | (0.0250) | (0.0249) | (0.0249) | (0.0248) | (0.0249) |
| ln($AveNbACit3_t$) | 0.0403 ** | 0.0389 ** | 0.0432 *** | 0.0491 *** | 0.0294 |
| | (0.0161) | (0.0163) | (0.0163) | (0.0167) | (0.0192) |
| 1+ln($AveCliqA3_t \times 10^3$) | -0.2180 *** | -0.2191 *** | -0.2052 *** | -0.1849 *** | -0.2069 *** |
| | (0.0396) | (0.0401) | (0.0393) | (0.0396) | (0.0392) |
| 1+ln($AveCliqP3_t \times 10^3$) | 0.0289 ** | 0.0287 ** | 0.0205 * | -0.2331 * | 0.0052 |
| | (0.0118) | (0.0118) | (0.0119) | (0.1266) | (0.0144) |
| [1+ln($AveCliqP3_t \times 10^3$)]$^2$ | | | | 0.0285 ** | |
| | | | | (0.0142) | |
| 1+ln($AveBtwP3_t$) | 0.0092 | 0.0100 | 0.1555 *** | 0.1982 *** | 0.1598 *** |
| | (0.0110) | (0.0111) | (0.0442) | (0.0506) | (0.0448) |
| [1+ln($AveBtw3_t$)]$^2$ | | | -0.0160 *** | -0.0177 *** | -0.0166 *** |
| | | | (0.0048) | (0.0050) | (0.0049) |
| $propAcInv_t$ | -0.2723 *** | -0.2873 ** | -0.3716 *** | -0.3803 *** | -0.4005 *** |
| | (0.1047) | (0.1123) | (0.1109) | (0.1113) | (0.1129) |
| $dAcAssignee_t$ | 0.0396 | 0.0418 | 0.0550 | 0.0535 | -0.2076 |
| | (0.0653) | (0.0658) | (0.0661) | (0.0661) | (0.1784) |
| *dNanoEx* | 0.4024 *** | 0.4032 *** | 0.3949 *** | 0.3839 *** | 0.3933 *** |
| | (0.0790) | (0.0791) | (0.0794) | (0.0797) | (0.0791) |
| $dPPP_t$ | -0.2059 *** | -0.2029 *** | -0.1769 ** | -0.1754 ** | -0.4335 *** |
| | (0.0715) | (0.0719) | (0.0722) | (0.0723) | (0.1225) |
| $dAcAssignee_t$ x ln($AveNbACit3_t$) | | | | | -0.0123 |
| | | | | | (0.0326) |
| $dAcAssignee_t$ x [1+ln($AveCliqP3_t \times 10^3$)] | | | | | 0.0446 * |
| | | | | | (0.0248) |
| $dPPP_t$ x ln($AveNbACit3_t$) | | | | | 0.1032 *** |
| | | | | | (0.0393) |
| Constant | 3.3677 *** | 3.3880 *** | 3.1826 *** | 3.3015 *** | 3.3250 *** |
| | (0.2139) | (0.2246) | (0.2132) | (0.2254) | (0.2264) |
| Nb observations | 1063 | 1063 | 1063 | 1063 | 1063 |
| $\chi^2$ | 97.1628 *** | 97.0718 *** | 103.565 *** | 105.634 *** | 112.092 *** |
| $R^2$ | 0.0155 | 0.0154 | 0.0068 | 0.0047 | 0.0171 |
| Adjusted $R^2$ | 0.0042 | 0.0032 | -0.0055 | -0.0086 | 0.0020 |

Notes: ***, **, * show significance at the 1%, 5%, and 10% levels. Standard errors are presented in parentheses.

In contrast, the cliquishness of the co-publication network surrounding the academic inventors has the opposite effect on the number of citations (positive and significant), and on the number of claims and on our originality index (negative and significant). Attracting citations appear to benefit from an integrated scientific team around the academic inventor, but this comes at the price of a less ambitious patent in terms of number of claims.

## 6 Discussion and conclusion

This brings us to what the number of citations, the number of claims or our originality index actually measure. Is it really "quality" as it is often claimed in the literature? On the one hand, patent citations measure the "use" of the patent as prior art cited in other patents, i.e. how other technologies build upon a specific patent. Claims on the other hand define the extent or the scope of the protection granted by the patent. The third element of our originality index, 1 minus the Herfindahl index of the number of backward citations, measure the diversity in citing patents in other international patent classes.

Our results find a negative impact of patent-paper pairs on the patent quality including number of forward citation, number of claims and an originality index. Prior studies have concluded that there is no significant impact of patents that have paper counterparts on patent citation flows. Our results corroborate these findings and further suggest that patents owned by universities and where the academic inventors are highly cited in the science world, yield a smaller number of citations in the technology world (lesser number of patent citations). The proximity between science and technology however crucial for knowledge transfer in general does not seem to have an equal importance when it comes to having an impact in the technology world. This is further exemplified by the lack of significance of academic assignees, but more importantly by the negative impact of a greater proportion of academics amongst the inventors. We however found mitigating factors, such academic inventors holding a prestigious chair, or having published highly cited articles, which contribute positively to the "quality" of the patents to which they contribute.

Because academics rarely work alone, their collaboration networks play an important role in outputs, outcomes and impacts resulting from their research. Prior research has found that patents occupying better brokerage positions prevent other competitors to capture the market (Blind et al. 2009). In terms of patent "quality" our results seem to indicate that better brokerage positions will only go so far to improve patent "quality". Too central an intermediary position is eventually associated with declining "quality". Further research is clearly need to disentangle the role of academic inventors in the technology world in regard to their position within the scientific and technological networks. What usage is being made of university patents and of the patents to which academics contribute and are they used more than the average patent? Which brings us back to the question of "quality". What is a quality patent? What indicators are relevant at the individual patent level? Our research has shown that all indicators are not interchangeable and imply very different concepts, which are sometimes used as "proxies" for "quality" for lack of better indicators. Results obtained are therefore highly dependent on the type of "proxy" used to measure a particular concept, we as empirical researchers must therefore tread with care in the vast realm of patent quality indicators. To quote Hagedoorn and Cloodt (2003), given "the variety in constructs, measurements, samples, databases, industries and country settings and inconsistency in definitions, it is of no surprise that there appears to be hardly any clear understanding of the concept and measurement of innovative performance" (pp. 1365-1366). Although a clear application of the concerns they raise, our results find common grounds in the lack of positive impact of patent-paper pairs and of the lack of importance of academic ownership of patents to which academic inventors have contributed.

Like every research ours has some limitations. First, taking into consideration the data to which we have access we could only analyze the quality of patents measured by the number of forward and backward citations, and the number of claims. There are however more indicators to measure patent quality including the number of IPC-subclasses, patent renewals, patent families, and number of applicants, to name a few. We postulate that in order to truly measure patent quality, a more complex indicator than the one used in this paper would have to be derived from a number of measures. The second limitation is related to the data. Our data only covers biotechnology and nanotechnology in Quebec and cannot be generalised to other disciplines or to other parts of the world.

## 7 References

Adams, C. W., 1946. The age at which scientists do their best work. Isis 36(3–4), 166–169.

Almeida, P., 1996. Knowledge sourcing by foreign multinationals: patent citation analysis in the US semiconductor industry. Strategic Management Journal 17(S2), 155–165.

Argyres, N. S., Liebeskind, J. P., 1998. Privatizing the intellectual commons: Universities and the commercialization of biotechnology. Journal of Economic Behavior & Organization 35(4), 427–454.

Arora, A., Merges, R. P., 2004. Specialized supply firms, property rights and firm boundaries. Industrial and Corporate Change 13(3), 451–475.

Baba, Y., Shichijo, N., Sedita, S. R., 2009. How do collaborations with universities affect firms' innovative performance? The role of "Pasteur scientists" in the advanced materials field. Research Policy 38(5), 756–764.

Balconi, M., Laboranti, A., 2006. University–industry interactions in applied research: the case of microelectronics. Research Policy 35(10), 1616–1630.

Bassecoulart E., Zitt, M., 2004. Patents and Publications. The Lexical Connection. In Moed H.F., Glänzel W., Schmoch U., Handbook of Quantitative Science and Technology Research. Kluwer. Dordrecht, Ch. 30.

Beaudry, C., Schiffauerova, A., 2011. Impacts of collaboration and network indicators on patent quality: The case of Canadian nanotechnology innovation. European Management Journal 29(5), 362–376.

Beaudry, C., Allaoui, S., 2012. Impact of public and private research funding on scientific production: The case of nanotechnology. Research Policy, 41(9), 1589–1606.

Beaudry, C., Kananian, R., 2013. Follow the (industry) money–The Impact of science networks and industry-to-university contracts on academic patenting in nanotechnology and biotechnology. Industry and Innovation 20(3), 241–260.

Blind, K., Cremers, K., Mueller, E., 2009. The influence of strategic patenting on companies' patent portfolios. Research Policy 38(2), 428–436.

Bonaccorsi, A., Thoma, G., 2007. Institutional complementarity and inventive performance in nano science and technology. Research Policy 36(6), 813–831.

Bowie, N., 1994. University–Business Partnerships: An Assessment. Rowman and Littlefield, Lanham, MD.

Branstetter, L., Nakamura, Y., 2003. Is Japan's Innovative Capacity Decline? NBER Working Paper Series 9438.

Breschi, S., Lissoni, F., Montobbio, F., 2005. From publishing to patenting: do productive scientists turn into academic inventors? Revue d'Economie Industrielle 110(1), 75–102.

Burke, P. F., Reitzig, M., 2007. Measuring patent assessment quality—Analyzing the degree and kind of (in)consistency in patent offices' decision making. Research Policy 36(9), 1404–1430.

Calderini, M., Franzoni, C., 2004. Is academic patenting detrimental to high quality research? An empirical analysis of the relationship between scientific careers and patent applications. CESPRI Working Paper No. 162. Università Commerciale Luigi Bocconi, CESPRI, Milano.

Calderini, M., Franzoni, C., Vezzulli, A., 2007. If star scientists do not patent: the effect of productivity, basicness and impact on the decision to patent in the academic world. Research Policy 36(3), 303–319.

Carpenter, M., Cooper, M., Narin, F., 1980. Linkage between basic research literature and patents. Research Management 23(2), 30–35.

Carrington, P.J., Scott, J., Wasserman, S., 2005. Models and Methods in Social Network Analysis, Cambridge University Press.

Cohen, W. M., Nelson, R. R., Walsh, J. P., 2002. Links and impacts: the influence of public research on industrial R&D. Management Science 48(1), 1–23.

Cole, S., Cole, J. R., 1967. Scientific output and recognition: a study in the operation of the reward system in science. American Sociological Review 32, 377–390.

Cole, J. R., Cole, S., 1973. Social Stratification in Science. The University of Chicago Press, Chicago, IL.

Crespi, G., Geuna, A., Nomaler, O., Verspagen, B., 2010. University IPRs and knowledge transfer: is university ownership more efficient? Economics of Innovation and New Technology 19(7), 627–648.

Czarnitzki, D., Hussinger, K., Schneider, C., 2012. The nexus between science and industry: evidence from faculty inventions. The Journal of Technology Transfer 37(5), 755–776.

Dasgupta, P., David, P., 1987. Information disclosure and the economics of science and technology. In: Feiwel, G. (Ed.), Arrow and the Ascent of Modern Economic Theory. New York University Press, New York, pp. 519–542.

Dasgupta, P., David, P., 1994. Toward a new economics of science. Research Policy 23(5), 487–522.

David, P., 2001. Will building good fences really make good neighbors in science. In Report to European Commission (DG-Research) STRATA-ETAN workshop on IPR aspects of internal collaborations.

Ducor, P., 2000. Coauthorship and coinventorship. Science 289(5481), 873–875.

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

Gans, J., Hsu, D., Stern, S., 2002. When does start-up innovation spur the gale of creative destruction? Rand Journal of Economics 33(4), 571–586.

George, G., Zahra, S. A, Wood, D. R., 2002. The effects of business–university alliances on innovative output and financial performance: a study of publicly traded biotechnology companies. Journal of Business Venturing 17(6), 577–609.

Gieryn, T. F., 1981. The aging of a science and its exploitation of innovation: Lessons from X–ray and radio astronomy. Scientometrics 3(4), 325–334.

Giuri, P., Mariani, M., Brusoni, S., Crespi, G., Francoz, D., Gambardella, A., Garcia-Fontes, W., Geuna, A., Gonzales, R., Harhoff, D., Hoisl, K., Le Bas, C., Luzzi, A., Magazzini, L., Nesta, L., Nomaler , O., Palomeras, N., Patel , P., Romanelli, M., Verspagen, B., 2007. Inventors and invention processes in Europe: Results from the PatVal-EU survey. Research Policy 36(8), 1107–1127.

Goetze, C., 2010. An empirical enquiry into co-patent networks and their stars: The case of cardiac pacemaker technology, Technovation 30(7) 436–446.

Grant, R. M., Baden-Fuller, C., 1995. A knowledge-based theory of inter-firm collaboration. Academy of Management Proceedings 1995(1), 17–21.

Guellec, D., van Pottelsberghe de la Potterie, B., 2000. Applications, grants and the value of patents. Economics Letters 69(1), 109–114.

Gulati, R., Gargiulo, M., 1999. Where do interorganizational networks come from? American Journal of Sociology 104, 1398–1438.

Haeussler, C., Jiang, L., Thursby, J., Thursby, M., 2011. Sharing among competing researchers. Working Paper.

Haeussler, C., Sauermann, H., 2013. Credit where credit is due? The impact of project contributions and social factors on authorship and inventorship. Research Policy 42(3), 688–703.

Hagedoorn, J., Cloodt, M. (2003). Measuring innovative performance: is there an advantage in using multiple indicators? Research Policy 32(8), 1365–1379

Hall, B.H., Jaffe, A., Trajtenberg, M., 2005. Market value and patent citations. RAND Journal of Economics 36, 16–38.

Hall, B., Trajtenberg, M., 2004. Uncovering GPTs with Patent Data. In: NBER Working Paper No. 10901. Cambridge, MA.

Harhoff, D., Narin, F., Scherer, F. M., Vopel, K., 1999. Citation frequency and the value of patented inventions. Review of Economics and Statistics 81(3), 511–515.

Hausman, J. A., Hall, B., Griliches, Z., 1984. Econometric models for count data with an application to the patents-R&D relationship. Econometrica 52(4), 909–938.

Heisey, P. W., Adelman, S. W., 2011. Research expenditures, technology transfer activity, and university licensing revenue. The Journal of Technology Transfer 36(1), 38–60.

Heller, M., Eisenberg, R., 1998. Can patents deter innovation? The anticommons in biomedical research. Science 280(5364), 698–701.

Hellman, T., 2007. Bridging the science to market gap: the role of patents. Journal of Economic Behavior and Organization 63, 624–647.

Henderson, R., Cockburn, I., 1994. Measuring competence? Exploring firm effects in pharmaceutical research. Strategic Management Journal 15(S1), 63–75 (Special issue).

Henderson, R., Jaffe, A. B., Trajtenberg, M., 1998. Universities as a source of commercial technology: a detailed analysis of University Patenting, 1965–1988. Review of Economics and Statistics 80(1), 119–127.

Hirschey, M., Richardson, V. J., 2004. Are scientific indicators of patent quality useful to investors? Journal of Empirical Finance 11(1), 91–107.

Hsu, D., Bernstein, T., 1997. Managing the university technology licensing process: findings from case studies. Journal of the Association of University Technology Managers 9(9), 1–33.

Jaffe, A.B., Trajtenberg, M., Henderson, R., 1993. Geographic localization of knowledge spillovers as evidenced by patent citations. The Quarterly Journal of Economics 108, 577–598.

Jaffe, A., 1998. Real effects of academic research. The American Economic Review 79(5), 957–970.

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

Kang, K. N., Ryu, T. K., Lee, Y. S., 2009. Effects of research tool patents on biotechnology innovation in a developing country: A case study of South Korea. BMC Biotechnology 9(1), 25.

Kitch, E. W., 1977. The nature and function of the patent system. Journal of Law and Economics 20, 265–290.

Kyvik, S., Olsen, T. B., 2008. Does the aging of tenured academic staff affect the research performance of universities? Scientometrics 76(3), 439–455.

Lanjouw, J.O., Schankerman, M., 2001. Characteristics of patent litigation: a window on competition. RAND Journal of Economics 32(1), 129–151.

Lanjouw, J. O., Schankerman, M., 2004. Patent quality and research productivity: Measuring innovation with multiple indicators. The Economic Journal 114(495), 441–465.

Leopold, E., May, M., Paass, G., 2004. Data Mining and Text Mining for Science & Technology Research. In Moed H.F., Glänzel W., Schmoch U., Handbook of Quantitative Science and Technology Research. Kluwer. Dordrecht 8.

Liebeskind, J. P., Oliver, A. L., Zucker, L., Brewer, M., 1996. Social networks, learning and flexibility: sourcing scientific knowledge in new biotechnology firms. Organization Science 7(4), 428–443.

Lissoni, F., Montobbio, F., 2006. Inventorship attribution in academic patents: quantitative analysis of patent-publication pairs.

Lissoni, F., Montobbio, F., 2008. Inventorship and authorship in patent-publication pairs: an enquiry into the economics of scientific credit. CESPRI – University Bocconi. Working Paper.

Lissoni, F., Montobbio, F., Seri, R., 2010. Ownership and impact of European university patents. Paper presented at EAEPE Conference 2010, Universitè Bordeaux IV, October 28–30.

Lubango, L. M, Pouris, A., 2010. Is patenting of technical inventions in university sectors impeding the flow of scientific knowledge to the public? A case study of South Africa, Technology in Society 32(3), 241–248.

Magerman, T., Van Looy, B., Debackere, K., 2011. In search of anti-commons: patent-paper pairs in biotechnology. An analysis of citation flows. DRUID Society, Copenhagen Business School, Denmark.

Merton, R.K., 1973. The Sociology of Science: Theoretical and Empirical Investigations. Chicago University Press, Chicago/London.

Meyer, M., 2006. Knowledge integrators or weak links? An exploratory comparison of patenting researchers with their non-inventing peers in nano-science and technology. Scientometrics 68(3), 545–560.

Motohashi, K., 2005. University–industry collaborations in Japan: The role of new technology-based firms in transforming the National Innovation System. Research Policy 34(5), 583–594.

Murray, F., 2002. Innovation as co-evolution of scientific and technological networks: exploring tissue engineering. Research Policy 31(8), 1389–1403.

Murray, F., Stern, S., 2007. Do formal intellectual property rights hinder the free flow of scientific knowledge? A test of the anti-commons hypothesis. Journal of Economic Behavior and Organization 63(4), 648–687.

Narin, F., Noma, E., Perry, R., 1987. Patents as indicators of corporate technological strength. Research Policy 16(2), 143–155.

Narin, F., Olivastro, D., 1992. Status report: linkage between technology and science. Research Policy 21(3), 237–249.

Narin, F.S., Hamilton, K.S., Olivastro, D., 1997. The increasing linkage between U.S. technology and public science. Research Policy 26(3), 317–330.

Niosi, J., Reid, S., 2007. Biotechnology and Nanotechnology: Science-based Enabling Technologies as Windows of Opportunity for LDCs? World Development 35(3), 426–438.

Nooy, W., Mrvar, A., Batagelj, V., 2005. Exploratory Social Network Analysis with Pajek, Cambridge University Press.

Oliver, A. L., 2004. Biotechnology entrepreneurial scientists and their collaborations. Research Policy 33(4), 583–597.

Oliver, A. L., Liebeskind, J. P., 1998. Three levels of networking: implications for the study of interorganizational networks. International Studies in Management and Organization 27(4), 76–103.

Peters, L., Groenewegen, P., Fiebelkorn, N., 1998. A comparison of networks between industry and public sector research in materials technology and biotechnology. Research Policy 27(3), 255–271.

Podolny, J. M., Stuart, T. E., 1995. A role-based ecology of technological change. American Journal of Sociology 100(5), 1224–1260.

Schilling, M. A., Phelps, C. C., 2007. Interfirm collaboration networks: the impact of large-scale network structure on firm innovation. Management Science 53(7), 1113–1126.

Stephan, P., 2012. How Economics Shapes Science. Harvard University Press, Cambridge.

Stern, S., 1978. Age and achievement in mathematics: A case–study in the sociology of science. Social Studies of Science 8(1), 127–140.

Sterzi, V., 2013. Patent quality and ownership: An analysis of UK faculty patenting. Research Policy 42(2), 564–576.

Terza, J. V., Basu, A., Rathouz P. J., 2008. Two-stage residual inclusion estimation: Addressing endogeneity in health econometric modelling. Journal of Health Economics 27(3), 531–543.

Thompson, N. C., Mowery, D. C., Ziedonis, A. A., 2011. Patents, Material Transfer Agreements (MTAs), and the Flow of Scientific Knowledge.

Trajtenberg, M., 1990. A penny for your quotes: patent citations and the value of innovations. RAND Journal of Economics 21(1), 172–187.

Van Looy, B., Ramga, M., Callaert, J., Debackere, K., Zimmermann, E., 2004. Combining entrepreneurial and scientific performance in academia: towards a compounded and reciprocal Matthew-effect? Research Policy 33(3), 425–441.

Wagner-Doebler, R., 2001. Continuity and discontinuity of collaboration behaviour since 1800 - from a bibliometric point of view. Scientometrics 52(3), 503–517.

Wang, J. C., Chiang, C. H., Lin, S. W., 2010. Network structure of innovation: Can brokerage or closure predict patent quality? Scientometrics 84(3), 735–748.

Zahra, S., 1996. Technology strategy and performance: a study of corporate-sponsored and independent biotechnology ventures. Journal of Business Venturing 11(4), 289–321.

Zucker, L., Darby, M., Brewer, M., 1998. Intellectual human capital and the birth of US. Biotechnology Enterprises. American Economic Review 88, 290–306.

Zuckerman, H. A., 1977. Scientific elite: Nobel laureates in the United States. New York: Free Press.

# 8 Appendix

**Table 8 – Variable description**

| Variable | Description |
|---|---|
| **Dependent Variables** | |
| *$NbFCit5_t$* | Number of forward citations in citing patents, during the 5 years following the granting year |
| *$NbClaims_t$* | Number of claims in the patent document |
| *$InOriginality_t$* | Principal components of the number of forward citations, the number of claims and of 1 minus the Herfindahl index of backward citations of the patent. |
| **Independent Variables** | |
| *$dPPP_t$* | Dummy variable taking the value 1 if the patent is part of a patent-paper pair (i.e. when the similarity measure between the patent and paper documents is greater than or equal to 0.30), and 0 otherwise |
| *$AveGrant3_t$* [a] | Average value amongst the academic-inventors of the patent of the amount of grants received by the academic inventors of the patent over the three years prior to the patent application |
| *$AveAge_t$* [a] | Average 'career' age of the academic inventors of the patent |
| *$MaxChair_t$* [c] | Maximum value amongst the academic inventors of the patent of the ordinal variable representing the 'best' chair occupied by an academic (0 = no chair, 1 = industrial chair, 2 = NSERC or CIHR chair, 3 = Canada Research Chair) |
| *$AveNbACit3_t$* [a] | Average value amongst the academic-inventors of the patent of the number of article citations received by their publications |
| *$AveCliqA3_t \times 10^3$* [a,c] | Average value amongst the academic inventors of the patent of the 3-year **co-publication** (articles) network individual clustering coefficient (cliquishness) |
| *$AveCliqP3_t \times 10^3$* [a,c] | Average value amongst the academic inventors of the patent of the 3-year **co-invention** (patents) network individual clustering coefficient (cliquishness) |
| *$AveBtwP3_t$* [a,c] | Average value amongst the academic inventors of the patent of the 3-year **co-invention** (patents) network individual betweenness centrality |
| *$propAcInv_t$* | Proportion of academic inventors over total number of inventors of the patent |
| *$dAcAssignee_t$* | Dummy variable taking the value 1 if the assignee of the patent is an academic institution |
| *dNanoEx* | Dummy variable taking the value 1 if the domain of the patent is exclusively nanotechnology (i.e. excluding nanobiotechnology) |
| **Endogenous Variables** | |
| $\mathrm{AveCont3}_t$ [a] | Average value amongst the academic-inventors of the patent of the amount of contracts received by the academic inventors of the patent over the three years prior to the patent application |
| **Instrumental Variables** | |
| *$AveNbLoopPast_t$* [a] | Average value amongst the academic inventors of the patent of the number of innovation loops (such a loop exists when the research of the named academic inventors of the patent has been funded by the assignee of the patent) to which they have contributed |
| *$AveCont3U_t$* [a] | Average value amongst the academic-inventors of the patent of the total amount of contracts received by their university over the past three years |
| *$AveGrantEI3_t$* [a] | Average value amongst the academic-inventors of the patent of the amount of grants for equipment and infrastructure over the past three years |

Notes: *t* is the application year. [(a)] All the variables have been averaged over all academic inventors that contributed to a given patent; [(b)] Only this variable uses the maximum value of all academic inventors (MaxChair); [(c)] We added 1 to all these variables to insure that their minimum value is greater than 1.

**Table 9 – Correlation table**

| Variable | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|---|---|---|---|---|
| $\ln(NbFCit5_t)$ | **1** | 1 | | | | | | | | |
| $\ln(NbClaims_t)$ | **2** | 0.1417* | 1 | | | | | | | |
| *HerfBCit5t* | **3** | 0.0533* | 0.1754* | 1 | | | | | | |
| $InOriginality_t$ | **4** | 0.5030* | 0.6524* | 0.6096* | 1 | | | | | |
| $\ln(AveGrant3_t)$ | **5** | -0.1384* | -0.0286 | 0.0603* | -0.0491 | 1 | | | | |
| $\ln(AveCont3_t)$ | **6** | 0.0309 | 0.0870* | 0.1404* | 0.1434* | 0.3270* | 1 | | | |
| $AveAge_t^{\,a}$ | **7** | -0.3207* | -0.0115 | 0.0189 | -0.1010* | 0.2292* | 0.0575* | 1 | | |
| $MaxChair_t^{\,c}$ | **8** | 0.0100 | 0.0450 | -0.0586* | 0.0213 | 0.2294* | 0.2849* | 0.1679* | 1 | |
| $\ln(AveNbACit3_t)$ | **9** | 0.0258 | -0.0119 | -0.0532* | -0.0208 | 0.1817* | 0.0464 | 0.2453* | 0.0246 | 1 |
| $1+\ln(AveCliqA3_t \times 10^3)$ | **10** | 0.1670* | -0.1690* | -0.1584* | -0.1345* | -0.0685* | -0.2720* | -0.0785* | -0.1637* | 0.2207* |
| $1+\ln(AveBtwP3_t)$ | **11** | -0.0725* | -0.0108 | -0.0564* | -0.0745* | 0.0851* | -0.1582* | 0.1955* | -0.1551* | 0.1069* |
| $1+\ln(AveCliqP3_t \times 10^3)$ | **12** | -0.0215 | 0.0975* | 0.1606* | 0.1413* | 0.0679* | 0.0899* | 0.0764* | -0.0221 | -0.0677* |
| $propAcInv_t$ | **13** | 0.0461 | -0.0623* | 0.0185 | -0.0042 | 0.2612* | 0.1326* | -0.0206 | 0.1745* | 0.1906* |
| $dPPP_t$ | **14** | -0.0801* | -0.1086* | -0.1160* | -0.1481* | -0.0016 | -0.0070 | -0.0403 | -0.0405 | 0.1221* |
| $dAcAssignee_t$ | **15** | 0.0169 | 0.0286 | -0.0080 | 0.0542* | -0.0094 | 0.1416* | -0.1216* | 0.1321* | 0.0620* |
| *dNanoEx* | **16** | 0.1219* | 0.1950* | 0.2146* | 0.2685* | -0.0147 | 0.0370 | -0.0738* | -0.0087 | -0.0022 |
| $\ln(AveNbLoopPast_t)$ | **17** | -0.0569* | 0.0072 | 0.2139* | 0.0889* | 0.0564* | 0.3278* | -0.0355 | 0.0471 | 0.0640* |
| $AveCont3U_{t-2}$ | **18** | -0.1146* | 0.0767* | 0.1145* | 0.0610* | 0.1593* | 0.2105* | 0.2273* | 0.1298* | -0.0096 |
| $\ln(AveGrantEI3_{t-1})$ | **19** | -0.0628* | 0.0123 | 0.0526* | 0.0184 | 0.3732* | 0.2760* | 0.1008* | 0.3034* | 0.0615* |

| Variable | | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $1+\ln(AveCliqA3_t \times 10^3)$ | **10** | 1 | | | | | | | | | |
| $1+\ln(AveBtwP3_t)$ | **11** | 0.3532* | 1 | | | | | | | | |
| $1+\ln(AveCliqP3_t \times 10^3)$ | **12** | -0.1274* | 0.0673* | 1 | | | | | | | |
| $propAcInv_t$ | **13** | -0.1477* | -0.2989* | -0.1271* | 1 | | | | | | |
| $dPPP_t$ | **14** | 0.0263 | -0.0106 | -0.0399 | 0.0183 | 1 | | | | | |
| $dAcAssignee_t$ | **15** | -0.2295* | -0.1897* | 0.0085 | 0.2141* | 0.1743* | 1 | | | | |
| *dNanoEx* | **16** | -0.1341* | -0.0933* | 0.0335 | 0.0146 | -0.1320* | 0.0382 | 1 | | | |
| $\ln(AveNbLoopPast_t)$ | **17** | -0.3001* | -0.0272 | 0.1237* | 0.0578* | 0.0328 | 0.1552* | 0.0505* | 1 | | |
| $AveCont3U_{t-2}$ | **18** | -0.2280* | -0.0282 | 0.0891* | 0.0773* | -0.1251* | -0.0679* | 0.2187* | 0.0639* | 1 | |
| $\ln(AveGrantEI3_{t-1})$ | **19** | -0.1903* | -0.0062 | 0.1380* | 0.1883* | -0.0819* | -0.0157 | 0.0994* | 0.0434 | 0.3223* | 1 |

Note: * corresponds to a 1% significance level.

**Table 10 – Factor loadings of the originality index (*InOriginality*$_t$) – Principal component analysis**

| Component | Eigenvalue | Difference | Proportion | Cumulative |
|---|---|---|---|---|
| Comp1 | 1.2454 | 0.2938 | 0.4151 | 0.4151 |
| Comp2 | 0.9516 | 0.1485 | 0.3172 | 0.7323 |
| Comp3 | 0.8030 | . | 0.2677 | 1 |
| **Rotated components** | | | | |
| **Variable** | **Comp1** | **Unexplained** | | |
| $NbClaims_t$ | 0.6671 | 0.4458 | | |
| $NbFCit5_t$ | 0.5067 | 0.6802 | | |
| *HerfBCit5t* | 0.5461 | 0.6286 | | |