

Farys, R., and Wolbring, T. (2021) Matthew Effects in Science and the Serial Diffusion of Ideas Testing Old Ideas with New Methods. Quantitative Science Studies. Advance Publication. https://doi.org/10.1162/qss_a_00129

Matthew Effects in Science and the Serial Diffusion of Ideas Testing Old Ideas with New Methods

Rudolf Farys

Tobias Wolbring

Abstract

The Matthew effect has become a standard concept in science studies and beyond to describe processes of cumulative advantage. Despite its wide success, a rigorous quantitative analysis for Merton's original case for Matthew effects – the Nobel Prize – is still missing. This paper aims to fill this gap by exploring the causal effect of the Sveriges Riksbank Prize in Economic Sciences in Memory of Alfred Nobel (hereafter the Nobel Prize in Economics). Furthermore, we test another of Merton's ideas: successful papers can draw attention to cited references leading to a serial diffusion of ideas. Based on the complete Web of Science 1900–2011, we estimate the causal effects of Nobel Prizes compared to a synthetic control group which we constructed by combining different matching techniques. We find clear evidence for a Matthew effect upon citation impacts, especially for papers published within five years before the award. Further, scholars from the focal field of the award are particularly receptive to the award signal. In contrast to that, we find no evidence that the Nobel Prize causes a serial diffusion of ideas. Papers cited by future Nobel laureates do not gain in citation impact after the award.

Keywords: Matthew effect; Nobel Prize; Citations; Diffusion; Matching; Longitudinal Modeling

Contact details

Corresponding author:

Prof. Dr. Tobias Wolbring
FAU Erlangen-Nürnberg
School of Business, Economics and Society
Findelgasse 7/9
D-90402 Nürnberg
Germany
tobias.wolbring@fau.de
+49 911 5302 950
ORCID: 0000-0002-9053-7965

First author:

Dr. Rudolf Farys
University of Bern
Institute of Sociology
Fabrikstr. 8
S-3012 Bern
Switzerland
rudolf.farys@soz.unibe.ch
+41 31 631 4826
ORCID: 0000-0002-9667-694X

Author contributions: Rudolf Farys: Conceptualization; Data curation; Methodology; Formal analysis; Investigation; Methodology; Project Administration; Software; Validation; Visualization; Writing—original draft; Writing—review & editing. Tobias Wolbring: Conceptualization; Data curation; Methodology; Formal analysis; Investigation; Methodology; Project Administration; Software; Validation; Visualization; Writing—original draft; Writing—review & editing.

Competing interests: The authors have no competing interests.

Funding information: No funding has been received for this research.

Copyright: © 2021 Rudolf Farys and Tobias Wolbring. Published under a Creative Commons Attribution 4.0 International (CC BY 4.0) license.

Peer Review: https://publons.com/publons/10.1162/qss_a_00129

Data availability: We acknowledge the use of ISI Web of Science data of Clarivate Analytics for our citation analysis. We thank the library of the Swiss Federal Institute of Technology Zurich for providing the WoS raw data. Data used in this manuscript are subject to strict requirements and cannot be made available in a data repository. In order to enable replications, paper identifiers and code are publicly and permanently available at the Harvard Dataverse (Wolbring & Farys, 2021).

Acknowledgements: We would like to thank Hans-Dieter Daniel, Neha Gondal, Michael Hechter, Debra Hevenstone, Ben Jann, Marc Keuschnigg, Rüdiger Mutz, Omar Lizardo, Edgar Treischl, Arnout van de Rijt, and Ezra Zuckerman and the reviewers who substantially improved the paper with many helpful comments. John Cirilli provided valuable language editing. We are also grateful for valuable input which we received in research colloquia at the University of Bern, the University of Bielefeld, University of Cologne, FAU Erlangen-Nürnberg, Goethe University Frankfurt/Main, INCHER Kassel, MZES Mannheim, Utrecht University, and ETH Zurich, the session “Analytical Sociology” at the ASA meeting 2015 in Chicago, IL, and the session “Social Networks” of the “Social Interactions and Society Conference” at ETH Zurich.

1 Introduction

In 1968, Robert K. Merton published a seminal paper in *Science* which has become one of the most cited references on the sociology of science and beyond. Based on previous research on the success of Nobel laureates after elevation, Merton coined the term Matthew effect¹ to describe “the accruing of greater increments of recognition for particular scientific contributions to scientists of considerable repute and the withholding of such recognition from scientists who have not yet made their mark.” (1968; p. 58). While Merton was well aware of the very advantageous career opportunities of many Nobel laureates and the accumulation of various forms of peer recognition such as the reception of other awards, outstanding citation impact, and external funding prior to being awarded the Nobel Prize, he emphasized that receiving the Nobel Prize elevated the research of laureates among other work of “prize-winning calibre” (p.57). As a consequence, the “crowning” of scientific careers with a Nobel Prize leads to a further accumulation of scientific rewards such as assigning priorities in independent multiple discoveries and attributing individual contributions in collaborative research.

¹ While Merton’s paper in *Science* has become the standard reference on Matthew effects, Merton himself acknowledged in the reprinting of the paper that the research of his wife Harriet Zuckerman (1977) was essential for developing the concept: “It is now [1973] belatedly evident to me that I drew upon the interview and other materials of the Zuckerman study to such an extent that, clearly, the paper should have appeared under joint authorship” (Merton, 1988, p.607).

Merton's paper has not only become *the* core reference in the rich literature on cumulative advantages in academia (Allison, Long, & Krauze, 1982; Cole & Cole, 1973; Price, 1976), but also in the broader literature on rich-getting-richer phenomena in other areas of social life (DiPrete & Eirich, 2006; Salganik, Dodds, & Watts, 2006; van de Rijt, Kang, Restivo, & Patil 2014). Thereby, the concept of Matthew effects has proven its explanatory value in a broad range of areas including research on health inequalities, cultural markets, educational success, and labour market trajectories (for reviews see Rigney, 2010; Zuckerman, 2011).

Despite this wide use of the concept, a rigorous quantitative analysis for Merton's original case of the Nobel Prize is still missing. Indeed, the ideas of Merton and Zuckerman have inspired further scholarship on the Nobel Prizes (e.g., Boettke, Fink, & Smith, 2012; Bjork, Offer, & Söderberg, 2014; Cole, 1970; Diamond, 1988; Karier, 2010). For example, research has shown that the number of awards (Chan, Gleeson, & Torgler, 2014) as well as citation impacts steadily increase ahead of the event (Garfield & Welljams-Dorof, 1992; Mazlounian, Eom, Helbing, Lozano, & Tortunato, 2011).

Similarly, Merton's and Zuckerman's pioneering work has marked the starting point for rigorous causal analyses of the effects of other positive status shocks in science. Analyzing decisions for early-career grant funding in the Netherlands as a sort of natural experiment, Bol, Vaan, and van de Rijt (2018) find that grantees just above the funding threshold receive substantially more funding in the following years and are significantly more likely to become full professors than applicants just below the threshold. Focusing on prestigious midcareer awards in medicine and economics, Azoulay, Stuart, and Wang (2014) and Chan, Frey, Gallus, and Torgler (2013) find evidence for a citation boost caused by the honoring, although the studies disagree how strong and lasting such an effect is. Moreover, numerous studies document that status markers such as author prestige (e.g., Wang, 2014), lead articles in journal volumes (e.g., Michayluk &

Zurbregg, 2014), and designation of a paper by the editor as very important (Mutz, Wolbring, & Daniel, 2017) affect future citation impact.²

To sum up, the literature has clearly corroborated the idea that status affects future rewards and career opportunities. However, to the best of our knowledge, no study exists which provides a rigorous analysis of the causal effect of Nobel Prize reception on the accumulation of further citations for a group of laureates. An exception is our case study on the honoring of Robert J. Aumann with the Sveriges Riksbank Prize in Economic Sciences in Memory of Alfred Nobel (hereafter the Nobel Prize in Economics)³ which finds no Matthew effect at all on citation impact (see Farys & Wolbring, 2017). However, these results are unlikely to generalize to other Nobel laureates, because Aumann's work has been rarely cited already before the award due to its high degree of mathematical abstraction.

Building on our previous work and based on the complete *Web of Science 1900–2011*, we aim to fill this gap by exploring the causal effect of a Nobel Prize in Economics on citation impacts and its dynamic over time. Using a combination of different matching techniques and longitudinal modeling, we not only control for differences in intrinsic quality and unobserved variables affecting citation impact, but also go beyond average effects in two ways. On the one hand, we explore potential heterogeneity for different Nobel Prize publications with respect to publication date, pre-Nobel citation impact, and journal reputation. On the other hand, we investigate audience-specific reactions to the awards by distinguishing citation impact among scholars of the focal field of the award (such as in business, economics, and management) and scholars of the neighboring social and behavioral sciences.

² A related literature also investigates the effects of negative status shocks. Taking the case of article retractions, Lu, Jin, Uzzi, and Jones (2013) report marked negative effects of non-self-reported retractions on citation impact of authors' recent and earlier papers. In addition, Azoulay, Zivin, and Wang (2010) highlight that negative status shocks can spillover: collaborators in the "invisible college" suffer from the death of a superstar by markedly lower quality-adjusted publication rates.

³ We are aware of the cultural and political dimensions of the Nobel Prize and the widespread criticism in of the Sveriges Riksbank Prize in Economic Sciences in Memory of Alfred Nobel as being treated as a "Nobel Prize in Economics," legitimating economics as a "science" comparable to other Nobel fields (see Offer & Söderberg, 2017). It is further worth noting that only one woman, Elinor Ostrom, received the award. These issues are beyond the scope of this study but certainly worth exploring.

In addition, we want to explore whether another mechanism is at work which might cause spillover effects of the Nobel Prize on publications cited by the laureate. Merton (1995, p.388) mentioned in later publications such a possibility, dubbing it the “serial diffusion of ideas” through “mediated references”. The basic idea is that papers written by future Nobel laureates receive more attention after the reception of the prize (see also Frandsen & Nicolaisen, 2013). This might indirectly raise scholars’ awareness of Nobel Prize winners’ cited references (see Peterson, Press, & Dill 2010 for the distinction between direct and indirect mechanisms for citations) and in that sense the social status of a Nobel laureate might leak down the citation network, even publications cited by Nobel Prize winners’ cited references might gain, to perhaps a lesser extent, in citation impact.

2 Matthew Effects on Citation Impacts

Peer recognition for scientific achievements can come in various forms ranging from more or less prestigious awards over memberships in scientific societies and external research grants to the maybe most elementary level of using and citing one’s work (Merton, 1988, p.620). In this paper, we focus on the effects of a Nobel Prize in Economics on citation impacts and the potential serial diffusion of ideas in the citation network. One reason for our focus on citations is that they are one of the most elementary forms of peer recognition in the science system. Another reason is that citations are “one of the micro-level stratifying mechanisms in science” (Baldi, 1998, p.830), since citation impact can positively affect other forms of peer recognition. For example, bibliometric analyses have become an integral part of most research evaluations and can have consequences for hiring, tenure, and funding decisions.

Citations are the building blocks of knowledge claims in modern science. They are located at the level of publications and connect the argument in one publication with the content of another paper creating a complex network of directed references among publications. Citations can thereby serve very different functions (Bornmann & Daniel, 2008; Leydesdorff, 1998; Nicolaisen, 2007;

Tahamtan & Bornmann, 2019). Two positions have emerged in the literature which conflict in their interpretation of the role of citations in science: the normative and the social constructivist view. Both views help to provide insights into the potential reasons why awards might affect citation impacts.

Proponents of the normative citation theory such as Merton (1988, p.621) argue that “the institutionalized practice of citations and references in the sphere of learning is [...] central to the incentive system and underlying sense of distributive justice” of modern science, since they serve two functions. On the one hand, citations have an instrumental cognitive function by making readers aware of the sources of knowledge and put them in a position to follow-up on ideas and claims formulated in the literature. From this perspective, a Nobel Prize could increase citation impact, since it raises awareness for the existence of a laureate’s knowledge claims. Such an attention boost caused by a Nobel Prize appears to be especially likely among less well informed scholars such as those coming from a different, though related field of inquiry. Similarly, cited references in a laureate’s publications could indirectly profit from this attention boost causing a serial diffusion of ideas.

On the other hand, citations also serve a symbolic institutional function according to the normative citation theory. Citations mark the origin of ideas, recognize authors’ original contributions, and accrue social esteem. As such, they acknowledge property rights, signal intellectual debt, and reward scientific achievements, in short: they are supposed to give credit where credit is due (Kaplan, 1965). Thereby, in an ideal world of science, scholars should accrue peer recognition solely based on the worth of a contribution, e.g. the importance, content, and quality of a publication, and regardless of other non-meritocratic criteria such as authors’ status or affiliation (Merton, 1973).

However, as the case of the Matthew effect shows, scientific practice sometimes deviates from this norm of universalism in systematic ways. In particular, authors might prefer to read and cite publications of a Nobel laureate as compared to other equally relevant references due to different

mechanisms. Merton (1968) himself already sketched one potential mechanism why scholars might deviate from the norm of universalism: in face of an increasing amount of scholarship as well as limited reading time – an argument nowadays even more important than back then (Falkinger, 2008; Franck, 2002) – scholars might rely on author status as a potential signal for underlying quality of a publication. As Bothner, Podolny, and Smith (2011) show in a simulation study, employing such a strategy can be rational in case of incomplete information as long as the association between status signal and intrinsic quality is sufficiently strong. However, such an approach becomes dysfunctional and leads to the neglect of other more relevant publications and ideas, if status and quality are only weakly correlated.

Proponents of social constructivist sociology of science (Callon, Law, & Rip, 1986; Knorr-Cetina, 1981; Latour 1987), develop a different view on science and the role of citations therein. Instead of assuming that science is governed by a certain set of internal norms and a recognition-driven reward system, they contend that science in practice is shaped by processes of social influence, political and financial interests, and power relationships. The constructivist view, hence, frames science as a “war of words” in which “publications are weapons in a struggle among scientists to persuade each other of the validity of knowledge claims, and thereby to establish dominant positions in the community” (Cozzens, 1989, p.440). Therefore, scientific claims are not mere objective facts, but socially constructed and deconstructed (Latour & Woolgar, 1979). In order to reach the status as objective facts, scholars need to convince readers, reviewers, and editors about the validity of their claims.

Against that background, proponents of social constructivist citation theory emphasize that citations do often not merely serve a cognitive instrumental or symbolic institutional function but are used as “tools of persuasion” (Gilbert, 1977; MacRoberts & MacRoberts, 1987). As rhetorical devices in the publication game, citations can mark novelty and relevance of one’s work, signal allegiance to certain intellectual traditions, or help to backup arguments. As scientific “defense lines”, references might be also misquoted on purpose to strengthen one’s position or be cited

without being actually read (Latour, 1987; Luukkonen, 1997).

In contrast to the normative citation theory, the actual relevance and intrinsic quality of a publication should only matter for citation behavior to the extent that it can positively influence the credibility of one's claim. Hence, authors will try to draw on "codified" knowledge and cite "authoritative" references to create the impression of "facticity" (Gilbert, 1977; Moed & Garfield, 2004). It appears likely that Nobel Prize decisions trigger such strategic citation behavior,⁴ since the award puts the laureate in a special position for convincing others about scientific claims (see also Strevens, 2006). While Nobel publications are likely to receive such "ceremonial" citations (see Adatto & Cole, 1981) according to the constructivist view, cited references do only matter for strategic behavior under certain conditions. For example, incentives for strategic citations might exist to cite references which were fundamental for the contribution of the Nobel laureate and hence also gain in authoritativeness by the award. To sum up, citations can serve very different functions. Normative theories highlight the role of citations as part of the scientific system of property rights and rewards, whereas social constructivist theories point out the often strategic nature of citations as rhetorical device of persuasion. Both theoretical accounts have proven their heuristic and explanatory value in empirical research (e.g., Baldi, 1998; Collins, 1999; Cronin, 2005; Safer & Tang, 2009; Shadish, Tolliver, Gray, & Sengupta, 1995; Thornley et al, 2015; White, 2004). Hence, in practice a mixture of these and other processes is likely to be at work simultaneously (for a comprehensive framework see Tahamtan & Bornmann, 2018).

While it is undisputed that a Nobel Prize confers peer recognition and raises the professional standing of the laureate, it is not completely clear which mechanisms cause Matthew effects in citation impacts. According the normative view, the work of the laureate might receive more attention due to the Nobel Prize especially by those less well informed prior to the honoring.

⁴ There also other forms of strategic citation behavior. For example, citations can be used to repay scientific debts, bribe potential referees, or to outsource responsibilities for errors (see Wang, 2014, p.331). All of these other forms of strategic citation behavior can also foster Matthew effects, as they usually occur in favor of citing a high status author or paper.

Thereby, beyond a mere attention effect, the award might also work as a signal helping scholars to identify particular important high-quality research. Both mechanisms might also cause a serial diffusion of ideas. However, according to the social constructivist view, the Nobel Prize could also create incentives to cite laureate's publications not because of their exceptional quality, but due to their authoritative status. We would expect such strategic citation behavior especially among those from the focal field of the award who should be already well-informed about the laureates' research before the honoring. A serial diffusion of ideas would also be compatible with a social constructivist view of science, but such a prediction requires additional assumptions and likely only holds for a restricted set of publications among the references cited in Nobel publications.

While our analytical approach does not allow us to fully disentangle the mechanisms behind the observed citation pattern, the analyses will give at least some hints which processes are at work against the background of these theoretical considerations.

3 Data and Analytical Approach

3.1 Database and Treatment Group

To dissect the effect of the Nobel Prize on the citation impacts of laureates' publications, we employ raw data from Clarivate Analytics' *Web of Science (WoS) 1900–2011*, including the *Science Citation Index Expanded*, the *Social Sciences Citation Index*, and the *Arts & Humanities Citation Index*, but excluding other sources such as the *Emerging Sources Citation Index* and the *Book Citation Index*. The raw data comprise over 250 files amounting to over 150 GB, originally managed by Clarivate Analytics in a databank system. We drew the necessary citation information directly from the raw data of the *WoS Core Collection* which does not cover books and publications in edited volumes and conference proceedings on the basis of unique article identifiers using Perl and R scripts. As the raw data also contain correction and gap files, which replace existing or

which add new entries, we generated a tailor-made correction and doublet filter to reproduce citation counts one-to-one as reported in the web version of the *WoS*.

We focused on the 23 winners of the Nobel Prize in Economics for the years 2000–2010 (see Table 1). One important reason for choosing the Nobel Prize in Economics for the years 2000–2010 was that coverage of publications in the *WoS* is much more comprehensive for Nobel laureates who received the award from 2000 onwards than for previous Nobel winners. Although going back in time would be definitely interesting from a substantive point of view, a more comprehensive coverage of publications improves the chances of detecting Nobel Prize effects and potential interactions should they actually exist. In addition, data for 184 publications of the 23 Nobel laureates yield a sufficient sample size for statistical analysis and stratification by publication characteristics and audience.⁵

Table 1: List of Nobel Laureates in Economics for the years 2000–2010

Year	Laureate	Year	Laureate	Year	Laureate
2000	James J. Heckman	2003	Clive W.J. Granger	2007	Roger B. Myerson
2000	Daniel McFadden	2004	Finn E. Kydland	2008	Paul Krugman
2001	George A. Akerlof	2004	Edward C. Prescott	2009	Elinor Ostrom
2001	A. Michael Spence	2005	Robert J. Aumann	2009	Oliver E. Williamson
2001	Joseph E. Stiglitz	2005	Thomas Schelling	2010	Peter A. Diamond
2002	Daniel Kahneman	2006	Edmund S. Phelps	2010	Dale Mortensen
2002	Vernon L. Smith	2007	Leonid Hurwicz	2010	Christopher Pissarides
2003	Robert F. Engle	2007	Eric S. Maskin		

⁵ To measure citation impact beyond short-term effects, a citation window of at least 3 years is desirable. Conducting a bibliometric analysis of all papers published in the year 1980 in the Web of Science, Wang (2013) has found correlations of .27, .75, .87 and .95 between the cumulative citation counts in years 1, 3, 5, and 10 after publication on the one hand and the total citations 31 years later on the other hand. Therefore, we conducted a robustness check only using Nobel laureates 2000–2008 with a minimum citation window of at least 3 years.

Next, we referred to the “Scientific Background Reports” of the Royal Swedish Academy of Sciences (www.nobelprize.org) to identify the recipients’ most important contributions. Using only those “Nobel publications” instead of all publications of the laureate offers the advantage to reduce variance in quality judgments of works and helps to build a strong case for a context with relative quality certainty. We further restricted the sample to full articles, excluding other publications by the Nobel Prize winners listed in the *Web of Science*, such as responses and corrections. Having defined the set of treated papers, we then searched for all 283 publications in the raw data of the *WoS*, collected yearly citation data for each of the 184 available Nobel publications (65%) in the raw data of the *WoS*, and linked further information regarding document, author, and publishing journal.

3.2 Construction of Synthetic Control Groups

Simple comparison of the numbers of annual citations for treated papers before and after the event is inadequate for estimating the causal effects of a Nobel Prize on citation impact because of a number of factors (for a detailed discussion see Farys & Wolbring, 2017). First, the number of citations in the *WoS* follows a strong positive time trend. Clarivate Analytics (and formerly Thomson Reuters) has substantially increased its coverage of journals over time and in 2005 added a new database, the *Book Citation Index*, to the *WoS Core Collection* (Testa, 2011). Second, modern science has expanded considerably. As a consequence, the number of publications and the average length of articles’ reference lists is nowadays considerably larger than in the past (Bornmann & Mutz, 2015). Third, citation paths of articles usually follow field-specific citation life cycles. The citation rates of most articles (disregarding Sleeping Beauties or citation classics) typically peak depending on the field several years after publication and then steadily decline (Burton & Kebler, 1960; De Solla Price, 1970). Confounding due to such time trends and

maturation effects problematizes any causal interpretation of changes in annual citations after Nobel Prize receipt.

Further strengthening these concerns for our current application is the fact that the set of Nobel Prize papers is a highly selective and highly cited subgroup which does not follow the typical citation life cycle and usually increases in citation impacts steadily ahead of the event (Garfield & Welljams-Dorof, 1992; Mazloumian et al, 2011). Hence, although a random sample of untreated papers from the *WoS* would probably suffice to control for general time trends in the citation frequency and for the growth of the global science system, this approach is not suited to adjust for biases due to selection on citation growth.

We therefore constructed tailor-made synthetic control groups which approximate the treated papers as regards publication date and yearly citations before the event (see Azoulay et al, 2014; Chan et al, 2013; Lu et al, 2013 for similar approaches).⁶ We proceeded in three steps:

First step: We generated a full list of publications in the *WoS 1900–2011*. This provides over 100 million papers as potential controls. We excluded all treated papers from this donor pool for the control group.

Second step: We performed a coarsened exact matching (CEM) procedure (Iacus, King, and Porro 2012, 2015). Unlike propensity score matching, CEM ensures that imbalances in covariates between matched observations from the treatment and control group do not exceed a certain threshold level defined *ex ante* by the specified coarsening of variables. CEM offers a good trade-off between bias reduction and the curse of dimensionality, provided that variables with numerous values are matched. In our case we use the (partially) coarsened publication year, a

⁶ Another approach to construct a control group would be to use short-listed scholars. This design would exploit the positional nature of status and the sharp discontinuities in success (Frank & Cook, 1995; Hirsch, 1977; Goode, 1978). While such analyses of Matthew effects at the author level are interesting and important when focusing on scholarly careers (e.g., Bol et al, 2018; Chan et al, 2014), several reasons led us to decide against this approach. First, the nomination list is top secret since many decades meaning that we would have to rely on public rumors. Second, some of the candidates won the prize a few years later limiting the use of this case as a control to the years between the first and the second award. Third, it is unlikely that short-listed scholars are good controls for the pre-event citation path. However, approximating the citation path of Nobel publications for the counterfactual scenario that the laureate had not received the award is essential to avoid biased estimates of the causal effect.

categorization of the cumulative number of citations and the WoS subject categories as matching criteria (for limitations of these categories see Leydesdorff & Bornmann, 2016). A match only occurs if a control paper has the treatment's exact same combination of field tags, publication year, and categorized number of cumulative citations prior to the Nobel Prize receipt. For all treated publications we matched on publication year dummies ranging from 1981 to the year of Nobel Prize receipt. For papers published before 1981, we had to be less restrictive: For papers published between 1950 and 1980 we also matched papers that did not appear in exactly the same year but in the same decade. For paper published before 1950 we searched for matches that appeared 1900–1949. Finding matches to paper with extraordinarily high citation numbers and sometimes step citation paths is especially difficult. To categorize the citation numbers, 20 percentile groups of 5% each were formed. For example, if a Nobelist paper is among the 5% most cited, then the paper of the control group must also belong to the top 5%. From this CEM procedure we derive weights for the control group as follows: if a control paper is the only possible match, it gets weight 1; if there are n matches for a paper, each of these controls gets weight $1/n$, thus forming a pool of controls for the treated paper. All unmatched papers get the weight 0 and do not appear in the further analysis.

Third step: Based on these weights, we used Entropy Balancing (Abadie, forthcoming; Abadie, Diamond, & Hainmueller, 2010; Hainmueller, 2012) to align the pre-Nobel citation life-cycle of the control group with that of the treatment group. Entropy Balancing in general relies on a reweighting scheme that calibrates weights in a way that the reweighted control group satisfies a potentially large set of prespecified balance conditions (Hainmueller, 2012). In our case we balanced the means of citations for all the single years between 1991 and the year of Nobel Prize receipt, the four decades from 1950 to 1990 as well as the time window 1900–1949. We further included the scientific field and publication date (as before) in order to preserve the previous restrictions. The control group is therefore equivalent to the treatment group in terms

of publication date, scientific field, and recent citation history up to the date of Nobel Prize receipt. Although our matching procedures do not use many variables, the strength of the approach lies in the fact that the pre-Nobel citation path controls a multitude of unobserved heterogeneity. As Abadie et al. (2010, 2015) emphasize, such a synthetic control group constructed can capture confounding unobserved characteristics, even allowing those influences to vary with time such as the reception of other awards. Because the distribution of yearly citations skews strongly to the left, the logged number of annual citations will serve as the outcome variable in the following multivariate models. We thus repeated the entropy balancing procedure for means of logged citations instead of unlogged citations. In the following, we will use weights balancing unlogged citations for a graphical inspection and weights balancing logged citations for the estimation of statistical models. Both approaches lead to the same substantive conclusions.

For the sake of transparency and in order to enable replications, paper identifiers and code are publicly and permanently available at the Harvard Dataverse (Wolbring & Farys, 2021).

3.3 Evaluation of Matching Quality

Table 2 contains descriptive statistics on the composition of the treatment and the control group prior to award announcement. The statistics illustrate that the combination of CEM and Entropy Balancing achieves covariate balance among the included variables annual citations, publication year, and subject category. Moreover, the synthetic control group closely approximate the treatment group as regards citations in the years before Nobel Prize receipt.

Table 2: Descriptive statistics for treatments and controls (weighted), prior to award announcement

	Variable	Mean	Median	SD	Min	Max
C	Logged annual citations	1.719	1.609	1.364	0	8.546
	Publication year	1977	1977	9.631	1951	2004
	Subject category "economics"	0.589	1	0.492	0	1
	Nobel Prize year	2004	2005	2.527	2000	2008
T	Logged annual citations	1.732	1.609	1.242	0	5.509
	Publication year	1977	1978	9.202	1956	2004
	Subject category "economics"	0.595	1	0.491	0	1
	Nobel Prize year	2004	2005	2.524	2000	2008

As can be seen in Figure A1, for some Nobel Prize laureates balancing is not perfect for the period of 20 to 10 years prior to the event, indicating that, in a few instances, it is difficult to find exact matches for Nobel laureates' outstanding publications as regards pre-award citation impact. This especially holds for highly cited publications by Nobel Prize winners in the years 2000 (James Heckman; Daniel McFadden) and 2004 (Finn E. Kydland; Edward C. Prescott). However, even though Nobel Prize winners' publications are already a very selective set of articles, entropy balancing ensures that citation paths of treatment and control group overlap perfectly for the 10 years before the event. As a robustness check, we dropped Nobel years with insufficient balances, but all substantive findings remained unchanged.

Moreover, some readers might worry that balancing treatment and control groups with respect to only three variables is insufficient. For example, one could additionally adjust for article length, author number, and length of reference list (see Mutz et al, 2017), since these variables also affect citation impact (Bornmann & Daniel, 2008). However, balancing for yearly citations in a large number of preintervention periods is a powerful tool to control for unobserved heterogeneity (Abadie et al, 2010, 2015) capturing those additional effects. In particular, including the flow of citations in the years before the award announcement in a rather fine-grained way helps to rule out reverse causality issues if a paper is "on the rise". Further, the chosen approach also takes into

account field-specific differences in average citations (caused by the size and hotness of a field). Because of this, the use of synthetic control groups is closely related to the normalization of citation counts by field and publication year, which is common in bibliometrics (for overviews, see Bornmann & Marx, 2015; Waltman, 2016a). However, the former approach addresses additional methodological problems (such as reverse causality and selection on citation trends; see Leszczensky & Wolbring, 2019).

3.4 Statistical Analysis

To quantify the effects of the Nobel Prize treatment, to control for confounders, and to explore potential interactions of the treatment effect with publication characteristics, we estimate linear panel regression models with the logged number of yearly citations as outcomes.⁷

To take into account the possibility of autocorrelation and heteroscedasticity, we use robust standard errors clustered around Nobel laureate for statistical inference (Angrist & Pischke, 2009). In addition to an idiosyncratic error term ε_{it} and a vector of covariates, \mathbf{X}_{it} we include paper fixed effects α_i into the model to control for time-constant influences of time-constant unobserved heterogeneity (Allison, 2009; Brüderl & Ludwig, 2015):

$$\log(Y + 1) = \beta \mathbf{X}_{it} + \beta_1 T + \alpha_i + \varepsilon_{it}$$

Including paper fixed effects avoids confounding due to time-constant effects of article features, author characteristics, publication outlet, and discipline. Consequently, the fixed effects approach removes remaining differences in the average levels of citations between treatment and synthetic control group. We first estimate a baseline model which only contains paper fixed effects and a

⁷ Annual citations are count data with overdispersion. It is state-of-the-art in bibliometrics to use negative binomial regression models (Ajiferuke & Famoy, 2015; Bornmann, Mutz, Neuhaus, & Daniel, 2008; Schubert & Glänzel, 1983). In the negative binomial regression, a logarithmic function links model regressors and outcome, but in a more complicated way than simply taking the log of Y . Because of this, first matching on the log transformed variable and then running a negative binomial regression would still provide biased estimates, because the second step would impair the balancing achieved in the first step. Thus, for the current application, we decided to use linear regression models with logged $Y+1$, which do not experience such problems.

binary treatment indicator T which changes from 0 to 1 for publications belonging to the treatment group if the current year is greater than the year of Nobel Prize receipt (model 1).⁸ Thus, although we include information on the control group in all models, we calculate point estimates and standard errors for the treatment effect in model 1 solely on the basis of the within change in annual citations in the treatment group. To take into account maturation effects in the control group and overall time trends in citations, we include in the further regression models linear, quadratic, and cubic terms for demeaned publication age (model 2) and fixed effects for calendar year (model 3). To further explore the dynamics of Nobel Prize effects across time, model 4 contains a dummy impact function for the years after the event (see Allison, 1994). This approach, which is also known as distributed fixed effects, allows to control for potential anticipation effects and to explore how the effects develops over time without imposing strong parametric restrictions on the exact functional form. Despite the non-random nature of our sample of Nobel laureates and Nobel publications, we will provide results from significance testing.^{9, 10}

4 Results

In this section, we present results on the overall effect of a Nobel Prize in Economics on citation impact, then we explore potential effect heterogeneity concerning publication characteristics and audience, and finally we test Merton's proposition of a serial diffusion of ideas.

⁸ We decided to classify the year after the Nobel announcement as first year of treatment. Press releases of the Nobel Prize in economics appear in mid-October. Publication lag due to peer review makes it unlikely that many SSCI-listed publications in that year experienced influences due to the event. We decided to classify the Nobel year as a control case. Our robustness checks corroborate this decision (see especially model 4 in Table 2).

⁹ While we are aware of the ongoing discussion in bibliometrics on the use of statistical inference in citation analysis and agree with some of the arguments pointing at conceptual difficulties (Schneider, 2016; Waltman, 2016b; Williams & Bornmann, 2016), we still believe that significance testing helps to quantify the degree of uncertainty and to get an idea how effects look like in a hypothetical super population of Nobel publications from which our sample comes from (see Berk, Western, & Weiss, 1995; Cochran, 1953; see also Abadie, Athey, Imbens, & Wooldridge, 2020 for an alternative design-based rationale).

¹⁰ Please also note that while samples sizes in the following analyses might at first glance suggest substantial statistical power and might raise questions about the value added from reporting standard errors, p-values, and confidence intervals, the effective sample size is much lower than this first impression might suggest. First, the analyses contain a large number of reweighted controls as compared to a relatively small number of 184 Nobel publications. However, for statistical inference the number of treated observations is an important determinant. Second, standard errors are clustered around Nobel laureates. This further reduces the effective sample size entering significance testing (see Snijders & Bosker, 2012). For that reasons, we decided to stick to standard thresholds of significance testing, but will keep in mind the difference between statistical and practical significance.

4.1 Matthew Effects for Nobel Laureates

The solid line in Figure 1 plots average yearly citations for Nobel Prize publications. As is apparent, the mean number of annual citations of these publications increases substantially over time and it appears that the growth in yearly citations accelerates after Nobel Prize receipt. Estimates from model 1 in Table 3, which contains only paper fixed effects and a binary treatment indicator, corroborate this conclusion. Average yearly citations increase by 89% ($1 - e^{0.637}$; $p < 0.001$) after Nobel Prize receipt. However, for the abovementioned reasons, simple pre-post comparisons are insufficient to identify causal effects in citation data and may be misleading (see also Farys & Wolbring, 2017).

Figure 1: Mean number of annual citations of Nobel Prize publication and the synthetic control group

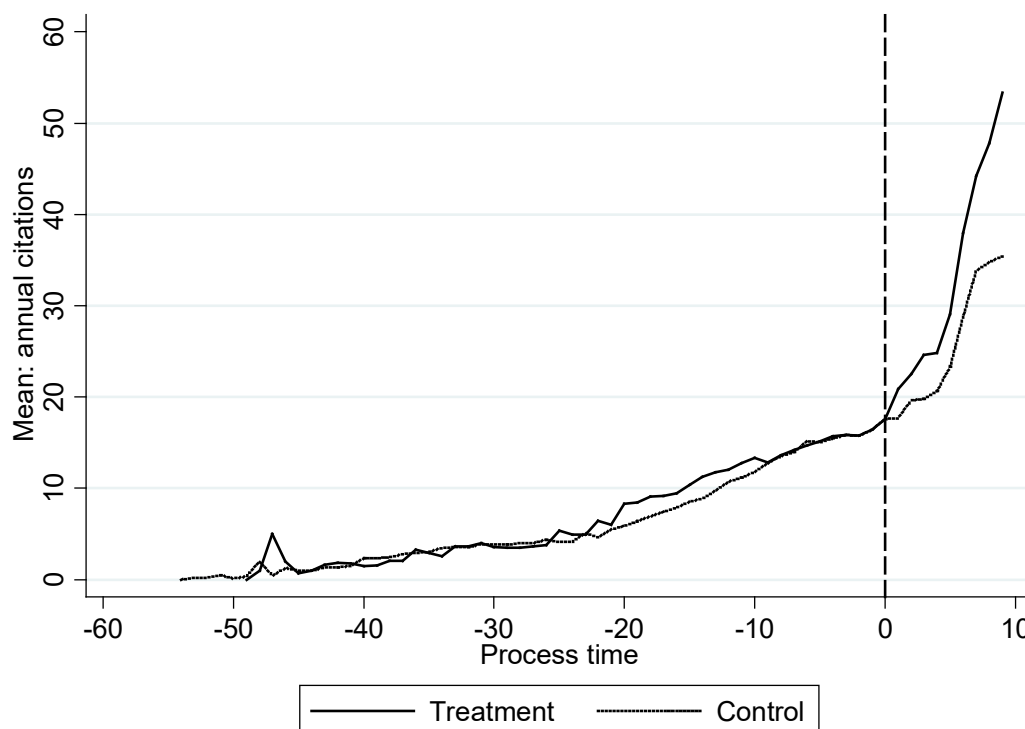


Table 3: Fixed effects linear regressions for logged annual citations of Nobel Prize publications

Outcome: log (citations +1)	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8a	Model 8b
Nobel Prize Treatment (1 if year > Nobel year)	0.637*** (9.67)	0.323*** (4.94)	0.255*** (4.78)						
Dummy Impact Function									
Year of receipt				0.074* (1.74)					
1 year after receipt				0.278*** (5.28)					
2 year after receipt				0.219** (3.39)					
3 year after receipt				0.258*** (6.20)					
4 year after receipt				0.198** (3.69)					
5 or more years after				0.303** (3.39)					
Treatment effect for publications within 5 years before the event					0.701*** (4.21)				
Nobel Prize treatment for publications 6 or more years before the event					0.236*** (4.27)				
Nobel Prize treatment for highly cited publications (top 5%)						0.260*** (3.84)			
Nobel Prize treatment for non-highly cited publications						0.252** (4.10)			
Nobel Prize treatment for publications in high impact journals (top 5%)							0.250*** (4.25)		
Nobel Prize treatment for publications in non-high impact journals (top 5%)							0.254* (2.64)		
Nobel Prize treatment by audience (m8a: econ; m8b: other SSCI journal)								0.233*** (4.39)	0.110* (2.66)
Publication age: 2nd & 3rd polynomial		<i>included</i>	<i>included</i>	<i>included</i>	<i>included</i>	<i>included</i>	<i>included</i>	<i>included</i>	<i>Included</i>
Year fixed effects			<i>included</i>	<i>included</i>	<i>included</i>	<i>included</i>	<i>included</i>	<i>included</i>	<i>Included</i>
Constant	1.792*** (256.20)	2.048*** (62.13)	2.369*** (21.08)	2.359*** (20.89)	2.325*** (20.02)	2.297*** (19.54)	2.402*** (23.08)	1.980*** (20.97)	1.023*** (15.58)
Publication-years	1,876,508	1,876,508	1,876,508	1,876,508	1,876,508	1,876,508	1,876,508	1,876,508	1,876,508
Publications	76,626	76,626	76,626	76,626	76,626	76,626	76,626	76,626	76,626
-LL	2,015,577	1,818,957	1,806,169	1,805,640	1,803,902	1,767,022	1,803,793	1,687,653	1,316,051
AIC	4,031,156	3,637,922	3,612,466	3,611,419	3,607,859	3,534,088	3,607,630	3,375,358	2,632,157
BIC	4,031,169	3,637,972	3,613,262	3,612,278	3,608,208	3,534,362	3,607,903	3,375,682	2,633,506

Note: Fixed effects regression model with robust standard errors clustered around Nobel laureates. Unstandardized coefficients; t statistics in parentheses. + p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001

It is thus necessary to compare the citation paths of the treatment and the tailor-made synthetic control group (dashed line). As becomes clear from visual inspection, the synthetic control group closely approximates the treatment group as regards citations in the years before Nobel Prize receipt. However, after Nobel Prize receipt citation paths for the treatment and the control group diverge: average differences in citation impacts amount to 5.7 annual citations per publication five years after the announcement and 11.5 annual citations per publication ten years after the announcement.

Models 2 and 3 in Table 3 shed further light on the Matthew effect while taking into account maturation effects in the control group and overall time trends in citations by including the first, second, and third polynomials of publication age (model 2) and year fixed effects (model 3). In consequence of this covariate adjustment, the treatment effect estimate for treated publications decreases considerably, particularly when we control for both sources of confounding in model 3. However, with an increase of 29% in annual citations (model 3; $1 - e^{0.255}$; $p < 0.001$) the increase remains significant both from a statistical and from a substantive point of view.¹¹

To further explore the dynamics of Nobel Prize effects across time, model 4 contains a dummy impact function for the years after the event. As can be seen, the annual number of citations of Nobel publications increases by 32% ($1 - e^{0.278}$; $p < 0.001$) in the year after receipt. This effect is remarkably stable across time and is still present five years after the event and later. With an increase of 35% ($1 - e^{0.303}$; $p < 0.001$), the effect is even slightly, although not significantly stronger five or more years after Nobel Prize receipt providing further suggestive evidence on the rich-

¹¹ This result is remarkably robust with respect to direction and strength if we drop 2000 and 2004 laureates, for whom we could achieve only imperfect balance. As another sensitivity analysis, we estimated triple and quadruple difference in differences model (Lee, 2016), which both demeans and (linearly or quadratically) detrends the data and hence provides another way to control for selection on citation impact and for selection on citation growth for Nobel Prize publications. The effects are remarkable similar to our results using a synthetic control group. Separate analyses for each Nobel Prize year further support our conclusions, but also illustrate heterogeneity with respect to average citation levels and strength of treatment effects (see Figure A1 in the appendix). Visual inspection indicates considerable treatment effects for publications of Nobel Prize winners in the years 2000, 2002–2004, 2006, and 2008, but not for laureates in the years 2001, 2005, and 2007.

getting-richer phenomenon in academia. In addition, model 4 serves as a robustness check for the correct specification of the timing of the event. The fact that the increase in annual citations is much smaller for the year of the Nobel Prize receipt corroborates our assumption of a delayed treatment effect on citations due to publication lag.

4.2 Interaction with Publication Characteristics and Audience

Next, we ran three models containing interaction effects with dummies for (a) publication age (published within 5 years before Nobel Prize receipt), (b) journal impact (top 5% in the subject category according to journal impact factor), and (c) pre-Nobel citation impact (top 5% according to the cumulative number of citations before Nobel Prize receipt). To test for variation in treatment effects by audience, we analyzed two different citation outcomes in separate models: logged yearly citations from “insiders” of the focal scientific field of economics (citations from publications in the *WoS* subject categories “economics”, “business”, “business, finance”, and “management”) and from “outsiders” (citations from publications in all other *WoS* subject categories covered by the Social Science Citation Index) (see Lynn, 2014 for a similar approach; for a more fine-grained approach to measure within-field and out-of-field citations see Reschke et al, 2018).¹²

Model 5 in Table 3 shows that considerable heterogeneity in the strength of treatment effects exists with regard to publication year. The treatment effect on citation impact for papers published up to 5 years before Nobel Prize receipt is much stronger as compared to less recent publications. The latter also receive a considerable attention boost but to a far lesser extent. Even after controlling for maturation effects using polynomials for publication age and calendar year fixed effects, more recent publications enjoy greater benefits from the Nobel Prize as regards citation impact. Annual citations of papers published up to five years before the event increased by 102%

¹² As a robustness check, we restricted our analyses to publication-years within ten years before the event and publications-years following Nobel Prize receipt. Sufficiently close balance between the treatment and the control group could be achieved for the publications-years within ten years before the event but not prior to that time period. The following results are robust to this sensitivity analysis.

$(1 - e^{0.701}; p < 0.001)$, whereas citations of publications appearing more than five years before the event only grew by 27% $(1 - e^{0.236}; p < 0.001)$.

In contrast to the results by publication year, the other two interactions in models 6 and 7 turn out to be not relevant as regards both substantive and statistical significance. Both highly cited (30%; $1 - e^{0.260}; p < 0.001$) and non-highly cited papers (29%; $1 - e^{0.252}; p < 0.01$) experience similar growth in citations after the prize as do publications in journals with very high field-specific impact factor (28%; $1 - e^{0.250}; p < 0.001$) and publications in all other journals (29%; $1 - e^{0.254}; p < 0.001$).

Finally, models 8a and 8b show that the Nobel Prize affects the citation behavior of both “insiders” and “outsiders” but has stronger effects on the former. Annual citations by publications in “economic” journals increase by 26% $(1 - e^{0.233}; p < 0.001)$, whereas citations by publications in other SSCI-listed journals increase by only 12% $(1 - e^{0.110}; p < 0.05)$.¹³ While the citation boost caused by outsiders might be due to their lower degree of familiarity with the work of the laureate before the award, we interpret the stronger effect for better-informed “insiders” from the focal field of research as an indication that citation impact not only increases because awards raise awareness for the work of Nobel laureates. Instead, the social recognition of the scientific achievement seems to additionally cause scholars to increasingly cite Nobel Prize publications.

4.3 Is there a Serial Diffusion of Ideas?

For the sake of analytical clarity, we distinguish among works by Nobel laureates (publications of *first degree* in the citation network), works they cite (*second degree*), and further works cited by works in the Nobel laureates’ cited references but not by the laureates themselves (*third degree*). To test for a “serial diffusion of ideas”, we extracted the reference lists of the Nobel Prize

¹³ The effects for “insiders” and “outsiders” remain statistically significant and become slightly larger if we omit post-Nobel publications by the psychologist Daniel Kahneman from our analyses. The reason for this slight change in results is the different pattern of audience-specific reactions to his receiving the prize: Citations of his work in economics journals increased by 32%, while citations in other SSCI journals increased by 23%. The latter increase is not restricted to psychological publication outlets, but reflect a more diverse growth in citations. Thus, due to the rather surprising decision of the Nobel Committee to honor a disciplinary “outsider,” Kahneman’s research program became more visible and gained in citation intensity both inside and outside economics.

publications and searched for papers of second degree (59% found; 1,380 out of 2,349). We repeated the step for publications of third degree (74% found; 12,134 out of 16,483) and generated synthetic control groups the same way we did for the first degree as described in section 3. Figure A2 in the appendix shows that treatment and control groups are almost perfectly balanced as regards pre-award citation paths.

Fixed effects models in Table 4 reveal that – after controlling for citation life cycles and general increases in citations – the Nobel Prize has no effect on citation impact of publications of second and third degree in the citation networks. Hence, we find no evidence of a serial diffusion of ideas: While publications of Nobel laureates receive more attention due to the award, cited references do not profit, but also do not suffer from the honoring as regards citation impacts.

Table 4: Fixed effects linear regressions for publications of second and third degree

	Model 1		Model 2		Model 3	
Outcome: log (citations +1)	Degree 2	Degree 3	Degree 2	Degree 3	Degree 2	Degree 3
Nobel Prize Treatment (1 if year > Nobel year)	0.244*** (11.17)	0.053*** (8.68)	-0.058** (-3.19)	-0.009+ (-1.67)	-0.016 (-0.77)	-0.001 (-0.16)
Publication age: 2nd & 3rd polynomial			<i>included</i>	<i>included</i>	<i>included</i>	<i>included</i>
Year fixed effects					<i>included</i>	<i>included</i>
Constant	1.421*** (701.43)	1.091*** (2071.73)	1.505*** (230.19)	1.141*** (649.89)	1.916*** (68.17)	1.336*** (128.08)
Publication-years	11,375,716	62,515,257	11,375,716	62,515,257	11,375,716	62,515,257
Publications	415,308	1,707,153	415,308	1,707,153	415,308	170,7153
-LL	12,070,423	58,660,654	11,410,798	58,008,944	11,235,676	57,440,026
AIC	24,140,848	117,321,309	22,821,604	116,017,896	22,471,579	114,880,282
BIC	24,140,863	117,321,325	22,821,661	116,017,960	22,473,204	114,882,116

Note: Fixed effects regression model with robust standard errors clustered around publications. Unstandardized coefficients; t statistics in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

5 Discussion

Using the case of the reception of the Sveriges Riksbank Prize in Economic Sciences in Memory of Alfred Nobel, we investigated, on the basis of the complete *Web of Science 1900–2011*, Nobel Prize effects upon citation impacts. Thus, this study provides the first rigorous analysis of the

Matthew effect in science using Merton's and Zuckerman's original example, Nobel Prize laureates. In a nutshell, we found clear evidence for a Matthew effect and hence for the existence of cumulative advantages in this supposedly meritocratic field. This finding is well in line with previous studies on the effects of other positive status shocks in the mid-career stage on citation impacts (Azoulay et al, 2014; Chan et al, 2013) as well as the likelihood to receive research funding and to become a full professor (Bol et al, 2018). Our study contributes to this literature by empirically showing that these processes are not restricted to the early and mid-career stage. The “crowning” of scientific careers with a Nobel Prize causes such Matthew effects with respect to citation impacts even among already well-established and usually highly-cited scholars.

Moreover, our analyses revealed that scholars from the focal field of the award are more receptive to decisions of the Nobel committee. While we can only speculate about the exact reasons for this finding, our results suggest that the substantial gain in legitimacy is the key mechanism for Nobel Prize effects upon citation impacts. In line with the social constructivist theory of citations, scholars in the focal field of the prize might try to exploit this increased credibility of the laureate to their advantage or feel compelled—due to expectations within the scientific community—by citing Nobel Prize to bolster their own arguments and to profit from the laureates' prestige. In the extreme scenario of “ceremonial” citations (see Adatto & Cole, 1981), scholars may cite Nobel Prize publications without personally believing in a high quality or without having actually read the papers in detail. Against this background, it seems likely that honoring a laureate with the Nobel Prize causes to some degree strategic citations in the focal field, while the mechanisms proposed by the normative theory of citations are likely simultaneously at work.

These findings have broader implications for science. First, our findings corroborate previous research showing that science is a social system which is not only driven by meritocratic considerations (Cole & Cole, 1970; Merton, 1973) but also by issues of persuasion, social expectations, and peer pressure (Callon et al, 1986; Knorr-Cetina, 1981; Latour, 1987). Such social

influence creates strategic incentives for scholars to use symbolic acts of recognition such as “ceremonial” citations and to float with the current instead of acting purely upon what they think is best from a scientific point of view. Second, awards and other forms of social recognition can cause concentration processes in science by providing focal points (Frank & Cook, 1995; Frey & Gallus, 2014; van Dalen & Henkens, 2005). This can have negative side effects for other scholars and can undermine the innovation potential of science (Bothner et al, 2011; Merton, 1968, 1988). Important contributions standing in the shadow of Nobel laureates might remain uncited and might be finally forgotten. Third, we have shown that reactions to awards can be audience-specific and are often limited to certain fields (for related arguments on audience- specificity see Ertug et al, 2016; Keuschnigg, 2015; Lynn, 2014). An award does not uniformly raise the legitimacy of a scholar’s research, but does so to different degrees among different audiences. Future research should further explore under what conditions awards cause a relevant status shift for an audience.

Limitations and Outlook for Future Research

These results and conclusions should be interpreted cautiously in light of a few limitations, which future research must address. First, citations are not only building blocks of scientific claims and mark the origin of certain ideas, but they can also serve very different functions (Baldi, 1998; Bornmann & Daniel, 2008; Leydesdorff, 1998). Our study suggests that considerations of legitimacy, persuasion, and peer pressure also drive citation. To provide a more direct test of these considerations, future research might extend our approach by distinguishing positive from negative citation or even use topic modelling techniques to enrich citations with context (see Ding et al, 2014; Tahamtan & Bornmann, 2019; Yan, Chen, & Li, 2020; Zhu, Turney, Lemire, & Vellino, 2015).

Second, using the raw data of the *Web of Science 1900–2011 (WoS)*, we had to exclude other sources such as the *Emerging Sources Citation Index* and the *Book Citation Index* from our analysis. Our estimates might hence not map the average treatment effect for all relevant

publications. However, the fact that only a few of the Nobel laureates in economics published their central insights and research findings in books or unlisted journals limits the potential impact of this pitfall upon our results. Another important consequence of the restriction to certain types of publications is that we have to assume that citation data are missing at random. A violation of this assumption would not affect the internal validity of our results, but would limit the generalizability of our results.

Third, we balanced the treatment and control groups on observable covariates. Due to the large number of potential control cases, except for a few outstanding Nobel Prize winners' publications (which we excluded in sensitivity analyses), common support was not an issue. Still, publications might differ in terms of inherently difficult-to-measure aspects, such as "quality". However, matching on the pre-event citation impact, a fixed effects approach, and higher-order difference in differences models capture a substantial portion of such unobserved heterogeneity (see also Abadie et al, 2010, 2015). While this helps to minimize the uncertainty in our causal inferences, such models still rely on assumptions and only indirectly control for field-specific dynamics and the hotness of a field. An approach using keyword matching or topic modeling would get closer to this, though this invites the curse of dimensionality in matching (Abadie & Imbens, 2006).

Fourth, we decided to study Matthew effects upon citation impacts at the level of individual publications. Taking into account selection effects by matching Nobel publications with publications of similar citation impact, we estimated the increase in citation numbers caused by the honoring. While this approach recognizes the fact that cumulative advantages are already at play for future Nobel laureates before Nobel receipt by controlling for their often already exceptionally high pre-Nobel citation impact, we were not able to disentangle the direct effects of the Nobel Prize upon citation impacts from its indirect effects in the form of further cumulative advantages. However, access to generous research funding, additional awards, and prestigious memberships in scientific academies might further increase citation numbers. Finally, it remains an open question whether the findings generalize to other Nobel Prize winners in economics and,

more importantly, to Nobel laureates in other disciplines and to other awards. Future research should hence on the one hand concentrate on the question of how different disciplinary citation cultures moderate effects due to the Nobel Prize in different research areas. On the other hand, it might be well worth the effort to further investigate the effects of awards for younger, less-established scholars (see Azoulay et al, 2014; Bol et al, 2018; Chan et al, 2013). It appears likely that Matthew effects of early and mid-career awards are stronger for two reasons. On the one hand, these scholars are much less well-known than future Nobel laureates, increasing the importance of status signals. On the other hand, status advantages have more time to work and can accumulate over a longer period of time.

References

- Abadie, A. (forthcoming). Using synthetic controls: Feasibility, data requirements, and methodological aspects. *Journal of Economic Literature*.
- Abadie, A., Athey, S., Imbens, G.W., & Wooldridge, J.M. (2020). Sampling-based versus Design-based Uncertainty in Regression Analysis. *Econometrica*, 88(1), 265–296.
- Abadie, A., Diamond, A., & Hainmueller, J. (2010). Synthetic control methods for comparative case studies: Estimating the effect of California's tobacco control program. *Journal of the American Statistical Association*, 105, 493–505.
- Abadie, A., Diamond, A., & Hainmueller, J. (2015). Comparative politics and the synthetic control method. *American Journal of Political Science*, 59(2), 495–510.
- Abadie, A., & Imbens, G.W. (2006). Large sample properties of matching estimators for average treatment effects. *Econometrica*, 74(1), 235–267.
- Adatto, K., & Cole, S. (1981). The functions of classical theory in sociological research: The case of Max Weber. In R. A. Jones & H. Kuklick (Eds.). *Knowledge and Society: Studies in the Sociology of Culture. Past and Present*, Vol. 3 (pp.137–162). Greenwich, CT: JAI Press.
- Ajijferuke, I., & Famoye, F. (2015). Modelling count response variables in informetric studies: Comparison among count, linear, and lognormal regression models. *Journal of Informetrics*, 9(3), 499–513.

- Allison, P.D. (1994). Using panel data to estimate the effects of events. *Sociological Methods and Research*, 23(2), 174–199.
- Allison, P.D. (2009). *Fixed Effects Regression Models*. Thousand Oaks, CA: Sage.
- Allison, P.D., Long, J.S., & Krauze, T.K. (1982). Cumulative advantage and inequality in science. *American Sociological Review*, 47(5), 615–625.
- Angrist, J.D., & Pischke, J.-S. (2009). *Mostly Harmless Econometrics. An Empiricist's Companion*. Princeton/Oxford: Princeton University Press.
- Azoulay, P., Stuart, T.E., & Wang, Y. (2014). Matthew: Effect or Fable? *Management Science*, 60(1), 92–109.
- Azoulay, P., Zivin, J.G., & Wang, J. (2010). Superstar extinction. *Quarterly Journal of Economics*, 125(2), 549–589.
- Baldi, S. (1998). Normative versus social constructivist processes in the allocation of citations: A network-analytic model. *American Sociological Review*, 63(6), 829–846.
- Berk, R.A., Western, B., & Weiss, R.E. (1995). Statistical inference for apparent populations. *Sociological Methodology*, 25, 421–458.
- Bjork, S., Offer, A., & Söderberg, G. (2014). Time series citation data: the Nobel Prize in economics. *Scientometrics*, 98(1), 185–196.
- Boettke, P.J., Fink, A., & Smith, D.J. (2012). The impact of Nobel Prize winners in economics: Mainline vs. mainstream. *American Journal of Economics and Sociology*, 71(5), 1219–1249.
- Bol, T., de Vaan, M., & van de Rijdt, A. (2018). The Matthew effect in science funding. *Proceedings of the National Academy of Sciences of the United States of America*, 115(19), 4887–4890.
- Bornmann, L., & Daniel, H.-D. (2008). What do citation counts measure? A review of studies on citing behavior. *Journal of Documentation*, 64(1), 45–80.
- Bornmann, L., & Marx, W. (2015). Methods for the generation of normalized citation impact scores in bibliometrics: Which method best reflects the judgements of experts? *Journal of Informetrics*, 9(2), 408–418.
- Bornmann, L., & Mutz, R. (2015). Growth rates of modern science: A bibliometric analysis based on the number of publications and cited references. *Journal of the Association for Information Science and Technology*, 66(11), 2215–2222.
- Bornmann, L., Mutz, R., Neuhaus, C., & Daniel, H.-D. (2008). Citation counts for research evaluation: Standards of good practice for analyzing bibliometric data and presenting and interpreting results. *Ethics in Science and Environmental Politics*, 8(1), 93–102.

- Bothner, M.S., Podolny, J., & Smith, E.B. (2011). Organizing contests for status: The Matthew effect vs. the Mark effect. *Management Science*, 57(3), 439–457.
- Brüderl, J., & Ludwig, V. (2015). Fixed-effects panel regression. In H. Best & C. Wolf (Eds.). *Regression Analysis and Causal Inference* (pp. 327–357). Thousand Oaks, CA: Sage.
- Burton, R.E., & Kebler, R.W. (1960). The “half-life” of some scientific and technical literatures. *American Documentation*, 11(1), 18–22.
- Callon, M., Law, J., & Rip, A. (Eds) (1986). *Mapping the Dynamics of Science and Technology*. *Sociology of Science in the Real World*. The MacMillan Press: London.
- Chan, H.F., Frey, B.S., Gallus, J., & Torgler, B. (2013). Does the John Bates Clark Medal boost subsequent productivity and citation success? *UZH Working Paper ECON* 111.
- Chan, H.F., Gleeson, L., & Torgler, B. (2014). Awards before and after the Nobel Prize: A Matthew effect and/or a ticket to one’s funeral? *Research Evaluation*, 23(3), 210–220.
- Cochran, W.G. (1953). *Sampling Techniques*. New York: Wiley.
- Cole, J.R., & Cole, S. (1973). *Social Stratification in Science*. Chicago: University of Chicago Press.
- Cole, S. (1970). Professional standing and the reception of scientific discoveries. *American Journal of Sociology*, 76(2), 286–306.
- Collins, H.M. (1999). Tantalus and the aliens: publications, audiences and the search for gravitational waves. *Social Studies of Science*, 29(2), 163–197.
- Cronin, B. (2005). A hundred million acts of whimsy? *Current Science*, 89(9), 1505–1509.
- De Solla Price, D.J. (1970). Citation measures of hard science, soft science, technology and nonscience. In C.E. Nelson & D.K. Pollack (Eds.). *Communication among Scientists and Engineers* (pp. 3–22). Lexington, MA: D.C. Heath and Company.
- Diamond, A.M. (1988). Citation counts for Nobel Prize winners in economics. *History of Economics Science Bulletin*, 10(1), 67–70.
- Ding, Y., Zhang, G., Chambers, T., Song, M., Wang, X., & Zhai, C. (2014). Content-based citation analysis: The next generation of citation analysis. *Journal of the Association for Information Science and Technology*, 65(9), 1820–1833.
- DiPrete, T., & Eirich, G. (2006). Cumulative advantage as a mechanism for inequality: A review of theory and evidence. *Annual Review of Sociology*, 32, 271–297.

- Ertug, G., Yogeve, T., Lee, Y.G., & Hedström, P. (2016). The art of representation: How audience-specific reputations affect success in the contemporary art field. *Academy of Management Journal*, 59(1), 113–134.
- Falkinger, J. (2008). Limited attention as a scarce resource in information-rich economies. *Economic Journal*, 118(8), 1596–1620.
- Farys, R., & Wolbring, T. (2017). Matched control groups for modeling events in citation data: An illustration of Nobel Prize effects in citation networks. *Journal of the Association for Information Science and Technology*, 68(9), 2201–2210.
- Franck, G. (2002). The scientific economy of attention: A novel approach to the collective rationality of science. *Scientometrics*, 55(1), 3–26.
- Frandsen, T.F., & Nicolaisen, J. (2013). The ripple effect: Citation chain reactions of a Nobel Prize. *Journal of the American Society for Information Science and Technology*, 64(3), 437–447.
- Frank, R.H., & Cook, P.J. (1995). *The Winner-take-all Society*. New York: The Free Press.
- Frey, B.S., & Gallus, J. (2014). Awards are a special kind of signal.” *CREMA Working Paper* No. 2014-04.
- Garfield, E., & Welljams-Dorof, A. (1992). Of Nobel class: A citation perspective on high impact research authors. *Theoretical Medicine*, 13(2), 117–135.
- Goode, W.J. (1978). *The Celebration of Heroes: Prestige as a Social Control System*. Berkeley, CA: University of California Press.
- Hainmueller, J. (2012). Entropy balancing for causal effects: A multivariate reweighting method to produce balanced samples in observational studies. *Political Analysis*, 20(1), 25–46
- Hainmueller, J., & Xu, Y. (2013). ebalance: A Stata package for entropy balancing. *Journal of Statistical Software* 54(7), 1–18.
- Hirsch, F. (1977). *The Social Limits of Growth*. London and Henley: Routledge & Kegan Paul.
- Iacus, S.M., King, G., & Porro, G. (2012). Causal inference without balance checking: Coarsened exact matching. *Journal of Political Analysis*, 20(1), 1–24.
- Iacus, S.M., King, G., & Porro, G. (2014). A theory of statistical inference for matching methods in applied causal research. Working Paper.
- Karier, T. (2010). *Intellectual Capital: Forty Years of the Nobel Prize in Economics*. New York: Cambridge University Press.

- Keuschnigg, M. (2015). Product success in cultural markets: The mediating role of familiarity, peers, and experts. *Poetics*, 51, 17–36.
- Knorr-Cetina, K. (1981). *The Manufacture of Knowledge: An Essay on the Constructivist and Contextual Nature of Science*. Oxford: Pergamon Press.
- Lee, M.-j. (2016). Generalized difference in differences with panel data and least squares estimator. *Sociological Methods and Research*, 45(1), 134–157.
- Leszczensky, L., & Wolbring, T. (2019). How to deal with reverse causality using panel data? Recommendations for researchers based on a simulation study. *Sociological Methods & Research*. doi: 10.1177/0049124119882473.
- Leydesdorff, L. (1998). Theories of citation? *Scientometrics*, 43(1), 5–25.
- Leydesdorff, L., & Bornmann, L. (2016). The operationalization of “fields” as WoS subject categories (WCs) in evaluative bibliometrics: The cases of “Library and Information Science” and “Science & Technology Studies”. *Journal of the Association for Information Science and Technology*, 67(3), 707–714.
- Lu, S.F., Jin, G.Z., Uzzi, B., & Jones, B. (2013). The retraction penalty: Evidence from the Web of Science. *Nature: Scientific Reports* 3, doi:10.1038/srep03146.
- Luukkonen, T. (1997). Why has Latour’s theory of citations been ignored by the bibliometric community? Discussion of sociological interpretations of citation analysis. *Scientometrics*, 38(1), 27–37.
- Lynn, F.B. (2014). Diffusing through disciplines: Insiders, outsiders, and socially influenced citation behavior. *Social Forces*, 93(1), 355–382.
- MacRoberts, M.H., & MacRoberts, B.R. (1987). Another test of the normative theory of citing. *Journal of the American Society for Information Science*, 38(4), 305–306.
- Mazlounian, A., Eom, Y.-H., Helbing, D., Lozano, S., & Fortunato, S. (2011). How citation boosts promote scientific paradigm shifts and Nobel Prizes. *PLoS ONE*, 6(5), e18975.
- Merton, R.K. (1968). The Matthew effect in science. *Science*, 159, 56–63.
- Merton, R.K. (1973). *The Sociology of Science. Theoretical and Empirical Investigations*. Chicago: University of Chicago Press.
- Merton, R.K. (1988). The Matthew effect in science, II: Cumulative advantage and the symbolism of intellectual property. *ISIS*, 79(4), 606–623.
- Merton, R.K. (1995). The Thomas Theorem and the Matthew effect. *Social Forces*, 74(2), 379–424.

- Michayluk, D., & Zurbregg, R. (2014). Do lead articles signal higher quality in the digital age? Evidence from finance journals. *Scientometrics*, 98(2), 961–973.
- Moed, H.F., & Garfield, E. (2004). In basic science the percentage of ‘authoritative’ references decreases as bibliographies become shorter. *Scientometrics*, 60(3), 295–303.
- Mutz, R., Wolbring, T., & Daniel, H.-D. (2017). Signaling high quality papers in Angewandte Chemie International Edition and its effect on citation impact: A propensity score matching analysis of the designation ‘Very Important Paper’ (VIP). *Journal of the Association for Information Science and Technology*, 68(9), 2139–2153.
- Nicolaisen, J. (2007). Citation analysis. *Annual Review of Information Science and Technology*, 41(1), 609–641.
- Offer, A., & Söderberg, G. (2017). *The Nobel Factor: The Prize in Economics, Social Democracy, and the Market Turn*. Princeton: Princeton University Press.
- Petersen, A.M., Fortunato, S., Pan, R.K., Kaski, K., Penner, O., Rungi, A., Riccaboni, M., Stanley, H.E., & Pammolli, F. (2014). Reputation and impact in academic careers. *Proceedings of the National Academy of Sciences of the United States of America*, 111(43), 15316–15321.
- Price, D.J. de S. (1976). A general theory of bibliometric and other cumulative advantage processes. *Journal of the American Society for Information Science*, 27(5), 292–306.
- Reschke, B.P., Azoulay, P., & Stuart, T.E. (2018). Status spillovers: The effect of status-conferring prizes on the allocation of attention. *Administrative Science Quarterly*, 63(4), 819–847.
- Rigney, D. (2010). *The Matthew Effect. How Advantage Begets Further Advantage*. New York: Columbia University Press.
- Safer, M.A., & Tang, R. (2009). The psychology of referencing in psychology journal articles. *Perspectives on Psychological Science*, 4(1), 51–53.
- Salganik, M.J., Dodds, P.S., & Watts, D.J. (2006). Experimental study of inequality and unpredictability in an artificial cultural market. *Science*, 311, 854–856.
- Schneider, J.W. (2016). The imaginary of statistical inference when data are the population: Comments to Williams and Bornmann. *Journal of Informetrics*, 10(4), 1243–1248.
- Schubert, A., & Glänzel, W. (1983). Statistical reliability of comparisons based on the citation impact of scientific publications. *Scientometrics*, 5(1), 59–74.
- Shadish, W.R., Tolliver, D., Gray, M., & Sengupta, S.K. (1995). Author judgments about works they cite – three studies from psychology journals. *Social Studies of Science*, 25(3), 477–498.

- Snijders, T.A.B., & Bosker, R.J. (2012). *Multilevel Analysis: An Introduction to Basic and Advanced Multilevel Modeling* (2nd edition). London: Sage.
- Strevens, M. (2006). The role of the Matthew effect in science. *Studies in History and Philosophy of Science*, 37(2), 159–170.
- Tahamtan, I., & Bornmann, L. (2018). Core elements in the process of citing publications: Conceptual overview of the literature. *Journal of Informetrics*, 12(1), 203–216.
- Tahamtan, I., & Bornmann, L. (2019). What do citation counts measure? An updated review of studies on citations in scientific documents published between 2006 and 2018. *Scientometrics*, 121(3), 1635–1684.
- Testa, J. (2011). *The Globalization of Web of Science, 2005-2010*. Philadelphia: Thomson Reuters.
- Thornley, C., Watkinson, A., Nicholas, D., Volentine, R., Jamali, H.R., Herman, E., Allard, S., Levine, K.J., & Tenopir, C. (2015). The role of trust and authority in the citation behaviour of researchers. *Information Research*, 20(3), paper 677.
- Van Dalen, H.P., & Henkens, K.E. (2005). Signals in science – On the importance of signaling in gaining attention in science. *Scientometrics*, 64(2), 209–233.
- Van de Rijdt, A., Kang, S.M., Restivo, M., & Patil, A. (2014). Field experiments of success-breeds-success dynamics. *Proceedings of the National Academy of Sciences of the United States of America*, 111(19), 6934–6939.
- Waltman, L. (2016a). A review of the literature on citation impact indicators. *Journal of Informetrics*, 10(2), 365–391.
- Waltman, L. (2016b). Conceptual difficulties in the use of statistical inference in citation analysis. *Journal of Informetrics*, 10(4), 1249–1252.
- Wang, J. (2013). Citation time window choice for research impact evaluation. *Scientometrics*, 94(3), 851–872.
- Wang, J. (2014). Unpacking the Matthew effect in citations. *Journal of Informetrics*, 8(2), 329–339.
- White, H.D. (2004). Reward, persuasion, and the Sokal Hoax: a study in citation identities. *Scientometrics*, 60(1), 93–120.
- Williams, R., & Bornmann, L. (2016). Sampling issues in bibliometric analysis. *Journal of Informetrics*, 10(4), 1225–1232.
- Wolbring, T., & Farys, R. (2021). Replication code for “Matthew Effects in Science and the Serial Diffusion of Ideas: Testing Old Ideas with New Methods”. *Harvard Dataverse*, doi:10.7910/DVN/SPYRIG, V1.

- Yan, E., Chen, Z., & Li, K. (2020). Authors' status and the perceived quality of their work: Measuring citation sentiment change in nobel articles. *Journal of the Association for Information Science and Technology*, 71(3), 314–324.
- Zhu, X., Turney, P., Lemire, D., & Vellino, A. (2014). Measuring academic influence: Not all citations are equal. *Journal of the Association for Information Science and Technology*, 66(2), 408–427.
- Zuckerman, H. (1977). *Scientific Elite: Nobel Laureates in the United States*. New York: The Free Press.
- Zuckerman, H. (2011). The Matthew effect writ large and larger: A study in sociological semantics. In E. Yehuda, S. András, & L. György (Eds.). *Concepts and the Social Order: Robert K. Merton and the Future of Sociology* (pp. 121–164). Budapest: Central European University Press.

Appendix

Figure A1: Mean number of annual citations of publications separately by Nobel year

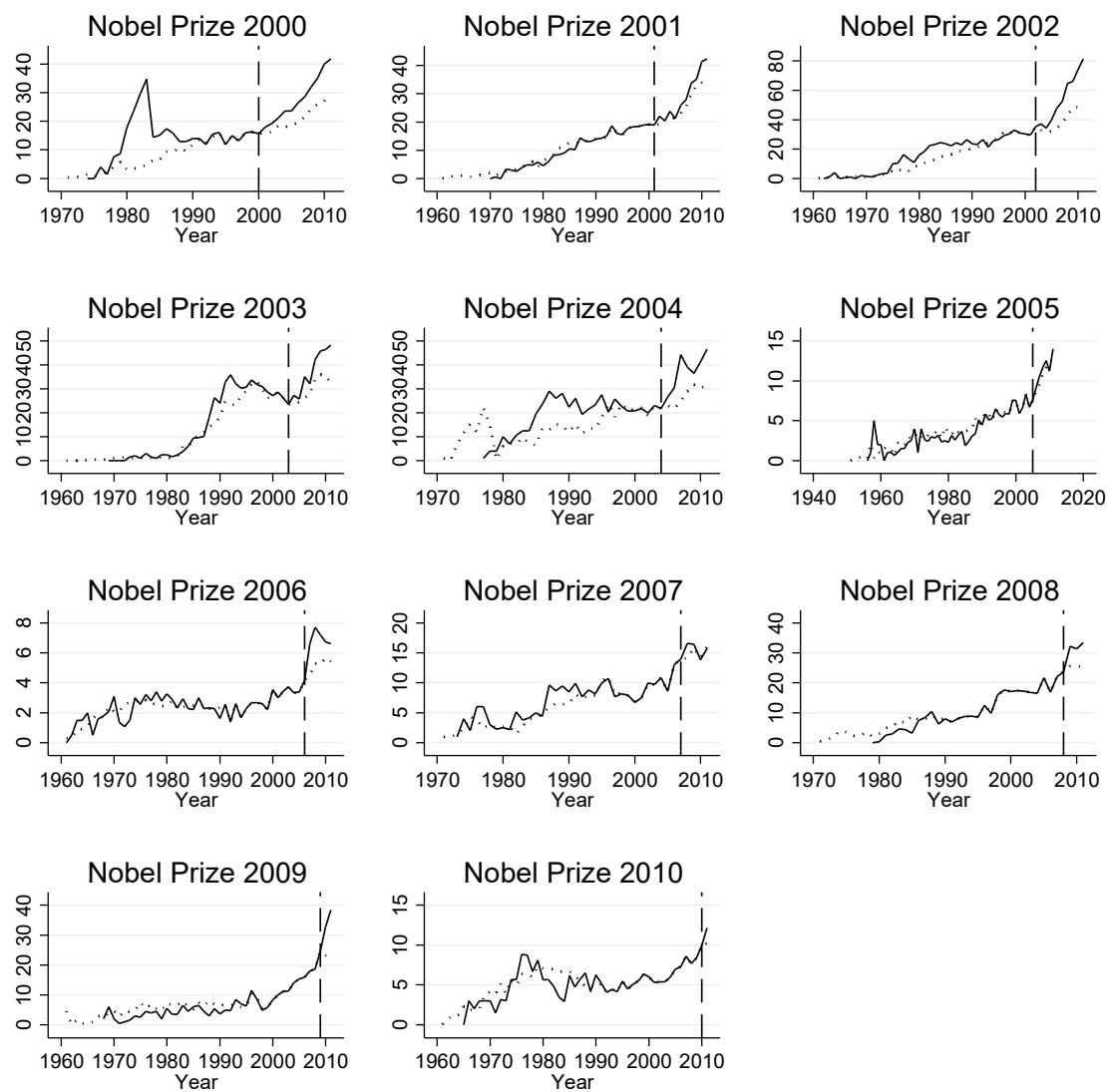


Figure A2: Mean number of annual citations of publications of second and third degree

