Centre de Recherche en Économie et Management
Center for Research in Economics and Management

University of Rennes 1

University of Caen Normandie

Working Paper

Cheat or Perish?
A Theory of Scientific Customs

Benoît Le Maux
CREM-CNRS and Condorcet Center, University of Rennes 1, France

Sarah Necker
University of Freiburg, Walter-Eucken Institute, Deutschland

Yvon Rocaboy
CREM-CNRS and Condorcet Center, University of Rennes 1, France

November 2016 - WP 2016-17

UNIVERSITÉ CAEN NORMANDIE

UNIVERSITÉ DE RENNES 1

Cheat or Perish? A Theory of Scientific Customs*

Benoît LE MAUX
benoit.le-maux@univ-rennes1.fr
University of Rennes 1, CREM-CNRS, Condorcet Center

Sarah NECKER
necker@eucken.de
University of Freiburg, Walter-Eucken Institute

Yvon ROCABOY
yvon.rocaboy@univ-rennes1.fr
University of Rennes 1, CREM-CNRS, Condorcet Center

December 1, 2016

Abstract

We develop a theory of the evolution of scientific misbehavior. Our empirical analysis of a survey of scientific misbehavior in economics suggests that researchers' disutility from cheating varies with the expected fraction of colleagues who cheat. This observation is central to our theory. We develop a one-principal multi-agent framework in which a research institution aims to reward scientific productivity at minimum cost. As the social norm is determined endogenously, performance-related pay may not only increase cheating in the short run but can also make cheating increasingly attractive in the long run. The optimal contract thus depends on the dynamics of scientific norms. The premium on scientific productivity should be higher when the transmission of scientific norms across generations is lower (low marginal peer pressure) or the principal cares little about the future (has a high discount rate). Under certain conditions, a greater probability of detection also increases the optimal productivity premium.

JEL-Classification: A11, A13, K42.
Keywords: Economics of Science, Contract Theory, Scientific Misbehavior, Social Norms.

*The authors thank Roland Kirstein, David Masclet and Fabien Moizeau for helpful suggestions. Earlier versions of this paper were presented at the University of Magdeburg, 2014; the European Public Choice Society Meeting, Groningen, 2015; the Max Planck Institute Bonn, 2015; The Spring Meeting of Young Economists, Ghent, 2015; the Beyond Basic Questions Workshop, Hannover, 2015, the University of Lisboa, 2015; the University of Minho, 2015; and the University of Hamburg, 2016.

1. Introduction

Scientific misbehavior imposes high costs on society. It has been estimated that every paper retracted due to scientific misbehavior costs almost $400,000 in funds from the US National Institutes of Health, for a total of $58 million between 1992 and 2012 (Stern et al., 2014). These investments are at best wasted; misbehavior also misleads and delays scientific progress. To minimize scientific misbehavior and the resulting costs, it is fundamental to understand to what extent scientific misbehavior results from the rewards proposed to researchers, and in particular the incentives to publish. Can rewards increase misbehavior? Do rewards imply not only a short-run rise but also an increasing level of misbehavior over time, by changing the social norms guiding research misbehavior? What is the best policy of research institutions if their rewards induce misbehavior? These are the questions that we address in this paper.

While incentives to publish have a long history in the US and Canada, they have become increasingly important in other countries over the last two decades. Research evaluation agencies have been established (e.g., in the UK and France), bibliometric indicators have been included in the formulae used to allocate research funds (e.g., in Belgium, Denmark, and Portugal), researchers' salaries have been linked to performance (e.g., in Germany, Spain, adn China), and researchers' contract duration has been modified (e.g., in Finland and Italy). The aim of these reforms was to improve the efficiency of public research by increasing the quantity and quality of research output while taking the public budget constraint into account. Franzoni et al. (2011) show that country incentives to publish are indeed positively related to its scientists' (successful) submissions to the journal Science.

Unfortunately, these reforms have also been cited as an explanation for the increasing prevalence of unethical behavior in academia, such as the disproportionately high number of positive findings (Fanelli, 2010) and statistically-significant results (Brodeur et al., 2016), and articles being retracted due to fraud (Fang et al., 2012).[1] The idea is analogous to the traditional economic model of crime, in which unethical behavior rises with the associated monetary benefit (Becker, 1968). More specifically, a growing body of literature has shown that competitive pressure is positively related to cheating (e.g., Schwieren and Weichselbaumer, 2010, Gilpatric, 2011, Cartwright and Menezes, 2014). Necker (2014) provides evidence for the same link in academia.

However, the previous literature is static in nature, and does not investigate the dynamics of scientific cheating. As such it may miss important incentive effects. Recent research has suggested that moral costs (the costs arising from the desire to do the right thing) are at least as important for honest behavior as financial incentives (e.g., Mazar et al., 2008a; Gneezy et al., 2013). As in the theory of social customs, a deviation from a social norm by a few people can produce the erosion of the norm in the long run (e.g., Akerlof 1980, Corneo 1995, Lindbeck et al. 1999, Fischer and Huddart 2008). This idea is central to our work here. If rewards influence the perception of what is morally acceptable in the scientific community,

[1] We define unethical behavior as actions that are "either illegal or morally unacceptable to the larger community" (Jones, 1991, p. 367).

misbehavior may increase over time.

To our knowledge, the existing literature has not looked at the extent to which remuneration policies influence social customs, and how contracts should be designed to steer the social custom towards honest behavior. An exception is Fischer and Huddart (2008), but their analysis does not consider the evolution of the optimal contract over time. We fill this gap by proposing a dynamic principal multi-agent model of scientific misbehavior. Academia is a particularly interesting example for analyzing the link between rewards, norms, and misbehavior. Although employers (government, universities or research centers) can provide incentives for performance, they cannot directly control monitoring. The design of the reward policy is therefore fundamental.

In our model, researchers (the agents) are heterogeneous in terms of labor productivity. The research institution (the principal) aims to reward performance at the lowest possible cost. The principal, however, cannot observe researchers' productivity. For less-productive researchers, this asymmetric information creates incentives to cheat in order to obtain the same rewards as more-productive researchers. Following theories of social customs, a central assumption of our model is that the researcher's decision to cheat depends on the fraction of colleagues who cheat. We back this assumption up with evidence from a survey of economists. Our empirical analysis suggests that cheating is negatively related to utility; however, the disutility of cheating falls with beliefs over the fraction of colleagues who cheat.

Our theoretical model reveals that the introduction of a premium on research productivity (performance-related pay) not only leads to fraudulent behavior in the short run, but may also change the norms of the scientific community in the long run (towards a "cheat or perish" culture). Research institutions have to bear this in mind when introducing productivity premia. The optimal contract depends on the dynamics of scientific norms (fraud). The optimal scientific productivity premium is higher when there is less transmission of scientific norms across generations (lower marginal peer pressure) or the principal cares little about the future (has a high discount rate). Under certain conditions, a higher detection probability (so that the expected reward from fraud is lower) also increases the optimal productivity premium. We emphasize that a productivity premium, by fostering scientific fraud, may produce negative externalities on scientific journals and their editors, while more effective peer review creates positive externalities for the principal. The two actors should hence cooperate.

Our work contributes to the economics literature in a number of ways. First, we contribute to the literature on the evolution of social customs. To our knowledge, we are the first to model the principal-agent relationship in a dynamic setting in which behavior is contagious. We not only investigate the long-run equilibrium, as do, e.g., Akerlof (1980), Corneo (1995) and more recently Fischer and Huddart (2008), but also the convergence process during which agents modify their decision to obey the code of good conduct. Second, only a few pieces of work have theoretically analyzed researcher misbehavior. None of those considers the principal-agent relationship or the moral costs of cheating.[2] Third, in contrast to

[2] The economic theory of scientific fraud in Wible 1998 assumes that researchers optimally split their time between fraudulent and legitimate activities. Hoover (2006) uses a game-theoretic model to show that it is rational to engage in academic plagiarism if the probability of prosecution is low. Lacetera and Zirulia (2009) study malfeasance in the research and publication process through a dynamic game of incomplete information.

much of the related literature on rank-order tournaments (see e.g. Gilpatric, 2011), we allow for agent heterogeneity.

The remainder of the paper is organized as follows. Section 2 discusses the literature emphasizing the importance of the internal costs of cheating, and presents our empirical results. Section 3 develops the theoretical framework and Section 4 simulates the model. Last, Section 5 concludes.

2. The moral costs of cheating

2.1. Traditional theory in the light of recent evidence

The traditional economic theory of crime (proposed by Becker, 1968) focuses on individuals' extrinsic motivations to cheat. Cheating results from a purely rational cost-benefit analysis of the associated expected external rewards and costs. In a tournament, the benefits from cheating increase with the intensity of competition (Gilpatric, 2011). However, recent findings from psychology and behavioral economics indicate that rewards, monitoring, and sanctions have a less clear-cut influence than that assumed in traditional theory (e.g., Rosenbaum et al., 2014). A number of contributions have shown that individuals cheat much less than predicted (e.g., Gneezy, 2005; Mazar et al., 2008b; Gneezy et al., 2013; Abeler et al., 2014). A few people cheat as much as they can, a few people are completely honest, and most people cheat a little. Conrads et al. (2014) find such heterogeneous reactions to competitive incentives. There then seems to be some variation in moral costs, i.e. the costs that arise from the desire to do the right thing.

In line with the theory of social customs, the extent to which others follow (or are believed to follow) the social norm has been shown to play a role. A classic example is given in Cialdini et al. (1990), where individuals' littering in public places depends crucially on the littering behavior of other people. Experimental participants steal more from each other according to others' stealing (Falk and Fischbacher, 2002). Gino et al. (2009) show that observing a confederate cheating is contagious. Abeler et al. (2014) find that participants who believe that others cheat are also more likely to cheat themselves. Bailey et al. (2001) and Necker (2014) provide evidence that economists' beliefs about others' behavior are related to their own admitted misbehavior. The social context is therefore essential for misbehaviour, as stressed in the theory of crime in Funk (2005).

2.2. Evidence on researchers' moral costs

The above literature review suggests that individuals experience moral costs from cheating that vary with the belief about or observation of what others are doing. We here consider whether this also applies to researchers by using a unique survey of economists. In 2010 and 2011, an anonymous online survey was conducted among the members of the European Economic Association (EEA), the German Economic Association, and the mailing lists from

Last, Kiri et al. (2014) show that increasing the benefit from confirmatory results is a way of improving the reliability of scientific research.

Table 1. List of questions/practices

1	Have you ever copied parts from work of others without citing?
2	Have you ever copied from your own previous work without citing?
3	Have you ever refrained from citing results or opinions that are not in line with your own analysis?
4	Have you ever refrained from checking the contents of the works cited?
5	Have you ever suffered from incorrectly being excluded as co-author?
6	Have you ever cited strategically to raise publication prospects (e.g. to please editors or possible referees)?
7	Have you ever refrained from citing work in lower ranked journals, which in a ranking from A+ to C rank lower than A?
8	Have you ever refrained from citing work from other disciplines?
9	Have you ever maximized the number of publications by dividing the work to the smallest publishable unit, meaning several individual articles covering similar topics and differing from each other only slightly?
10	Have you ever complied with suggestions by referees or editors when you thought that they are wrong?
11	Have you ever defined the research question according to data availability?
12	Have you ever excluded part of the data (e.g. outliers) without reporting this?
13	Have you ever corrected data to fit the theory?
14	Have you ever fabricated some data?
15	Have you ever presented empirical findings selectively so that they confirm one's argument
16	Have you ever used tricks to increase t-value, R2, or other statistics?
17	Have you ever searched for control variables until you got the desired results?
18	Have you ever stopped statistical analysis when you had a desired result?

the French Economic Association and the Journées de Microéconomie Appliquée. The survey requested information on economists' norms, (mis-)behavior, professional situation, and life satisfaction. A summary of the responses can be found in Necker (2014). Feld et al. (2014) use the survey to show that the professional situation helps determine economists' satisfaction. This information allows us to see whether reported misbehavior and satisfaction are related, and whether the link depends on beliefs over peer behaviour.

One concern regarding self-reports is that they underestimate the true frequency of misbehavior. However, John et al. (2012) find that the bias seems to be smaller with regard to questionable research practices, and the survey focuses on these kinds of practices. The sample has been shown to be representative of the population.[3] We include all respondents for whom information is available ($n = 934$). To account for item-non response, missing values were replaced via multiple imputation.[4] At the minimum, this survey yields suggestive evidence on researchers' moral costs from misbehavior.

The survey asked economists to assess the justifiability of 18 research practices on a 6-point Likert scale, and to report their own engagement in these practices (see Table 1). We create a misbehavior index M_i based on the answers. We first calculate the average justifiability of each practice. Second, for each respondent, we sum the average justifiability

[3]The response rate was 17% in the first wave (EEA) and 11% in the second (German) and third (French) waves. Survey participants are representative of members of the association in terms of gender and location of workplace. The comparison of responses from early and late respondents, and from participants who continued until the last page and those that dropped out, do not indicate unit-non response bias. A detailed description of the methodology, the representativeness analysis, and the multiple imputation can be found in Necker (2014).

[4]The data are imputed using multiple imputation using chained equations. The methodology is described, e.g., in Cameron and Trivedi (2005).

Table 2. Misbehavior, beliefs and satisfaction[a]

	(1)	(2)	(3)	(4)
Own misbehavior M_i	-0.296***	-0.448*	-0.264**	-0.445*
	(0.106)	(0.235)	(0.108)	(0.239)
Belief B_i	-0.029***	-0.037**	-0.027***	-0.037**
	(0.007)	(0.016)	(0.008)	(0.016)
$M_i * B_i$		0.013		0.015
		(0.020)		(0.020)
Other controls	NO	NO	YES	YES
F	13.7	9.5	3.2	3.2
R2	0.03	0.03	0.12	0.12
N	934	934	934	934

[a] These are OLS coefficients. Standard errors between brackets. All five imputations are used, with the results combined using Rubin's rule. Hypothesis tests are based on robust standard errors. The measures of fit are the lowest statistic among results from the five imputations. Significance levels: $* = 10\%$ $** = 5\%$ $*** = 1\%$.

of the practices that he or she admits to having employed and divide this by the total number of practices.[5] A higher index value indicates that the respondent admits to having employed a greater number of practices or of less well-accepted practices.

Economists' beliefs regarding others' behavior are captured by responses to the question on what fraction of research in the top journals the respondent believes to be affected by four categories of misbehavior (e.g., "incorrect application of empirical methods"). Respondents answered on a 10-point scale from "up to 10%" to "up to 100%." We calculate a variable B_i as the sum of these four responses (which we treat as continuous): a higher value of B_i implies a greater perceived prevalence of misbehavior. The satisfaction question is "Generally speaking, how satisfied are you with the life you lead?" Responses were given on a 6-point Likert scale ("highly dissatisfied" to "highly satisfied"). In line with the literature, reported satisfaction is employed as a measure of utility U_i (e.g., Layard et al. (2008)).

To see whether researchers' utility U_i is affected by misbehavior M_i, and whether this link depends on the fraction of researchers who are believed to cheat B_i, we estimate the following ordinary least squares (OLS) equation. Much of the happiness literature treats the dependent variable as cardinal and uses OLS (Ferrer-i Carbonell and Frijters, 2004).

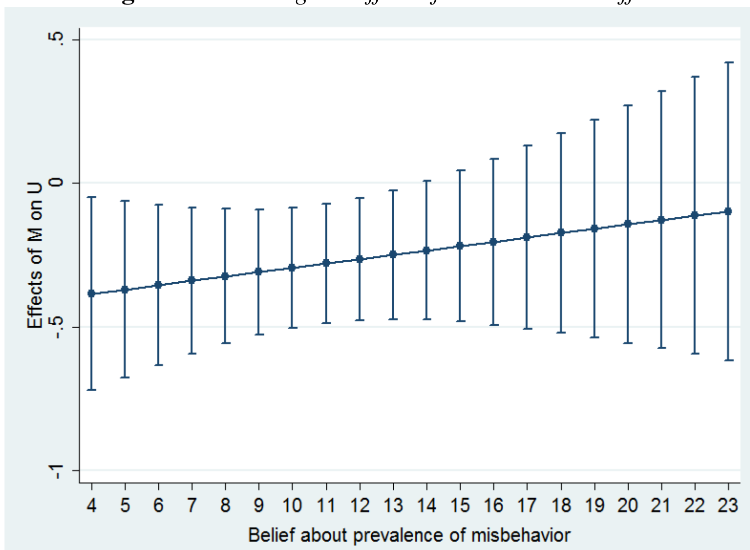
$$U_i = \gamma_0 + \gamma_1 M_i + \gamma_2 B_i + \gamma_3 M_i \times B_i + \mathbf{x}_i \gamma_4 + \varepsilon_i. \tag{1}$$

The interaction $M_i \times B_i$ captures whether the effect of misbehavior on satisfaction is lower when others are believed to misbehave more. We include gender, year of birth, citizenship, academic position, and whether the respondent has tenure or not in the controls, $\mathbf{x}_i$. We estimate the model with and without the interaction $M_i \times B_i$ and with and without the controls.

Table 2 shows the regression results. Own misbehavior M_i is negatively correlated with satisfaction. A one standard-deviation (SD) rise in this index reduces satisfaction by 1/10 of a SD (column (1)). This effect is largely unchanged when control variables are included (column (3)). Although these results cannot prove causality, they are consistent

[5] The summary statistics of the variables used appear in Necker (2014).

Figure 1. *The marginal effect of the interaction effect*



Note: The figures are calculated from the regression shown in column (4). These are average marginal effects $\frac{\partial U_i}{\partial M_i} = \gamma_1 + \gamma_3 B_i$ with 95%-confidence intervals at different values of B_i.

with individuals facing moral costs from cheating. While the relationship seems small in size, we should remember that participants report their overall satisfaction with life. The belief that misbehavior is widespread is also negatively related to satisfaction. A rise in this belief by one SD implies a fall in satisfaction of 1/7 of a SD.

The second and fourth columns present the estimates including the interaction term $M_i \times B_i$. The coefficient on this interaction is insignificant. However, Brambor et al. (2006) stress that the standard error of interest is not the standard error of γ_3 but that on the value of $\frac{\partial U_i}{\partial M_i} = \gamma_1 + \gamma_3 B_i$. We thus calculate the derivative and its standard error at different levels of B_i: the results are shown in Figure 1. These indicate that, for instance, respondents who believe that on average up to 10% of the research in journals is subject to misbehavior ($B_i = 4$) experience a utility loss of 0.4. By way of contrast, respondents who believe that on average up to 30% of the research in journals is subject to misbehavior ($B_i = 12$) experience a utility loss of 0.3. The effect of M_i on U_i is insignificant, i.e., no significant utility loss, for economists who believe that there is a high prevalence of misbehavior ($B_i \geq 15$). The results are consistent with the notion that researchers' moral costs of cheating vary with their beliefs about what their peers do.

3. A dynamic theory of optimal scientific misconduct

A fundamental goal in the reward structure of science is to establish priority of discovery by being the first to communicate an advance in knowledge. Publication is a necessary step in establishing this priority and gaining recognition from the scientific community (Stephan, 1996). The number of published articles and the ranking of the journals in which articles are published are thus outcomes of interest for the researcher. The central question here is whether these articles will be produced in a fraudulent way. The considerations in Section 2 emphasize the importance of including three different types of motivations in our theoretical model.

1. *Extrinsic motivation*. In our setting, this corresponds to the contracts offered to researchers. In line with Becker (1968), promotion, tenure, and compensation based on publication are assumed to provide incentives to misbehave.

2. *Unconditional (intrinsic) commitment to honesty*. We assume that agents are heterogeneous in the utility loss they experience from dishonestly reporting research. A similar assumption is made by Funk (2005) in her theory of crime and by Fischer and Huddart (2008) in their theory of optimal contracting.

3. *Conditional (norm-based) commitment to honesty*. We assume that individuals' utility is affected by others' obedience to a code of good conduct (Akerlof, 1980, Corneo, 1995). As a result of this conditional honesty, performance incentives may have multiplier effects.

3.1. The dynamics of scientific cheating

Consider a scientific community in which an infinite number of researchers (the *agents* in our model) are divided up into identical non-overlapping generations, each living for one period t. The agents are heterogeneous in their job performance θ (see also Lacetera and Zirulia, 2012), having either low θ^L ("normal researchers") or high θ^H ("leading researchers") scientific productivity, with $\theta^L < \theta^H$. These differences may reflect multiple factors: ability, teaching and administrative duties, research budgets, spillover effects due to the presence of colleagues who are specialized in the field, or family life. The distribution of θ is characterized by:

$$p = \mathbb{P}(\theta = \theta^L) \in (0,1) \ \text{ and } \ 1 - p = \mathbb{P}(\theta = \theta^H), \tag{2}$$

Researchers are hired by a single principal, e.g. the government, a university, or a research institute.[6] The principal aims to reward productivity by paying researchers with low productivity y_t^L and researchers with high productivity y_t^H, with $y^H > y^L$. High-type researchers receive higher remuneration, which can take the form of a higher salary or of rewards, prizes, and other monetary compensations. The core of the problem here is that the principal cannot observe agents' productivity θ, although it is common knowledge that θ takes the value of θ^L with probability p. In other words, the principal only observes output (e.g., publications) without knowing whether these were produced unethically.

We distinguish between researcher utility functions u^L and u^H. However, these functions are similar, depending only on lifetime income y_t and being concave (so that individuals are risk-averse). We assume that $u(0) = 0$. The agents will accept a contract offer only if their utility there is greater than their reservation utility, denoted by $u^L_{\min}$ and $u^H_{\min}$. Researchers with

[6] One justification for the focus on one principal is that it reflects a national government's choice of policies that provide incentives for researchers to publish, as described in the introduction or in Franzoni et al. (2011). These incentives affect every researcher in the country in a similar way.

a contract offer that produces utility lower than the reservation level will leave the scientific community (the participation constraint). We have $u^L_{\min} < u^H_{\min}$ and we let $y^k_{\min}$, $k = L,H$, denote the reservation income such that $u^k(y^k_{\min}) = u^k_{\min}$.

We assume that there is a code of good conduct in the scientific community. High-type researchers reach high levels of output, earning y^H_t, without having to infringe this code. In contrast, due to information asymmetries, low-type researchers can improve their publication record by not obeying the code. Since our interest is in cheating, which is by assumption only carried out by low-type researchers, we focus on them in the following.[7]

Low-type researchers choose an action $a \in A = \{\text{Obey}, \text{Disobey}\}$. If they obey the code, they receive y^L_t. If they disobey, it is impossible for the principal to distinguish low-type from high-type agents. The agent receives y^H_t. We assume that the growing demand for publication slots is satisfied by a growing supply of publication slots/journals.[8]

All researchers are audited with certainty but this auditing process is imperfect, with a probability π of being detected when cheating. If a low-type researcher obeys the code, he/she is not punished and receives y^L_t. If a low-type disobeys and is not detected, he/she obtains the salary of the high-type researcher y^H_t without being punished. However, if the misbehavior is detected, he/she is punished by a sanction s_t, which reduces his/her lifetime income.

Researchers face moral costs. These vary by the individual's commitment to honesty (integrity), denoted by α^i hereafter, which is independent of what others are doing and is private information. This has a continuous uniform distribution on the unit interval I, i.e. $F(\alpha) = \alpha$. In other words, α percent of researchers have an integrity coefficient less than or equal to α.

Moral costs are also conditional on peer behavior. Past behaviors, whether they are observed or not, create a working atmosphere in which agents form certain attitudes of approval or disapproval, affecting the disutility of cheating agents. The fraction of low-productivity researchers who cheated in the past period is denoted by β_{t-1}. The peer-pressure function v translates past cheating β_{t-1} into peer pressure: $v(\beta_{t-1})$. We assume $v > 0$ and $v' = -\mu$ with $\mu \geq 0$, i.e. peer pressure is a constantly decreasing function of the share of past cheaters. The moral costs of cheating then fall with the fraction of cheaters in the past period. A higher value of the scalar μ (the marginal peer pressure) implies that v decreases more steeply with β_{t-1}, which increases the utility from cheating more.

Agent i will not cheat if the utility from choosing $a =$ Obey (the left-hand side of equation 3) is larger than the expected utility from choosing $a =$ Disobey weighted by the

[7] Schwieren and Weichselbaumer (2010) show experimentally that less-able individuals cheat more. Charness et al. (2013) find that experimental participants who have lower rank in the performance distribution cheat more.

[8] Although acceptance rates are falling at the top journals, the number of available publication slots has increased extensively. Between 1997 and 2006, the number of journal articles covered by the Science Citation Index (SCI) increased by 2.2% per year. The number of journals covered in the extended SCI increased from 5,467 in 1998 to 8,060 in 2009 (from 600 in 1964). The increase in the number of journals may even be underestimated, as the SCI covers a falling proportion of the traditional scientific literature (Larsen and Von Ins, 2010).

unconditional (α) and conditional (β) commitment to honesty (the right-hand side)[9]:

$$u^L(y_t^L) \geq \frac{(1-\alpha^i)}{v(\beta_{t-1})} \times \left[\pi u^L(y_t^L - s_t) + (1-\pi)u^L(y_t^H)\right], \tag{3}$$

Equation (3) shows the incentive-compatibility constraint, in which peer pressure and the unconditional commitment to honesty (α_i) interact. Since low-type agents vary only with respect to the integrity coefficient, α will separate those who misbehave from those who do not. Equation (3) can be rewritten as:

$$\alpha^i \geq 1 - \frac{u^L(y_t^L)v(\beta_{t-1})}{\pi u^L(y_t^L - s_t) + (1-\pi)u^L(y_t^H)}, \tag{4}$$

The researchers who refrain from cheating are those with a higher value of the integrity coefficient, as illustrated in Figure 2(a). As researchers are distributed on the unit interval, the share of dishonest scientists among the low-types (β_t) in the short-run is directly given by Equation (4). We have:

$$\beta_t = 1 - \frac{u^L(y_t^L)v(\beta_{t-1})}{\pi u^L(y_t^L - s_t) + (1-\pi)u^L(y_t^H)}. \tag{5}$$

The fraction of cheaters will fall with y_t^L, π and s_t, but rise with y_t^H, as stated in Proposition 1. These are the predictions of the traditional economic model of crime. The condition $v' \leq 0$ implies that β_t rises with β_{t-1}.

Proposition 1. *The model leads to the following comparative-static derivatives:*

$$\frac{\partial \beta_t}{\partial \beta_{t-1}} \geq 0, \quad \frac{\partial \beta_t}{\partial y_t^L} \leq 0, \quad \frac{\partial \beta_t}{\partial y_t^H} \geq 0, \quad \frac{\partial \beta_t}{\partial \pi} \leq 0, \quad \frac{\partial \beta_t}{\partial s_t} \leq 0.$$

The influence of β_{t-1} on β_t is central in our theory, as it determines the extent to which the $t-1$ generation affects that at period t. One important result is that the marginal impact of β_{t-1} depends on the contract, as shown in the following equation:

$$\frac{\partial \beta_t}{\partial \beta_{t-1}} = \frac{u^L(y_t^L)}{\pi u^L(y_t^L - s_t) + (1-\pi)u^L(y_t^H)} \times -v'(\beta_{t-1}) \tag{6}$$

Past social customs will matter more for current social customs as (1) marginal peer pressure is higher and (2) current incentives to cheat are lower. The higher the utility from honesty, $u^L(y_t^L)$, or the lower the expected utility from cheating, $\pi u^L(y_t^H - s_t) + (1-\pi)u^L(y_t^H)$, the

[9] It is possible that researchers derive disutility from others' cheating even if they do not cheat themselves, as suggested by our empirical results. If we consider that the norm $(1-\beta_t)$ is a public good from which agents derive utility, β_t would appear on both sides of the equation in the multiplicative utility function, thereby canceling each other out. Another possibility is to assume that disutility falls with the share of deviators in the case of an additively-separable utility function (as in Lindbeck et al., 1999). We consider a multiplicative function for the main reason that it allows us to assume that even the most amoral researcher (with $\alpha = 0$) is affected by peer pressure. This implies that $\beta_t = 0$ is a potential long-run equilibrium (if peer pressure is strong).

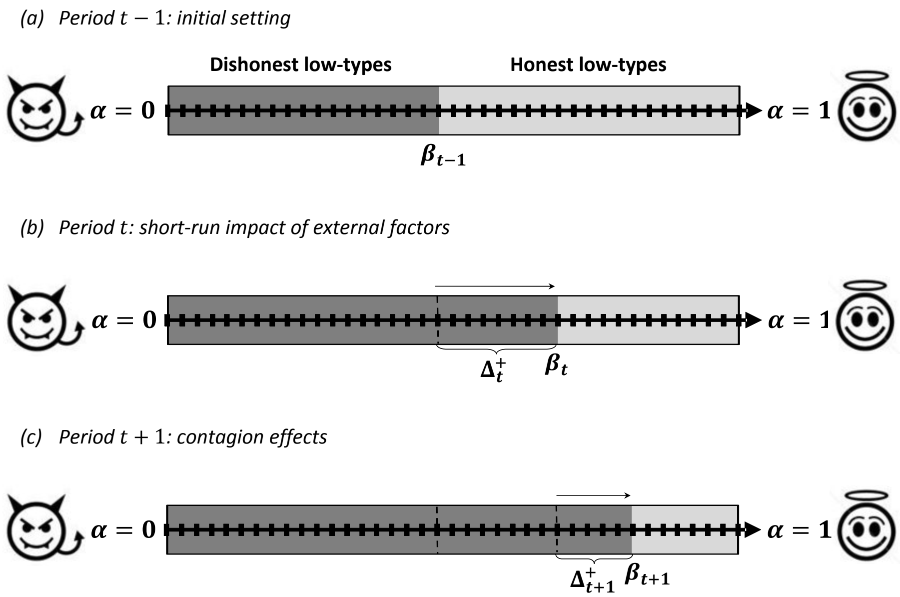


Figure 2. *The link between the contract, individual morality and the social norm.*

greater is the impact of β_{t-1} on β_t. One possible interpretation is that people pay more (less) attention to norms when current incentives encourage (discourage) honesty.

A key feature of the model is that a change in any of the variables y_t^L, y_t^H and s_t or the parameter π affects not only β_t but also β_{t+1}, i.e. the level of future fraud:

$$\frac{\partial \beta_{t+1}}{\partial(.)} = \underbrace{\frac{\partial \beta_t}{\partial(.)}}_{\text{initial}} \times \underbrace{\frac{\partial \beta_{t+1}}{\partial \beta_t}}_{\text{contagion}} \tag{7}$$

Equation (7) shows that the derivative consists of an initial and a contagion effect, as illustrated in Figures 2(b) and 2(c). Greater publication incentives not only encourage fraud at t but also reduce the conditional commitment to honesty at $t+1$. The latter effect can be interpreted as the "crowding out" of social norms by external incentives. As we will see below, performance-related pay can have multiplier effects. Publication pressure not only leads to fraudulent behavior in the short run but can also make cheating increasingly attractive in the long run.

3.2. The optimal remuneration of researchers in the short run

We describe the contracts by adopting the following time line within each period: (1) β_{t-1} is known; (2) researchers learn their type (productivity θ and integrity α); (3) the principal designs the salary policy; (4) each researcher accepts or rejects a contract (i.e. decides to fraud or not); and (5) researchers who reject both contracts receive their reservation utility. We assume that in each period the principal cannot withdraw the offer once it is accepted. Any withdrawal would destroy the reputation for honouring agreements.

Table 3. Components of the Principal's cost function

	Percentage of low-type researchers: p		Percentage of high-type researchers: $(1-p)$
	Percentage of dishonest researchers: β_t	Percentage of honest researchers: $(1-\beta_t)$	
Probability of being discovered: π	$y_t^H - s_t$	y_t^L	y_t^H
Probability of not being discovered: $(1-\pi)$	y_t^H	y_t^L	y_t^H

The principal aims to identify the remuneration policy or remuneration contracts $(y_t^{L*}, y_t^{H*}, s_t^*)$ that reward differences in productivity at the lowest possible expected cost per capita (denoted ec_t), given the incentive-compatibility and participation constraints.[10] Expected costs per capita are given by the following expression:

$$\begin{aligned} ec_t &= \pi p \left[\beta_t \times (y_t^H - s_t) + (1-\beta_t) \times y_t^L\right] \\ &\quad + (1-\pi) p \left[\beta_t \times y_t^H + (1-\beta_t) \times y_t^L\right] + (1-p) \times y_t^H \end{aligned} \tag{8}$$

Table 3 sets out the components of the principal's cost function. The principal bears costs of $(y_t^H - s_t)$ if the researcher is a cheating low-type and the misconduct is detected, i.e. with a probability of $[\pi \times p \times \beta_t]$. The costs are y_t^H in the case of a non-detected cheating low-type, with a probability of $[(1-\pi) \times p \times \beta_t]$, or in the case of a high-type researcher, with a probability of $[1-p]$. The costs are y_t^L in the other states of nature.

The decision to pay based on publication implies that the principal cares about performance. The principal cares about fraud if it affects costs, but not about fraud per se. This assumption is motivated by the observation that universities and research centers take the credit for publications, but are generally not held liable for inaccuracies in the results. Moreover, better performance may allow the principal to obtain a higher operating budget. Performance and the related budget are thus a realistic objective for the principal.

A central assumption is that the probability of detecting misconduct cannot be controlled by the principal, i.e., it is exogenous. In academia, quality control is exercised by peer review. An important aim of peer review is to discourage and detect fraud (Stephan, 2012). The principal can send scientific work to qualified peers to assess whether the work is non-fraudulent. However, the reliability of the peer review process cannot be controlled by the principal. The optimal remuneration policy is thus given by:

$$\begin{aligned} \min_{\{y_t^L, y_t^H, s_t\}} \quad & ec_t = ec_t(y_t^L, y_t^H, s_t) \\ \text{s.t.} \quad & u^L(y_t^L) \geq u_{\min}^L, \\ & u^H(y_t^H) \geq y_{\min}^H, \end{aligned} \tag{9}$$

Disobedience can be eliminated by setting $y_t^L = y_t^H = \tilde{y}$, i.e. a contract without any performance-based incentives. However, if $\tilde{y}$ is too low, high-type researchers will leave the

[10]Note that, by construction, we focus on expected cost per capita due to the infinite number of agents.

scientific community. This is a standard adverse-selection problem. On the other hand, if $\tilde{y}$ is set such that $u^H(\tilde{y}) \geq u^H_{\min}$, i.e. such that high-types accept the contract, the remuneration scheme will be costly to the principal. Were information about researcher type to be perfectly known by the principal, she/he would give $y^H_{\min}$ to the high-types and $y^L_{\min}$ to the low-types. However, information is asymmetric.

An important feature of the model is the nonlinear relationship between low-type income (y^L_t) and expected costs per capita (ec_t). At low values of y^L_t, the incentives to cheat are high for low-types, which implies that the principal has to pay y^H_t to many researchers, producing high ec_t. If the principal increases y^L_t, disobedience is reduced. However, ec_t increases due to the increase in y^L_t. Hence, there exists an optimal level of remuneration (of fraud) at which the differences in productivity are rewarded at minimum cost.

To solve the optimization problem, the principal first chooses the lowest value of y^H_t such that the participation constraint of the high-types is satisfied, i.e., $y^{H*}_t = y^H_{\min}$. Second, given that punishment is free, the sanction must be as high as possible. By assumption, therefore, $s^*_t = y^H_{\mathrm{t}}$, which is the well-known penalty “à la Becker.” The optimization problem for an interior solution thus becomes:

$$\begin{array}{ll} \min\limits_{\{y^L_t\}} & ec_t = [p(1-\beta_t)] \times y^L_t + [p(1-\pi)\beta_t(y^L_t) + (1-p)] \times y^H_{\min} \\ \text{where} & \beta_t = 1 - \frac{v(\beta_{t-1})u^L(y^L_t)}{(1-\pi)u^H_{\min}}. \end{array} \tag{10}$$

The optimal income of low-types (y^{L*}_t) is determined by substituting $\beta_t(y^L_t)$ into the expected-cost function ec_t and solving the following first-order condition:

$$u^L(y^{L*}_t) - u^{L'}(y^{L*}_t)((1-\pi)y^H_{\min} - y^{L*}_t) = 0. \tag{11}$$

Interestingly, the probability of being detected, π, and the reservation income of the high-types, $y^H_{\min}$, are the only variables which affect short-run optimal income. Using the implicit-function theorem, we have:

$$\frac{\partial y^{L*}_t}{\partial \pi} < 0, \quad \frac{\partial y^{L*}_t}{\partial y^H_{\min}} > 0. \tag{12}$$

First, there exists a tradeoff between the probability of detection and the optimal value of y^{L*}_t. Second, the higher is the (reservation) income of the high-types, the greater are the incentives to cheat, and the higher should be low-type income.

Once the optimal value of y^{L*}_t is determined, we can calculate the optimal level of fraud β^*_t (from equation (5)) and optimal expected costs ec^*_t (from equation (8)). This allows us to determine the sign of the derivative of ec^*_t with respect to β_{t-1}:

Proposition 2. *From the optimal value of ec^*_t in the short run, we have* $\frac{\partial ec^*_t}{\partial \beta_{t-1}} \geq 0$.

A rise in the share of past cheaters, β_{t-1}, will reduce moral costs, $v(\beta_{t-1})$, increasing the incentive to cheat. As stated previously, in the short run the optimal value of y^{L*}_t is independent of β_{t-1}. From Proposition 1, a rise in β_{t-1} will produce higher β_t and, from equation (8), greater expected costs ec_t^*. This yields Proposition 2.

3.3. The optimal remuneration of researchers in the long-run

A change in remuneration may not only affect short-run expected costs but, due to the effect on moral costs, those in the long run as well. Figure 3 provides an illustration of this phenomenon in a two-period setting. As shown in panel *a*, the principal chooses the optimal contract at point A to minimize expected costs ec_t (equation (11)). This yields the level of fraud in t (equation (5)), as shown in panel *b*. The level of fraud determines future costs ec_{t+1}, as shown in panel *c*, given the remuneration policy at $t+1$ (Proposition 2). In contrast, if the principal chooses contract B, the expected costs in t are higher than the optimal level, as shown in panel *a*. However, the level of fraud is lower (panel *b*), which reduces future expected costs (panel *c*). There is thus a tradeoff between present and future expected costs (panel *d*). In the long-run, it might be optimal to pay a higher wage to low-types, and thus reduce fraud, in order to minimize total expected costs.

The aim of this section is to determine the optimal contract that minimizes total expected cost in a T-period setting. We consider two types of long-run contracts: the first-best and the second-best. In both cases, the principal wishes to identify the optimal remuneration policy $(y_t^L, y_t^H, s_t, \forall t = 1 \dots T)$ that minimizes the present value of expected costs (*PEC*, hereafter). Following the discussion above, the principal sets y_t^H to y_{min}^H and s_t to y_t^L in both types of contracts.

First-best optimal contract. Under this contract the incomes of low-type researchers across all periods minimize the present value of the expected cost of the remuneration policy. Unlike in Section 3.2, the principal takes into account that today's salary has an effect on the future fraction of cheaters, via the social norm. The principal is assumed to be non-myopic and can set a different salary in each t. From equation (10), the expected cost ec_t at period t can be written as a function of y_t^L and β_{t-1}. The optimization problem is

$$\begin{array}{lrcl} \min\limits_{\{y_1^L, y_2^L, \dots, y_T^L\}} & PEC = \sum_{t=1}^T \frac{ec_t(y_t^L, \beta_t(y_t^L, \beta_{t-1}))}{\delta^{t-1}} & & \\ \text{where} & \beta_t(y_t^L, \beta_{t-1}) & = & 1 - \frac{v(\beta_{t-1}) u^L(y_t^L)}{(1-\pi) u_{\min}^H}, \quad t = 1 \dots T, \end{array} \tag{13}$$

and $\delta \geq 1$ is the discount factor. In this setting, the principal chooses at $t = 1$ the contract menu for all periods. The problem is convex. The levels of y_t^L, $t = 1, \dots, T$, are given by the following first-order conditions:

$$\frac{\partial PEC}{\partial y_t^L} = \frac{1}{\delta^{t-1}} \left(\frac{\partial ec_t}{\partial y_t^L} + \frac{\partial ec_t}{\partial \beta_t} \frac{\partial \beta_t}{\partial y_t^L} \right) + \sum_{k=t+1}^{T} \frac{1}{\delta^{k-1}} \frac{\partial ec_k}{\partial \beta_k} \frac{\partial \beta_k}{\partial y_t^L} = 0, \quad t = 1 \dots T. \tag{14}$$

Equation 14 shows that y_t^L has two effects on the present value of the expected costs. The first term represents the effect of y_t^L on present expected cost (via the direct change in the

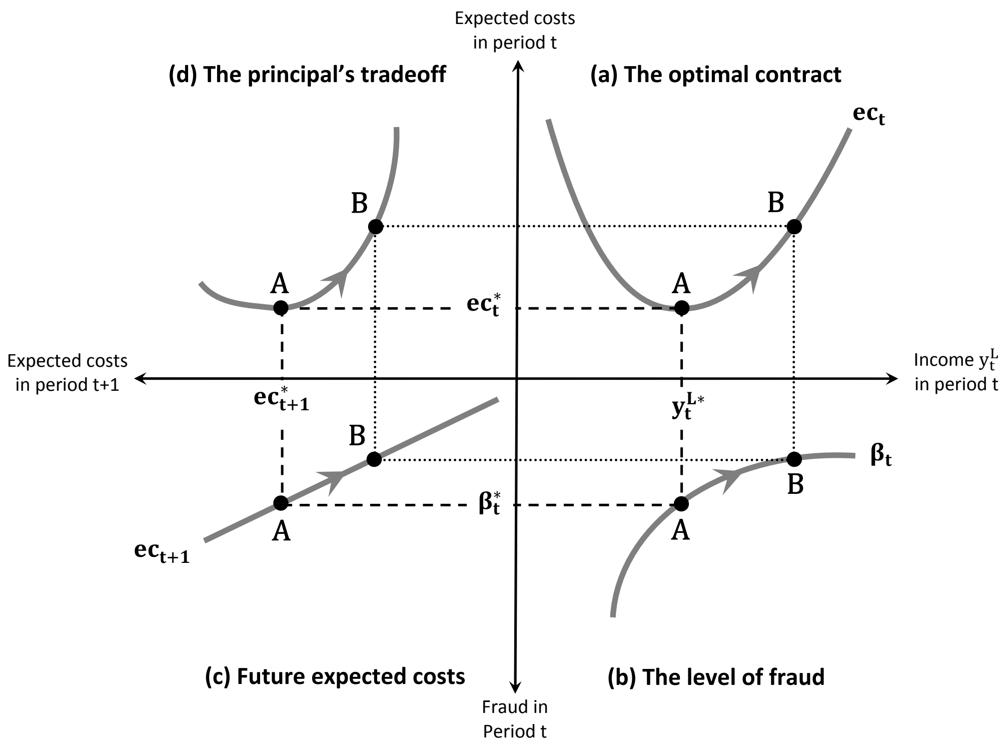


Figure 3. *Optimal contract, misbehavior and consequences for the future*

remuneration cost, and the indirect effect via the present population of cheaters). The second term corresponds to the impact of y_t^L on the expected costs over future periods through the change in social customs. The effect of y_t^L on β_k in equation (14) can be expressed as:

$$\frac{\partial \beta_k}{\partial y_t^L} = \left[\frac{\partial \beta_k}{\partial \beta_{k-1}} \times \frac{\partial \beta_{k-1}}{\partial \beta_{k-2}} \times ... \frac{\partial \beta_{t+1}}{\partial \beta_t}\right] \frac{\partial \beta_t}{\partial y_t^L} = \prod_{m=t+1}^{k} \frac{\partial \beta_m}{\partial \beta_{m-1}} \frac{\partial \beta_t}{\partial y_t^L}, \tag{15}$$

i.e., the income y_t^L chosen in period t will affect the level of fraud β_t, which in turn will affect the level of fraud in $t+1$ and so on. The term $\prod_{m=t+1}^{k} \frac{\partial \beta_m}{\partial \beta_{m-1}}$ is a *multiplier*:

$$\prod_{m=t+1}^{k} \frac{\partial \beta_m}{\partial \beta_{m-1}} = \prod_{m=t+1}^{k} \frac{\mu u^L(y_m^L)}{(1-\pi) u_{\min}^H}. \tag{16}$$

This multiplier rises with marginal peer pressure μ, future incomes y_m^L, and the probability of detection π. While the principal takes μ and π to be exogenous, contagion can be influenced by the choice of y_m^L. The multiplier is directly connected to the contagion effects in equation (7), i.e. the extent to which past generations influence future generations. Equation (14) can now be rewritten as:

$$\begin{aligned} 0 &= \left[u^L(y_t^L) - ((1-\pi)y_{\min}^H - y_t^L) \times u^{L'}(y_t^L)\right] - \\ &\sum_{k=t+1}^{T} \frac{1}{\delta^{k-t}} \left((1-\pi)y_{\min}^H - y_k^L\right) \prod_{m=t+1}^{k} \frac{\mu u^L(y_m^L)}{(1-\pi)u_{\min}^H} u^{L'}(y_t^L), \quad t = 1 \dots T. \end{aligned} \tag{17}$$

The first term is equivalent to the first-order condition for the short-run equilibrium (equation (11)). The second term represents the impact of y^L on the principal's costs in future periods (the contagion effects). The first-best optimal salary is higher than the short-run optimal salary in all periods except for the last, as it brings about lower future costs. If the parameters of the model do not change, the optimal salary falls over time and, in the last period, the optimal salary is set at its short-run optimal level.

Second-best optimal contract. A policy under which income falls over time (at least up to the last period, as in the first-best optimal contract) may not be socially acceptable. We now consider that the principal is not totally free to set the optimal income in each period, and thus has to design a second-best optimal policy. We suppose that the non-myopic principal knows that remuneration affects the future size of the population of cheaters. However, the principal is restricted in designing the policy, since income cannot change over time. The optimization problem is

$$\begin{aligned} &\min_{\{y^L\}} \quad PEC = \textstyle\sum_{t=1}^{T} \frac{ec_t(y^L, \beta_t(y^L, \beta_{t-1}))}{\delta^{t-1}} \\ &\text{where} \quad \beta_t(y^L, \beta_{t-1}) \quad = \quad 1 - \frac{v(\beta_{t-1})u^L(y^L)}{(1-\pi)u_{\min}^H}, \quad t = 1 \dots T. \end{aligned} \tag{18}$$

The first-order condition is:

$$\frac{dPEC}{dy^L} = \sum_{t=1}^{T} \frac{1}{\delta^{t-1}} \left(\frac{\partial ec_t}{\partial y^L} + \frac{\partial ec_t}{\partial \beta_t} \frac{\partial \beta_t}{\partial y^L} \right) + \sum_{t=2}^{T} \frac{1}{\delta^{t-1}} \frac{\partial ec_t}{\partial \beta_t} \sum_{k=1}^{t-1} \frac{\partial \beta_k}{\partial y^L} \prod_{m=k}^{t-1} \frac{\partial \beta_{m+1}}{\partial \beta_m} = 0, \tag{19}$$

which can be rewritten as:

$$\begin{aligned} 0 &= \sum_{t=1}^{T} \frac{1}{\delta^{t-1}} v(\beta_{t-1}) \left[u^L(y^L) - ((1-\pi)y_{\min}^H - y^L) \times u^{L'}(y^L)\right] - \\ &\sum_{t=1}^{T-1} \frac{1}{\delta^t} \sum_{k=1}^{t} \left[\frac{\mu u^L(y^L)}{(1-\pi)u_{\min}^H}\right]^k v(\beta_{t-k}) \left[((1-\pi)y_{\min}^H - y^L) \times u^{L'}(y^L)\right]. \end{aligned} \tag{20}$$

As in the first-best case, the second-best optimal salary is higher than the short-run optimal salary as the principal takes into account that a higher low-type salary reduces the future population of cheaters, and so future costs. However, remuneration does not change over time, implying higher costs for the principal.

Let y_t^{First} and y_t^{Second} denote the optimal incomes of low-type researchers in the first- and second-best contracts respectively. As can be seen from equations (17) and (20), marginal peer pressure, the discount factor and the probability of detection determine the principal's efficient choice under both contracts. Using the implicit-function theorem to calculate the partial derivatives yields the following proposition:

Proposition 3. *The first-best and second-best optimal contracts have the following comparative-static derivatives:*

$$\frac{\partial y_t^{\text{First}}}{\partial \mu} > 0, \quad \frac{\partial y_t^{\text{Second}}}{\partial \mu} > 0, \quad \frac{\partial y_t^{\text{First}}}{\partial \delta} < 0, \quad \frac{\partial y_t^{\text{Second}}}{\partial \delta} < 0, \quad \frac{\partial y_t^{\text{First}}}{\partial \pi} \lesseqgtr 0, \quad \frac{\partial y_t^{\text{Second}}}{\partial \pi} < 0.$$

From Proposition 3, the lower is μ or the higher is δ, the closer the optimal incomes are to that under the short-run contract (until we have $y_t^{\text{First}} = y_t^{\text{Second}} = y_t^{L*}$). On the other hand, as these variables fall the principal has to increase the low-type salary above the level in the short-run contract. We call this a *prevention policy* that aims to reduce fraud in future periods. This is a central finding of our analysis. If $\mu = 0$, the actions of others do not affect the individual decision to cheat. Contagion effects are then absent (the multiplier is zero). Any prevention policy is cost-ineffective since it cannot influence the social norm and implies high remuneration costs. On the other hand, when $\mu > 0$, a change in y_t^L also affects future researchers, making a prevention policy more cost-effective. The higher is μ, the lower the optimal performance premium. The optimal incomes y_t^{First} and y_t^{Second} also fall with the principal's patience. When $\delta = 0$, the principal does not care about about future expected costs and so does not implement a prevention policy. For the second-best optimal contract, there exists a tradeoff between the probability of detection and the optimal value of the low-type salary. However, for the first-best optimal contract, the sign of the derivative with respect to π is ambiguous. Section 4 offers a more detailed analysis of these relationships.

4. The main characteristics of optimal policies

This section identifies the main characteristics of optimal policies (short-run vs. long-run, and first-best vs. second-best) by using simulation results. In contrast to the conventional literature on social norms (e.g., Akerlof 1980 and Corneo 1995), we are not only interested in the long-run equilibrium (Section 4.1) but also, and more importantly, in how fraud evolves over time (Sections 4.2 and 4.3). The dynamics of social norms are important as they affect the principal's expected costs and the choice of the optimal contract.

4.1. The long-run equilibrium level of fraud

Which fraud equilibrium will prevail in the long run when the low-type salary does not change over time, i.e. in the case of the short-run second-best contract? We analyze this question by appealing to the concept of "stationary points" that we define as an equilibrium which continues to hold in each period (see, e.g., Corneo, 1995). Assuming that $y_t^H = y_{\min}^H$, $s_t = y_{\min}^H$, and $v(\beta_{t-1}) = k - \mu\beta_{t-1}$, i.e. constant marginal peer pressure, equation (5) can be rewritten as:

$$\beta_t = 1 - \frac{ku^L(y_t^L)}{(1-\pi)u^L(y_{\min}^H)} + \frac{\mu u^L(y_t^L)}{(1-\pi)u^L(y_{\min}^H)}\beta_{t-1}, \tag{21}$$

which is the equation of a line with positive slope.

The interior stationary value of β is attained when $\beta_t = \beta_{t-1} = \beta$, that is when:

$$\beta = \frac{1 - \frac{u^L(y_t^L)}{(1-\pi)u^L(y_{\min}^H)}k}{1 - \frac{u^L(y_t^L)}{(1-\pi)u^L(y_{\min}^H)}\mu}. \tag{22}$$

By definition, an interior stationary point β is stable if the absolute value of the derivative of β_t with respect to β_{t-1} in equation (21) is strictly less than 1, and unstable if it is strictly greater than 1. This derivative is the slope of the line. Hence, the equilibrium is stable if

$$\mu < \frac{(1-\pi)u^L(y_{\min}^H)}{u^L(y_t^L)} = k_0 \tag{23}$$

In addition, note that, since $\beta \in [0,1]$, two corner stationary equilibria may exist:

$$\beta = 0 \ \text{ if } \ k > \frac{(1-\pi)u^L(y_{\min}^H)}{u^L(y_t^L)} = k_0 \quad \text{and} \quad \beta = 1 \ \text{ if } \ k > \mu. \tag{24}$$

Figure 4 shows the possible sets of equilibria for different values of the peer-pressure function k and μ. Equation (21) is depicted by the black lines. On these equilibrium curves, β_t is only a function of β_{t-1}, and all other variables are held constant. The model parameters were chosen for graphical convenience: $y_{\min}^H = 100$, $y_t^L = 37.5$, $u(y) = y^{0.6}$, $\pi = 0.05$, and $p = 0.5$. The intersections with the 45-degree line indicate the interior stationary point (equation (22)).

If $\mu = 0$, or equivalently $v(\beta_{t-1}) = k$, the moral costs of cheating are independent of past cheating. The slope is zero and the equilibrium level of fraud is reached immediately. This is shown in Figure 4a. From equation (21), fraud is equal to $1 - ku^L(y_t^L)/[(1-\pi)u^L(y_{\min}^H)]$. Fraud is then zero only if k exceeds the following value:

$$k_0 = \frac{(1-\pi)u^L(y_{\min}^H)}{u^L(y_t^L)} \tag{25}$$

For instance, in Figure 4a, $\beta = 0$ holds for $k_0 = 1.7$. When k is lower, the moral costs exogenously fall for the whole scientific community and the level of fraud becomes non-zero.

When $0 < \mu < k_0$ the slope of the line is less than one. If $\mu < k < k_0$ (Figure 4b), there is a unique stable interior solution at the intersection of the 45-degree line and the equilibrium curve. From equation (24), the conditions that $\mu < k$ and $k < k_0$ imply that the stationary point is less than 1 and greater than 0, respectively, i.e. we obtain an interior solution. On the other hand, if $k > k_0$ (Figure 4c), the unique stable equilibrium is zero.

Last, if $0 < k_0 < \mu$ the slope of the line is over one. Two corner equilibria coexist (see

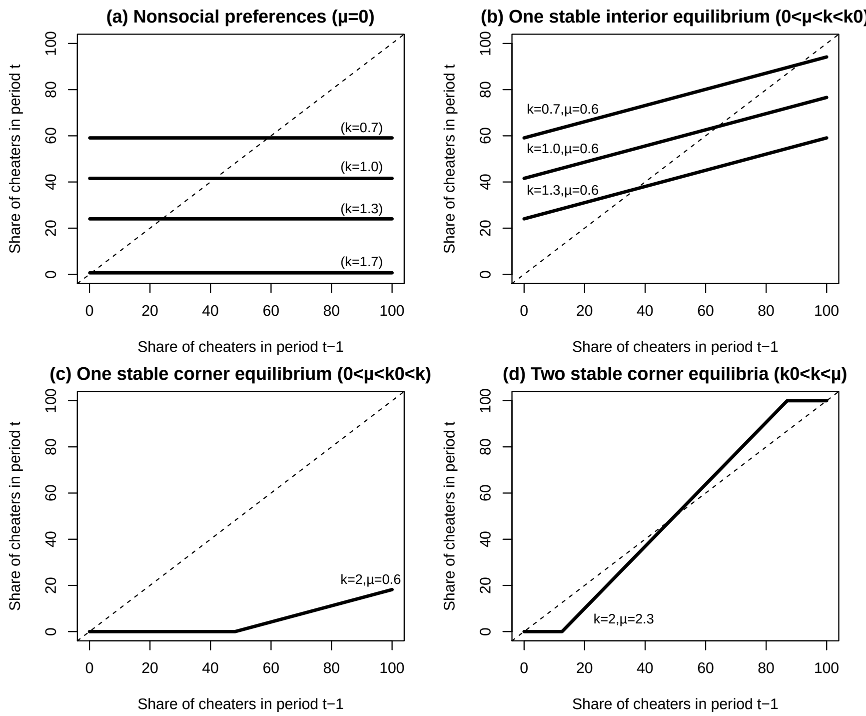


Figure 4. *Set of stationary points under linear peer pressure.*

Figure 4d). This situation is described by Akerlof (1980), p.751: "*in one of these equilibria the custom is obeyed, and the values underlying the custom are widely subscribed to by members of the community. In the other equilibrium the custom has disappeared, no one believes in the values underlying it, and it is not obeyed.*"

4.2. Short-run vs. long-run optimal policy

This section illustrates the main characteristics of the three different contracts (long-run first-best, long-run second-best and short-run optimal). For the sake of simplicity, we choose parameter values $0 < \mu < k < k_0$ such that we obtain an interior and stable optimal stationary percentage of cheaters β (shown by the dashed line in the figures). Figure 5 illustrates the case with low marginal peer pressure ($\mu = 0.4$). The difference between the three contracts is rather small since a change in y^L has little impact on the dynamics of β_t, i.e. on the contagion of fraud across periods. However, the present value of the expected cost of the first-best contract (panel a) is slightly lower than that of the second-best contract (panel b), which is in turn lower than that of the short-run contract (panel c).

The first-best optimal contract implies falling y^L over time, mostly over the final periods. The first-best salary in the very last period is the same as the short-run optimal salary, namely 37.3125 (since the future no longer matters then). While the optimal salaries are higher in the long-run than in the short-run contracts, they also produce a lower percentage of cheaters. This illustrates the main difference between the short- and long-run contracts. Due to the dynamic effects of remuneration, the principal sets a higher salary today to reduce the future percentage

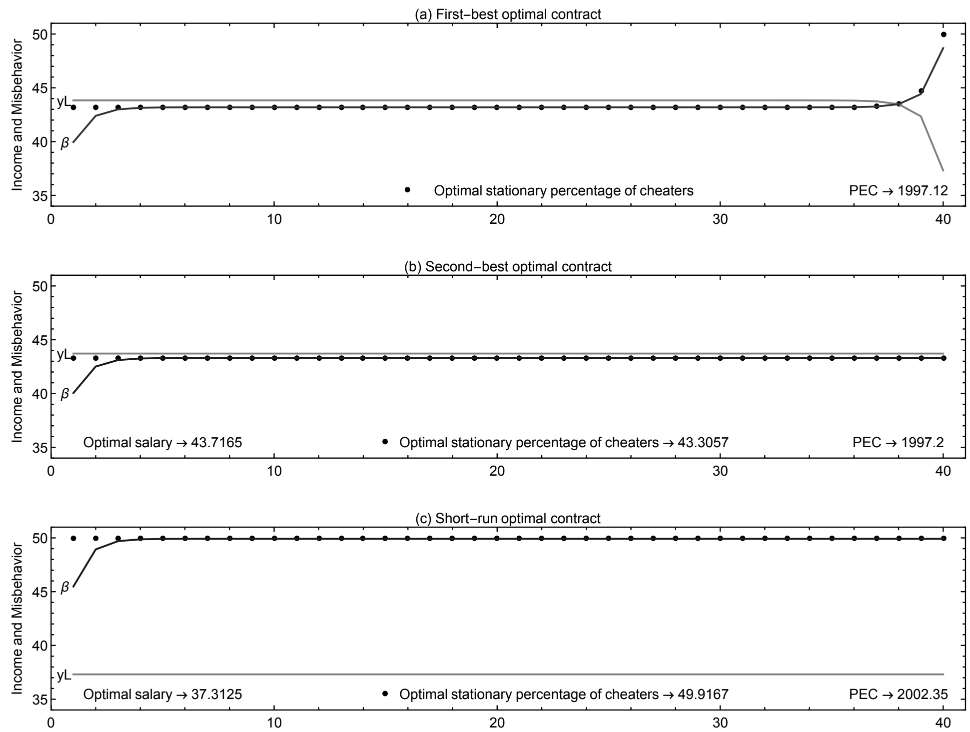


Figure 5. *Optimal policy with low marginal peer pressure.*

of cheaters and future expected costs. As shown in Figure 6, the differences between the three optimal contracts are much more pronounced with higher peer pressure (μ=1.007). This latter implies that the social norm is more sensitive to changes in remuneration.

Figure 7 shows the difference between the short-run and second-best optimal salaries, and illustrates two situations with different values of marginal peer pressure μ. The income of low-type researchers is on the horizontal axis and the present value of expected cost on the vertical axis. The present value of expected cost is seen to be minimized when income is at its second-best optimal level. Income here is higher than at the short-run optimal level. The gap between the two optimal income levels depends on $v(\beta_t)$. When μ is high, the share of cheaters is more sensitive to income. The principal then sets the income of low-type researchers at a higher level. It should be noted that the first-best optimal contract cannot be shown in this figure as optimal remuneration differs in each period.

4.3. The probability of detection as a positive externality for the principal

The disclosure of scientific misconduct may affect the reputation of journals in which the falsified research appeared. The implementation of a performance-based remuneration policy that fosters scientific fraud may thus produce negative externalities for journals. Reciprocally, the peer-review process organized by editors may yield a positive externality for the principal if it deters researchers from cheating.

Positive externalities arise if, by increasing the probability of being detected, peer review

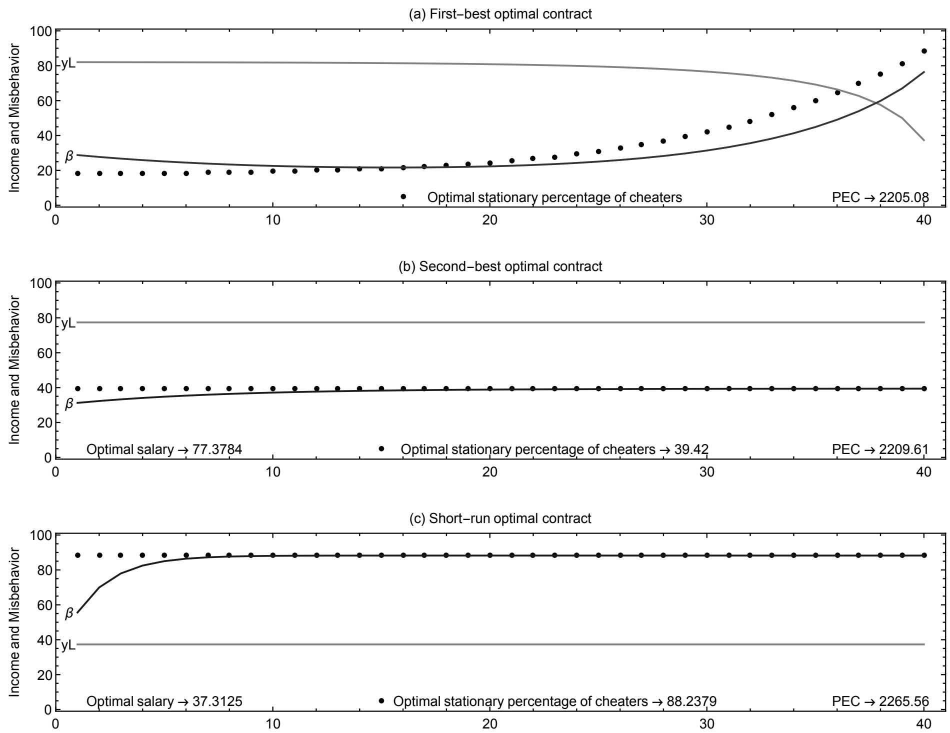


Figure 6. *Optimal policy with high marginal peer pressure.*

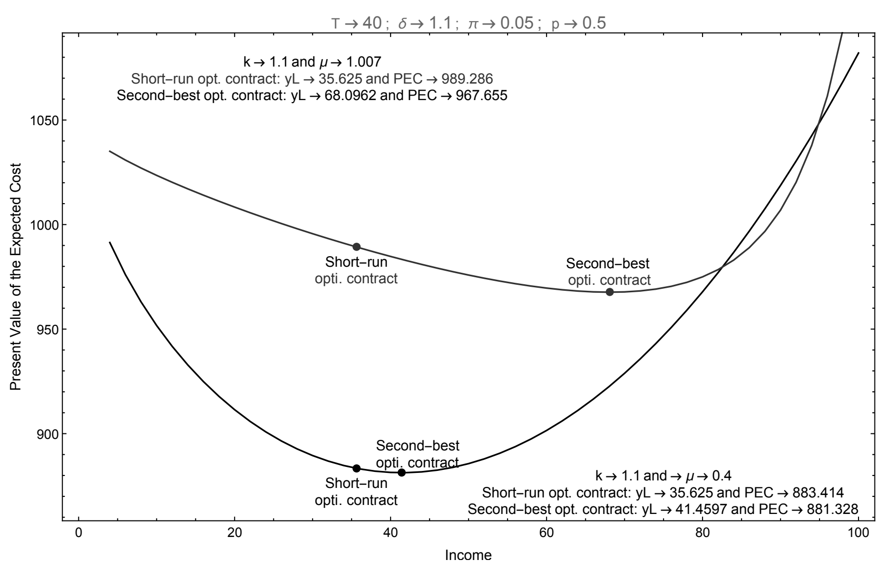


Figure 7. *Short-run and second-best optimal contracts for two different specifications of $v(\beta_{t-1})$.*

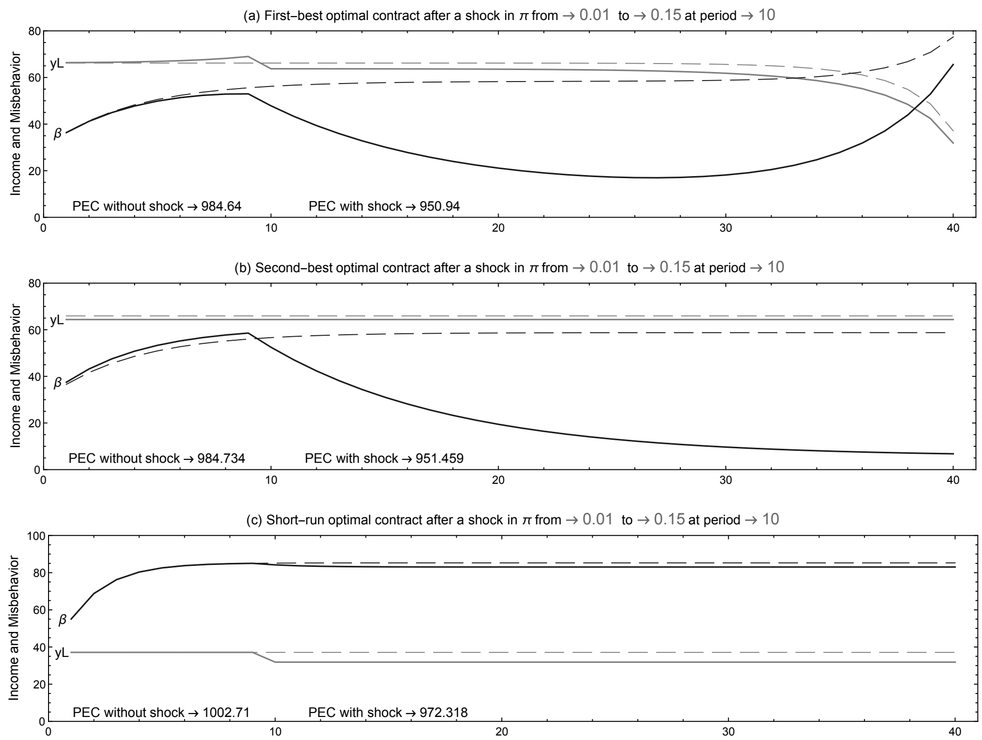


Figure 8. *A positive shock to the probability of detection.*

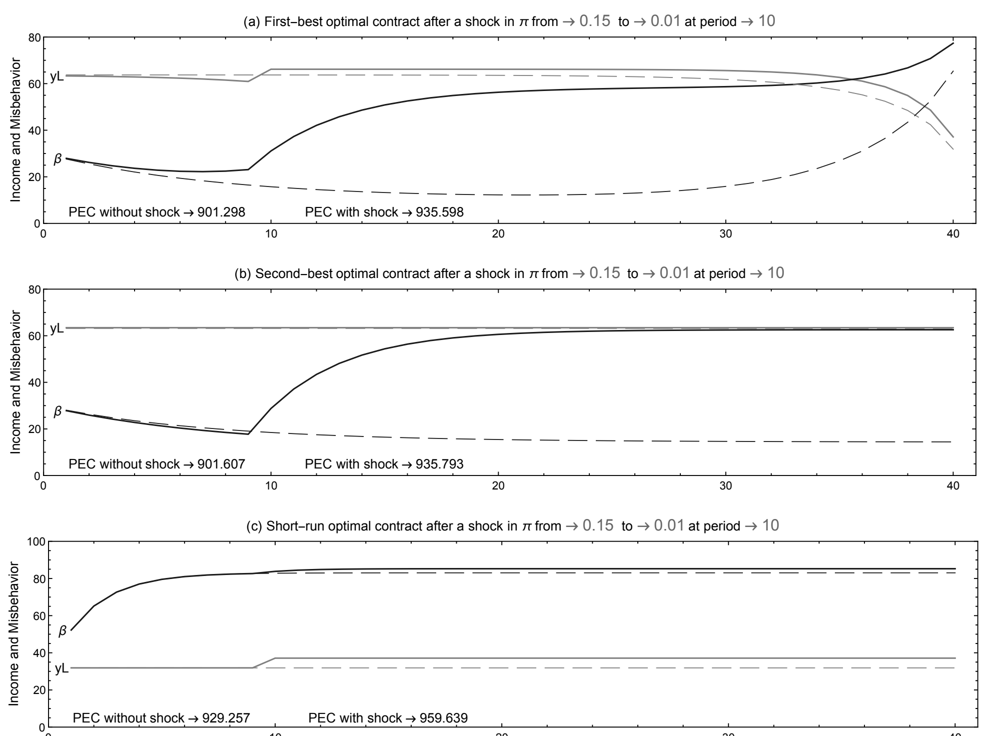


Figure 9. *A negative shock to the probability of detection.*

reduces the expected benefit of fraud, thereby reducing the share β of cheaters (equation (5)). Ceteris paribus there will be fewer low-type researchers to whom the principal pays y^H and the salary of honest low-types will be lower (equation (8)).

The prospect of a rise in π in the future may also affect the optimal contract, as shown in Figure 8. This shows what happens if the probability of detection exogenously rises from 1% to 15% in period 10. The dashed lines show the optimal values of y^L and β in the absence of a π shock, whereas the solid lines show these values with this shock (which is common knowledge and taken into account by the principal). In the latter case, y^L in the first-best optimal contract (panel a) rises until one period before the change in π. This is followed by the usual pattern of a decreasing y^L over time. Setting a high value of y_t^L in the periods prior to the shock increases the shock's negative impact on the percentage of cheaters. Formally, the reduction in the percentage of cheaters β following the rise in π depends on the value of y^L when the shock occurs. The change in β following a change in π is

$$\frac{\partial \beta_t}{\partial \pi} = -\frac{v(\beta_{t-1})u_t^L(y_t^L)}{(1-\pi)^2 u_{\min}^H} < 0 \tag{26}$$

The impact of y_t^L here depends on the sign of the cross-partial derivative.

$$\frac{\partial^2 \beta_t}{\partial \pi \partial y_t^L} = -\frac{v(\beta_{t-1})u_t^{L\prime}(y_t^L)}{(1-\pi)^2 u_{\min}^H} < 0 \tag{27}$$

The higher is y_t^L when the jump in π occurs, the lower is β afterwards. In other words, it is worth increasing the salary of low-type researchers prior to the shock as this increases the reduction in the percentage of cheaters after the shock.

Figure 8a shows that the first-best salary after the positive shock is lower than that in the absence of the shock (the dashed line). Mathematically, this can easily be explained. The short-run optimal salary y_t^{L*} is a lower bound of the first-best salary y_t^{First}. Since a rise in π produces lower y_t^{L*} (see equation (12)), y_t^{First} will fall accordingly. In the second-best optimal contract, the prospect of a future rise in π slightly reduces low-type salaries over the whole period. In contrast, the increase in π reduces salaries only after the shock for the short-run optimal contract. Figure 9 depicts the symmetric case in which there is a fall in π from 15% to 1%. Compared to the no-shock situation, the first-best salary resulting from the shock is lower before the shock and greater afterwards. A low probability of detection thus forces the principal to reward scientific productivity after the shock.

5. Conclusion

The decision to misbehave in our theory is determined by three types of motivation: an unconditional commitment to honesty, a commitment to honesty which is conditional on what others are doing, and extrinsic incentives. The novelty of the model is that each motivation interacts with the others and plays a role in the decision to cheat. If some researchers are attracted by the pecuniary gain from breaking the code of good conduct, this can affect the perception of what is good behavior and in turn generate more disobedience.

The main purpose of our theory is to show which contract the principal (the government, a university, or a research institute) will choose if the aim is to reward differences in productivity at the lowest possible expected cost per capita. If the information about the researcher type was perfectly known, the principal would remunerate the less productive researchers at the lowest possible salary. However, if this information is unknown, low-productivity types may cheat: the larger the difference in salary, the greater the incentives to fraud. Moreover, the principal has to consider that the performance-related pay not only encourages fraud at t but also lowers the conditional commitment to honesty at $t+1$. Due to the rising share of cheaters, the expected costs of the salary policy rise over time. The principal thus faces a tradeoff between higher expected costs today (by reducing the productivity premium) and in the future (by increasing the share of futur cheaters). We investigate this tradeoff and show that, inter alia, the size of the productivity premium depends on three of our model parameters: the discount rate, marginal peer pressure and the probability of detection. Table 4 provides a summary of these results.

Our policy implications are threefold. First, in a time in which most researchers' salaries and funds are related to performance, we show that a slight reduction in the performance premium may actually be efficient. This may not only reduce scientific fraud in the short-run, and thereby improve the quality of scientific production, but also change the social norm guiding research misbehavior in the long run. This policy can also reduce the size of the negative externality on scientific journals caused by the performance premium.

Second, our analysis stresses the role of marginal peer pressure for the adverse effects of performance-based pay. Greater marginal peer pressure increases the importance of offering a lower productivity premium. A principal who wants to offer a high premium to remunerate performance should first make sure that social norms (i.e. fraud behaviors) are only weakly transmitted. Moreover, moral reminders that increase the unconditional commitment to honesty may prove to be useful.

Third, it is well known that the likelihood that scientific misconduct be detected is low. An example is Alan Sokal's fake submission that was accepted for publication in a journal. Our model shows that the dynamics of fraud and its cost to the principal strongly depend on the detection probability. Since employers in academia cannot directly control monitoring, a low-premium policy is fundamental unless the probability of detection exogenously increases as editors try to save the reputation of their journals.

Although academia is a particularly interesting example for the analysis of the link between rewards, norms, and misbehavior, our theoretical framework also has other applications. This is the case when the principal is interested in performance but less in how performance is achieved. An example is given by the banking industry, and in particular the case of Jerome Kerviel, a French trader who was convicted for breach of trust, forgery and unauthorized use of bank computers. His type of trading behavior was widespread in the profession, and also highly profitable. When the principal does not care about fraud but remunerates performance, the consequences for agents' behavior may be devastating in the long run.

Table 4. Summary of the theoretical results

Variable	Definition	Low value	High value
Marginal peer pressure μ	Extent to which present behaviors influence future ones.	When μ is low, agents have nonsocial preferences. An increase in y_L will not change the norm much. A prevention policy is inefficient.	When μ is high, behaviors, whether they are good or bad are passed on to the next generations. An increase in y_L will change the norm. A prevention policy is efficient.
Probability of detection π	Probability of being detected when cheating, leading to punishment	When π is low, cheating is attractive. The principal has to increase the income of low-types, thereby reducing the productivity premium.	When π is high, the expected benefit of fraud is low. The principal has a greater incentive to reward performance.
Discount factor δ	Extent to which the principal cares about future costs.	If δ is low, the principal cares a lot about future costs. A prevention policy is efficient.	If δ is high, the principal cares less about future costs. The principal has less incentive to increase the income of low-types, which yields a higher level of fraud.

References

Abeler, J., Becker, A., and Falk, A. (2014). Representative evidence on lying costs. *Journal of Public Economics*, 113:96–104.

Akerlof, G. A. (1980). A theory of social custom, of which unemployment may be one consequence. *Quarterly Journal of Economics*, 94:749–775.

Bailey, C., Euzent, P., Martin, T., and List, J. (2001). Academic economists behaving badly? a survey on three areas of unethical behavior. *Economic Inquiry*, 39:162–170.

Becker, G. S. (1968). Crime and Punishment: An Economic Approach. *Journal of Political Economy*, 76(2):169–217.

Brambor, T., Clark, W. R., and Golder, M. (2006). Understanding Interaction Models: Improving Empirical Analyses. *Political Analysis*, 14(1):63–82.

Brodeur, A., Lé, M., Sangnier, M., and Zylberberg, Y. (2016). Star wars: The empirics strike back. *American Economic Journal: Applied Economics*, 8(1):1–32.

Cameron, A. and Trivedi, P. (2005). *Microeconometrics: Methods and Applications*. Cambridge University Press.

Cartwright, E. and Menezes, M. L. (2014). Cheating to win: Dishonesty and the intensity of competition. *Economics Letters*, 122(1):55–58.

Charness, G., Masclet, D., and Villeval, M. C. (2013). The dark side of competition for status. *Management Science*, 60(1):38–55.

Cialdini, R. B., Reno, R. R., and Kallgren, C. A. (1990). A focus theory of normative conduct: Recycling the concept of norms to reduce littering in public places. *Journal of Personality and Social Psychology*, 58:1015.

Conrads, J., Irlenbusch, B., Rilke, R. M., Schielke, A., and Walkowitz, G. (2014). Honesty in tournaments. *Economics Letters*, 123(1):90–93.

Corneo, G. (1995). Social custom, management opposition, and trade union membership. *European Economic Review*, 39:275–292.

Falk, A. and Fischbacher, U. (2002). "(c)rime" in the lab - detecting social interaction. *European Economic Review*, 46:859–869.

Fanelli, D. (2010). "Positive" results increase down the hierarchy of the sciences. *PloS one*, 5(4):e10068.

Fang, F. C., Steen, R. G., and Casadevall, A. (2012). Misconduct Accounts for the Majority of Retracted Scientific Publications. *Proceedings of the National Academy of Sciences*, 109(42):17028–17033.

Feld, L. P., Necker, S., and Frey, B. S. (2014). Happiness of Economists. CESifo Working Paper Series No. 5099.

Ferrer-i Carbonell, A. and Frijters, P. (2004). How Important is Methodology for the Estimates of the Determinants of Happiness? *Economic Journal*, 114(497):641–659.

Fischer, P. and Huddart, S. (2008). Optimal contracting with endogenous social norms. *American Economic Review*, 98(4):1459–1475.

Franzoni, C., Scellato, G., and Stephan, P. (2011). Changing incentives to publish. *Science*, 333:702–703.

Funk, P. (2005). Governmental action, social norms, and criminal behavior. *Journal of Institutional and Theoretical Economics*, 161(3):522–535.

Gilpatric, S. M. (2011). Cheating In Contests. *Economic Inquiry*, 49:1042–1053.

Gino, F., Ayal, S., and Ariely, D. (2009). Contagion and Differentiation in Unethical Behavior The Effect of One Bad Apple on the Barrel. *Psychological Science*, 20:393–398.

Gneezy, U. (2005). Deception: The role of consequences. *American Economic Review*, 95(1):384–394.

Gneezy, U., Rockenbach, B., and Serra-Garcia, M. (2013). Measuring lying aversion. *Journal of Economic Behavior & Organization*, 93(0):293 – 300.

Hoover, G. A. (2006). A game-theoretic model of plagiarism. *Atlantic Economic Journal*, 34:449–454.

John, L. K., Loewenstein, G., and Prelec, D. (2012). Measuring the prevalence of questionable research practices with incentives. *Psychological Science*, 23(5):524–532.

Jones, T. M. (1991). Ethical decision making by individuals in organizations: An issue-contingent model. *Academy of Management Review*, 16(2):366–395.

Kiri, B., Lacetera, N., and Zirulia, L. (2014). Above a Swamp: A Theory of High-Quality Scientific Production. NBER Working Paper No. w21143.

Lacetera, N. and Zirulia, L. (2009). The economics of scientific misconduct. *Journal of Law, Economics, and Organization*, 27:568–603.

Lacetera, N. and Zirulia, L. (2012). Individual preferences, organization, and competition in a model of R&D incentive provision. *Journal of Economic Behavior and Organization*, 84:550 – 570.

Larsen, P. and Von Ins, M. (2010). The rate of growth in scientific publication and the decline in coverage provided by science citation index. *Scientometrics*, 84(3):575–603.

Layard, R., Nickell, S., and Mayraz, G. (2008). The marginal utility of income. *Journal of Public Economics*, 92:1846–1857.

Lindbeck, A., Nyberg, S., and Weibull, J. W. (1999). Social Norms And Economic Incentives In The Welfare State. *Quarterly Journal of Economics*, 114:1–35.

Mazar, N., Amir, O., and Ariely, D. (2008a). More ways to cheat – expanding the scope of dishonesty. *Journal of Marketing Research*, 45:651–653.

Mazar, N., Amir, O., and Ariely, D. (2008b). The Dishonesty of Honest People: A Theory of Self-Concept Maintenance. *Journal of Marketing Research*, 45:633–644.

Necker, S. (2014). Scientific Misbehavior in Economics. *Research Policy*, 43(10):1747–1759.

Rosenbaum, S. M., Billinger, S., and Stieglitz, N. (2014). Let's be honest: A review of experimental evidence of honesty and truth-telling. *Journal of Economic Psychology*, 45:181–196.

Schwieren, C. and Weichselbaumer, D. (2010). Does competition enhance performance or cheating? A laboratory experiment. *Journal of Economic Psychology*, 31(3):241–253.

Stephan, P. (2012). *How Economics Shapes Sciences*. Harvard University Press.

Stephan, P. E. (1996). The economics of science. *Journal of Economic Literature*, 34:1199–1235.

Stern, A. M., Casadevall, A., Steen, R. G., and Fang, F. C. (2014). Financial costs and personal consequences of research misconduct resulting in retracted publications. *Elife*, 3:e02956.

Wible, J. (1998). *The Economics of Science: Methodology and Epistemology as if Economics Really Mattered*. Routledge.