Paper to be presented at
the DRUID16 20th Anniversary Conference
Copenhagen, June 13-15, 2016

# Guilt by Association

**Katrin Hussinger**
**University of Luxembourg**
**CREA**
**katrin.hussinger@uni.lu**

**Maikel Pellens**
**Centre for European Economic Research (ZEW)**
**Industrial Economics and International Management**
**pellens@zew.de**

**Abstract**

Scientific misconduct harms scientific progress. One of the channels through which this happens is distrust: afraid to build on unreliable scientific results, scientists tend to avoid the work of authors which they associate with unethical research behaviour. While prudent, this strategy risks the misjudged ignoring of good science. Here we show that the scientific community moves away from the work of past collaborators of misconducting scientists, even though they were not involved in any misconduct. Applying a difference-in-differences framework, we analyse the citations received by the past collaborators of 36 scientists who were found guilty of scientific misconduct by the Office of Research Integrity and compare them to a random control group of scientists not associated with fraudulent peers. We estimate that publication of the misconduct case by the ORI results in an 8 to 13% drop in citations. Scientists who worked with fraudulent scientists in the past thus undeservedly lose trust of their peers. As scientific team sizes are ever increasing, this is an important source of collateral damage which suggests that the costs of scientific misconduct are being underestimated.

Jelcodes:O33,-

# Guilt by Association

**Early draft – please do not quote without permission**

May 2016


### Abstract

Scientific misconduct harms scientific progress. One of the channels through which this happens is distrust: afraid to build on unreliable scientific results, scientists tend to avoid the work of authors which they associate with unethical research behaviour. While prudent, this strategy risks the misjudged ignoring of good science. Here we show that the scientific community moves away from the work of past collaborators of misconducting scientists, even though they were not involved in any misconduct. Applying a difference-in-differences framework, we analyse the citations received by the past collaborators of 36 scientists who were found guilty of scientific misconduct by the Office of Research Integrity and compare them to a random control group of scientists not associated with fraudulent peers. We estimate that publication of the misconduct case by the ORI results in an 8 to 13% drop in citations. Scientists who worked with fraudulent scientists in the past thus undeservedly lose trust of their peers. As scientific team sizes are ever increasing, this is an important source of collateral damage which suggests that the costs of scientific misconduct are being underestimated.

## 1. Introduction

Scientific misconduct is bad for both society and scientific progress.[i] One example is the case of Andrew Wakefield and co-authors and their publication about a possible link between the measles, mumps and rubella (MMR) vaccine and autism(Wakefield et al., 1998). Even though the publication has been retracted on ground of falsification, it sowed doubts among parents whether they should have their children vaccinated. Vaccination rates been suffering ever since (Godlee, 2011a, 2011b). The work also led to wasted research effort: a Google Scholar search of the terms ‘vaccine’, ‘autism’, and “Wakefield” (while excluding articles authored by Wakefield) returns approximately 4,500 results. The article is cited more than 2,000 times in Google Scholar.

This article is about yet a different cost of misconduct: the damage done to the reputation of scientists who are associated with scientific misconduct. Prior research has shown that scientists who have to retract a paper are cited 7 to 10% less by other scientists (Azoulay, Bonatti, & Krieger, 2015). The retraction thus leads their work to be perceived as of lesser quality by their peers, which has profound implications for future career opportunities. What is more, this also means that the community ignores potentially valid research results.

While it can be expected that scientists lose the trust of their peers when they become known for committing misconduct, it is perhaps surprising to which extent this extends beyond them. Co-authors of fraudulent work who were not aware of any misconduct also experience a

[i] In this paper, we employ the Office of Research Integrity’s (ORI) definition of scientific misconduct: "Research misconduct means fabrication, falsification, or plagiarism in proposing, performing, or reviewing research, or in reporting research results. Fabrication is making up data or results and recording or reporting them; falsification is manipulating research materials, equipment, or processes, or changing or omitting data or results such that the research is not accurately represented in the research record. Plagiarism is the appropriation of another person's ideas, processes, results or words without giving appropriate credit. Research misconduct does not include honest error or differences of opinion" (Office of Research Integrity, 2011).

citation penalty (Mongeon & Larivière, 2014). There is anecdotal evidence suggesting that entire institutes can be harmed in the wake of an important misconduct case. A recent scandal at the Riken Center for Development Biology (CBD) in Kobe, involving only one scientist, led to severe budget cuts and downsizing of the institute Employees of the institute expressed experiencing an immense loss of trust in the wake of the events (Cyranoski, 2015). The impact likewise seems to spread to articles which only relate thematically to retracted work (Azoulay, Furman, Krieger, & Murray, 2014). As there is no reason to believe that these researchers are not trustworthy, this represents a wasteful ignoring of most likely valid and valuable scientific knowledge.

In this article, we offer yet another direction: researchers who worked in the past with misconducting scientists experience a citation drop when the misconduct – in which they were not involved – is revealed. To show this we investigate the citation rate of scientists who worked with 36 scientists who were found guilty of scientific misconduct in the five years before this misconduct was publicized (but who were not involved in the fraudulent work). We draw these from Findings of Research Misconduct issued between 1993 and 2008 by the U.S. Office of Research Integrity (ORI). The ORI publishes findings of misconduct for cases funded by the National Institutes of Health and the Public Health Service. These documents list details of the case, including the name and affiliation of the scientist, the nature of the misconduct, any administrative sanctions, and any publications which have to be corrected or retracted because of the misconduct.

We specifically analyse citation patterns of a group of 856 scientists who worked with the 36 misconducting authors in the five years before the misconduct was publicized, comparing them with a matched control sample of 1,149 scientists who were funded by thematically and structurally similar NIH grants as the scientists who collaborated with misconducting researchers.

Our results, robust to a number of specifications, show that the associated researchers' citations drop 8 to 13% more after the Findings of Misconduct were published than the citations of the control scientist. It is noteworthy that the citation penalty for innocent collaborators reaches about the same level as the penalty for authors of retracted publications (Azoulay et al., 2015). Our findings thus show that the scientific cost of scientific misconduct – as measured in loss of trust, which slows scientific progress and harms careers – is much broader than previously estimated.

The remainder of the article is organized as follows. In the next section we make some remarks on the role played by trust in scientific research. Second three presents the data, followed by methodology and results. Section six concludes.

## 2. Scientific Misconduct and Trust

Apart from monetary rewards and intellectual challenges, peer recognition – or fame, or reputation - is the key reward of the scientific profession (Stephan & Levin, 1992). Scientists build their careers on reputation. Reputation triggers prestigious jobs offers, contacts to star scientists and affluent industry partners, as well as other lucrative opportunities inside and outside the academic sector.

Scientists gain reputation by being the first to publish new discoveries in scientific journals (Merton, 1973). Being second to make an invention is not rewarded at all, leading to a winner-takes-all nature of the competition among scientists. This feature of science is societally beneficial, as it forms an incentive for scientists to share their discoveries with the world as soon as possible, thus pushing science forwards, allowing follow-up research in a speedy manner, and avoiding the duplication of research streams (Stephan, 2012).

The downside of the winner-takes-all system of science is that it creates incentives to cheat. Classic game theory predicts that it can be a rational strategy to cheat in this setting, at least when the chances of being caught are relatively low (Nalebuff & Stiglitz, 1983). Kiri, Lacetera & Zirulia (2014) as well as Lacetera & Zirulia (2009) have applied game theoretic concepts to scientific misconduct. Key results of this research include a certain positive equilibrium level of misconduct in all settings, and the observation that more stringent verification procedures do not necessarily lead to less misconduct.

Reputation is also crucial for the transmission of scientific knowledge. As it is not possible to personally verify all research results, scientists need to rely on their colleagues to be reliable and honest. When a scientist is caught behaving unethically, this trust is violated and others will naturally be more suspect of the misbehaving scientist’s work, or avoid it altogether.

Fraud detection – as well as the detection of honest mistakes - in science is in principle based on peer review: scientific results proposed for publication in scientific journals are screened by peer reviewers and an editorial office who evaluate the discovery and the quality of the research. After publication, journal readers have, in principle, the opportunity to replicate published findings. Replication is however not always feasible, as it can be costly and might require specific equipment and data. Moreover, the rewards for replicating previous results are low and the replicating scientists lose time in the race of being the first to make a new discovery (Dewald, Thursby, & Anderson, 1986). As a consequence, replication occurs in practice rather rarely (Hamermesh, 2007), so that it is not unimaginable that doctored results slip by undetected (Lacetera & Zirulia, 2009).

# 3. Data

## 3.1. Data sources and treated authors

The analysis is based on a database created from the Findings of Research Misconduct as published by the Office of Research Integrity (ORI). The Findings of Research Misconduct are concerned with misconduct cases of National Institute of Health (NIH) grant recipients and grant recipients at the Public Health Service (PHS). We identified 177 such cases between 1993 and 2008.[ii] Among them we selected those cases in which at least one scientific publication was affected as the most severe and most visible cases to the scientific community. Other cases include for instance faking of credentials, plagiarism of grant proposals or scientific misconduct that could be stopped before the results were published in a scientific journal. This resulted in a list of 36 cases for which at least one scientific publication was affected (through correction or retraction) as listed in the proceedings.

We used the Scopus publication database to retrieve all of the misconducting author's publications. Through their publications we identified their collaborators in a five-year period prior to the publication of the case in the Findings of Research Misconduct.[iii] Authors who also collaborated on work mentioned in the Findings of Research Misconduct were excluded from the set of collaborators even if they were cleared from any suspicion. In the analysis we also removed all articles co-authored with the misconducting scientist, as citations to these articles

[ii] We only consider cases up to 2008 to allow for the observation of the associated researchers for five years, and to have a minimum of three years in which every publication can accumulate citations.
[iii] Throughout the analysis, we take articles, conference papers, notes, reviews, and short surveys into account. Letters, books, and other document types are not taken into consideration.

could be affected by negative attention directed to the misconducting author (instead of the innocent co-author). Furthermore, we removed authors who were co-authors of more than one misconducting scientist, as these are subject to multiple treatments, as well as authors with no publications before or after the treatment, as observation before and after treatment is required in the differences-in-differences framework.

After this cleaning we arrived at 929 unique co-authors whose publication and citation record we want to analyse before and after the publication of the Findings of Research Misconduct. However, our final sample is somewhat smaller, consisting of 856 treatment co-authors (cf. the method section), as we wish to observe publication outputs at least four years before the misconduct was published so that we can apply a pre-sample correction for unobserved ability. We compare the citation rates of these authors to that of a control group of 1,149 co-authors, who we select based on the procedure described below.

## 3.2. Construction of control group

Our approach to constructing a control group hinges on identifying comparable co-authors to the treated co-authors by finding collaborators of control authors funded by a similar grant as the treatment authors (who have been found guilty of scientific misconduct). Selection based on grant receipt is essential since we observe a positive selection when focusing on scientists that receive grants. To do this, we used the NIH grant database made available by Pierre Azoulay.[iv] This database covers NIH grants since 1971, and lists grant numbers, general information about the grant, and any publications (indexed by Pubmed identification numbers) which list the grant as source of funding.

[iv] http://pazoulay.scripts.mit.edu/Data.html

The selection process was conducted as follows (see Appendix 1 for a more detailed description of the process). We selected a NIH grant which was listed as funding on the corrected or retracted publication mentioned in the Findings of Research Misconduct ('treatment grant'). We then randomly choose a control grant (using the NIH grant database), which has similar characteristics (in terms of medical research area, type of grant, grant year, and grant duration) as the treatment grant ('control grant'). We match on research area by selecting grants issued by the same NIH institute. The National Institutes of Health consist of 27 institutes and centers, which are divided thematically. Thus, grants issued by a center should be in comparable fields. Examples of institutes are the National Cancer Institute (NCI), the National Human Genome Research Institute (NHGRI), and the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK).

We then selected a random author of a random publication supported by the control grant to serve as the control author. This author is the counterpart to the author who engaged in misconduct. We then used Scopus to identify who this author's co-authors were in the five years before the Findings of Research Misconduct associated with the corrected or retracted publication were published. This group of scientists forms the control co-authors for this case. We repeat this for all 36 relevant cases, resulting in a control group of 1,149 co-authors.

## 3.3.Method

We will estimate an equation of the form:

$$\mathrm{Citations_{it}} = \mathrm{f}(\psi_1 \mathrm{T_t} + \psi_2 \mathrm{Pt_i} + \psi_3 \mathrm{T_i} * \mathrm{Pt_t} + \beta \Gamma_\mathrm{i} + \varphi_\mathrm{t} + \xi_i + \ln(Publications)) + \varepsilon_\mathrm{it}$$

Where $\mathrm{Citations_{it}}$ represents the number of citations received by author i in year t. We offset citations by publication output by including the natural logarithm of the number of publications issued by the author in the year, with coefficient constrained to one, to account for differences in publication outputs. Without this correction our results could be biased by heterogeneous publication outputs.

$T_t$ and $Pt_i$ represent the core of the model. The first is an indicator variable that takes value one if the author is in the treatment group, and zero otherwise. The second takes value zero in pre-treatment years, and one after treatment. $\psi_1$ and $\psi_2$ thus capture any systematic differences in citations received between the treatment and control group, and between any shared differences pre- and post-treatment. The main result of the model is provided by $\psi_3$, which captures the average difference in change in citations between control and treatment observations after the misconduct was discovered. If co-authors in the treatment group experience a drop in citation rates after being associated with scientific misconduct, while co-authors in the control group do not, $\psi_3$ takes a negative and significant coefficient.

$\Gamma_i$ represents a vector of individual-specific factors that affect citations. One such factor is career age, as measured by the time since the authors' first publication. This accounts for the fact that authors have changing levels of commitment to publishing as their career progresses (e.g. Stephan and Levin, 1992). We include career age in linear and squared terms.

Another factor is talent, or ability. As this is usually impossible to observe directly by the econometrician, we control for inherent differences in citation rates either through unobserved (fixed) effects models or by applying a pre-sample average estimator (Blundell, Griffith, & Van Reenen, 1999).[v] Since we need to observe scientists at least four years before the Findings of Research Misconduct are published, the application of the pre-sample mean estimator results in

[v] The pre-sample mean estimator suggested by Blundell et al. (1999) accounts for unaccounted heterogeneity due to unobservables by including an additional parameter in the model which contains the pre-sample average outcome.

dropping some co-authors who we do not observe for a sufficient period of time to calculate these measures. The final dataset consists out of 832 treatment co-authors and 1,157 control co-authors.

Finally, $\varphi_t$ captures common time trends through a set of year dummies, and $\xi_i$ captures that misconduct can have a quite heterogeneous impact through a set of case dummies. As publication and citation counts can be considered events and tend to follow a count distribution, we estimate the model as a Poisson model.

# 4. Results

## 4.1. Descriptive statistics

Table 1 provides summary statistics and correlations. As our models condition on researchers having at least one publication per year, we present summary statistics under the same conditioning. The average co-author in the sample publishes 5.40 papers per year, which were cited 347.24 times by other papers. The average co-author enters the sample relatively late in their career, 17.63 years after their first publication. Table 2 compares means between the treatment and control group for the key variables. Treated co-authors publish slightly more as control authors (treated: 5.73, control 5.14 difference significant at $p<0.01$), but control authors are cited more often (treatment: 331 cites per year, control: 361, difference significant at $p<0.01$). In terms of control variables, treated authors are approximately half a year younger than control authors (treatment: 17.31, control: 17.89, significant at $p<0.01$) and tend to have lower measures of pre-sample citations per publication (treatment: 40.92, control:53.25, significant at $p<0.01$).

------------------------
Tables 1 and 2 here
----------------------

Figure 1 shows the evolution of the citations received by treatment authors proportional to those received by control authors. A value of one indicates that the two groups are on average cited equally often, a value lower than one means that treated authors are cited less often than control authors. 3 to 2 years before the Findings of Research Misconduct are published, both groups are cited at equal rates. In the year before the misconduct is revealed, citations to treated authors drop to 93% of those to control authors. In the year in which the misconduct is revealed, treated authors receive 84% of the citations control authors receive. In the years after the Findings of Research Misconduct were published, the citation rate of threated authors increases again but stays below that of control authors. Articles published by treated authors three years after the reveal receive 89% of the citations that control authors receive for articles at the same time.

----------------
Figure 1 here
----------------

## 4.2. Estimation results

Table 3 presents the main results. In all specifications we find that the number of citations received per publication by treatment authors drops compared to co-authors in the control group in the years after the Findings of Misconduct were published. The point estimates of all specifications are significant at at least $p < 0.05$ and show an average drop between 8% on the low end and 13% on the high end. Considering that the co-authors in the treatment group are only involved in the misconduct case by virtue of having a past collaboration with the person found guilty, this is a highly significant effect. Especially, if compared to past results that show that the citation penalty for authors on retracted publications was of a comparable magnitude (Azoulay et al. report 10%). One possible explanation for this is that the publications under analysis here stem from NIH misconduct cases, while Azoulay et al. (2015) uses a sample of all retractions listed in Pubmed, which also cover honest mistakes. Azoulay et al. report a much stronger citation penalty (17.6%) for authors when only considering retractions due to misconduct.

----------------
Table 3 here
----------------

Column one shows the model without taking into account any observed or unobserved scientist-specific controls, only offset citation counts by publication counts to control for the volume of scientific output. Additionally, we control for case and common year effects through a set of case and year dummies. The former is justified by the idea that cases of misconduct bring about highly heterogeneous amounts of (media) attention in the scientific community and severity. Therefore, the effect on the careers of former collaborators is also likely to be partly case-specific. The results show that treated co-authors are not cited significantly differently from control co-authors, but all authors tend to be cited more often over time which is captured by the dummy indicating the after-treatment period. The treatment effect of being a treated co-author (as inferred from the interaction between being in the treated group and the after-treatment indicator) is negative and highly significant, with a marginal effect of 7.7 %.[vi]

In model 2, we enrich the baseline model by controlling for the career age of the scientist through the time since first publication in linear and quadratic form. We additionally include a proxy for unobserved ability in the form of the pre-sample citation rate. We find a positive quadratic relation between career age and citation counts, and a positive significant relationship

[vi] Calculated as the difference in expected incidence rate ratio at the mean, i.e. 1-exp(-0.08).

between pre-sample citations per publications and current citation outputs. Including these factors slightly increases the treatment effect from 7.7% to 8.6%.

The results reported above depend partly on the assumption that there is a one-to-one relationship between citations and publications. That is, the model assumes that the coefficient of the natural logarithm of publications is one. In model three, we relax this assumption. We find that the estimated treatment effect is the same as before, and the coefficient of the publication offset is highly significant ($p<0.01$) and estimated at 1.02.

Figure 2 showed an initial drop in citations among treatment authors, as compared to control authors, at t-1. We speculate that this can be a result of inaccuracy of the treatment effect: while we are certain that the misconduct has been published with the publication of the Findings of Research Misconduct, the misconduct could also have been made public through other channels before then.

We test whether this concern explains our findings through model 4, where we move the treatment one year forward, to t-1. While the results are qualitatively the same, the treatment effect is estimated slightly higher at 11.3% in this specification. As an alternative robustness test, we provide results when disregarding years t-1 and t0 in column 5. We thus compare citations three and two years before the findings of misconduct to one to three years after. This shows an even higher estimate of the treatment effect at 13.1%.

Lastly, we present the results of fixed effects Poisson estimation with robust standard errors as described in Wooldridge (1999). Using fixed effects instead of pre-sample estimators does not yield different conclusions. As in model 2 and 3, the treatment effect is estimated at 8.6%.

## 5. Discussion

In this paper, we investigate how scientific misconduct negatively impacts scientists that collaborated in the past with the fraudulent scientist. Specifically, we investigate the citations received by the past collaborators of 36 scientists who were found guilty of scientific misconduct by the Office of Research Integrity. Compared to a carefully constructed control sample, and conditional on past citation performance, career stage, and publication output, these scientists are cited 8 to 13% less after the misconduct by the past collaborator was published.

These findings have worrying implications: scientific misconduct leads to a non-negligible disregard of the work of innocent bystanders who are only associated to the misconduct through prior collaboration. Scientific misconduct thus does not only directly waste money and effort; it also slows science indirectly by tainting the credibility of innocent scientists who are only marginally related to the events. This is especially relevant given the large teams in modern medical science: Every fraudulent scientist collaborated with an average of 30 scientists in the five years before the Findings of Research Misconduct were published.

It is thus clear that scientific misconduct does more damage to scientific progress than the dissemination of false science. The discovery of false science has the unintended side effects of tainting the credibility of misconducting author’s prior work (which might or might not be false) (Azoulay et al., 2015; Lu, Jin, Uzzi, & Jones, 2013), their co-authors (Mongeon & Larivière, 2014), the institutes in which they work (Cyranoski, 2015), thematically related articles (Azoulay et al., 2014), the entire field (Azoulay et al., 2014), and, as shown in this article, the people they worked with in the past.

Once false science has made it into scientific literature, it is hard to remove. Retractions do not always help. Using matched control samples, Azoulay et al. (2014) and Furman, Jensen, and Murray (2012) report drops in citations to retracted articles of respectively 69 and 60% - in other words, some keep citing false science as valid. Other authors, using more descriptive results, report that retracted papers continue to be mostly cited as valid, with only a small share of citations making note of the retraction (Budd, Sievert, Schultz, & Scoville, 1999; Neale, Dailey, & Abrams, 2010; Neale, Northrup, Dailey, Marks, & Abrams, 2007; Redman, Yarandi, & Merz, 2008). On top of that, the financial costs of misconduct investigations are non-negligible: one case study estimated the total direct costs of a misconduct investigation at USD 525,000 (Michalek, Hutson, Wicher, & Trump, 2010).

Taken together, the disastrous effects of scientific misconduct are undeniable. However, it is not clear whether policy makers can take any clear action to reduce the damage done to innocent bystanding scientists when retractions need to happen. But before any action can be taken, the fact that collaborators suffer consequences of scientific collaboration should be acknowledged by the scientific community and the public. Some rare exceptions aside,(Mongeon & Larivière, 2014) collaborators have been neglected in discussions about the personal consequences of scientific misconduct.

While there are rehabilitation programs for fraudulent scientists, such as the RePAIR program at the Saint Louis University in Missouri,(Cressey, 2013) prior collaborators do not have access to such initiatives since they cannot rehabilitate themselves of an association. What to do about this issue thus remains an open question.

## References


Azoulay, P., Bonatti, A., & Krieger, J. L. (2015). The career effects of scandal: evidence from scientific retractions. NBER Working Paper No. 21146.

Azoulay, P., Furman, J. L., Krieger, J. L., & Murray, F. E. (2014). Retractions. Review of Economics and Statistics. doi:10.1162/REST_a_00469

Blundell, R., Griffith, R., & Van Reenen, J. (1999). Market share, market value and innovation in a panel of British manufacturing firms, 66(3), 529–554. doi:10.1111/1467-937X.00097

Budd, J. M., Sievert, M., Schultz, T. R., & Scoville, C. (1999). Effects of article retraction on citation and practice in medicine. Bulletin of the Medical Library Association, 87(4), 437–443. Retrieved from http://www.ncbi.nlm.nih.gov/pmc/articles/PMC226618/

Cressey, D. (2013). “Rehab” helps errant researchers return to the lab. Nature, 493(7431), 147. doi:10.1038/493147a

Cyranoski, D. (2015). Collateral Damage: how a case of misconduct brought a leading Japanese biology institute to its knees. Nature, 520, 600–603.

Dewald, W. G., Thursby, J. G., & Anderson, R. G. (1986). Replication in Empirical Economics: The Journal of Money, Credit and Banking Project. The American Economic Review, 76(4), 587–603. doi:10.2307/1806061

Furman, J. L., Jensen, K., & Murray, F. (2012). Governing knowledge in the scientific community: Exploring the role of retractions in biomedicine. Research Policy, 41(2), 276–290. doi:10.1016/j.respol.2011.11.001

Godlee, F. (2011a). The fraud behind the MMR scare. BMJ, 342, d22. doi:10.1136/bmj.c5347

Godlee, F. (2011b). Wakefield’s article linking MMR vaccine and autism was fraudulent. BMJ, 342, c7452. doi:http://dx.doi.org/10.1136/bmj.c7452

Hamermesh, D. S. (2007). Replication in Economics. The Canadian Journal of Economics, 40(3), 715–733.

Hamilton, K. (2003). Subfield and level classification of journals. CHI Report 2012-R.

Kiri, B., Lacetera, N., & Zirulia, L. (2014). Above a Swamp: A Theory of High-Quality Scientific Production. Working Paper. Retrieved from http://www2.dse.unibo.it/dsa/seminari/953/Paper.Kiri_Lacetera_Zirulia.pdf

Lacetera, N., & Zirulia, L. (2009). The Economics of Scientific Misconduct. Journal of Law, Economics, and Organization, 27(3), 568–603. doi:10.1093/jleo/ewp031

Lu, S. F., Jin, G. Z., Uzzi, B., & Jones, B. (2013). The retraction penalty: evidence from the Web of Science. Scientific Reports, 3, 3146. doi:10.1038/srep03146

Merton, R. K. (1973). The sociology of science. Chicago, IL: The University of Chicago Free Press.

Michalek, A. M., Hutson, A. D., Wicher, C. P., & Trump, D. L. (2010). The costs and underappreciated consequences of research misconduct: a case study. PLoS Medicine, 7(8), e1000318. doi:10.1371/journal.pmed.1000318

Mongeon, P., & Larivière, V. (2014). Costly Collaborations : The Impact of Scientific Fraud on Co- authors ' Careers. arXiv Preprint.

Nalebuff, B. J., & Stiglitz, J. E. (1983). Prizes and Incentives: Towards a General Theory of Compensation and Competition. The Bell Journal of Economics, 14(1), 21–43. doi:10.2307/3003535

Neale, A. V., Dailey, R. K., & Abrams, J. (2010). Analysis of citations to biomedical articles affected by scientific misconduct. Science and Engineering Ethics, 16(2), 251–61. doi:10.1007/s11948-009-9151-4

Neale, A. V., Northrup, J., Dailey, R. K., Marks, E., & Abrams, J. (2007). Correction and use of biomedical literature affected by scientific misconduct. Science and Engineering Ethics, 13(1), 5–24.

Office of Research Integrity. (2011). Definition of Research Misconduct. Retrieved January 21, 2016, from http://ori.hhs.gov/definition-misconduct

Redman, B. K., Yarandi, H. N., & Merz, J. F. (2008). Empirical developments in retraction. Journal of Medical Ethics, 34(11), 807–9. doi:10.1136/jme.2007.023069

Stephan, P. (2012). How Economics Shapes Science. Cambridge, MA: Harvard University Press.

Stephan, P. E., & Levin, S. G. (1992). Striking the mother lode in science: the importance of age, place, and time. Oxford University Press.

Wakefield, A. J., Murch, S. H., Anthony, A., Linnell, J., Casson, D. M., Malik, M., … Walker-Smith, J. A. (1998). Ileal-lymphoid-nodular hyperplasia, non-specific colitis, and pervasive developmental disorder in children [retracted]. The Lancet, 351, 637–641.

Wooldridge, J. M. (1999). Distribution-free estimation of some nonlinear panel data models. Journal of Econometrics, 90(1), 77–97. doi:10.1016/S0304-4076(98)00033-5

# Figures

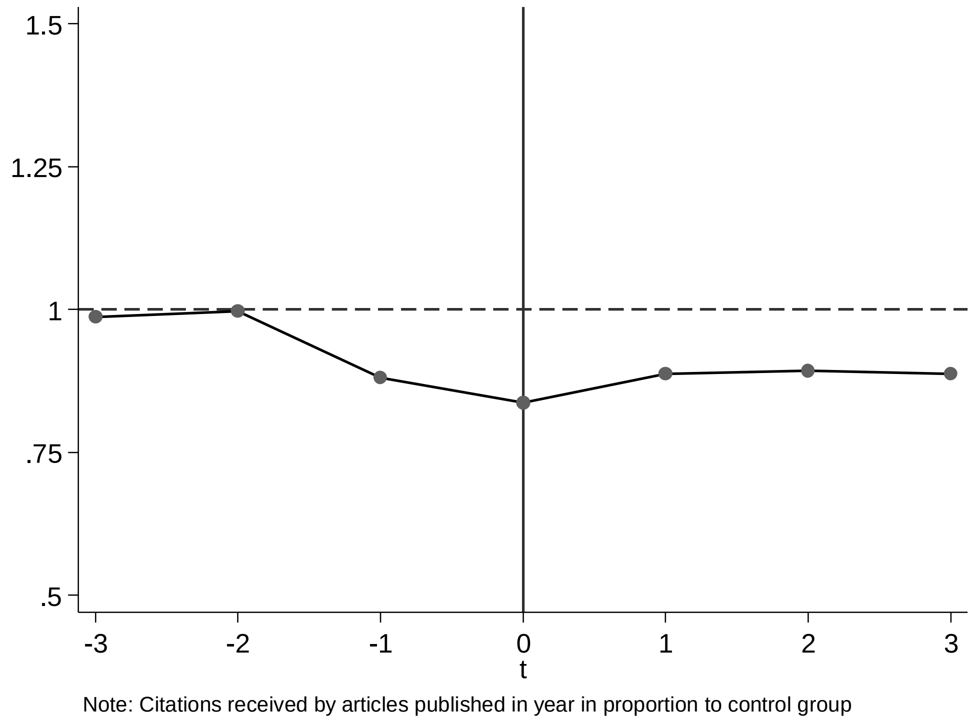


**Figure 1: Annual citations received before and after publishing of Findings of Misconduct**

**Notes:** This graph shows annual citations received by articles authored by scientists who are associated with an author who committed scientific misconduct, proportional to the control group. t: year relative to publication of Findings of Misconduct (t0). 3 to 2 years before the Findings are published, both groups are cited at equal rates. In the year before the misconduct is revealed, citations to treated authors drop to 93% of those to control authors. In the year in which the misconduct is revealed, treated authors receive 84% of the citations control authors receive. In the years after the Findings of Misconduct were published, the citation rate of threated authors increases again but stays below that of control authors. Articles published by treated authors three years after the reveal receive 89% of the citations that control authors receive for articles at the same time.

# Tables

**Table 1 : Summary Statistics (2005 authors, 10758 observations)**

| | Summary Statistics | | | | | Correlation Matrix | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | Mean | St. Dev | Min | Median | Max | 1 | 2 | 3 | 4 | 5 | 6 |
| 1. Publications | 5.40 | 6.38 | 1 | 3 | 121 | 1.00 | | | | | |
| 2. Citations | 347.24 | 527.20 | 0 | 146 | 4318 | 0.66 | 1.00 | | | | |
| 3. Years since first publication | 17.63 | 9.38 | 1 | 16 | 56 | 0.20 | 0.15 | 1.00 | | | |
| 4. Pre-sample citations per publication average | 47.76 | 57.35 | 0 | 32 | 1046.5 | 0.03 | 0.18 | -0.14 | 1.00 | | |
| 5. Treatment dummy | 0.44 | 0.50 | 0 | 0 | 1 | 0.05 | -0.03 | -0.03 | -0.11 | 1.00 | |
| 6. After treatment indicator | 0.55 | 0.50 | 0 | 1 | 1 | 0.02 | 0.01 | 0.17 | -0.02 | 0.02 | 1.00 |

Note: Sample restricted to author-years with at least one publication

**Table 2: Mean Comparison Treated and Control samples**

| | **Treated** | | **Control** | | **Diff** |
|---|---|---|---|---|---|
| | Mean | St. Dev | Mean | St. Dev | |
| Publications | 5.73 | 7.41 | 5.14 | 5.40 | *** |
| Citations | 330.58 | 501.13 | 360.59 | 546.88 | *** |
| Years since first publication | 17.31 | 9.36 | 17.89 | 9.39 | *** |
| Pre-sample citations per publication average | 40.92 | 52.37 | 53.25 | 60.49 | *** |
| After treatment indicator | 0.56 | 0.50 | 0.54 | 0.50 | ** |
| Authors | 856 | | 1149 | | |
| Observations | 4787 | | 5971 | | |

**Notes:** Sample restricted to author-years with at least one publication.
Diff: two-sample t-test.
Stars indicate significance level of difference:
*: $p<0.10$
**: $p < 0.05$
***: $p< 0.01$

**Table 3: Poisson regression estimates of citation rates by authors associated with misconducting authors and control authors**

| Dependent: Citations | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| Model | Baseline | Incl. Controls | Relaxed offset assumption | Treatment shifted forward one period | Disregard t-1 and t0 | QML Fixed Effects |
| Treatment Group | -0.09* | -0.03 | -0.04 | 0.00 | 0.00 | |
| | (0.05) | (0.05) | (0.05) | (0.05) | (0.05) | |
| After-treatment period | 0.09** | 0.10*** | 0.10*** | 0.04 | 0.04 | 0.08** |
| | (0.04) | (0.04) | (0.04) | (0.03) | (0.08) | (0.04) |
| Treatment* After-treatment period | -0.08** | -0.09** | -0.09** | -0.12*** | -0.14*** | -0.09*** |
| | (0.04) | (0.04) | (0.04) | (0.04) | (0.05) | (0.04) |
| Years since first publication | | 0.02*** | 0.02*** | 0.02*** | 0.02*** | |
| | | (0.01) | (0.01) | (0.01) | (0.01) | |
| Years since first publication$^2$/100 | | -0.03** | -0.03** | -0.03** | -0.04*** | |
| | | (0.01) | (0.01) | (0.01) | (0.01) | |
| Pre-sample citations per publication/10 | | 0.03*** | 0.03*** | 0.03*** | 0.03*** | |
| | | (0.00) | (0.00) | (0.00) | (0.00) | |
| Ln(number of publications) | | | 1.02*** | | | 0.91*** |
| | | | (0.02) | | | (0.02) |
| Intercept | 3.95*** | 3.71*** | 3.71*** | 3.70*** | 3.75*** | |
| | (0.37) | (0.37) | (0.37) | (0.37) | (0.39) | |
| Year effects | YES | YES | YES | YES | YES | YES |
| Case effects | YES | YES | YES | YES | YES | NO |
| Individual Fixed Effect | NO | NO | NO | NO | NO | YES |
| Number of observations | 10758 | 10758 | 10758 | 10758 | 7528 | 10681 |

**Notes:** Poisson estimation of yearly citation counts. Citation count estimates offset by publication counts by including ln(number of publications) with coefficient fixed at one in the model. This assumption is relaxed in model 3 and 6, where ln(number of publications) is included as regressor. Citation regressions only include author-years where at least one publication occurred. Cluster-robust standard errors in parentheses. Years since first publication² and pre-sample citations per publication scaled down for readability. Results indicate robust lower publications and citations for treatment coauthor after treatment, compared to difference in publications and citations before and after treatment for control coauthors.

Stars indicate significance level of coefficient: *: p<0.10, **: p < 0.05, ***: p< 0.01

## Appendix 1: Data Gathering Protocol

The construction of the control group proceeded as follows. See also Figure A.1 for a graphical summary of the process. The process requires four main steps:

1. For each of the scientists mentioned as guilty of scientific misconduct in the Findings of Research Misconduct: Select a NIH grant which funded one of the corrected or retracted papers :
    - This step was necessary because not all cases reported in the Findings of Research Misconduct refer to the NIH grant number.
    - If there was more than one publication associated with the specific misconduct case, the earliest published paper was selected, in order to identify the associated grant.
    - If the earliest published paper did not list grants in its acknowledgement field, the second earliest paper was taken (and so on until a paper which referred to a NIH grant was found).
    - If the paper referred to more than one NIH grant, the earliest grant was selected.
2. Match the so retrieved 36 'treatment' NIH grants to 36 NIH 'control' grants with similar characteristics as the treatment grants;
    - Matching criteria were the granting institute, grant type, grant duration, and grant year.[7]
    - This step ensures that the group of scientists from whom we generated the control group of co-authors is working in the same broadly defined field as the

[7] Time of granting was not used as a hard matching condition. Instead, the closest grant was chosen in case of multiple options.

misconducting scientist from which we generated the treatment group of co-authors.

- For treatment and control groups, the individuals under analysis are the scientists who collaborated with respectively the control or treated scientist in the five years before the Findings of Misconduct were published.

3. Select 36 'control publications' which were funded by the 36 control grants;
    - In case more than one paper was supported by the control grant, the publication published as closely as possible to the treatment publication was chosen.[8]
4. Select 36 control authors, which form the counterparts of the misconducting authors.
    - A random author from this publication was chosen as control scientist.

The 36 control authors are then used to find the control co-authors, with whom the control authors collaborated in the five years before the Findings of Research Misconduct were published.

[8] For 60% of the cases, the treatment publication occurred in the same year as the control publication. For all but three cases, the control publication falls in a three-year window of the treatment publication. The maximum time difference between control and treatment publication was six years.

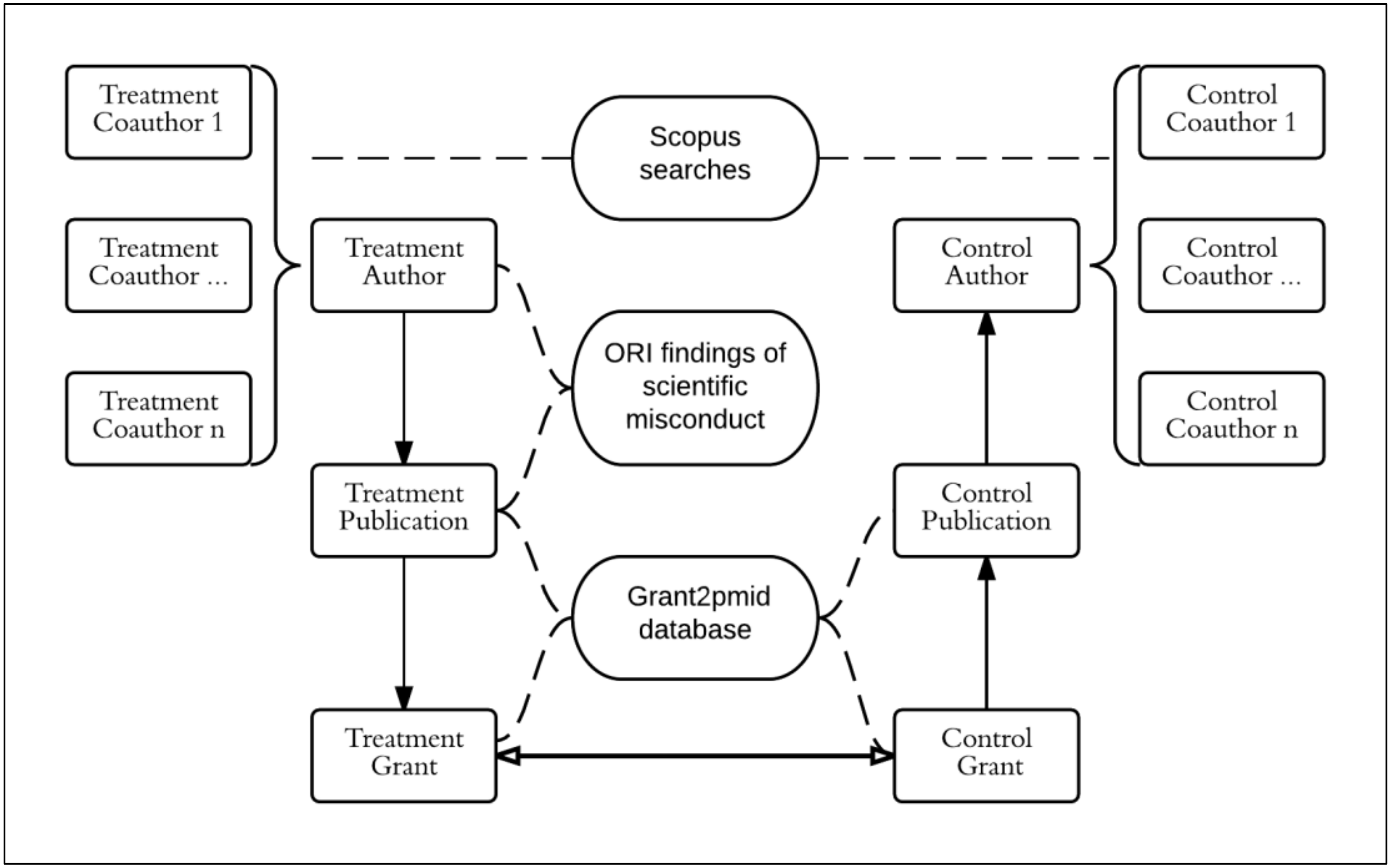


**Figure A.1: Visual representation of data collection protocol**