Preliminary and Incomplete – Do Not Cite

The Reverse Matthew Effect: Collaboration and Consequence in Scientific Misconduct

Ginger Jin

University of Maryland and NBER

Benjamin Jones

Northwestern University, Kellogg School of Management

Susan Feng Lu

University of Rochester, Simon Graduate School of Business

Brian Uzzi

Northwestern University, Kellogg School of Management

May 14, 2013

Abstract

How is credit shared in teams? This question extends across many production settings but is of long-standing interest in science and innovation, where the "Matthew Effect" suggests that eminent team members garner the credit for great works at the expense of less eminent team members. In this paper, we study this question in reverse, asking how credit is shared in cases of scientific misconduct. We examine article retractions in the Web of Science and study their effects on citations to the authors' prior publications. We find that the Matthew Effect works in reverse – namely, scientific misconduct imposes no observable citation penalty to the prior publications of eminent coauthors. By contrast, less eminent coauthors face substantial declines in citations to their prior work. These findings are of increasing relevance as both scientific collaboration and retraction rates are sharply on the rise. More broadly, these finding suggest that a good reputation can have protective properties, but at the expense of those with unformed reputations.

I. Introduction

Team production is pervasive in modern economies, intimately related to the division of labor and benefits therein.[1] Yet team production raises challenges, from free riding problems and communication breakdowns to the sharing of credit that governs the evolution of careers. In situations where the output of the individual is not directly observed, reputation can become a cornerstone in providing incentives while also becoming a central tool for how the community assigns credit across a team.

In a classic study, Robert K. Merton suggested the "Matthew Effect" as a fundamental issue in an important team production context: science (Merton 1968). Merton argued that more eminent coauthors tend to receive disproportionate credit for team-authored work.[2] In effect, teamwork leads to a "rich get richer" phenomenon, where, faced with a great paper, the scientific community assumes that the more eminent coauthor was the key producer while less well-known coauthor(s) were subordinate contributors who deserve little credit. Arguably, such a credit assignment mechanism, if it operates, could have large effects on reputations, on the dynamics of individual careers, on incentives to work in teams, and efficient matching of team members.

This paper considers a natural experiment to assess credit sharing in teams. Our question, however, is not how team members share credit for "good" events, but rather how they share credit for catastrophe. Namely, we look at the effect of scientific misconduct in team production settings. We examine whether eminent coauthors attract or repel blame compared to their less eminent coauthors. On the one hand, one

[1] See, e.g., classic observations in Bacon (1620) and Smith (1776) or modern analyses such as Becker and Murphy (1992), Hamilton et al. (2003), Jones (2010), Mas and Moretti (2011).

[2] Merton coined the Matthew Effect after the biblical passage "For unto every one that hath shall be given, and he shall have abundance: but from him that hath not shall be taken even that which he hath" (Matthew 25: 29, King James Version).

might imagine that eminent authors receive disproportionate credit for the output, whether good or bad, as the presumed leader of the research enterprise. On the other hand, one may imagine that eminent authors have such positive reputations that they escape blame for bad events, leaving any blame to accrue to junior coauthors. Thus we may imagine a "Reverse Matthew Effect", which might be phrased as "To whom much has been given, much less will be taken away."

In this paper, we first provide a simple Bayesian updating model to consider the reputational effect of scientific misconduct in joint-authored work. In the model, each publication is a noisy signal of author quality and the scientific community takes each new signal to update its belief on author quality. When retraction occurs, the signal of the retracted paper is adjusted downward and the posterior belief is updated for each author. Our model yields several testable predictions: first, the posterior update is less negative for an author if that author already has an established publication record before the retraction. This finding follows because the community is more certain about the quality of an established researcher and it is more difficult to update on a stronger prior. Second, conditional on an author's own reputation, the retraction can bring her more reputational harm if her coauthor is more established in the profession. Low-status authors are especially vulnerable to this negative spillover. Third, the reputation harm from retraction is greater if the retracted work was deemed important before retraction, and this effect is of a lesser magnitude for established authors.

We test these predictions in the Web of Science database from Thomson Reuters. We focus on 376 retractions that were authored in teams and then study the citation behavior to the prior publications of each author. The content of prior work does not change by definition, but their citations could change after the retraction if the scientific community has changed its view of the authors. To examine the effect of retraction, we match each of these prior publications (the treated papers) with a set of other publications (the control papers) that were published in the same field-year and received similar citations every year before the retraction. This approach allows us to identify the effect of retraction via differences-in-differences estimation.[3]

Using various measures of reputation, we find that the less eminent coauthors receive substantially fewer citations to their prior work upon retraction. By contrast, eminent coauthors appear able to deflect blame for misconduct episodes, showing no citation consequences for their prior work. These findings suggest that the Matthew Effect does work in reverse.

The rest of the paper is organized as follows. Section II reviews the literature. Section III presents the Bayesian model and derives testable predictions. Section IV summarizes the data. Section V elaborates our identification strategy. Section VI presents empirical results and evidence against alternative explanations. A brief conclusion is offered in Section VII.

II. Literature Review

Our work is related to several literatures. In the innovation literature, researchers have documented the rising importance of teamwork in scientific publication, and the advantage of teamwork in citation (Wuchty, Jones and Uzzi 2007), both of which raise a question of why scientists are more willing to participate in teamwork over time. One explanation attributes it to the burden of knowledge: because knowledge accumulates over time, successive generations of innovators become increasingly specialized, and teamwork becomes necessary to tap differential knowledge (Jones 2009). Similar productivity gain from teamwork is observed in a garment plant (Hamilton, Nickerson

[3] Using citations to prior scientific work to assess the effects of information shocks was pioneered as an identification strategy in Furman and Stern (2011).

and Owan 2003), in supermarket cashiers (Mas and Moretti 2011), and in the experimental lab (Falk and Ichino 2006, Fehr and Gächter 2000).[4]

These beneficial features of teamwork run against agency and information problems in team production. One interesting question about teamwork is how high- and low-status team members fare asymmetrically from the teamwork, with Merton's Matthew Effect presenting a seminal point of departure. The Matthew Effect, which may dissuade junior coauthors from collaborating with eminent authors, is tangled with other complicated career incentives where collaboration with an established name may advantage further career development through other channels (Peterson et al. 2011, Merton 1988). For example, coauthorship with an eminent scientist may bestow a signal that junior coauthor is worth working with, and eminent coauthors may open doors for junior collaborators through the power structures of science. In this paper, we contribute to this literature focus on the question of credit sharing, and by looking at credit sharing from reputation-damaging events.

Our work is also related to the vast literature of reputation and learning. Theorists have long argued that reputation can be an effective mechanism to ensure high quality production although product quality is neither observable nor contractible before transaction (Klein and Leffler 1981, Shapiro 1983). In that literature, seller reputation is defined as the posterior belief about a seller's type given the history of transaction outcomes involving that seller (Bar-Isaac and Tadelis 2008). Following this literature, we model author reputation as the posterior belief about an author's research ability given the observed quality of all her publications. When one publication is discovered to be low quality, this new information leads to an update of author reputation.

Empirical researchers attempted to test the reputation theory in several ways. Using eBay feedback score as a proxy for seller reputation, studies find that sellers with

---

[4] These studies often attribute the positive spillover between team members to social reasons such as peer pressure.

more positive feedback enjoy a higher probability of sale and/or higher transaction price (Bajari and Hortacsu 2004). When it is difficult to measure seller reputation directly, evidence suggests that the market does respond to a piece of good or bad news about a seller. For example, Borenstein and Zimmerman (1988) document stock market response to air plane crash, Pope (2009) describes consumer response to the US News ranking of hospitals, Azoulay, Tuarte and Wang (2012) show citation increase after a researcher won the status of Howard Hughes Medical Investigator (HHMI); and Dranove, Ramanarayanan and Watanabe (2012) find that patient composition changes in response to doctor's malpractice litigation. In theory, market response to reputation change should create incentives for sellers to obtain a good reputation and shed a bad reputation. These seller incentives are confirmed in medical service (Johnson 2011), e-commerce (Cabral and Hortacsu 2004), and restaurant hygiene quality (Jin and Leslie 2009).

Our work differs from all these studies because we emphasize that a single signal of teamwork can have different reputation effects on different team members. Differential update of reputation implies different incentives to pursue authentic research and different incentives to engage in teamwork, both of which are important factors for the progress of science.

III. A Bayesian Model of Author Reputation

As a baseline, we first set up a model for solo-authored papers only. Though we do not have enough data to test predictions from the solo model, it helps establish fundamental forces in Bayesian updating. We then extend the model to teamwork and derive testable predictions regarding the effect of retraction on author reputation.

III.1 Basic Setup for Solo-authored Papers

Let q_i denote researcher i's inherent ability. In reality, researcher ability may have multiple dimensions, for example, creativity, honesty, diligence, etc. Since our data have only one dimension of outcome (citation), it is difficult to distinguish multiple dimensions of ability. Given this data limit, we simplify q_i into one dimension and assume it ranges from the most blatant cheating type ($q_i = -\infty$) to the most honest and creative type ($q_i = \infty$). In the middle there might be diligent-but-non-creative people or creative-but-sloppy people. We assume q_i does not change over time and is not directly observed by the public.

The scientific society tries to learn about q_i from published papers. As a starting point, the society believes that the inherent ability of a rookie researcher i conforms to a normal distribution:

$$\tilde{q}_{i0} \sim N(\bar{q}_{i0}, \sigma_{i0}^2).$$

One can interpret $(\bar{q}_{i0}, \sigma_{i0}^2)$ as i's academic credential: for example, a graduate from a top Ph.D. program may initially be believed to have higher ability.

Publication sends out a signal of continuous value. For example, being published in a prestigious journal may send out a high-value signal while being published in an obscure journal corresponds to a low-value signal. Alternatively, a paper that solves an important question may send out a high-value signal, and a paper that provides a trivial robustness check to an existing solution sends out a low-value signal. In the model, we treat signal as a real number that is known to the public upon publication. Our basic setup assumes every paper is solo-authored. This assumption will be relaxed in the next subsections.

In particular, the publication outcome of paper j by author i sends out a signal:

$$s_{ij} = q_i + \varepsilon_j \text{ where } \varepsilon_j \sim N(0, \sigma_\varepsilon^2), \text{ iid.}$$

If we interpret signal value as a function of journal prestige, σ_ε^2 captures the overall noise in the mapping from author ability to, say, the particular quality of a given piece of output, or to journal prestige.

After publication, if the paper is found to be false and therefore retracted, signal s_{ij} will be degraded to a large negative number, which we denote as $\underline{s}$. For simplicity, we treat retraction from a prestigious journal to be the same as retraction from an obscure journal, assuming that a false study is equally useless regardless of where it was published first. However, a retraction of a paper with a higher signal (e.g. from publication in a prestigious journal) implies a bigger *change* in signal value.[5]

A new signal s_{ij} helps to update the belief about q_i. Define N_{it} as the total number of papers that i has published up to time t and $I_{it} = \{s_{i1}, s_{i2}, \dots, s_{iN_{it}}\}$ as the publication outcomes of all these papers. To keep notations tractable, we define the precision of the prior and signals as:

$$v_{i0} = \frac{1}{\sigma_{i0}^2}$$

$$v_\varepsilon = \frac{1}{\sigma_\varepsilon^2}.$$

The posterior belief of q_i can be written as:

$$\tilde{q}_{it}|I_t \sim N(\bar{q}_{it}, \sigma_{it}^2) \qquad \text{where}$$

$$\bar{q}_{it} = \frac{v_{i0} \cdot \bar{q}_{i0} + v_\varepsilon \cdot \sum_{j=1}^{N_{it}} s_{ij}}{v_{i0} + v_\varepsilon \cdot N_{it}},$$

$$\sigma_{it}^2 = \frac{1}{v_{it}} = \frac{1}{v_{i0} + v_\varepsilon \cdot N_{it}}.$$

Clearly, the mean of the posterior belief is a weighted average of the prior and the signals, where the weights are the precisions. The precision of the posterior, defined as $v_{it} = \frac{1}{\sigma_{it}^2}$, is the sum of the prior precision and the signal precisions.

Let paper *j* be published at time τ_j. The number of times this paper is cited at time t, denoted as C_{ijt}, depends on the paper's signal s_{ij}, the current belief about researcher i's quality, $\bar{q}_{it}$, and the duration since publication, $t - \tau_j$.

$$C_{ijt} = f\left(s_{ij}, \bar{q}_{it}, t - \tau_j\right).$$

We assume both s_{ij} and $\bar{q}_{it}$ have a positive effect on C_{ijt}. Meanwhile, the dependence on duration since publication, $t - \tau_j$, will follow non-linear dynamics as seen throughout the scientific literature. Typically, citation rates first rise with duration since

[5] To keep the model simple, we treat no publication as no signal. If one believes that no publication in a given period sends a negative signal about a researcher's quality, it is possible to incorporate no publication as a draw of low value in s_{ij}, just as if the researcher has written a hypothetical paper but that paper is published in a very obscure journal.

publication and then fall, with different temporal patterns in different fields (Stringer et al. 2010).

Now suppose a paper k written by author i gets retracted. The value of its signal will degrade from s_{ik} to $\underline{s}$. Thus, the mean of the posterior belief of q_i will change by:

$$\Delta\bar{q}_{it}|s_{ik} \to \underline{s} = \frac{v_\varepsilon \cdot (\underline{s} - s_{ik})}{v_{i0} + v_\varepsilon \cdot N_{it}} < 0.$$

Obviously, upon retraction, citation on the retracted paper itself will drop because two determinants of new citations, s_{ik} and $\bar{q}_{it}$, decline. For clean prior work that was published before k was retracted, its citation will decline as well because the society has adjusted $\bar{q}_{it}$ downwards. In particular, the citation change on this prior work is:

$$\Delta C_{ijt}|s_{ik} \to \underline{s} = f\left(s_{ij}, \bar{q}_{it} + \Delta\bar{q}_{it}, t - \tau_j\right) - f\left(s_{ij}, \bar{q}_{it}, t - \tau_j\right) \leq 0, \forall j \neq k.$$

Note that the negative spillover on a clean prior work is driven by the updated belief about author's quality. Thus citation changes on prior works give us an indirect measure of reputation change in the event of retraction.

To summarize, retraction of one paper will have a negative spillover impact on the citation of the author's prior work. The degree of the negative impact is greater if (1) the retracted paper was published in a more prestigious journal (i.e. $\frac{d|\Delta C_{ijt}|s_{ik} \to \underline{s}|}{ds_{ik}} > 0$), (2) the author has fewer publications at the time of retraction (i.e. $\frac{d|\Delta C_{ijt}|s_{ik} \to \underline{s}|}{dN_{it}} < 0$), (3) the author graduated from a high-variance school (i.e. $\frac{d|\Delta C_{ijt}|s_{ik} \to \underline{s}|}{d\sigma_{i0}} > 0$), or (4) publication quality depends less on luck (i.e. $\frac{d|\Delta C_{ijt}|s_{ik} \to \underline{s}|}{d\sigma_\varepsilon} < 0$). Note that all the above predictions do not vary by the publication time of the retracted paper. Every signal is treated equally and thus the order of publication does not matter.

III.2 Teamwork

Consider two researchers, i_1 and i_2. Before they collaborate on paper k, the ability of each researcher is believed to follow a normal distribution:

$$\tilde{q}_{i_1 t} \sim N\left(\bar{q}_{i_1 t}, \sigma^2_{i_1 t}\right),$$
$$\tilde{q}_{i_2 t} \sim N\left(\bar{q}_{i_2 t}, \sigma^2_{i_2 t}\right).$$

These two distributions represent the two researchers' professional reputation at time t, which depends on the number and quality of publications of each researcher before t. Following the notation in II.1, if each researcher does solo works only before the collaboration, we have:

$$\bar{q}_{i_1 t} = \frac{v_{i_1 0} \cdot \bar{q}_{i_1 0} + v_\varepsilon \cdot \sum_{j=1}^{N_{i_1 t}} s_{i_1 j}}{v_{i_1 0} + v_\varepsilon \cdot N_{i_1 t}},$$

$$\sigma^2_{i_1 t} = \frac{1}{v_{i_1 t}} = \frac{1}{v_{i_1 0} + v_\varepsilon \cdot N_{i_1 t}}.$$

$$\bar{q}_{i_2 t} = \frac{v_{i_2 0} \cdot \bar{q}_{i_2 0} + v_\varepsilon \cdot \sum_{j=1}^{N_{i_2 t}} s_{i_2 j}}{v_{i_2 0} + v_\varepsilon \cdot N_{i_2 t}},$$

$$\sigma^2_{i_2 t} = \frac{1}{v_{i_2 t}} = \frac{1}{v_{i_2 0} + v_\varepsilon \cdot N_{i_2 t}}.$$

The publication of their joint work sends out a signal that depends on the ability of both authors:

$$s_{i_1 + i_2, k} = \alpha\left(q_{i_1} + q_{i_2}\right) + \varepsilon_k.$$

For simplicity, we assume the two authors contribute equally to the joint paper but the quality of the joint paper can be a simple average of the two authors' true quality ($\alpha = \frac{1}{2}$), higher than the average ($\alpha > \frac{1}{2}$), or lower than the average ($\alpha < \frac{1}{2}$). Empirical evidence has shown that joint works are cited significantly more than solo papers, which could be interpreted as $\alpha > \frac{1}{2}$ (Wuchty et al. 2007). The assumption of equal contribution is adopted for easy illustration; all the key predictions remain valid if the two authors contribute differentially to the joint work, as long as their contribution is public knowledge.

Given $s_{i_1+i_2,k}$, the societal belief on q_{i_1} and q_2 will be updated to:

$$\begin{bmatrix}\tilde{q}_{i_1t+1}\\ \tilde{q}_{i_2t+1}\end{bmatrix} \sim N\left(\begin{bmatrix}\bar{q}_{i_1t+1}\\ \bar{q}_{i_2t+1}\end{bmatrix}, \Sigma\right) \text{ where}$$

$$\Sigma^{-1} = \begin{bmatrix} v_{i_1t} + v_\varepsilon \cdot \alpha^2 & v_\varepsilon \cdot \alpha^2 \\ v_\varepsilon \cdot \alpha^2 & v_{i_2t} + v_\varepsilon \cdot \alpha^2 \end{bmatrix}$$

$$\bar{q}_{i_1t+1} = w_1 \cdot \bar{q}_{i_1t} + (1 - w_1) \cdot \left[\frac{s_{i_1+i_2,k}}{\alpha} - \bar{q}_{i_2t}\right]$$

$$\bar{q}_{i_2t+1} = w_2 \cdot \bar{q}_{i_2t} + (1 - w_2) \cdot \left[\frac{s_{i_1+i_2,k}}{\alpha} - \bar{q}_{i_1t}\right]$$

$$w_1 = \frac{v_{i_1t}\left(v_{i_2t} + v_\varepsilon \cdot \alpha^2\right)}{|\Sigma^{-1}|}$$

$$w_2 = \frac{v_{i_2t}\left(v_{i_1t} + v_\varepsilon \cdot \alpha^2\right)}{|\Sigma^{-1}|}.$$

Each posterior mean is a weighted average of own prior and the part of the signal that reflects own quality. From i_1's point of view, the weight on own prior, w_1, increases with the precision of own prior (v_{i_1t}), but decreases with the precision of coauthor's prior (v_{i_2t}), the precision of the signal of the joint paper (v_ε), and the scale economy of collaboration (α).

Thus, i_1's posterior mean update can be written as:

$$\bar{q}_{i_1t+1} - \bar{q}_{i_1t} = (1 - w_1) \cdot \left[\frac{s_{i_1+i_2,k}}{\alpha} - \bar{q}_{i_2t} - \bar{q}_{i_1t}\right] = \frac{v_{i_2t} \cdot v_\varepsilon \cdot \alpha^2}{|\Sigma^{-1}|}\left[\frac{s_{i_1+i_2,k}}{\alpha} - \bar{q}_{i_2t} - \bar{q}_{i_1t}\right].$$

In other words, if the signal of the joint paper turns out to be better (worse) than what one would expect from the prior means of the two researchers, the societal belief will adjust upwards (downwards) for both researchers.

The effect of interest is how the posterior updating process differs upon retraction when the signal of the joint work is $\underline{s}$ instead of $s_{i_1+i_2,k}$. In the real data, there may be a time lag between the initial publication and retraction. This time lag can be ignored in our model because the noise of each paper is assumed to be i.i.d. and the

Bayesian updating at the time of retraction considers all information available regardless of the order of signal arrival. Thus, to keep notation simple, we assume retraction happens immediately after the initial publication and compare that situation to no retraction. It is not difficult to show that:

$$\Delta[\bar{q}_{i_1t+1} - \bar{q}_{i_1t}]|_{s_{i_1+i_2,k}\to\underline{s}} = (1-w_1)\cdot\left[\frac{\underline{s} - s_{i_1+i_2,k}}{\alpha}\right] = \frac{v_{i_2t}\cdot v_\varepsilon\cdot\alpha^2}{|\Sigma^{-1}|}\left[\frac{\underline{s} - s_{i_1+i_2,k}}{\alpha}\right] < 0,$$

Because retraction is a negative shock for each author, each posterior mean will update more downwards with retraction than without retraction.

A key question is how the retraction-triggered adjustment varies by author status. Although author status can be proxied by both prior mean and prior precision, the above expression does not vary by prior means ($\bar{q}_{i_1t}, \bar{q}_{i_2t}$) because the term related to $\bar{q}_{i_1t}$ and $\bar{q}_{i_2t}$ cancels out when we contrast the cases with and without retraction. Hence, we can focus on prior precision as author status. We have:

$$\frac{\partial\Delta[\bar{q}_{i_1t+1} - \bar{q}_{i_1t}]|_{s_{i_1+i_2,k}\to\underline{s}}}{\partial v_{i_1t}} = \frac{\partial(1-w_1)}{\partial v_{i_1t}}\cdot\left[\frac{\underline{s} - s_{i_1+i_2,k}}{\alpha}\right] > 0,$$

$$\frac{\partial[\bar{q}_{i_1t+1} - \bar{q}_{i_1t}]|_{s_{i_1+i_2,k}\to\underline{s}}}{\partial v_{i_2t}} = \frac{\partial(1-w_1)}{\partial v_{i_2t}}\cdot\left[\frac{\underline{s} - s_{i_1+i_2,k}}{\alpha}\right] < 0.$$

Because retraction always generates downward adjustment as compared to non-retraction (i.e. $\Delta[\bar{q}_{i_1t+1} - \bar{q}_{i_1t}]|_{s_{i_1+i_2,k}\to\underline{s}} < 0$), these two expressions suggest that retraction will have a smaller negative effect on i_1 if i_1 herself already has a long publication record besides the joint paper (i.e. higher v_{i_1t}), or if her coauthor i_2 has a short publication record besides the joint paper (i.e. lower v_{i_2t}).

The above discussion focuses on the absolute magnitude of $\Delta[\bar{q}_{i_1t+1} -$ *qi1t]|si1+i2,k→s*. We can also compare this effect in relative terms between the two authors, which amounts to:

$$\frac{\Delta[\bar{q}_{i_1t+1} - \bar{q}_{i_1t}]|_{s_{i_1+i_2,k}\to\underline{s}}}{\Delta[\bar{q}_{i_2t+1} - \bar{q}_{i_2t}]|_{s_{i_1+i_2,k}\to\underline{s}}} = \frac{1-w_1}{1-w_2} = \frac{v_{i_2t}}{v_{i_1t}}.$$

This ratio is smaller than one if the prior precision of i_1 (v_{i_1t}) is greater than the prior precision of i_2 (v_{i_2t}). In other words, retraction of the joint work will have a relatively smaller effect on the more established member of the team. Again, this is consistent with a reverse Matthew effect.

Now consider i_1's solo paper j that has been published before the retraction of her joint paper with i_2. Recall that citation of this paper at time t depends on the paper's signal $s_{i_1 j}$, the current belief about the author's quality, $\bar{q}_{i_1 t}$, and the duration since publication, $t-\tau_j$.

$$C_{i_1 jt} = f\left(s_{i_1 j}, \bar{q}_{i_1 t}, t-\tau_j\right).$$

Once the teamwork is retracted, it changes $\bar{q}_{it}$ and this change triggers changes in citation. In other words,

$$\Delta C_{i_1,j,t}|_{s_{i_1+i_2,k}\to \underline{s}} \approx \frac{\partial f\left(s_{ij}, \bar{q}_{i_1 t}, t-\tau_j\right)}{\partial \bar{q}_{i_1 t}} \cdot \Delta[\bar{q}_{i_1 t+1} - \bar{q}_{i_1 t}]|_{s_{i_1+i_2,k}\to \underline{s}}.$$

Under the assumption that $\frac{\partial f\left(s_{ij}, \bar{q}_{i_1 t}, t-\tau_j\right)}{\partial \bar{q}_{i_1 t}}$ is a positive constant, changes in citation are proportional to changes in the posterior mean of i_1 in response to the retraction, which in turn depends on the prior precision of i_1 and i_2 right before the retraction. Under the same assumption, the relative changes in citation are:

$$\frac{\Delta C_{i_1,j,t}|_{s_{i_1+i_2,k}\to \underline{s}}}{\Delta C_{i_2,j,t}|_{s_{i_1+i_2,k}\to \underline{s}}} = \frac{\Delta[\bar{q}_{i_1 t+1} - \bar{q}_{i_1 t}]|_{s_{i_1+i_2,k}\to \underline{s}}}{\Delta[\bar{q}_{i_2 t+1} - \bar{q}_{i_2 t}]|_{s_{i_1+i_2,k}\to \underline{s}}} = \frac{1-w_1}{1-w_2} = \frac{v_{i_2 t}}{v_{i_1 t}} = \frac{v_{i_1 0} + v_\varepsilon \cdot N_{i_1 t}}{v_{i_2 0} + v_\varepsilon \cdot N_{i_2 t}}.$$

It is easy to show that this expression decreases with the prior precision of oneself ($v_{i_1 t}$) but increase with the prior precision of the coauthor ($v_{i_2 t}$). Assuming both authors start from the same prior when they enter the profession ($v_{i_1 0} = v_{i_2 0}$), the prior precision of each author depends on the number of papers that each has published before the retraction. Thus, retraction of a joint paper will have negative spillover impacts on the citation of all authors' prior work. We have three sets of predictions regarding the magnitude of this negative effect:

- *Prediction 1: the more papers an author has written before the retraction, the smaller is the negative spillover from the retraction to her own prior work.*

This prediction holds because $\frac{d|\Delta C_{i_1,j,t}|s_{i_1+i_2,k} \to \underline{s}|}{dN_{i_1 t}} < 0$. The intuition behind this prediction is the same as what is shown in the solo case: if an author has a large number of publications before the retraction, the prior on her ability is strong and it is difficult to update on a strong prior.

- *Prediction 2: The negative impact of retraction on an author's prior work also depends on the status of her coauthor in the retracted teamwork. The more papers her coauthor has written before the retraction, the bigger is the harm of the retraction on herself. This negative spillover is more severe if she has a short publication record by herself. Put it another way, within the team, the negative effect on prior work will be of a greater magnitude if an author has relatively fewer publications than other team member at the time of retraction.*

These predictions can be proved by $\frac{d|\Delta C_{i_1,j,t}|s_{i_1+i_2,k} \to \underline{s}|}{dN_{i_2 t}} > 0$, $\frac{d^2|\Delta C_{i_1,j,t}|s_{i_1+i_2,k} \to \underline{s}|}{dN_{i_2 t} dN_{i_1 t}} < 0$, and $\frac{\Delta C_{i_1,j,t}|s_{i_1+i_2,k} \to \underline{s}}{\Delta C_{i_2,j,t}|s_{i_1+i_2,k} \to \underline{s}} > 1$ if $N_{i_1 t} < N_{i_2 t}$. The main reason driving these predictions is simple: if the community has a stronger prior on one author, the new information delivered by the retraction will be used more to update on the ability of the other author.

- *Prediction 3: The negative impact from the retraction to an author's prior work is greater if the retracted paper was perceived of higher quality before retraction. This effect decreases with the number of publications by that author.*

The first half of this prediction is obtained from $\frac{d|\Delta C_{i_1 jt}|s_{i_1+i_2,k} \to \underline{s}|}{ds_{i_1+i_2,k}} > 0$, and the second half can be proved by $\frac{d^2|\Delta C_{i_1 jt}|s_{i_1+i_2,k} \to \underline{s}|}{ds_{i_1+i_2,k} dN_{i_1 t}} < 0$. Both parts are easy to understand: the higher the signal before the retraction, the greater is the drop of the signal in the event of retraction, thus the more update there is in the societal belief. However, belief update also depends on the strength of prior on each author: because it is more difficult to update from a stronger prior, a large drop of the signal leads to less update on a more established name.

Overall, our Bayesian model predicts that a retraction of teamwork should have negative impacts on an author's prior work because of reputation update. The negative effect depends on one's own status at the time of retraction, the status of her coauthor in the retracted work, as well as the perceived quality of the retracted work before retraction.

IV. Data

Our data comes from the largest known repository of scientific knowledge, the Web of Science (WOS) from Thomson Reuters, which includes 25 million publications published in over 15,000 journals worldwide from 1945 to 2011. This database includes detailed bibliographic information for each paper (authors, journal, publication year, etc.) and further defines the citation linkages between each paper. It provides retraction notices that describe the time and reasons for each retraction and whether the errors are reported by the authors.

IV.1 Sample construction

Lu et. al (2012) show that a single retraction spreads citation losses to an author's prior work, but these penalties disappear if the author(s) self-report the error. Therefore, to examine how retraction affects authors by social status, we limit our sample to the retraction cases where scientific errors were not self reported. To avoid spillovers across multiple retractions by the same author, we further restrict the sample to single retractions where each author is only involved in one retraction between 1993 and 2011. As of August 31, 2011, we located 667 single retractions and 95% of these retracted papers (634) were written by more than one author. Among these team-authored retractions, 59.3%(376) are non-self-reported, 31.4%(199) are self-reported and 9.3%(59) have unclear or unknown retraction reasons. For each of the teamwork cases where the retraction was not self-reported, we identified the authors' prior work published before the retraction. The exact procedure we use to track the same author is described in Lu et al. (2012). We focus on prior work because this work is in a fixed published form that itself does not change because of the retraction. By contrast, any articles written after the retraction can be affected by the retraction in terms of publication time, publication outlet, etc. We refer to each prior publication by authors involved in single retraction of teamwork as a treated paper.

Because different papers may be cited differently at different stages of their life cycles and in different disciplines, we must construct a control group to match each "treated" paper in the pre-retraction period. The underlying assumption is that both treated and control articles will continue the same course of citation patterns if there were no retractions on the treated. This methodology draws on the earlier work of Furman and Stern (2011).

For a treated paper *i* published in field *f* and year *p*, we search for its control papers within the same field and the same publication year, where field is defined by the 252 field categories in the WOS. In particular, for each non-treated paper j in this pool, we define the arithmetic distance between i and j as

$$\mathrm{AD_{ij}} = \sum_{\mathrm{t=p}}^{\mathrm{r}-1}(\mathrm{c_{it}} - \mathrm{c_{jt}}) \qquad (1)$$

and the Euclidean distance between i and j as:

$$\mathrm{ED_{ij}} = \left[\sum_{\mathrm{t=p}}^{\mathrm{r}-1}\left(\mathrm{c_{it}} - \mathrm{c_{jt}}\right)^2\right]^{1/2} \qquad (2)$$

where c_{it} denotes the citations paper *i* receives in year *t* and *r* is the year of retraction. Both distances attempt to measure the citation discrepancy between paper *i* and paper j, but arithmetic distance $\mathrm{AD_{ij}}$ indicates whether j were cited less or more than i up to the year before retraction while Euclidean distance $\mathrm{ED_{ij}}$ is direction-free.

The quality of control group matching is assessed in Figure 1. A crude choice of control papers focuses on the ten papers with the lowest Euclidean distance to a treated paper. The upper-left graph of Figure 1 shows that the average Euclidean distance of the ten controls has high density around zero, which suggests close matches to the treated papers. In the bins for greater-than-zero distances, the density drops gradually except for the bin of 50 or more (which is driven by some retracted papers that were exceptionally highly cited before retraction). Because Euclidean distance is direction-free, there is no guarantee that the arithmetic distance of these ten control papers are distributed evenly on the two sides of the treated paper. As shown in the bottom-left graph of Figure 1, the average arithmetic distance of the ten controls has substantially more density on the negative side, so that these controls on average underestimate the citation flow of the treated papers.

If we restrict choice of control to the one paper with the lowest Euclidean distance, we are able to find perfect match (with zero AD) for 34.6% of the treated papers. As shown in the bottom-middle graph of Figure 1, when we cannot find a perfect match, the arithmetic distance of the one control is still negative on average, though it is more evenly distributed on both sides of zero than the ten-control sample.

To achieve a more careful match, where control papers have low Euclidean distance and low average distance, we further consider the two nearest neighbors, one from above (with positive AD) and one from below (with negative AD). As shown in the bottom-right panel of Figure 1, the density of the average arithmetic distance of these two controls is either exactly zero or concentrated in the neighborhood of zero. In particular, the two nearest neighbors now yield an average of zero distance for a substantially larger share (67.4%) of our treated papers. This sample, with zero distance, is the main sample used in our analysis.

Overall, by focusing on 376 team-authored, non-self-reported single retractions, our sample consists of 976 authors and 17,316 prior works written by authors involved in any retraction.[6] On average, each retracted team article is comprised of 2.6 authors and each author has 17.7 prior publications. The whole sample includes 578,025 observations (some papers with one year observation may be dropped out due to the Poisson estimation) and the unit of observation is an author/paper/year.

IV.2 Definition of Author Status

[6] Some papers will be counted more than once if several authors collaborated more than once.

We construct three variables to measure the social status of an author at the point when his or her paper was retracted: paper counts (production), total citations (impact) and h-index.[7] The h-index, which bridges between the total number of publications and the average citation impact of these publications, is defined as follows: a scholar with an index of h has published h papers each of which has been cited in other papers at least h times at the time point right before retraction. All these measures are subject to papers and citations within journals covered by WOS.

Taking each treated author as an observation, Figure 2 plots the distribution of h-index at the time of retraction. Consistent with the previous literature, the distribution is positively skewed, with a long tail on high status. Similar skewness exists for paper counts and total citations. In the main part of our statistical analysis, we use the continuous measure of paper counts, total citations and h-index. In robustness checks, we define two dummies to indicate whether an author is at the top 5 or top 10 percentile of a status distribution, and test whether these dichotomous classifications of high- and low-status generate similar results as the continuous measure of status.

Because we focus on retractions of teamwork, we also define a relative measure of social status that equals to one if an author has the highest social status within the team at the time of retraction. These authors are referred to as "relative eminent." Compared to the absolute measure of author status, relative eminence help us tell the relative difference within a team, even if all team members have high status or low status in the absolute measures. The relative eminence measure also helps us rule out the heterogeneity of social status measures across different academic fields.

IV.3 Data summary

Table 1 provides two panels of summary statistics: the first panel, at the author level, considers the status of each treated author at the time of retraction; the second panel, at the paper-year-author level, considers basic data pooling both treated and

[7] In the regressions, we normalized the total prior publications by 1,000, the total prior citations by 10,000 and prior h-index by 100.

control papers. As shown in the first panel, the distribution of author status – whether it is measured by total counts of prior work, total counts of citation, or h-index – is dispersed and skewed. On average, each treated author had 26 prior publications, 1,149 citations, and an h-index of 10 before retraction. For relative status, nearly half the author sample had the highest status among authors of the retracted publication. For absolute measures of status, the top 5 percentile of treated authors (as measured by total counts of prior work) account for 40% of the treated papers, and the top 10 percentile of authors account for 60% of the treated papers.

Figure 3 shows the citation patterns before and after retraction by social status of treated authors. Zero on the horizontal axis indicates the year of retraction, the blue line is for treated papers, and the red line is for control papers. In the upper panel, we separate treated authors by whether they are in the top 10 percentile of h-index among all treated authors. It is clear that a treated paper is cited similarly as the control papers (and even better after retraction) if the treated author is in the top 10 percentile, but a treated paper is cited less after retraction if its author is not in the top 10 percentile. The bottom panel repeats this exercise by distinguishing the highest-status author in the retracted teamwork from the other team members. Compared with control papers, citation on a treated paper is only reduced after retraction if its author did not have the highest status in the retracted teamwork. These graphs suggest that the after-retraction decline of citation is more severe for ordinary authors than for eminent authors.

V. Identification and Estimation

Our identification strategy employs triple differences-in-differences. Pooling treated and control papers in one dataset, ideally we would like to perform the following regression to identify the treatment effect across different social status of authors:

$$\begin{aligned}\Pr(y_{irt}) = f(\alpha_i + \mu_t + \beta_{triple}\cdot Status_r \cdot Treat_i \cdot Post_{kt} + \beta_{dif}\cdot Treat_i \cdot Post_{kt} \\ + \beta_{stat}\cdot Status_r \cdot Post_{kt} + \beta_{post}\cdot Post_{kt})\end{aligned} \tag{3}$$

where *i* indexes article, *r* indexed author and *t* indexes how long the paper has been published up to the citation time, and the dependent variable, *y*, denotes counts of citations to article *i* at time *t* for author *r*. Fixed effects for each paper (α_i) and each year since publication (μ_t) capture the mean citation pattern of articles. $Treat_i$ is a dummy variable that equals 1 if article *i* is a treatment paper, and $Post_{kt}$ is a dummy variable that equals 1 if year *t* is after the retraction event for a given treatment and control group *k*. $Status_r$ is either a continuous measure of the status of the treated author, or a dummy variable that equals 1 if a treated author is eminent at the time of retraction. We use the social status of each author at the pre-retraction year to interact with $Treat_i \cdot Post_{kt}$ because the post-retraction social status of each author will be affected by the event of retraction.

The coefficient of interest (β_{triple}) captures any difference in citations to eminent authors on a treated paper, after the retraction event, compared to the non-eminent coauthors of the retracted paper. We estimate (3) using the standard Poisson model for count data and cluster it by a given treatment and control group.

The key identification assumption is that the retraction event is unrelated to unobservable features that drive the citation dynamics; therefore prior work written by authors of different social statuses will continue the same course of citation patterns as their control papers if there were no retractions on the treated. Because our selection of control papers is not conditional on author status, this is equivalent to assuming that control papers that closely match the year-to-year citation of a treated paper before retraction can predict the future citation patterns, regardless of author status. We expect this assumption to be less valid if the prior work is published close to the retraction time and therefore give us a shorter time window for the match of control papers. As robustness checks, we will exclude these treated papers and test whether the results change.

VI. Results

This section has three parts: we first test the theoretical predictions on the role of own and coauthor reputation, and then look for evidence on the importance of retraction. The third part focuses on alternative explanations and robustness checks.

VI.1 Main results on author reputation

We first confirm that retraction has a significant negative spillover effect on the citation of the same author's prior work, if the retraction was not triggered by self-reported error. Figure 4 shows the spillover effect of non-self-reported retraction on a team author's prior work. The annual flow of citations to a prior publication falls 4.8%(p<.0001) in the first two years after the retraction and 13.0%(p<0.0001) five or more years after the retraction, compared to the control group. This suggests that non-self-reported retractions lead to substantial citation declines to prior work of team authors, which is consistent with the results shown in Azoulay et. al (2012) and Lu et. al (2012).

(1). Author's Own Absolute Reputation

Table 2 reports results from our main specification. We highlight the differences-in-differences coefficient on *treated*post* retraction (t>=1) and the triple differences-in-differences coefficient on *author reputation * treated * post* (t>=1). The latter indicates whether a treated author that had a different absolute reputation at the time of retraction receives different citations on their prior work after the retraction. There are five columns in the table, differing by measures of absolute reputation. The first three columns use the continuous measures in total prior publications (production), total prior citations (impact) and h-index respectively. The last two define a binary variable equal to one if an author's h-index falls in the top 5 or top 10 percentile of all authors involved in the studied retractions. Authors with this dummy equal to one are referred to as "absolute eminent" and the other authors are referred to as "absolute ordinary." We separate the year of retraction from the rest of years because we do not know the exact time of retraction within the year.

All three continuous measures of reputation show similar patterns so we only discuss column (3) as an example. The coefficient in column (3) indicates that one standard deviation increase in prior h-index results in 4.9% less reduction in citations per year per paper due to retraction.[8] This suggests that higher status at the time of retraction may help to alleviate the reputation harm due to retraction, which is consistent with our first prediction from Section II.

To make sure that the above results are not driven by a few extremely productive authors, we separate authors in the top 5 or top 10 percentile of h-index from the other treated authors. As shown in Columns (4) and (5), results using these crude classifications are similar to those using continuous measures. For example, we find that compared to closely-matched control papers, citations fall by an average of 10.0% ($p<0.0001$) per year for each prior publication made by an ordinary author whose h-index is out of the top 10 percentile of all treated authors. However, the percentage reduction in citations for those in top 10 percentile is much smaller (2.1%). This percentage difference is highly significant (7.9% per year per paper, $p<0.0001$). A comparison of columns (4) and (5) further suggests that as more authors are classified as "eminence", the citation reduction increases for the ordinary authors and the differences between eminence and ordinary authors become smaller. Again, this is consistent with the first prediction from Section II.

(2). Author Reputation within a Team

Like Table 2, Table 3 reports results from the main specification but measures author reputation relative to other authors of the retracted teamwork. In particular, we define a dummy equal to one if a treated author has the highest number of prior work at the time of retraction within the team. Similar dummies are created when we compare authors on total number of prior citations and h-index. These three dummies are used separately in the three columns of Table 3. Because results are similar across the three columns, we will only discuss the coefficients of Column 3. Within a team,

[8] The standard deviation of h-index is 22.8. The change is 0.214*22.8/100=4.9%.

Column (3) shows that citations after retraction fall by 3.7% per year per paper for the author with the highest h-index within the team while the corresponding number is 11.1% for her lower-status teammates. This suggests that the reduction in citations for high status authors in teams is 7.4% smaller than their low status coauthors due to retraction. This result confirms the second prediction from Section II.

Another way to examine the role of team members is checking whether coauthor reputation influences the effect of retraction on one's own prior work. In particular, we separate our full analysis sample into two, according to whether a treated author is in the top ten percentile of a reputation measure (refereed to as eminent authors) or not (referred to as ordinary authors). Then within each subsample, we estimate the main specification but define the triple differences-in-differences term as *coauthor reputation*treated*post* (t>=1), where coauthor reputation is a dummy equal to one if at least one coauthor in the retracted teamwork has a better reputation than oneself at the time of retraction.

In every column of Table 4, the reputation measure we use to define whether oneself is ordinary or eminent is always the same as the reputation measure that defines whether there is a more reputable coauthor within the team. This leads to six columns, columns (1) to (3) for ordinary authors in the three measures of reputation, and columns (4) to (6) for eminent authors.

Comparing column (3) with column (6), we find that the prior work of an ordinary author suffers a greater decline of citation if some coauthor in the retracted teamwork has a higher h-index than herself at the time of retraction; in contrast, the effect of retraction on an eminent author is not sensitive to having an even more eminent coauthor in the retracted teamwork. This difference also holds when reputation is measured by number of prior work or number of prior citations. Again, the results confirm the second prediction from Section II.

VI.2 Importance of retraction

The third prediction from Section II says that the retraction will have a more negative spillover effect on prior work if the article was perceived to be a good paper before retraction, and this effect should be less when the author has a good reputation. To test this prediction, we construct two variables to proxy for paper importance: one is a dummy of whether the retracted article was published in one of the top three journals (Science, Nature and PNAS). These three are general journals, cross all fields and well-recognized, especially in the fields of science. Papers published in these three journals are retracted at an average rate of 0.91% over the 2000-2009 period, which is 9.6 times the background retraction rate (Lu et. al, 2012) A retraction in one of the three journals receives substantial attention from both the media and the research community, thus the shock of retraction should be greater and more recognizable in the scientific community.

Table 5, Column (1)-(6) reports the estimation on two subsamples, one includes the authors with a retraction in the top three journals, and the second subsample includes all the other authors. As the theory predicts, we find more negative impact from a retraction at the top three journals, and the hit is more protected in case of top-three retractions if one has a better reputation at the time of retraction.

The second proxy for paper importance is constructed by citations of the retracted papers till retraction for each retracted paper. We interact this proxy with our difference-in-difference conjecture *treated*post (t>=1)* and find that a one unit increase in citations of a retracted paper leads to 5.9% reductions in citations of prior work for the authors of this retracted paper, as Table 5, Column (7) shown. This finding echoes with our theoretical prediction that an author gets great reputation loss if a retracted paper is highly cited before the retraction.

VI.3 Alternative Hypotheses and Robustness Checks

(1) Self Citations

In general, eminent authors are more productive than their ordinary peers. Therefore, they may have more follow-up papers to cite their own prior work, which could generate a pattern observed in our main results. To address this alternative explanation, we exclude self citations from the citation counts and re-estimate the model. Results presented in Table 6 are very similar to Tables 2 and 3, no matter whether we use the absolute or relative reputation measures. Take Column (6) as an example: citations fall by 14.0% per year per paper for the low-reputation authors after retraction and 5.8% for their high-reputation coauthors. The magnitude in citation reduction becomes slightly larger for both types of researchers compared to the last column of Table 3, because in theory the negative spillover effect on prior work should mostly come from non-self citations.

(2) Old Papers

One may argue that old papers are close to the end of a paper's life cycle and cannot get worse than zero citations after retraction. Because eminence authors are more senior and may have more dying papers than ordinary authors, this could contribute to the smaller citation reductions for eminent authors.

Figure A1 shows the citation trajectories for our treated papers. The average citations for treated papers fall to two at the 10th years since publication and converge to one after the 15th year since publication. Suggested by Figure A1, we limit our sample by excluding those prior articles published more than 10 years earlier than the retraction year. As a result, 61.8% of treated papers and 43.7% of total observations are kept in the new subsample.

As shown in Table 7, results estimated on this subsample remain robust. Citations fall fewer for high-reputation authors than for low-reputation authors after retraction. According to Column (6), citations fall by 10.6% for low-reputation authors after retraction and the percentage difference between high- and low-reputation researchers is 7.5%.

If the old paper hypothesis holds, the coefficient of *Treated*Post(t>=1)* should be more negative and the differences between high- and low-reputation authors would be smaller after old papers are excluded from the citation counts. Both numbers shown in Column (6) of Table 7 are similar to the corresponding ones in Table 3. This is inconsistent with the old paper hypothesis.

(3) Task Allocation

The third alternative explanation for our main results is task allocation. One may argue that eminent authors contribute less to the tasks (e.g. coding or lab experiments) that led to the error in the retracted paper and therefore they receive less blame from the community. The underlying assumption of this argument is that authors with similar reputation take similar tasks within the team. If task allocation was the reason for our main findings, we would expect to see that authors with similar reputation when publishing the retracted papers should have the similar punishment regardless of whether he or she is eminent or not later when the paper gets retracted.

To test this hypothesis, we first construct the past-reputation measures using the status of an author when he or she published a retracted paper. Then we include both author reputation at the time of retraction and author reputation at the time of publication in the triple differences-in-differences estimation. For easy interpretation, both reputations are measured by a dummy of whether the reputation is in the top 10 percentile of all treated authors. As shown in the first three columns of Table 7, the reduced effect on high-reputation authors is driven by reputation at the time of retraction, not the time of publishing the retracted work. This is against the task allocation hypothesis.

In the last three columns of Table 8, we further restrict the sample to ordinary authors only, where ordinary authors refer to those that are out of top 10 percentile at time of publishing the retracted work.[9] Some of these authors became famous later and

[9] In terms of sample size, the low-reputation sample includes more authors with fewer papers per author than the high-reputation sample. Nevertheless, the sample size of the two subsamples is quite identical.

others remained ordinary at the time point of retraction. If the reputation effect observed in our main results is primarily driven by reputation at the time of publication, we should see little effect from the reputation at time of retraction, once we focus on ordinary authors at time of publication. Results shown in the last three columns of Table 8 are clearly again this prediction, suggesting that those authors that became famous late on still got less reputation harm than those remaining ordinary.[10] Above all, we are confident that task allocation is not the key explanation for our main findings.

(4) Other Robustness Checks

We conduct a series of robustness checks by estimating different samples and different models. First, we use the full sample with two control papers, regardless of whether the two control papers have an average zero arithmetic distance to the treated paper or not. Estimates from this sample are shown in Table A1. Second, we replace our poisson estimation with OLS estimation. The OLS results are reported in Table A2. Third, we consider the possibility that citation counts are left censored by zero and therefore a paper that already receives zero citations cannot have further reduction in citations. This issue is different from the old paper hypothesis because zero citations could occur soon after publication, especially for ordinary authors who do not have many high quality publications. To deal with this issue, we exclude all prior work that has zero citations in the year before retraction. As shown in Table A3, results remain robust in those still-cited papers. Fourth, prior work that was published close to the retraction time may have a short citation history before retraction, which could hurt our ability to find the best control papers. We address this issue by excluding all prior work published within three years before retraction. Results are shown in Table A4.

[10] We run the same regressions on the high-reputation sample. In this sample, all authors were high-reputation when they published the papers. Some of them remain to be famous and others are unable to maintain their reputation as time goes. Consistent with the findings from the low-status sample, we find that those authors remaining to be famous got less reputation harm than those whose reputation was depreciated over time.

Overall, the results remain robust and support our main finding that ordinary authors receive more blame than eminent authors when a paper is retracted.

VII. Conclusion

We have considered a natural experiment to assess the relationship between scientific discredit and scientific misconduct in team production settings. Our results demonstrate asymmetry: Eminent authors show no change in citations to their prior work after a coauthored retraction, while less eminent coauthors experience large citation losses. These and other results are consistent with a simple Bayesian model updating model that operationalizes Merton's canonical Matthew Effect. Not only do the rich get richer, when riches are to be had, but the rich don't get poorer when catastrophe strikes.

Team production now comprises the vast majority of papers in the sciences and engineering. Therefore, issues of credit sharing become more acute. Especially for junior scientists, who increasingly establish their individual reputations exclusively through team-authored outputs, the Matthew Effect presents a difficult challenge. If established authors can both take credit for successes and avoid discredit from failures, the junior author may take substantially longer to develop their own reputation while facing greater career risks along the way. These features may act as entry barriers to scientific careers. More subtly, these concerns may influence how scientists choose collaborators, so that credit considerations turn scientists away from potentially productive teams. These issues are important areas for future work.

References

1. Azoulay, Pierre; Toby Stuart; and Yanbo Wang (2012) "Matthew: Effect or Fable" NBER working paper series #18625.
2. Azoulay, Pierre; Jeffrey L. Furman; Joshua L. Krieger, and Fiona E. Murray (2012): "Retractions" NBER working paper series #18499.
3. Bar-Isaac, Heski and Steve Tadelis (2008): "Seller Reputation" *Foundations and Trends in Microeconomics,* 2008, 4(4):273-351.
4. Borenstein, Severin and Martin B. Zimmerman (1988) "Market Incentives for Safe Commercial Airline Operation." American Economic Review 78: 913-35.
5. Cabral, Luis and Ali Hortacsu (2010): "The Dynamics of Seller Reputation: Evidence from eBay" The Journal of Industrial Economics 58(1): 54-78.
6. Dranove, David; Subramaniam Ramanarayanan and Yasutora Watanabe (2012): "Delivering Bad News: Market Responses to Negligence", Journal of Law and Economics, 55(1).
7. Falk, Armin and Andrea Ichino (2006): "Clean Evidence on Peer Effects" Journal of Labor Economics, 24(1): 39-57.
8. Fang, Ferric C.; R. Grant Steenc, and Arturo Casadevalld (2012): "Misconduct accounts for the majority of retracted scientific publications" PNAS.
9. Fehr, Ernst and Simon Gachter (2000): Cooperation and Punishment in Public Goods Experiments" American Economic Review, 90(4): 980-94.
10. Furman, Jeffrey L. and Scott Stern (2011): "Climbing atop the Shoulders of Giants: The Impact of Institutions on Cumulative Research" American Economic Review 101: 1933-1963.
11. Jin, Ginger Z. and Phillip Leslie (2009): "Reputation Incentives for Restaurant Hygiene" *American Economic Journal (Microeconomics)* February 2009, 1(1): 237-67.
12. Johnson, Erin (2012): "Ability, Learning and the Career Path of Cardiac Specialists" MIT Sloan working paper.

13. Jones, Benjamin F. (2009): "The Burden of Knowledge and the Death of the Renaissance Man: Is Innovation Getting Harder?" *Review of Economic Studies*. 76(1).
14. Klein, Benjamin and Keith B. Leffler (1981): "The Role of Market Forces in Assuring Contractal Performance" Journal of Political Economy 89(4): 615-641.
15. Lu, Susan Feng; Ginger Jin; Benjamin Jones; Brian Uzzi "The Prior Publication Penalty: An Investigation of Retraction and False Science", working paper.
16. Mas, Alexandre and Enrico Moretti (2009): "Peers at Work□" The American Economic Review, 99(1): 112-145.
17. Merton, Robert (1968) "The Matthew Effect in Science", Science 159 (3810): 56-63.
18. Merton, Robert (1988) "The Matthew Effect in Science, II: Cumulative Advantage and the Symbolism of Intellectual Property" ISIS 79: 606-623.
19. Peterson, Alexander M.; Woo-Sung Jung; Jae-Suk Yang; and H. Eugene Stanley (2011) "Quantitative and empirical demonstration of the Matthew effect in a study of career longevity", PNAS 108(1).
20. Pope, Devin (2009): "Reacting to Rankings: Evidence from 'America's Best Hospitals'" *Journal of Health Economics* 28(6), 1154-1165.
21. Shapiro, Carl (1983): "Premiums for High Quality Products as Returns to Reputations" The Quarterly Journal of Economics 98(4): 659-679.
22. Stringer, Michael, Marta Sales-Pardo, and Luis A. Nunes Amaral (2010): "Statistical Validation of a Global Model for the Distribution of the Ultimate Number of Citations Accrued by Papers Published in a Scientific Journal" *Journal of the American Society for Information Science and Technology* 61(7): 1377-1385.
23. Wuchty, Stefan, Benjamin F Jones and Brian Uzzi (2007a): "The Increasing Dominance of Teams in the Production of Knowledge" *Science*. 316(5827): 1036-1039.
24. Wuchty, Stefan, Benjamin F Jones and Brian Uzzi (2007b): "Why Do Team Authored Papers Get Cited More?" *Science (Letters)*. 317(5844): 1496-1498.

Figure 1: Matching quality of control papers

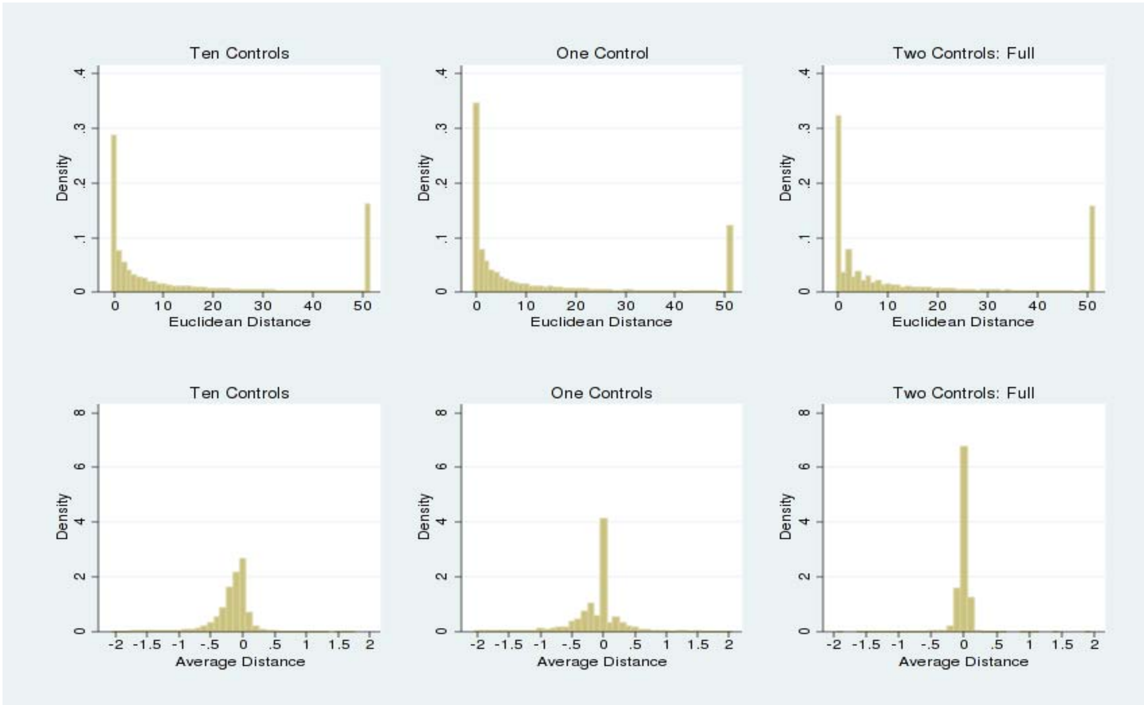
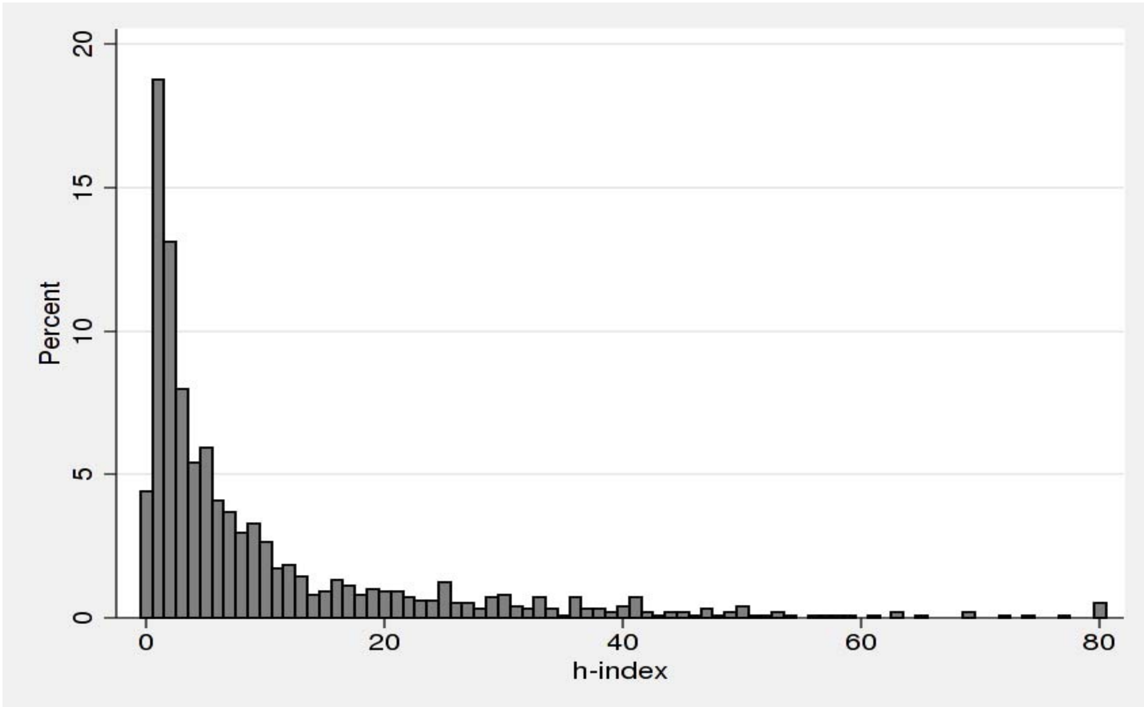
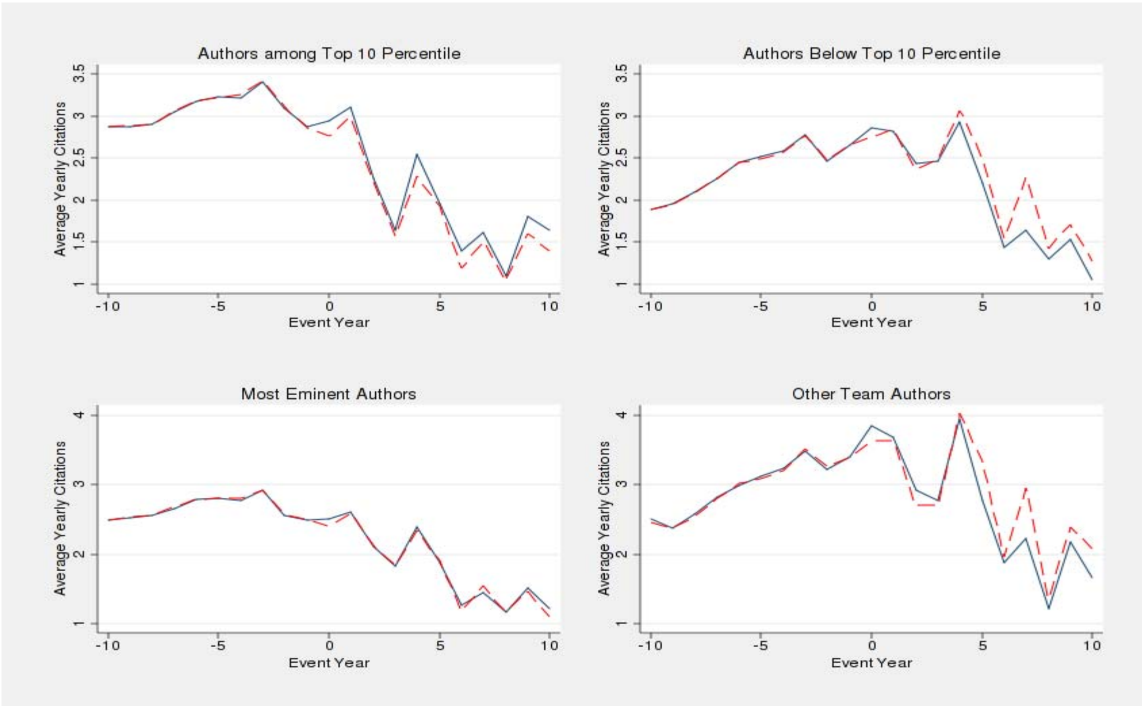


Figure 2: Distribution of h-index per treated author at the time of retraction



Note: we treated those authors with h-index greater than 80 as 80 in this figure.

Figure 3: Citation before and after retraction, by author status in the treated paper



Blue for treated papers, red for control papers.

Figure 4: The effect of retraction on the citation of a treated paper, by year since retraction

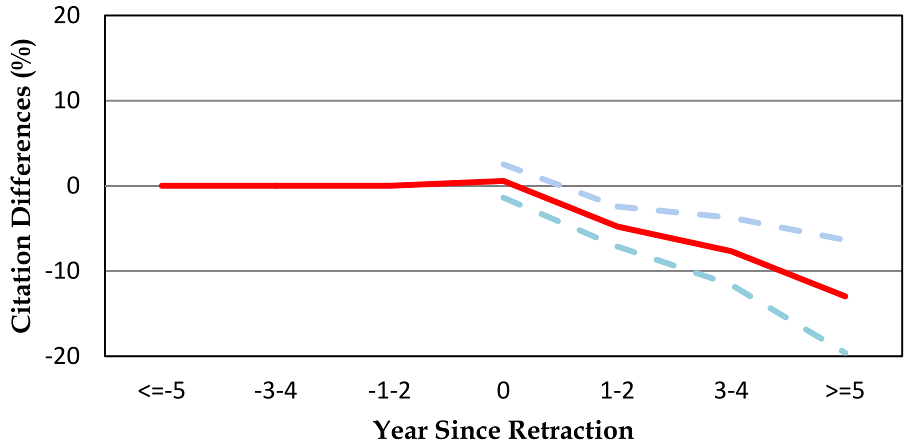


Figure A1: citation life cycle of control papers

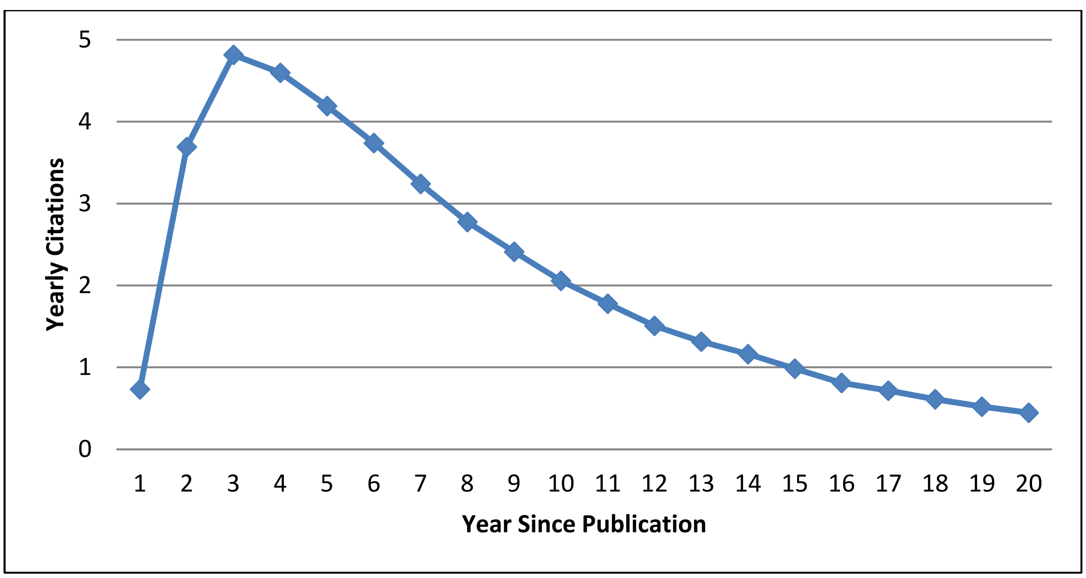


Table 1: Summary statistics

Panel A: unit of observation = author, treated only

Variables	Definition	Obs	MEAN	SD	Min	Max
Absolute Measure of Status						
Prior Publications	total prior paper counts	976	26	50	1	452
Prior Citations	total prior citations	976	1149	3538	0	67946
Prior h-index	prior h-index	976	10	14	0	132
Relative Measure of Status						
Production	1 if an author has the most prior papers within the team[a]	976	0.4	0.5	0	1
Impact	1 if an author has the most prior citations within the team[a]	976	0.5	0.5	0	1
h-index	1 if an author has the highest prior h-index within the team[a]	976	0.5	0.5	0	1

Panel B: unit of observation = paper-author-year

Variables	Definition	Obs	MEAN	SD	Min	Max
Yearly Citations	yearly citation counts	578025	2.6	5.1	0	129
Absolute Measure of Status						
Prior Publications	total prior papers	578025	122	106	1	452
Prior Citations	total prior citations	578025	5975	8220	0	67946
Prior h-index	prior h-index	578025	34	23	0	132
Production at Top 5	1 if an author falls into top 5 percentile of total prior papers	578025	0.4	0.5	0	1
Impact at Top 5	1 if an author falls into top 5 percentile of total prior citations	578025	0.3	0.5	0	1
h-index at Top5	1 if an author falls into top 5 percentile of prior h-index	578025	0.3	0.5	0	1
Production at Top 10	1 if an author falls into top 10 percentile of total prior papers	578025	0.6	0.5	0	1
Impact at Top 10	1 if an author falls into top 10 percentile of total prior citations	578025	0.5	0.5	0	1
h-index at Top 10	1 if an author falls into top 10 percentile of prior h-index	578025	0.5	0.5	0	1
Relative Measure of Status						
Production	1 if an author has the most prior papers within the team[a]	578025	0.8	0.4	0	1
Impact	1 if an author has the most prior citations within the team[a]	578025	0.8	0.4	0	1
h-index	1 if an author has the highest prior h-index within the team[a]	578025	0.8	0.4	0	1

[a] "within the team" refers to within the team-authored retracted paper of that author.

Table 2: Effect of retraction on citation of prior work, by absolute reputation measures of the treated author at the time of retraction

Absolute Reputation of the treated author	Continuous Measures			Discrete Measures	
	Total # of prior papers	Total # of prior citations	H-index	=1 if within Top 5% of H-index	=1 if within top 10% of H-index
	(1)	(2)	(3)	(4)	(5)
Treated*Post(t>=1)	-0.101***	-0.109***	-0.134***	-0.098***	-0.105***
	(0.022)	(0.019)	(0.026)	(0.018)	(0.021)
Author reputation*Treated*Post(t>=1)	0.347**	0.083***	0.214***	0.093***	0.076***
	(0.163)	(0.023)	(0.068)	(0.032)	(0.029)
Treated*Post(t=0)	0.001	-0.004	-0.013	-0.0003	-0.011
	(0.014)	(0.011)	(0.015)	(0.012)	(0.014)
Author reputation*Treated*Post(t=0)	0.036	0.014*	0.051	0.015	0.032
	(0.100)	(0.008)	(0.036)	(0.021)	(0.020)
Post(t>=1)	-0.080***	-0.043***	-0.026	-0.113***	-0.094***
	(0.015)	(0.013)	(0.019)	(0.012)	(0.014)
Post(t=0)	-0.121***	-0.101***	-0.096***	-0.158***	-0.159***
	(0.021)	(0.018)	(0.024)	(0.017)	(0.020)
Author reputation*Post(t>=1)	-0.547***	-0.154***	-0.337***	-0.065**	-0.085***
	(0.132)	(0.017)	(0.060)	(0.025)	(0.022)
Author reputation*Post(t=0)	-0.783***	-0.184***	-0.314***	-0.114***	-0.072***
	(0.147)	(0.022)	(0.063)	(0.028)	(0.026)
Paper Fixed Effects	Y	Y	Y	Y	Y
Year Since Publication Dummies	Y	Y	Y	Y	Y
Observations	549,928	549,928	549,928	549,928	549,928
Number of unique papers	47,999	47,999	47,999	47,999	47,999

Author reputation refers to the absolute reputation of a treated author at the time of retraction. Every paper in the same treatment-control group has the same value on author reputation. Total number of prior papers is measured in 1,000; total number of prior citations in measured in 10,000 and H-index is measured in 100. All regressions report coefficients from maximum likelihood estimation of a Poisson count model, errors clustered by each treatment-control group. Standard errors in parentheses, ***p<0.01, **p<0.05, *p<0.1.

Table 3: Effect of retraction on citation of prior work, by relative reputation measures of the treated author at the time of retraction

Relative reputation of a treated author within the team	Discrete measures		
	=1 if have the largest # of prior work	=1 if have the largest # of prior citations	=1 if have the highest h-index
	(1)	(2)	(3)
Treated*Post(t>=1)	-0.113***	-0.118***	-0.118***
	(0.025)	(0.024)	(0.025)
Author reputation*Treated*Post(t>=1)	0.065**	0.073**	0.071**
	(0.031)	(0.031)	(0.031)
Treated*Post(t=0)	-0.005	-0.013	-0.009
	(0.017)	(0.017)	(0.017)
Author reputation*Treated*Post(t=0)	0.014	0.026	0.020
	(0.021)	(0.021)	(0.021)
Post(t>=1)	-0.093***	-0.097***	-0.087***
	(0.017)	(0.017)	(0.017)
Post(t=0)	-0.206***	-0.206***	-0.211***
	(0.024)	(0.024)	(0.024)
Author reputation*Post(t>=1)	-0.066***	-0.061***	-0.074***
	(0.022)	(0.022)	(0.022)
Author reputation*Post(t=0)	0.022	0.022	0.03
	(0.028)	(0.028)	(0.029)
Paper Fixed Effects	Y	Y	Y
Year Since Publication Dummies	Y	Y	Y
Observations	549,928	549,928	549,928
Number of unique papers	47,999	47,999	47,999

Author reputation refers to the dummy of whether a treated author had the highest reputation within the team at the time of retraction. Every paper in the same treatment-control group has the same value on author reputation. All regressions report coefficients from maximum likelihood estimation of a Poisson count model, errors clustered by each treatment-control group. Standard errors in parentheses, ***p<0.01, **p<0.05, *p<0.1.

Table 4: Effect of retraction on citation of prior work, by own and coauthor reputation

	Ordinary Authors			Eminence Authors		
Absolute reputation of co-authors in the retracted teamwork	Total # of prior work	Total # of prior citations	h-index	Total # of prior work	Total # of prior citations	h-index
	(1)	(2)	(3)	(4)	(5)	(6)
Treated*Post(t>=1)	-0.050	-0.008	-0.055*	-0.047**	-0.065***	-0.040*
	(0.032)	(0.030)	(0.030)	(0.022)	(0.023)	(0.023)
Co-author has a better reputation than self *Treated*Post(t>=1)	-0.087**	-0.141***	-0.100**	0.089*	0.030	0.075
	(0.042)	(0.042)	(0.042)	(0.052)	(0.053)	(0.050)
Paper Fixed Effects	Y	Y	Y	Y	Y	Y
Year Since Publication Dummies	Y	Y	Y	Y	Y	Y
Observations	225,684	259,354	259,774	324,244	290,574	290,154
Number of unique papers	23,068	26,228	26,224	24,931	21,771	21,775

An author is defined ordinary (eminent) if her absolute reputation measure fell out of (into) the top 10 percentile of all treated authors at the time of retraction. "Coauthor has a better reputation" is a dummy equal to one if at least one coauthor in the retracted work has a better reputation than the author herself at the time of retraction. Self and coauthor reputations are always measured by the same metric. Total number of prior papers is measured in 1,000; total number of prior citations in measured in 10,000 and H-index is measured in 100. All regressions report coefficients from maximum likelihood estimation of a Poisson count model, errors clustered by each treatment-control group. All regressions include the same one-way and two-way interactions terms as in Table 2, we do not report their coefficients for illustration purpose only. Standard errors in parentheses, ***p<0.01, **p<0.05, *p<0.1.

Table 5: Effect of retraction on citation of prior work, by paper importance

	Retraction from Top Three			Retraction from Non-Top Three			Full sample
Absolute measure of author reputation	Total # of prior work	Total # of prior citations	h-index	Total # of prior work	Total # of prior citations	h-index	Total # of prior citations
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Treated*Post(t>=1)	-0.202**	-0.176**	-0.282***	-0.089***	-0.104***	-0.115***	-0.061***
	(0.085)	(0.083)	(0.106)	(0.022)	(0.019)	(0.027)	(0.016)
Author reputation*Treated*Post(t>=1)	1.396**	0.166	0.551**	0.210	0.078***	0.156**	-0.068*
	(0.655)	(0.108)	(0.234)	(0.166)	(0.024)	(0.070)	(0.038)
Paper Fixed Effects	Y	Y	Y	Y	Y	Y	Y
Year Since Publication Dummies	Y	Y	Y	Y	Y	Y	Y
Observations	41,407	41,407	41,407	508,521	508,521	508,521	549,928
Number of unique papers	2,762	2,762	2,762	45,237	45,237	45,237	47,999

Author reputation refers to the absolute reputation of a treated author at the time of retraction. Every paper in the same treatment-control group has the same value on author reputation. Total number of prior papers is measured in 1,000; total number of prior citations in measured in 10,000 and H-index is measured in 100. All regressions report coefficients from maximum likelihood estimation of a Poisson count model, errors clustered by each treatment-control group. All regressions include the same one-way and two-way interactions terms as in Table 2, we do not report their coefficients for illustration purpose only. Standard errors in parentheses, ***p<0.01, **p<0.05, *p<0.1.

Table 6: Effect of retraction on citation of prior work, excluding self-citations

	Absolute Reputation			Relative Reputation		
Measure of Author reputation	Total # of prior work	Total # of prior citations	h-index	=1 if have the largest # of prior work within the team	=1 if have the largest # of prior citations within the team	=1 if have the largest h-index within the team
	(1)	(2)	(3)	(4)	(5)	(6)
Treated*Post(t>=1)	-0.126***	-0.147***	-0.175***	-0.143***	-0.152***	-0.150***
	(0.024)	(0.020)	(0.028)	(0.026)	(0.026)	(0.027)
Author reputation*Treated*Post(t>=1)	0.319*	0.105***	0.254***	0.068**	0.082**	0.078**
	(0.175)	(0.025)	(0.073)	(0.033)	(0.033)	(0.033)
Paper Fixed Effects	Y	Y	Y	Y	Y	Y
Year Since Publication Dummies	Y	Y	Y	Y	Y	Y
Observations	548,817	548,817	548,817	548,817	548,817	548,817
Number of unique papers	47,798	47,798	47,798	47,798	47,798	47,798

When author reputation is measured in absolute terms, it refers to the absolute reputation of a treated author at the time of retraction. Every paper in the same treatment-control group has the same value on author reputation. Total number of prior papers is measured in 1,000; total number of prior citations in measured in 10,000 and H-index is measured in 100. When author reputation is measured in relative terms, it refers to a dummy equal to one if the treated author has the highest reputation within the team. All regressions report coefficients from maximum likelihood estimation of a Poisson count model, errors clustered by each treatment-control group. All regressions include the same one-way and two-way interactions terms as in Table 2, we do not report their coefficients for illustration purpose only. Standard errors in parentheses, ***p<0.01, **p<0.05, *p<0.1.

Table 7: Effect of retraction on citation of prior work, excluding old papers

	Absolute Reputation			Relative Reputation		
Measure of Author reputation	Total # of prior work	Total # of prior citations	h-index	=1 if have the largest # of prior work within the team	=1 if have the largest # of prior citations within the team	=1 if have the largest h-index within the team
	(1)	(2)	(3)	(4)	(5)	(6)
Treated*Post(t>=1)	-0.097***	-0.103***	-0.128***	-0.106***	-0.112***	-0.112***
	(0.024)	(0.020)	(0.028)	(0.026)	(0.025)	(0.026)
Author reputation*Treated*Post(t>=1)	0.363**	0.084***	0.214***	0.063*	0.073**	0.072**
	(0.181)	(0.026)	(0.074)	(0.033)	(0.033)	(0.033)
Paper Fixed Effects	Y	Y	Y	Y	Y	Y
Year Since Publication Dummies	Y	Y	Y	Y	Y	Y
Observations	252,354	252,354	252,354	252,354	252,354	252,354
Number of unique papers	31,672	31,672	31,672	31,672	31,672	31,672

When author reputation is measured in absolute terms, it refers to the absolute reputation of a treated author at the time of retraction. Every paper in the same treatment-control group has the same value on author reputation. Total number of prior papers is measured in 1,000; total number of prior citations in measured in 10,000 and H-index is measured in 100. When author reputation is measured in relative terms, it refers to a dummy equal to one if the treated author has the highest reputation within the team. All regressions report coefficients from maximum likelihood estimation of a Poisson count model, errors clustered by each treatment-control group. All regressions include the same one-way and two-way interactions terms as in Table 2, we do not report their coefficients for illustration purpose only. Standard errors in parentheses, ***p<0.01, **p<0.05, *p<0.1.

Table 8: Effect of retraction on citation of prior work, including author reputation at the time of publishing the retracted teamwork

	Full Sample			Ordinary Authors at Publishing		
Author reputation measures	=1 if total # of prior work is in top 10%	=1 if total # of prior citations is in top 10%	=1 if h-index is in top 10%	=1 if total # of prior work is in top 10%	=1 if total # of prior citations is in top 10%	=1 if h-index is in top 10%
	(1)	(2)	(3)	(1)	(2)	(3)
Treated*Post(t>=1)	-0.097***	-0.085***	-0.104***	-0.096***	-0.082***	-0.105***
	(0.021)	(0.021)	(0.021)	(0.021)	(0.021)	(0.022)
Author reputation at time of retraction *Treated*Post(t>=1)	0.179*	-0.031	0.091*	0.193*	-0.054	0.105*
	(0.099)	(0.056)	(0.047)	(0.112)	(0.071)	(0.060)
Author reputation at time of publication *Treated*Post(t>=1)	-0.124	0.065	-0.018			
	(0.099)	(0.056)	(0.046)			
Paper Fixed Effects	Y	Y	Y	Y	Y	Y
Year Since Publication Dummies	Y	Y	Y	Y	Y	Y
Observations	549,928	549,928	549,928	233,943	266,397	259,778
Number of unique papers	47,999	47,999	47,999	23,900	26,442	26,113

An author is defined ordinary at publishing if her absolute reputation measure fell out of the top 10 percentile of all treated authors at the time of publishing the retracted paper. Author reputation refers to the absolute reputation of a treated author if this author falls into top 10 percentile of all treated authors at the time of retraction. Every paper in the same treatment-control group has the same value on author reputation. All regressions report coefficients from maximum likelihood estimation of a Poisson count model, errors clustered by each treatment-control group. All regressions include the same one-way and two-way interactions terms as in Table 2, we do not report their coefficients for illustration purpose only. Standard errors in parentheses, ***p<0.01, **p<0.05, *p<0.1.

Table A1: Effect of retraction on citation of prior work, based on all of the two-control sample, including those that have on average non-zero arithmetic distance to the treated paper

	Absolute Reputation			Relative Reputation		
Author reputation measures	Total # of prior work	Total # of prior citations	h-index	=1 if have the largest # of prior work within the team	=1 if have the largest # of prior citations within the team	=1 if have the largest h-index within the team
	(1)	(2)	(3)	(4)	(5)	(6)
Treated*Post(t>=1)	-0.084***	-0.090***	-0.126***	-0.055***	-0.060***	-0.052**
	(0.018)	(0.016)	(0.023)	(0.021)	(0.021)	(0.021)
Author reputation *Treated*Post(t>=1)	0.520***	0.100***	0.278***	0.034	0.040	0.028
	(0.153)	(0.024)	(0.071)	(0.028)	(0.028)	(0.028)
Paper Fixed Effects	Y	Y	Y	Y	Y	Y
Year Since Publication Dummies	Y	Y	Y	Y	Y	Y
Observations	1,019,509	1,019,509	1,019,509	1,019,509	1,019,509	1,019,509
Number of unique papers	73,145	73,145	73,145	73,145	73,145	73,145

When author reputation is measured in absolute terms, it refers to the absolute reputation of a treated author at the time of retraction. Every paper in the same treatment-control group has the same value on author reputation. Total number of prior papers is measured in 1,000; total number of prior citations in measured in 10,000 and H-index is measured in 100. When author reputation is measured in relative terms, it refers to a dummy equal to one if the treated author has the highest reputation within the team. All regressions report coefficients from maximum likelihood estimation of a Poisson count model, errors clustered by each treatment-control group. All regressions include the same one-way and two-way interactions terms as in Table 2, we do not report their coefficients for illustration purpose only. Standard errors in parentheses, ***p<0.01, **p<0.05, *p<0.1.

Table A2: Effect of retraction on citation of prior work, OLS

	Absolute Reputation			Relative Reputation		
Author reputation measures	Total # of prior work	Total # of prior citations	h-index	=1 if have the largest # of prior work within the team	=1 if have the largest # of prior citations within the team	=1 if have the largest h-index within the team
	(1)	(2)	(3)	(4)	(5)	(6)
Treated*Post(t>=1)	-0.162***	-0.152***	-0.207***	-0.191***	-0.190***	-0.193***
	(0.043)	(0.031)	(0.048)	(0.058)	(0.054)	(0.056)
Papers*Treated*Post(t>=1)	0.695**	0.137***	0.388***	0.134**	0.136**	0.138**
	(0.272)	(0.043)	(0.130)	(0.063)	(0.060)	(0.062)
Paper Fixed Effects	Y	Y	Y	Y	Y	Y
Year Since Publication Dummies	Y	Y	Y	Y	Y	Y
Observations	578,025	578,025	578,025	578,025	578,025	578,025
R-squared	0.163	0.166	0.164	0.163	0.163	0.163
Number of unique papers	51,948	51,948	51,948	51,948	51,948	51,948

When author reputation is measured in absolute terms, it refers to the absolute reputation of a treated author at the time of retraction. Every paper in the same treatment-control group has the same value on author reputation. Total number of prior papers is measured in 1,000; total number of prior citations in measured in 10,000 and H-index is measured in 100. When author reputation is measured in relative terms, it refers to a dummy equal to one if the treated author has the highest reputation within the team. All regressions report OLS coefficients, errors clustered by each treatment-control group. All regressions include the same one-way and two-way interactions terms as in Table 2, we do not report their coefficients for illustration purpose only. Standard errors in parentheses, ***p<0.01, **p<0.05, *p<0.1.

Table A3: Effect of retraction on citation of prior work, excluding treated papers that had zero citation in the year before retraction

	Absolute Reputation			Relative Reputation		
Author reputation measures	Total # of prior work	Total # of prior citations	h-index	=1 if have the largest # of prior work within the team	=1 if have the largest # of prior citations within the team	=1 if have the largest h-index within the team
	(1)	(2)	(3)	(4)	(5)	(6)
Treated*Post(t>=1)	-0.106***	-0.111***	-0.140***	-0.123***	-0.125***	-0.126***
	(0.023)	(0.020)	(0.027)	(0.025)	(0.025)	(0.026)
Papers*Treated*Post(t>=1)	0.390**	0.084***	0.230***	0.079**	0.082***	0.082**
	(0.170)	(0.024)	(0.070)	(0.032)	(0.032)	(0.032)
Paper Fixed Effects	Y	Y	Y	Y	Y	Y
Year Since Publication Dummies	Y	Y	Y	Y	Y	Y
Observations	313,353	313,353	313,353	313,353	313,353	313,353
Number of unique papers	30,300	30,300	30,300	30,300	30,300	30,300

When author reputation is measured in absolute terms, it refers to the absolute reputation of a treated author at the time of retraction. Every paper in the same treatment-control group has the same value on author reputation. Total number of prior papers is measured in 1,000; total number of prior citations in measured in 10,000 and H-index is measured in 100. When author reputation is measured in relative terms, it refers to a dummy equal to one if the treated author has the highest reputation within the team. All regressions report coefficients from maximum likelihood estimation of a Poisson count model, errors clustered by each treatment-control group. All regressions include the same one-way and two-way interactions terms as in Table 2, we do not report their coefficients for illustration purpose only. Standard errors in parentheses, ***p<0.01, **p<0.05, *p<0.1.

Table A4: Effect of retraction on citation of prior work, excluding treated papers published within three years before retraction

	Absolute Reputation			Relative Reputation		
Author reputation measures	Total # of prior work	Total # of prior citations	h-index	=1 if have the largest # of prior work within the team	=1 if have the largest # of prior citations within the team	=1 if have the largest h-index within the team
	(1)	(2)	(3)	(4)	(5)	(6)
Treated*Post(t>=1)	-0.181***	-0.158***	-0.192***	-0.228***	-0.267***	-0.221***
	(0.046)	(0.040)	(0.052)	(0.054)	(0.054)	(0.055)
Papers*Treated*Post(t>=1)	0.786**	0.111***	0.280**	0.158**	0.212***	0.147**
	(0.314)	(0.041)	(0.121)	(0.066)	(0.066)	(0.067)
Paper Fixed Effects	Y	Y	Y	Y	Y	Y
Year Since Publication Dummies	Y	Y	Y	Y	Y	Y
Observations	163,865	163,865	163,865	163,865	163,865	163,865
Number of unique papers	13,040	13,040	13,040	13,040	13,040	13,040

When author reputation is measured in absolute terms, it refers to the absolute reputation of a treated author at the time of retraction. Every paper in the same treatment-control group has the same value on author reputation. Total number of prior papers is measured in 1,000; total number of prior citations in measured in 10,000 and H-index is measured in 100. When author reputation is measured in relative terms, it refers to a dummy equal to one if the treated author has the highest reputation within the team. All regressions report coefficients from maximum likelihood estimation of a Poisson count model, errors clustered by each treatment-control group. All regressions include the same one-way and two-way interactions terms as in Table 2, we do not report their coefficients for illustration purpose only. Standard errors in parentheses, ***p<0.01, **p<0.05, *p<0.1.