Anna Kiss

A Literature Review of Scientometric Methods to Model Academic Careers*

Introduction

The classic, prevailing question of Hirsch (2005): "For the few scientists that earn a Nobel prize, the impact and relevance of their research work are unquestionable. Among the rest of us, how does one quantify the cumulative impact and relevance of an individual's scientific research output?" The academic career is the product of the socio-economic-cultural background of a given society (or a set of societies, participating in the development of the personality and the career) and, on the other hand, an important vehicle of science. That's why this study is at a focal point of scientometrics, sociology, and policy analysis. Analysis of careers in the field of science is gaining in importance and popularity because the in-depth knowledge of mechanisms governing scientific career paths is important for planning and the realization of science policy, thereby increasing knowledge, economic and social output (Dietz 2000; Antonelli et al. 2011) and the S–T capacity as well as human capital (Bozeman–Rogers 2002). As Hirsch formulates it: "In a world of not unlimited resources such quantification (even if potentially distasteful) is often needed for evaluation and comparison purposes, e.g. for university faculty recruitment and advancement, award of grants, etc.". Nevertheless, it is hard to answer the question, how to measure academic performance.

Research questions

Academic careers can be characterized on the basis of different sciences and approaches. One of the research questions is what the methods and tools of measuring academic performance are. Traditionally, academic performance can be measured by the number of (quality) publications and their impact on science, which is manifested in the number of citations (Van Balen–Leydesdorff 2009). This view of academic careers can be contested, because in the more "application-oriented" fields of science the number of publications is just one measure of academic performance. In high-tech industries, the number of patents is a competing measurement dimension of academic performance. According to the traditional approach, there is a strong correlation between the number of publications and the number of patents,

* The first publication of this article: Anna Kiss (2018). Methods of scientometrics to model academic careers – A literature review. *Kultúra és Közösség*, 2018/4, 13–30.

but the cointegration analysis, focusing on some rapidly advancing field of technology (e.g. pharmaceutical industry) is not able to prove a statistically significant relationship between the yearly number of publications and the number of patent applications in a given nation or research group. A further, more complex question is the efficiency of using as a measurement the number of patent applications. It is well–documented that just a very low share of patents will be applied in practice. In some fields, e.g. in plant production or animal husbandry, the number of new varieties of breeds could be applied as a measure of academic productivity.

If we accept to measure the academic career on the basis of the number of publications, it is a further problem, how to measure academic performance: on the basis of the total number of papers during the lifetime, or on the basis of productivity per a given time interval. Both measures offer some advantages and disadvantages, the time-based measures of academic productivity are capable to quantify the regularity of the authors. The time of determination of end-point of academic career leaves opens one important question: whether the end of a career is the publication of the last paper in the lifetime of the researcher should be at the time of retirement, and all the additional papers should be considered as a product of some hobby activity.

We will focus on the number of academic papers produced during the lifetime of the researchers because these data lend themselves for a comparative approach, and if necessary a quantitative analysis can be performed. However, we have to take into consideration that this approach is a relatively narrow one: in future research, a more holistic approach should be applied, taking into consideration other outputs, e.g. teaching activity, preparation of textbooks, as well as such activities. consulting, running spin-off companies, or the popularization of the sciences (Enders 2005; Glänzel et al. 2007).

A considerable part of the publications on academic careers applies an ontological approach, emphasizing the importance of the roots of academic careers. There is a wide consensus that the academic career is a product of a complex set of socio-economic factors. Some studies apply a more qualitative approach to this problem and try to grasp the motivational base and early results of academic careers by measuring the cultural capital of the family as well as the effect of narrower and wider socio-economic environment, emphasizing the influence of culture on publication behavior and life strategy (Van Balen et al. 2012; Leahey 2006).

Another important research question is the role different "vehicles" play in the academic career. According to Van Balen et al. (2012) and Wells et al. (2011) such individual factors, like cultural and social capital, results of the effect of parents (Amarnani et al. 2016) and mentoring (Ehrich et al. 2004) as well as networking will exercise a considerable impact on the development of academic careers. Another important factor of career development is the organizational environment, which could be measured by the performance, prestige, or network position of the university (Van Balen et al. 2012). Besides contextual factors, like labor market fluctuations should be taken into account, too. The overwhelming majority of the relevant publications have been written in the US, where a relatively high level of financial stability and individual mobility is a general condition. According to the experiences of some other countries (e.g. in crisis-hidden European research centers or universities) these general conditions do not exist anymore, that's why the fluctuations in financial resources or the dry-up of some sources for a given research activity could lead to the termination to an academic career *(Figure 1)*.

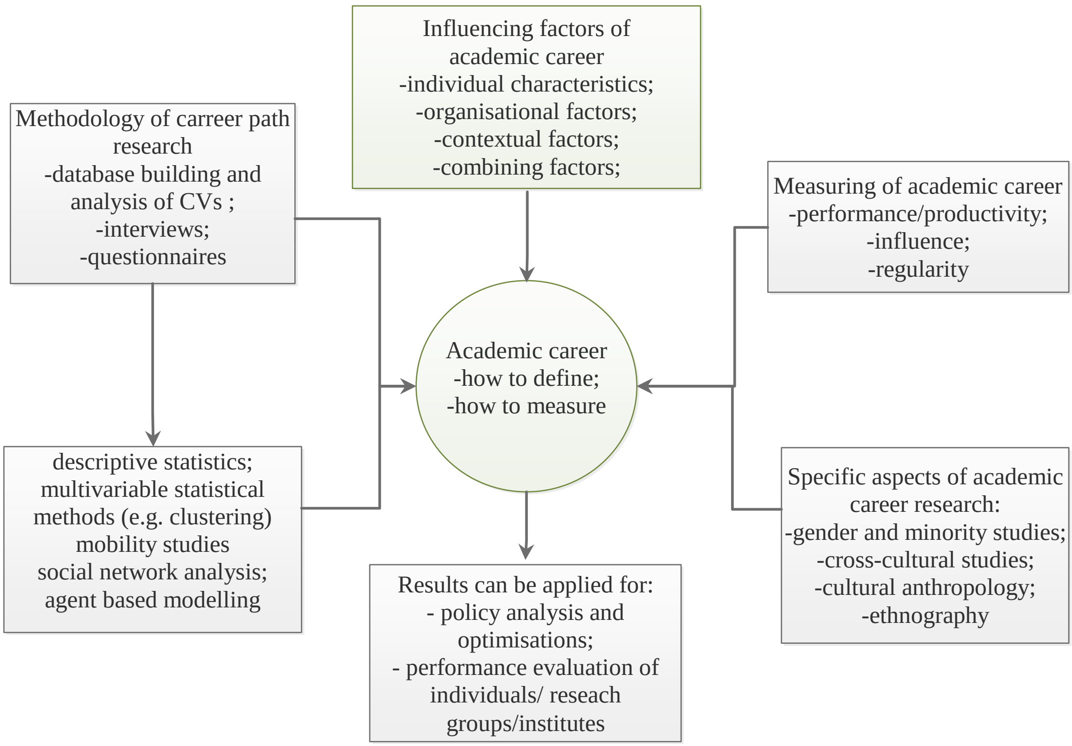

Figure 1: The "academic career puzzle"
Source: own construction

Methods

The current investigation generally followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines (Moher et al. 2009). This reporting standard is widely used in the field of medical and health-care, and it is commonly accepted as a useful reporting guideline in those disciplines to enhance the completeness of the reporting of systematic reviews. The PRISMA guideline is an appropriate method to include the relevant literature with adequate accuracy, and it can exclude any that is not relevant.

According to the PRISMA statement, a flow diagram needs to be provided to represent the steps of elimination, these being 'identification, screening, eligibility and final inclusion".

Information Sources and Search

Literature searches were conducted in PubMed, Scopus, Web of Science, ScienceDirect, and Google Scholar. No limitations were placed on the dates of the searches, and the final search was completed in December 2016. After reviewing Scopus social subject headings for 'academic career' and 'scientific career', keywords selected for the search included research productivity, performance, success, patents, curriculum vitae, mobility, citation and collaboration. These keywords were combined with bibliometric, mathematic, scientometrics, research value mapping, and social network analysis.

To find additional studies, the reference lists of the articles obtained were searched, as was the literature database of an investigator with extensive experience with academic career research.

Eligibility Criteria

Articles were selected for the review if they were (1) written in English, (2) involved bibliometric, mathematics, or expositive methods to describe academic careers, and (3) provided a quantitative assessment. Titles were first examined and abstracts were reviewed if the article appeared to involve academic career and either scientometrics or bibliometric. The full text of the article was retrieved if there was a possibility that scientometrics analysis had been included within the investigation. Quantitative data could be contained within the text of the article, in tabular form, or presented in graphs. Data presented in graphic form were estimated. If the authors did not specifically aim to measure academic career, but data were available in the article to calculate it, then the article and the data were included in the review. Abstracts, case studies, and case series were not included. Stand-alone abstracts (without full-text articles) were excluded because they were difficult to locate, were generally not included in reference databases, and in many cases were not peer-reviewed. Case studies and case series involved few individuals and were often published because they were atypical.

Results

Figure 2 shows the number of publications included and excluded at each stage of the literature search. The initial search identified 21,694 citations, 5339 of which were duplicate publications (from different databases) that were removed. Based on a review of titles and abstracts, 345 full articles were obtained for review, and subsequently, 135 were removed for not having relevance for research purposes or meeting the exclusion criteria. A total of 210 studies were further reviewed, but 127 of these did not contain either relevant or useful data. In total, 83 unique studies finally met the inclusion criteria.

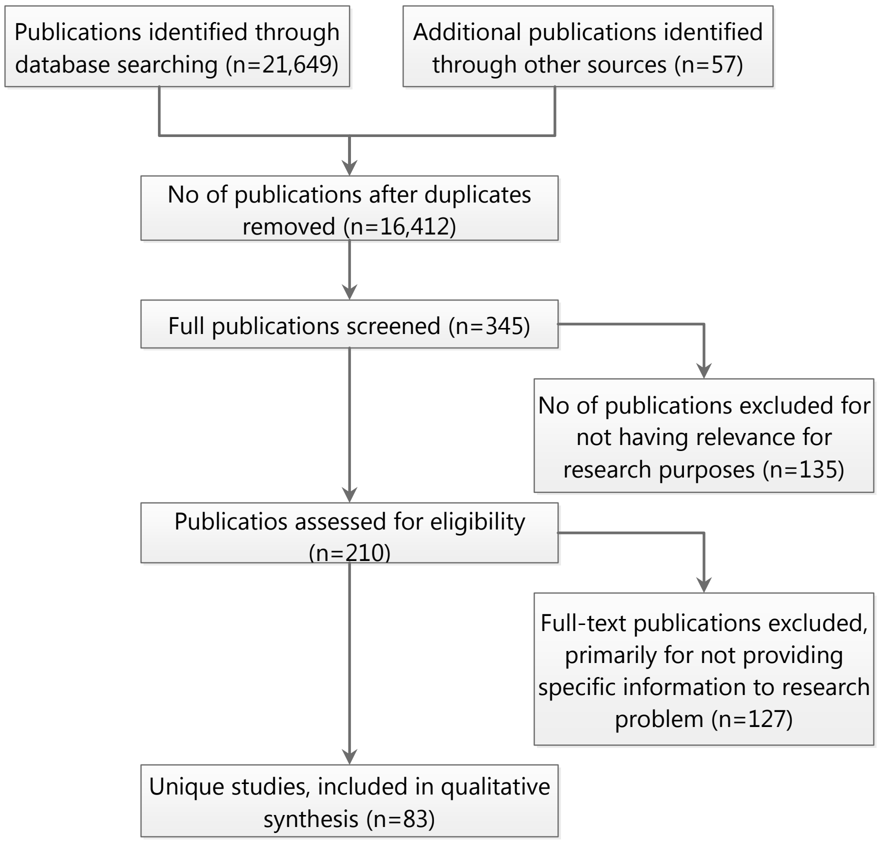

Figure 2: Publications included and excluded at each stage of literature review

The analysis of academic careers

It is widely recognized, that academic performance can be measured by two dimensions: overall productivity and the impact of works. According to Dietz and Boseman (2005), studies on academic careers often begin with the question as to why there seems to be a skewed distribution of research productivity across the population of an academic scientist. In his seminal paper Lotka (1928), cited by Seglen (1992) highlighted that the vast majority of papers are published by a small minority of researchers. The vast majority of papers on this topic up to the 90s of the last century had been focusing on different sociological aspects of this question (Keith–Babchuk 1998). A considerable part of papers had been focusing on the sociological structures of science (Merton 1961), analyzing science as a sociological entity. This approach considered science as an academic enterprise (Merton 1957, 1961), not taking into consideration the social embeddedness of science. In our opinion, this can be measured on base of publications, as opposed to some attempts (e.g. Dietz–Bozeman 2005) trying to involve into this topic the issue of patents. According to Baruch and Hall (2004), the academic career system has unique features, but empirical studies about academic careers are hardly available. Earlier studies have been made to model academic careers, but those were personal and introspective. Publications on academic career development are less focused on the development of the entire career. Van Balen et al. (2012) described which factors influence a successful academic career, the main question their paper aimed to answer was: Why do some talented researchers have a continued academic career, whereas others do not? The study was based on 42 semi-structured interviews, their results suggest that the academic careers of talented researchers are stimulated or inhibited by an accumulation of advantages or disadvantages.

In the last decades, as a result of the collaboration of bibliometricians, information scientists, sociologists, physicists, and computer scientists, comprehensive science maps have been developed (Börner et al. 2005). Guevara et al. (2016) developed the concept of *research space* as a more suitable approach for the evaluation of the performance of individual researchers, teams or nations because this is based on publication patterns of individuals. *Table 1* shows studies on academic careers separated by study design.

Table 1: Studies on academic career separated by study design

CV analyses – mobility	Dietz et al. 2000; Canibano–Bozeman 2009; Gaughan–Bozeman 2002; Wooley–Turpin 2009; Bonzi 1992; Dietz–Bozeman 2005; Fernandez-Zubieta et al. 2013; Corley et al. 2003; Gaughan–Ponomariov 2008; Mangematin 2001; Enders–Weert 2004; Enders 2005; Ackers 2005; Ackers–Oliver 2007; Gaughan–Robin 2004; Fernandez-Zubieta et al. 2015; Sandström 2009; Morano-Foadi 2005; Ackers 2005; Canibano et al. 2008
Bibliometric – mathematic	Hack et al. 2010; Chakraborty et al. 2014; Petersen 2015; Efron–Brennan 2011; You et al. 2015; Zhang–Glänzel 2012;

	Franceschini–Maisano, 2011; Burrel 2007; Matia et al. 2005; Liang 2006; Petersen et al. 2011; Petersen et al. 2010; Ding et al. 2011; Egghe 2010; Petersen et al. 2012
Gender studies	De Pater 2005; Leahey 2006; Cole–Zuckerman 1984; Xie–Shauman 2003, 1998; Fox 1983, 1985, 2001, 2005; Bentley 2011; McBrier 2003; Long–Fox 1995; Prpic 2012; Long 1992; Symonds et al. 2006; Teodorescu 2000; Kyvik–Teigen 1996; Probert 2005; Sonnert 1995; Symonds 2006; Duch et al. 2012; Sax et al. 2002; Ackers 2007
Cultural analyses	Leong–Leung 2004
Geography	Carvalho–Batty 2006

CV analysis

According to Dietz et al. (2000) CVs are particularly useful for the analysis of academic careers since they provide a complex picture of the life trajectory of researchers. The combined application of data collected from CVs and bibliographic measures improve data accuracy, helps to avoid mismatches, and offers valuable information to explain the changes in publication patterns and co-authors' space. At the same time, Dietz et al. (2000) state that the analysis of curriculum vita to study career path is an extremely difficult task, due to the hard quantification of different stages of individual lives Their article offers a detailed description of ways and means of elimination of intercoder errors, and present a model, describing the effect of different factors on publication rate. Results prove a significant, positive regression coefficient (determined by OLS) between the pre-Ph.D publications as well as a number of patents, and a negative coefficient in time of duration in the rank of assistant professor. The number of jobs has not been an important factor for productivity. In our opinion, the years spent as an assistant professor cannot be considered as an explanatory variable, because it could be rather a consequence of the relatively low academic performance.

Statistical methods to measure academic career

Analyzing the relevant literature, it is beyond doubt, that there is a wide and ever-increasing field of career research. This can be explained by the steadily increasing level of interest towards the problems of academic careers and the complexity of this question: this field of science lends itself to apply the tools and paradigms offered by different sciences. In *Figure 3* we have summarized the field of application of different methods in career research.

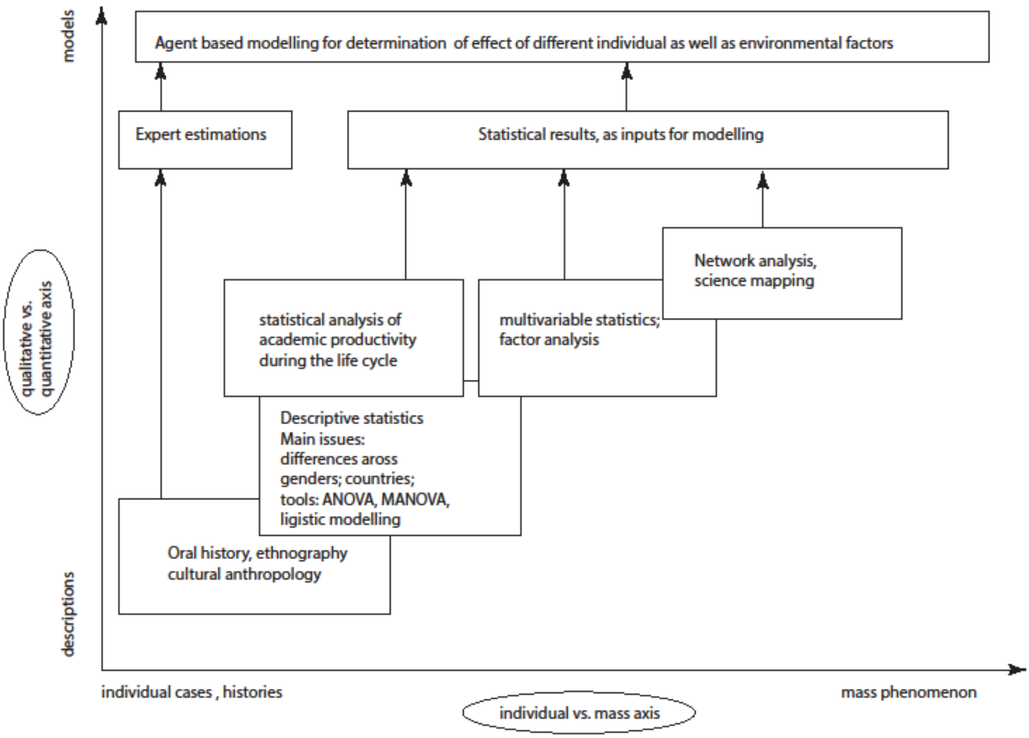

Figure 3: The field of application of different methods in career research
Source: own construction

Development of databases

Recently there has been an important emergence of complex, unified, large-scale databases, offering the possibility of inter-individual as well as for inter-institutional comparison in the analysis of academic careers on the basis of bibliometric data. As a result, we witness the birth of the science of science measurement (Nature 2010). Nowadays the two leading academic publication databases are the Web of Science and Scopus, but there is an increasing number of databases for geographic locations (e.g. Brasil: http://lattes.cnpq.br/; Hungary: https://www.mtmt.hu/).

Analysis of data on academic performance

The modern methods of scientometrical analysis are applying the statistical methods at an increasing rate. The rapid accumulation of information on citation patterns offers a favorable possibility to apply different statistical methods to citation patterns. Wallace et al. (2009) have proven that the citations can be characterized by a stretched-exponential function and a form of the Tsallis function to fit complete distributions over the 20th century. The Hirsch-core is well known for more than one decade (Glänzel 2006): Liang (2006) has introduced the h-index sequence for measuring the dynamics in the scientific career. According to her theory, the h index sequence h_k is the h-index of the papers published by the author in question in n-k+1, n time interval, where n is the most recent year. This is a logical continuation of Burrell's (2007) approach. According to their results (the h-index and its derivatives can be of great importance to track the life cycle of research teams. However, there are considerable differences between averages of citations for one paper in different fields (e.g. according to Iglesias and Pecharromán (2007) based on ISI the expected number of a citation for a paper in economics was 4,17 on average in the period between 1995-2005, the value of this indicator for molecular biology and genetics was 24,57). At the same time, the wide range of utilization of citation indices is fueled by the fact that - as Radicchi et al. (2008) have proven - there is a universality of citation distributions across disciplines and years.

Measurement of regularity in academic performance

Franceschini and Maisano (2011) claim that the regularity of academic performance is gaining importance. They define regularity as the ability to generate an active and stable research output over time, in terms of both quantity and diffusion. To measure regularity they suggest three tools: the PY/CY diagram depicts the distribution of publications and citations according to different years. The Ferrers diagram is a generalization of Hirsch's h-index in a two-dimensional space (Egghe 2010a). The Borda score is a simple sum of the yearly rank of different scientists' performance compared to his/her competitors.

All of the proposed methods have some positive effects and limitations: the PY/CY diagram lends itself of an indication on the temporal evaluation of a scientists career; Ferrers diagram makes it possible to identify the most important years for publication and their diffusion; Borda's method offers a favorable possibility to compare different scientists' performance according to two dimensions: publications and citations.

Modeling the patterns of individual academic trajectories

Petersen et al. (2010) offer normalized publication metrics to achieve a universal framework of analyzing and comparing scientific achievement across both time and discipline. They have determined, that the scaling exponent for individual papers ($\gamma \approx 3$) is larger than the scaling exponent for total citation shares ($\alpha \approx 2.5$) and that for total paper shares ($\alpha \approx 2.6$), which indicates that there is a higher frequency of stellar careers than stellar papers (Petersen et al. 2011). Zhi-Quiang You et al. (2015) claim that in the field of science, from the point of view of quantitative analysis, there are two basic fields: (1) network-theoretic analysis and (2) soft-modeling of large datasets. They have applied an agent-based model to capture the most important aspects of publication and citation networks. In the model, the agents were authors or research teams, and the nodes were the publications of citation networks. The inheritance process had been manifested through the spread of citation relationships. In a subsequent publication, Peresen et al. (2011) offer a piece of strong empirical evidence for universal statistical laws that describe career progress in competitive professions. The career paths often can be characterized by bimodal distributions: one class of careers is stunned by the difficulty in making progress at the beginning of a career. Based on the dynamics of publications they separate convex as well as concave progress.

Petersen et al. (2011) have introduced the $N_i(t) \approx A_i\ [t(\exp \alpha_i)]$ temporal scaling relation, where α_i is a scaling exponent that quantifies the career trajectory dynamics. The estimation of α shows a relatively large similarity across disciplines, its value is between 1.3 and 1.44. According to Petersen et al. (2011), there is a possibility that short-term contracts may reduce the motivation for a young scientist to invest in human and social capital accumulation. In a summary, it can be stated, that there is an urgent need to group productivity measures, too.

The analysis of researchers' mobility and academic career

As it is demonstrated in *Figure 4,* there are different approaches to career development analysis. A specific one is the analysis of *thematic mobility patterns*, based on the *scientific mapping*. In the last decade, there was an effort to introduce some more quality-oriented methods into the evaluation of bibliometric data. That's why the g–index has been introduced by Egghe (2006). This index is the highest number of g of articles (a set of articles ordered by decreasing citation counts) that together received 2 or more citations. However bibliometric has more than half a century traditions, its application shows considerable differences between disciplines and countries (Abbott et al. 2010). Notwithstanding the bibliometric, as science has anglo-saxon roots, many British, Commonwealth, and US institutes use this for the evaluation of performances of universities as well as research organizations, but in personal-related decisions, the „soft" factors of personality evaluation (e.g. recommendation letters) are considered as more important factors. Sahel (2011) claims that the professional analysis of bibliometric data is important, but – in line with the recommendations of the French National Academy (FAS) – he discourages the application of this data concerning personal decisions on the young scientist.*

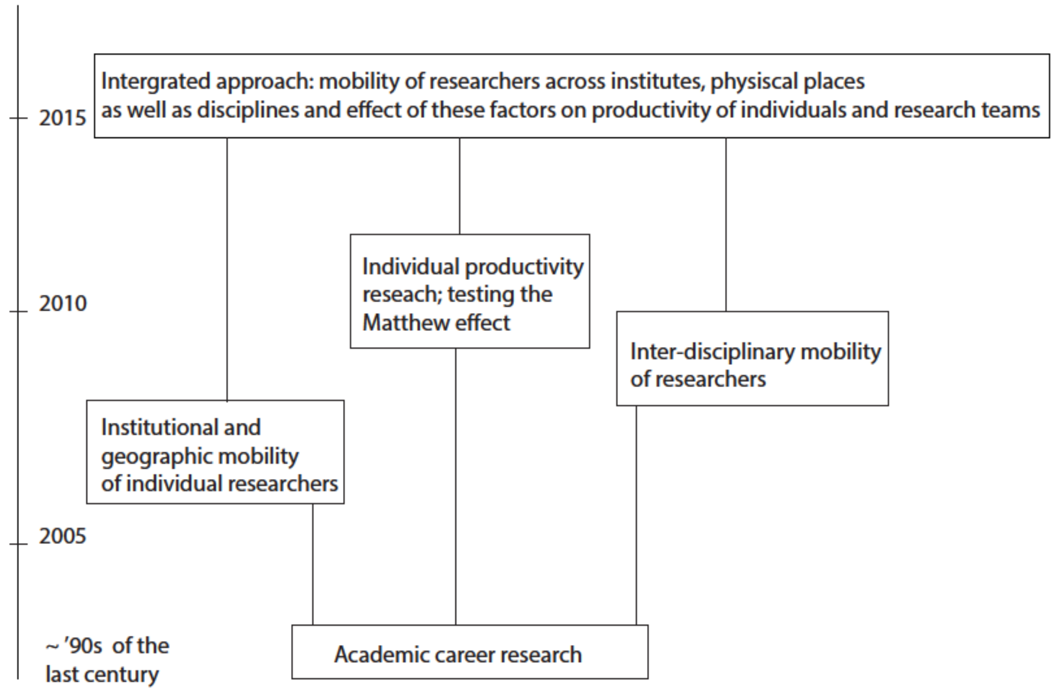

Figure 4: The different approaches of career development analysis
Source: own construction

Workforce mobility has become a mainstream economic, political, and media issue during the last decade in the world (Almeida–Kogut 1999; Nunn 2012). It is widely acknowledged, that there is a strong relationship between competitiveness and the flexibility of the workforce because the workforce mobility between different sectors is a key factor of institutional mobility. Toffler's (1970) prediction that the pace of change in the world is increasing at a faster rate, and this creates a more complex environment, creating a more complex atmosphere for individuals as well as organizations (Toffer 1970) is a reality today. It is well-proven that social and geographical mobility, as well as mobility within the firms, are necessary prerequisites for socio-economic analysis. Culié et al. (2014) have determined a conceptual model for the consequences of inter-firm collaborations on employment mobility. They emphasize

* FAS: L'Académie des sciences de l'Institut de France, Évaluation des chercheurs et des enseignants-chercheurs en sciences exactes et expérimentales: Les propositions de l'Académie des sciences. available at: http://www.academie-sciences.fr/archivage_site/activite/rapport/rapport080709.pdf

the role of inter-firm collaborations in career capital-building, psychological mobility as well as analyzing support. The mobility of academic staff was – and continues to be – of vital importance for the building of networks. According to Hauknes and Ekeland (2002) in the area of mobility research, we can apply different methods. The differences reflect whether the population is static or dynamic; the time scale used; the basis of unit used. The basic units of business demography are different. The most important categories are geographic location, ownership, employees, internal structure, and what it produces and how. The author remarks that "mainstream economic theory does not offer much help here. Most schools of economic thought generally take the firm as a given, unproblematic entity. Ladinsky (1967) have analyzed the geographic migration patterns of professional workers. According to his results, professions that require heavy investments in capital equipment and close cultivations of clients can be described by low migration rates, salaried professions with short analyzing hierarchies, and analyzing work units have high migration rates, unstandardized work conditions, and strong occupational communication networks led to salaried workers in highly professional occupations moving on the national and regional level rather than in local labor markets. Sullivan and Arthur (2006) have introduced the concept of *psychological mobility*, as "the perception of the capacity to make transitions". According to the findings of Geuna et al. (2015), there is a positive, significant effect of researcher's mobility on academic performance in the case of voluntary mobile researchers both in the US and GB. Mobility is a key factor in knowledge creation in different regions. Besides the favorable aspects of professional mobility, the "inevitable disclosure" (2001) in trade secrets is a negative aspect of this phenomenon (Lincicum 2001).

The European intellectual workforce mobility is promoted by the development of accreditation systems, the increasing role of multinational companies (Crescenzi et al. 2014), the emergence of new human resource management practices, and the decreasing importance of language barriers (Tenzer et al. 2014). Spilerman (1972) states that besides its computational simplicity the Markovian models is attractive because it is suitable for the description of different interrelationships as a system. Markovian chain models have been widely applied for the study of migration (Rogers 1966) and projecting growth in social mobility (Erola–Moisio 2007) and manpower supply planning (Zanakis–Maret 1980). The sequence of events can be considered as a Markov chain if the outcome of each event is one of a set of discrete states and the outcome of an event depends only on the present state and not on any past states. The matrix, describing the probabilities of transition from one state to another, is called the transition matrix (Craig–Sendi 2002).

Research productivity

A considerable number of publications aim to analyze the differences between individual career paths. The most important of these are the analyses related to gender differences as well as to cross-cultural differences. The effect of children on academic productivity has been analyzed by a linear growth model in the article of Hunter and Leahey (2010). They have determined that children have a negative effect on productivity over time. At the same, the authors acknowledge that their results are not generalizable.

Another measurement of academic productivity has been the application of the concept of *prestige*, applying different methods of *social network analysis* (Cole–Cole 1967; Reskin 1977; Long et al. 1979). A considerable number of papers have analyzed the academic careers as a series of state (position) changes, applying the approach of economic sociology and statistics (Markov models). Stephan and Levin (1992) applied an integrated model to research careers. Based on their work there are three drivers of academic careers: (1) intrinsic pleasure; (2) recognition and (3) reward. Put in another way: the triangle of the puzzle, the ribbon, and the gold will determine the academic path. Lee et al. (2012) determine two components of career success: extrinsic success and intrinsic one. In their seminal paper, Dietz and Bozeman (2005) analyzed the effects of job transformations and career patterns on productivity. The conceptual base of their research has been built on Scientific – Technical human capital theory (Bozeman et al. 2001; Bozeman–Corley 2004). Based on the analysis of 1200 scientists' and engineers' CVs and publications, they have set up a Tobit model in which the dependent variable has been the number of publications per career year starting the year after the doctorate. Independent variables were the job homogeneity, precocity (measured by the cumulative number of publications at the doctorate year, as well as numerous other characteristic features of academic career paths. According to their results, there is a slightly positive association between career pattern homogeneity and publication productivity. Precocity and homogeneity both had a weak, positive relationship with publication rates.

The interrelationship between productivity and scientific career has been widely analyzed in the US by Long (1978), who has proven, that (1) traditional cross-sectional research offers spurious results. (2) The productivity, expressed by the number of publications and citations had had an insignificant effect on both the prestige of a scientist's initial academic appointment and on the outcome of institutional changes in the later career. (3) The effect of departmental prestige on productivity increases steadily over time.

Lindahl and Danell (2016) have applied a machine-learning algorithm to test the hypothesis, that early career productivity can be an efficient predictor of productivity in the later career. Their results prove in the case of mathematicians, that early career productivity is a useful predictor of future academic productivity, but it is especially useful to predict the performance of the top 10% of academics. Ding et al. (2011) used bibliometric methods to explore the academic impact. The academic impact is produced during scientific communication; the two main forms of scientific communication are citation and collaboration. By a combined analysis of citation impact and collaboration impact, the authors were able to discover detailed information about authors' career status, such as career paths in Scientometrics. The limitation of the study is, that the result does not reflect the real-life career roles or paths for researchers as the dataset is limited to one journal. They offer a combined method of scientist evaluation, based on citation and collaboration impact. According to Petersen

(2015) and Lee and Bozeman (2005) research collaboration have a positive effect on publishing productivity, the authors suggest that developing effective strategies are necessary to understand the potential benefits of collaboration. *Table 2* shows the most important studies on academic productivity in the reviewed literature.

Table 2: Estimation of the importance of academic productivity in the reviewed literature

Author	Year of publication	Target group	Method	Results
Dietz–Bozeman	2005	1200 US scientists and engineers	Tobit regression	Significant influence of Career Homogenity index, Precociticy; year of graduation importantly, held the position, triple helix, first industry or governmental jobs; doctorate in biology or computer science have not been significant
Leahey	2006	Sociology (n=196) and linguistics (n=222) faculty members at US research universities	Structural equation modeling	Married family status (ever married) and affiliation to a public institution, as well as the number of former institutions and receipt of research funding, have a significant, positive effect on performance. Gender and PhD-granting institution ranking according to NRC is not significant
Chakraborty et al	2014	DBLP dataset of the computer science domain (702,973 valid papers and 495,311 authors)	stochastic model	The expertise of an author for a particular field is usually defined by the average number of citations received by the author by publishing papers in this field.
Fernández-Zubieta et al.	2013	171 UK academic researchers	negative binomial regressions	There is a positive albeit insignificant overall effect of mobility, and a negative weakly significant short-term effect. The mobility to a higher-ranked university has only a weakly positive impact on publications output, but not on citations. The authors find no evidence that mobility per se increases academic performance.
Lindahl–Danell	2016	451 authors in the mathemati-	Univariate ROC analysis	The authors conclude that early-career performance productivity has an information value in all tested decision

		cal sub-field number theory	with multiple logistic regression analysis	scenarios, but future performance is more predictable in some cases.
Bentley	2011	Academic staff in Australian public universities, during the periods 1991–3 and 2005–7. Two surveys: a sample of 1420 and 1252 respondents.	Linear multiple regression	The proportion of variation in publication productivity accounted for by the 12-variable model (adjusted R-square) was 32 and 42 percent among men and women in the 1993 data, and 44 and 47 percent respectively in the 2007 data. Academic rank, doctorate qualifications, research time, and international research collaboration were the strongest factors positively associated with publication productivity, but women typically reported significantly lower levels on each of these factors.
Petersen	2015	more than 166,000 collaboration records	Combination of descriptive and panel regression methods	Super ties contribute to above-average productivity and a 17% citation increase per publication, thus identifying these partnerships as a major factor in science career development. Strong collaborations have a significant positive impact on productivity and citations representing the advantage of "super" social ties characterized by trust, conviction, and commitment.
You et al.	2015	Two real-world citation datasets: The citation network of the American Physical Society (APS) journals and the condensed matter (Cond-mat) citation network of	A multi-agent modeling framework	The work efficiency strongly affects agents' academic outputs and impacts under a wide variety of conditions. Research direction selectivity plays a less important role since the results indicate that a selection of hot research topics alone cannot provide sustainable academic careers under intensely competitive conditions.

		the arxiv.orgon-line pre-print repository		
Carvalho–Batty	2006	A total of 116, 771 distinct authors with a U.S. address.		The productivity of U.S. research centers in computer science was highly skewed and the physical location of research centers in the U.S. formed a fractal set.

Conclusions and recommendations for future research

There is considerable knowledge on the effect of different factors (prestige of the university, pre-Ph.D. publications, work abroad, the birth of a child) on academic productivity. As a consequence, if we would like to evaluate the factors of academic career, we have to analyse not just these factors, on a one-by-one basis, but to take into account the combination of all of these influencing conditions. On this basis, some typical career paths could be constructed. An agent-based simulation would be a suitable tool to model the effect of different „events" on academic productivity. It is rather hard to obtain quantifiable pieces of information on this topic because there is great variability in individual "fate" and career, and it should be taken into consideration that there are considerable differences between different fields of sciences. That's why we suggest a series of expert interviews with the purpose to estimate the effect of different "events" on academic activity, based on the experiences of researchers. A convenient way of analysis of estimation results is the R-package "Expert" by Pigeon et al. (2009). Based on these estimations a set of statecharts could be constructed, serving as an input for agent-based modeling. Such high-level software (e.g. Anylogic) offers a favorable solution to the development of such a project aiming at forecasting the different events on academic productivity.

Scientometrics and career research is a rapidly evolving field of science. Rapidly developing information systems, as well as archives, system dynamics, computer sciences, network analysis offer new possibilities for researchers from different scientific backgrounds to form inter-and multidisciplinary research teams. Based on our literature review, the most important problems of scientometrics and academic career research are as follows:

1. Influence of different events and shocks on academic productivity. How the changes in intellectual and material institutional background influence productivity in science?
2. Participation of scholars in science, as a self-organizing network. It is widely acknowledged, that there are some institutional and topical "hot spots" in science. Some people, depending on their level of ambitions, the versatility

of their qualification, personal background are more willing and able to "jump up to these bandwagons", some remain attached to their original field. Who are these people? Is the change of field a promising possibility to enhance one's scientific production?

3. The role of research-group attachment in academic career: it is well known, that the dynamically changing world makes it necessary to become attached to some research groups which do some research together, then, in the framework of another project, a "recombination" takes place in the academic community, new teams are formed. Are there any patterns of these research team formations across countries and cultures?

References

Abbott, A. – Cyranoski, D. – Jones, N. – Maher, B. – Schiermeier, Q. – Van Noorden, R. (2010) Do metrics matter? Many researchers believe that quantitative metrics determine who gets hired and who gets promoted at their institutions. With an exclusive poll and interviews, Nature probes to what extent metrics are really used that way. *Nature,* 465/7300. 860–863.

Ackers, L. (2005a) Moving people and knowledge: Scientific mobility in the European Union1. *International Migration,* 43/5. 99–131.

Ackers, L. (2005b) Promoting scientific mobility and balanced growth in the European research area. *Innovation,* 18/3. 301–317.

Ackers, L. (2007) Legislating for equality? Working hours and progression in science careers. *European Law Journal,* 13/2. 169–185.

Ackers, L. – Oliver, L. (2007) From Flexicurity to Flexsecquality?: The Impact of the Fixed-Term Contract Provisions on Employment in Science Research. *International Studies of Management – Organization,* 37/1. 53–79.

Almeida, P. – Kogut, B. (1999) Localization of knowledge and the mobility of engineers in regional networks. *Management Science,* 45/7. 905–917.

Amarnani, R. K. – Garcia, P. R. J. M. – Restubog, S. L. D. – Bordia, P. – Bordia, S. (2016) Do You Think I'm Worth It? The Self-Verifying Role of Parental Engagement in Career Adaptability and Career Persistence Among STEM Students. *Journal of Career Assessment,* 1069072716679925.

Antonelli, C. – Franzoni, C – – Geuna, A. (2011) The organization, economics, and policy of scientific research: what we do know and what we don't known an agenda for research. *Industrial and Corporate Change,* 20/1. 201–213.

Baruch, Y. – Hall, D. T. (2004) The academic career: a model for future careers in other sectors? *Journal of Vocational Behavior,* 64/2. 241–262.

Bentley, P. (2011) Gender differences and factors affecting publication productivity among Australian university academics. *Journal of Sociology,* 48/1. 85–103.

Bonzi, S. (1992) Trends in research productivity among senior faculty. *Information Processing – Management,* 28/1. 111–120.

Boyack, K. W. – Klavans, R. – Börner, K. (2005) Mapping the backbone of science. *Scientometrics,* 64/3. 351–374.

Bozeman, B. – Corley, E. (2004) Scientists' collaboration strategies: implications for scientific and technical human capital. *Research Policy,* 33/4. 599–616.

Bozeman, B. – Dietz, J. S. – Gaughan, M. (2001) Scientific and technical human capital: an alternative model for research evaluation. *International Journal of Technology Management,* 22/7–8. 716–740.

Bozeman, B. – Rogers, J. D. (2002) A churn model of scientific knowledge value: Internet researchers as a knowledge value collective. *Research Policy,* 31/5. 769–794.

Burrell, Q. (2007) Hirsch index or Hirsch rate? Some thoughts arising from Liang's data. *Scientometrics,* 73/1. 19–28.

Canibano, C. – Bozeman, B. (2009) Curriculum vitae method in science policy and research evaluation: the state-of-the-art. *Research Evaluation,* 18/2. 86–94.

Canibano, C. – Otamendi, J. – Andujar, I. S. (2008) Measuring and assessing researcher mobility from CV analysis: the case of the Ramón y Cajal programme in Spain. *Research Evaluation,* 17/1. 17–31.

Carvalho, R. – Batty, M. (2006) The geography of scientific productivity: Scaling in US computer science. *Journal of Statistical Mechanics: Theory and Experiment,* 10. P10012.

Chakraborty, T. – Tammana, V. – Ganguly, N. – Mukherjee, A. (2014) Understanding and modeling diverse scientific careers of researchers. *Journal of Informetrics,* 9/1. 69–78.

Cole, J. R. – Zuckerman, H. (1984) The productivity puzzle. Advances in Motivation and Achievement. In *Women in Science.* Greenwich, CT.: JAI Press

Cole, S. – Cole, J. R. (1967) Scientific output and recognition: A study in the operation of the reward system in science. *American Sociological Review,* 31/6. 377–390.

Corley, E. – Bozeman, B. – Gaughan, M. (2003) Evaluating the impacts of grants on women scientist careers: the curriculum vitae as a tool for research assessment. In *Learning from Science and Technology Policy Evaluation: Experiences from the US and Europe.* Cheltenham, UK: Edward Elgar Publishing, 293–315.

Craig, B. A. – Sendi, P. P. /2002) Estimation of the transition matrix of a discrete-time Markov chain. *Health Economics,* 11/1. 33–42.

Crescenzi, R. – Pietrobelli, C. – Rabellotti, R. (2014) Innovation drivers, value chains and the geography of multinational corporations in Europe. *Journal of Economic Geography,* 14/6. 1053–1086.

Culie, J.-D. – Khapova, S. N. – Arthur, M. B. (2014) Careers, clusters and employment mobility: The influences of psychological mobility and organizational support. *Journal of Vocational Behavior,* 84/2. 164–176.

De Pater, I. E. (2005) *Doing things right or doing the right thing: a new perspective on the gender gap in career success.* Dissertation, Universitiet van Amsterdam.

Diamond, A. (1984) An economic model of the life-cycle research productivity of scientists. *Scientometrics,* 6/3. 189–196.

Dietz, J. S. (2000) Building a social capital model of research development: the case of the experimental program to stimulate competitive research. *Science and Public Policy,* 27/2. 137–145.
Dietz, J. S. – Bozeman, B. (2005) Academic careers, patents, and productivity: industry experience as scientific and technical human capital. *Research Policy,* 34/3. 349–367.
Dietz, J. S. – Chompalov, I. – Bozeman, B. – Lane, E. O. N. – Park, J. (2000) Using the curriculum vita to study the career paths of scientists and engineers: An exploratory assessment. *Scientometrics,* 49/3. 419–442.
Ding, J. – Yang, L. – Liu, Q. (2011) Measuring the academic impact of researchers by combined citation and collaboration impact. *Proceedings of the 14th International Conference on Scientometrics and Informetrics,* 1177–1187.
Duch, J. – Zeng, X. H. T. – Sales-Pardo, M. – Radicchi, F. – Otis, S. – Woodruff, T. K. et al. (2012) The possible role of resource requirements and academic career-choice risk on gender differences in publication rate and impact. *PloS one,* 7/12. e51332.
Efron, N. – Brennan, N. A. (2011) Citation analysis of Australia-trained optometrists. *Clinical and Experimental Optometry,* 94/6. 600–605.
Egghe, L. (2006) Theory and practise of the g-index. *Scientometrics,* 69/1. 131–152.
Egghe, L. (2010) Conjugate partitions in informetrics: Lorenz curves, h-type indices, Ferrers graphs and Durfee squares in a discrete and continuous setting. *Journal of Informetrics,* 4/3. 320–330.
Ehrich, L. C. – Hansford, B. – Tennent, L. (2004) Formal mentoring programs in education and other professions: A review of the literature. *Educational Administration Quarterly,* 40/4. 518–540.
Enders, J. (2005) Border crossings: Research training, knowledge dissemination and the transformation of academic work. *Higher Education,* 49/1–2. 119–133.
Enders, J. – De Weert, E. (2004) Science, training and career: Changing modes of knowledge production and labour markets. *Higher Education Policy,* 17/2. 135–152.
Erola, J. – Moisio, P. (2007) Social mobility over three generations in Finland, 1950–2000. *European Sociological Review,* 23/2. 169–183.
Fernandez-Zubieta, A. – Geuna, A. – Lawson, C. (2015) Mobility and Productivity of Research Scientists 1. *Global Mobility of Research Scientists: The Economics of Who Goes Where and Why,* 105.
Fernandez-Zubieta, A. – Geuna, A. – Lawson, C. (2013) Researchers Mobility and its Impact on Scientific Productivity. University of Turin Working paper No. 13/2013
Fox, M. F. (1983) Publication productivity among scientists: A critical review. *Social Studies of Science,* 13/2. 285–305.
Fox, M. F. (1985) Location, sex-typing, and salary among academics. *Work and Occupations,* 12/2. 186–205.
Fox, M. F. (2001) Women, science, and academia Graduate Education and Careers. *Gender – Society,* 15/5. 654–666.

Fox, M. F. (2005) Gender, family characteristics, and publication productivity among scientists. *Social Studies of Science,* 35/1. 131–150.
Franceschini, F. – Maisano, D. (2011) Proposals for evaluating the regularity of a scientist's research output. *Scientometrics,* 88/1. 279–295.
Gaughan, M. – Bozeman, B. (2002) Using curriculum vitae to compare some impacts of NSF research grants with research center funding. *Research Evaluation,* 11/1. 17–26.
Gaughan, M. – Ponomariov, B. (2008) Faculty publication productivity, collaboration, and grants velocity: using curricula vitae to compare center-affiliated and unaffiliated scientists. *Research Evaluation,* 17/2. 103–110.
Gaughan, M. – Robin, S. (2004) National science training policy and early scientific careers in France and the United States. *Research Policy,* 33/4. 569–581.
Geuna, A. – Kataishi, R. – Toselli, M. – Guzman, E. – Lawson, C. – Fernandez-Zubieta, A. et al. (2015) SiSOB data extraction and codification: A tool to analyze scientific careers. *Research Policy,* 44/9. 1645–1658.
Glänzel, W. (2006) On the opportunities and limitations of the H-index. *Science Focus,* 1/1. 10–11.
Glänzel, W. – Debackere, K. – Meyer, M. (2007) Triad or tetrad? On global changes in a dynamic world. *Scientometrics,* 74/1. 71–88.
Goulet, V. – Jacques, M. – Pigeon, M. (2009) expert: Modeling without data using expert opinion. *The R Journal,* 1/1. 31–36.
Guevara, M. R. – Hartmann, D. – Aristaran, M. – Mendoza, M. – Hidalgo, C. s. A. (2016) The research space: using career paths to predict the evolution of the research output of individuals, institutions, and nations. *Scientometrics,* 109/3. 1695–1709.
Hack, T. F. – Crooks, D. – Plohman, J. – Kepron, E. (2010) Research citation analysis of nursing academics in Canada: identifying success indicators. *Journal of Advanced Nursing,* 66/11. 2542–2549.
Hauknes, J. – Ekeland, A. (2002) Mobility of researchers policy, models and data. *STEP report,* 4, 2002.
Hirsch, J. E. (2005) An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences of the United States of America,* 16569–16572.
Hunter, L. – Leahey, E. (2010) Parenting and research productivity: New evidence and methods. *Social Studies of Science,* 40/3. 433–451.
Iglesias, J. E. – Pecharroman, C. (2007) Scaling the h-index for different scientific ISI fields. *Scientometrics,* 73/3. 303–320.
Jl, D. – Yang, L. – Liu, Q. Measuring the Academic Impact of Researchers by Combined Citation and Collaboration impact. *IEEE Access* PP/99. 1–1.
Keith, B. – Babchuk, N. (1998) The quest for institutional recognition: A longitudinal analysis of scholarly productivity and academic prestige among sociology departments. *Social Forces,* 76/4. 1495–1533.

Kyvik, S. – Teigen, M. (1996) Child care, research collaboration, and gender differences in scientific productivity. *Science, Technology – Human Values,* 21/1. 54–71.
Ladinsky, J. (1967) Occupational determinants of geographic mobility among professional workers. *American Sociological Review,* 253–264.
Lane, J. (2010) Let's make science metrics more scientific. *Nature,* 464/7288. 488–489.
Leahey, E. (2006) Gender differences in productivity: Research specialization as a missing link. *Gender – Society,* 20/6. 754–780.
Lee, L. S. – Pusek, S. N. – McCormack, W. T. – Helitzer, D. L. – Martina, C. A. – Dozier, A. M. et al. (2012) Clinical and translational scientist career success: metrics for evaluation. *Clinical and Translational Science,* 5/5. 400–407.
Lee, S. – Bozeman, B. (2005) The impact of research collaboration on scientific productivity. *Social Studies of Science,* 35/5. 673–702.
Leong, F. T. L. – Leung, K. (2004) Academic careers in Asia: A cross-cultural analysis. *Journal of Vocational Behavior,* 64/2. 346–357.
Liang, L. (2006) h-index sequence and h-index matrix: Constructions and applications. *Scientometrics,* 69/1. 153–159.
Lincicum, D. (2001) Inevitable Conflict: California's Policy of Worker Mobility and the Doctrine of Inevitable Disclosure. *S. Cal. L. Rev.,* 75, 1257.
Lindahl, J. – Danell, R. (2016) The information value of early career productivity in mathematics: a ROC analysis of prediction errors in bibliometricly informed decision making. *Scientometrics,* 109/3. 2241–2262.
Long, J. S. (1992) Measures of sex differences in scientific productivity. *Social Forces,* 159–178.
Long, J. S. – Allison, P. D. – McGinnis, R. (1979) Entrance into the academic career. *American Sociological Review,* 43/9. 816–830.
Long, J. S. – Fox, M. F. (1995) Scientific careers: Universalism and particularism. Annual *Review of Sociology,* 21/1. 45–71.
Mangematin, V. (2001) Individual careers and collective research: is there a paradox? *International Journal of Technology Management,* 22/7–8. 670–675.
Matia, K. – Nunes Amaral, L. A. – Luwel, M. – Moed, H. F. – Stanley, H. E. (2005) Scaling phenomena in the growth dynamics of scientific output. *Journal of the American Society for Information Science and Technology,* 56/9. 893–902.
McBrier, D. B. (2003) Gender and career dynamics within a segmented professional labor market: The case of law academia. *Social Forces,* 81/4. 1201–1266.
Merton, R. K. (1957) Priorities in scientific discovery: a chapter in the sociology of science. *American Sociological Review,* 22/6. 635–659.
Merton, R. K. (1961) Singletons and multiples in scientific discovery: A chapter in the sociology of science. *Proceedings of the American Philosophical Society,* 105/5. 470–486.
Moher, D. – Liberati, A. – Tetzlaff, J. – Altman, D. G. – Prisma, G. (2009) Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS med,* 6/7. e1000097.

Moranoâ-Foadi, S. (2005) Scientific mobility, career progression, and excellence in the european research area1. *International Migration,* 43/5. 133–162.
Nunn, A. (2012) The political economy of competitiveness and social mobility. *British Politics,* 7/2. 86–110.
Petersen, A. M. (2015) Quantifying the impact of weak, strong, and super ties in scientific careers. *Proceedings of the National Academy of Sciences,* 112/34. E4671–E4680.
Petersen, A. M. – Jung, W.-S. – Yang, J.-S. – Stanley, H. E. (2011) Quantitative and empirical demonstration of the Matthew effect in a study of career longevity. *Proceedings of the National Academy of Sciences,* 108/1. 18–23.
Petersen, A. M. – Riccaboni, M. – Stanley, H. E. – Pammolli, F. (2012) Persistence and uncertainty in the academic career. *Proceedings of the National Academy of Sciences,* 109/14. 5213–5218.
Petersen, A. M. – Wang, F. – Stanley, H. E. (2010) Methods for measuring the citations and productivity of scientists across time and discipline. *Physical Review E,* 81/3. 036114.
Probert, B. (2005) I just couldn't fit it in: Gender and unequal outcomes in academic careers. *Gender, Work – Organization,* 12/1. 50–72.
Prpi, K. (2002) Gender and productivity differentials in science. *Scientometrics,* 55/1. 27–58.
Radicchi, F. – Fortunato, S. – Castellano, C. (2008) Universality of citation distributions: Toward an objective measure of scientific impact. *Proceedings of the National Academy of Sciences,* 105/45. 17268–17272.
Reskin, B. F. (1977) Scientific productivity and the reward structure of science. *American Sociological Review*, 41/5. 491–504.
Rogers, A. A Markovian policy model of interregional migration. In *Papers of the Regional Science Association,* 1966 (Vol. 17, 205–224.), Springer
Sahel, J.-A. (2011) Quality versus quantity: assessing individual research performance. *Science Translational Medicine,* 3/84.
Sandström, U. (2009) Combining curriculum vitae and bibliometric analysis: mobility, gender and research performance. *Research Evaluation,* 18/2. 135–142.
Sax, L. J. – Hagedorn, L. S. – Arredondo, M. – DiCrisi, F. A. (2002) Faculty research productivity: Exploring the role of gender and family-related factors. *Research in Higher Education,* 43/4. 423–446.
Seglen, P. O. (1992) The skewness of science. *Journal of the American Society for Information Science,* 43/9. 628.
Sonnert, G. – Holton, G. J. (1995) *Who Succeeds in Science?: The Gender Dimension.* New Brunswick, NJ: Rutgers University Press.
Spilerman, S. (1972) The analysis of mobility processes by the introduction of independent variables into a Markov chain. *American Sociological Review,* 277–294.
Stephan, P. E. – Levin, S. G. (1992) *Striking the Mother Lode in Science: The Importance of Age, Place, and Time.* New York: Oxford University Press.

Sullivan, S. E. – Arthur, M. B. (2006) The evolution of the boundaryless career concept: Examining physical and psychological mobility. *Journal of Vocational Behavior,* 69/1. 19–29.

Symonds, M. R. E. – Gemmell, N. J. – Braisher, T. L. – Gorringe, K. L. – Elgar, M. A. (2006) Gender differences in publication output: towards an unbiased metric of research performance. *PloS one,* 1/1. e127.

Tenzer, H. – Pudelko, M. – Harzing, A.-W. (2014) The impact of language barriers on trust formation in multinational teams. *Journal of International Business Studies,* 45/5. 508–535.

Teodorescu, D. (2000) Correlates of faculty publication productivity: A cross-national analysis. *Higher Education,* 39/2. 201–222.

Toffler, A. – Nathan, M. (1970) *Future Shock: JSTOR.* New York: Bantam Books.

Van Balen, B. – Van Arensbergen, P. – Van der Weijden, I. – Van den Besselaar, P. (2012) Determinants of success in academic careers. *Higher Education Policy,* 25/3. 313–334.

Van den Besselaar, P. – Leydesdorff, L. (2009) Past performance, peer review and project selection: a case study in the social and behavioral sciences. *Research Evaluation,* 18/4. 273–288.

Wallace, M. L. – Lariviöre, V. – Gingras, Y. (2009) Modeling a century of citation distributions. *Journal of Informetrics,* 3/4. 296–303.

Wells, R. S. – Seifert, T. A. – Padgett, R. D. – Park, S. – Umbach, P. D. (2011) Why do more women than men want to earn a four-year degree? Exploring the effects of gender, social origin, and social capital on educational expectations. *The Journal of Higher Education,* 82/1. 1–32.

Woolley, R. – Turpin, T. (2009) CV analysis as a complementary methodological approach: investigating the mobility of Australian scientists. *Research Evaluation,* 18/2. 143.

Xie, Y. – Shauman, K. A. (1998) Sex differences in research productivity: New evidence about an old puzzle. *American Sociological Review,* 847–870.

Xie, Y. – Shauman, K. A. (2003) *Women in Science: Career Processes and Outcomes* (Vol. 26) Cambridge, MA.: Harvard University Press.

You, Z.-Q. – Han, X.-P. – Hadzibeganovic, T. (2016) The role of research efficiency in the evolution of scientific productivity and impact: An agent-based model. *Physics Letters A,* 380/7. 828–836.

Zanakis, S. H. – Maret, M. W. (1980) A Markov chain application to manpower supply planning. *Journal of the Operational Research Society,* 31/12. 1095–1102.

Zhang, L. – Glänzel, W. (2012) Where demographics meets scientometrics: Towards a dynamic career analysis. *Scientometrics,* 91/2. 617–630.