Professional identity and the gender gap in risk-taking. Evidence from field experiments with scientists

Drupp, Moritz A.; Khadjavi, Menusch; Riekhof, Marie Catherine; Voss, Rudi

published in
Journal of Economic Behavior and Organization
2020

DOI (link to publisher)
10.1016/j.jebo.2019.12.020

document version
Publisher's PDF, also known as Version of record

document license
Article 25fa Dutch Copyright Act

Journal of Economic Behavior and Organization 170 (2020) 418–432

Contents lists available at ScienceDirect

Journal of Economic Behavior and Organization

journal homepage: www.elsevier.com/locate/jebo

Professional identity and the gender gap in risk-taking. Evidence from field experiments with scientists

Moritz A. Drupp[a], Menusch Khadjavi[b,c,d], Marie-Catherine Riekhof[e,*], Rudi Voss[f,g]

[a] *Department of Economics, University of Hamburg, Germany*
[b] *Department of Spatial Economics, VU Amsterdam, the Netherlands*
[c] *Tinbergen Institute, the Netherlands*
[d] *Kiel Institute for the World Economy, Germany*
[e] *Department of Agricultural Economics, University of Kiel, Germany*
[f] *Biodiversity Economics, German Centre for Integrative Biodiversity Research (iDiv), Germany*
[g] *Department of Economics, University of Kiel, Germany*

ARTICLE INFO

Article history:
Received 5 April 2018
Revised 1 November 2019
Accepted 21 December 2019
Available online 18 January 2020

JEL classification:
J16
D81
C93

Keywords:
Risk-taking
Gender
Identity
Priming
Labor market
Field experiment

ABSTRACT

The gender gap in risk-taking is often used to explain differences in labor market outcomes. Yet, a number of studies suggest that this gap is larger in private contexts and is reduced in professional contexts. In two online field experiments with more than 1500 scientists we shed light on the causal role of the professional context by varying the salience of the professional or private identity. The main study finds that the gender gap in risk-taking is moderated—and vanishes for older scientists—when the professional identity is salient. The second study—designed to further explore mechanisms relating to non-professional identity—yields inconclusive results. While the results of the main study imply that gender gaps may be driven by the ability to switch between identities and adapt to prevailing norms, our second study suggest the need for further research to examine how prevailing norms are shaped and identity affects gender gaps.

1. Introduction

Risk-taking is part of many day-to-day decisions, in private life as well as in professional contexts. Experimental evidence summarized by Croson and Gneezy (2009) suggests that females act—on average—more risk-averse than males and that this gender gap in risk-taking reduces or vanishes in professional environments.[1] Differences in risk-taking may lead to unwanted

* Corresponding author.
E-mail address: mcriekhof@ae.uni-kiel.de (M.-C. Riekhof).

[1] While this finding appears widely accepted, Filippin and Crosetto (2016) and Nelson (2015, 2016) suggest that risk-taking behavior may not be so different across genders after all. Atkinson et al. (2003), Johnson and Powell (1994) and Schubert et al. (1999) find that risk-taking accross gender is more similar in a professional context. While Dwyer et al. (2002) find that females take fewer risks in mutual fund investments (see also Beckmann and Menkhoff 2008), the gap reduces when controlling for financial investment knowledge. Yet, Adams and Funk (2012) find for board members that female directors are even more risk-seeking.

https://doi.org/10.1016/j.jebo.2019.12.020

disparities. Specifically, the gender gap in risk-taking has been used to explain the gender gap on the labor market.[2] Indeed, females and males are neither equally paid nor equally represented in leadership positions. For instance, there are only 23 female CEOs in Fortune 500 companies (Bellstrom, 2015) and only 28% of all full professors in the United States are female (National center of education statistics, 2016). In spite of efforts made by many countries to increase gender equity, there is still a mean gender wage gap of approximately 15% in OECD countries (OECD, 2016). As appropriate policy approaches to dealing with gender gaps on the labor market hinge on what drives the observed effects, a better understanding of the underlying causes—such as what drives the gender gap in risk-taking—is required.

In this paper, we consider another explanation for the two stylized findings from Croson and Gneezy (2009) on the gender-gap in risk-taking. We draw on the theory of identity to examine how different contexts and associated identities—a "professional" and a "private" identity—influence risk-taking for females and males. Identity theory (Akerlof and Kranton, 2000) posits that individuals may form multiple identities that moderate behavior across different contexts. While there appear to be gender-related stereotypes in private life[3]—females are perceived to be more risk-averse (Ball et al., 2010) and to act more financially risk-averse (Charness and Gneezy, 2012) on average—there is usually a common idea of 'professional behavior'.[4] If behavior is affected by prescribed behavior, this may explain why gender differences in risk-taking are found to be smaller in professional environments. Based on a theoretical model, we explore different predictions with respect to labor market outcomes.

To examine the drivers of potential differences in the gender gap in risk-taking between professional and private settings, we performed two online field experiments. Our first, main study included 474 scientists, with 278 males and 196 females. Since scientific success is uncertain, risk-taking is usually warranted in science. Indeed, the European Research Council and the US National Science Foundation, for instance, specifically aim at supporting 'high risk, high gain' projects.[5] To activate the professional scientific identity and associated norms and behaviors, we draw on the priming technique that uses environmental cues to make a certain identity temporarily more salient (Benjamin et al., 2010).[6] We experimentally vary the salience of the private or professional identity in two treatments using nine questions relating to either context. For example, subjects in the professional identity treatment are asked "How large is your direct working team (yourself included)?", while subjects in the private identity comparison treatment were asked "How large is your direct family (yourself included)?". We use a standard, incentivized risk elicitation task (Binswanger, 1981; Dave et al., 2010; Eckel and Grossman, 2002), where subjects are asked to choose one out of six lotteries that increase in riskiness, from a safe option to a lottery that elicits risk-seeking behavior.[7] In our second study, we broaden the sample to a general science population and shed more light on the non-professional comparison by disentagling the private identity into a 'gender' and a 'non-gender' private identity.

In line with Croson and Gneezy (2009), we find that females take fewer risks compared to males on average. We show that the gender gap in risk-taking and in risk-seeking is smaller in the professional identity compared to the private identity, suggesting a causal role of the professional context and its associated identity in Study 1. In Study 2, priming was not as successful as in Study 1 and we do not fully replicate this second result. However, an exploratory analysis focussing on a sub-sample that mirrors our first sample more closely in terms of the success of priming, the share of females and the share of European and North American scientists yields no gender gap in risk-taking in the professional and private identity treatments but a gap in the 'gender' treatment, suggesting a likely role of gender-related identity aspects as a driver of the results in Study 1.

As competitiveness seems related to risk attitudes (van Veldhuizen, 2016), our findings provide some confirmation and extension of the results obtained by Cadsby et al. (2013), the study most closely related to ours. They consider the competitiveness of 132 MBA-students when their professional identity is salient compared to when their gender/family identity is salient. They find that females act less competitively when their gender/family identity is more salient, while males do not. While Cadsby et al. (2013) focus their analysis on competitiveness, they also use their participants' choices of entering either into chance pay or a piece rate payment mechanisms as an indication of their risk-seeking/risk-neutral or risk-neutral/risk-averse attitude. They find that, first, females are not less likely to enter chance pay as compared to males, second, professional identity priming has a positive effect on the likelihood of females entering chance pay, and, third, the effect of professional identity priming on the overall likelihood of entering chance pay is moderated by gender. Our analysis can be seen as an extension of their work to a larger, professional subject pool, and to study not only the binary distinction between risk-seeking and risk-aversion but to also consider different degrees of risk-aversion. In terms of the findings, both Study 1 and 2 do not confirm the first result of Cadsby et al. (2013). While our main Study 1 confirms the second finding of

[2] Bertrand (2011), Blau and Kahn (2017) and Boll et al. (2016) provide analyses of gender gaps. Recent studies focus on gender differences in preferences, especially on risk-taking and competitiveness (Azmat and Petrongolo, 2014; Croson and Gneezy, 2009), suggesting that these may help to explain observed gender gaps. Decomposing drivers of competitiveness, van Veldhuizen (2016) suggests that the gender gap in competitiveness may be largely driven by risk attitudes and overconfidence. This stresses the enduring importance of risk preferences for gender gaps.

[3] See Bordalo et al. (2016a,b) for general results on stereotypes and beliefs about gender.

[4] The findings of Bursztyn et al. (2017) even suggest that 'private' considerations may be the reason for still observed gender differences in professional behavior.

[5] See https://erc.europa.eu/news/2017-qualitative-evaluation-projects and https://www.nsf.gov/about/, last accessed: March 26, 2018.

[6] See Cohn and Maréchal (2016) for a recent review of identity priming in economics.

[7] Our design is closely related to Cohn et al. (2017), who find that priming the saliency of professional identity causally affects bankers' risk-taking. They employ a different task to elicit risk preferences (Gneezy and Potters, 1997) and find that the professional norm points towards more risk-averse behavior. They do not report an impact of gender.

Cadsby et al. (2013), Study 2 cannot confirm the second finding of Cadsby et al. (2013) for all observations, yet a sub-sample suggest that females take fewer risks in the gender treatment as compared to the professional identity treatment. While coefficients go in the expected direction, we do not find a statistically significant moderating role of gender on the priming effects on risk-taking.

We further make use of the much wider age distribution in our larger sample to show that the gender gap in risk-taking reduces with age and closes for senior scientists in the professional identity treatment. Risk-taking decreases with age for male scientists, independent of treatment. In contrast, risk-taking is higher for older females in the professional identity treatment, while risk-taking is lower for older females in the private identity treatment.

Overall, our results thus paint a more nuanced picture concerning the explanation that relatively risk-averse females are deterred from or leave risky professional contexts such as academia. Previous findings in the literature suggest that the observation of a diminished gender gap in risk-taking in professional contexts may be driven by a selection effect or by social learning and a general adaptation to prevailing norms. There is some experimental economic evidence on the selection or sorting effect (Buser et al., 2014; Flory et al., 2015; Niederle and Vesterlund, 2007), as well as evidence that the (work) environment shapes behavior (Cohn et al., 2017; Booth and Nolen, 2012). While the selection effect offers an explanation for both—the gender gap on the labor market as well as a reduced gender gap in risk-taking in a professional environment—the gender gap on the labor market is more difficult to explain if only the adaptation channel is considered. Our results from Study 1 suggest that female scientists still take fewer risks when their private identity is salient, a finding that can neither be explained by the pure selection nor by the pure adaptation channel. While Study 2 cannot confirm this in general, the exploratory analysis suggests that the result from Study 1 may be driven by gender-related identity aspects also captured by the private identity treatment in Study 1. In turn, if the professional identity is salient, gender-related differences in risk-taking are smaller or insignificant. Our findings thus suggest that if the gender gap in risk-taking is driven by selection, the selection is not (only) along general risk-aversion, but (also) along the ability to adapt to prevailing norms by switching between private (or gender) and professional identities.

Our study offers a further potential explanation why the gender gap in risk-taking is moderated in a professional environment. We replicate the two stylized findings from Croson and Gneezy (2009) that females take less risk than males, and that this gap is moderated in a (risky) professional context. We provide causal evidence that making either the private or the professional identity salient influences the gender gap in risk-taking. Thereby, we identify a new channel according to which females may select in or out of professional environments: the ability to switch between private and professional identities to better adapt to differences in the environments' prescribed behaviors. Our study thus provides new insights for the discussion on gender, risk-taking and labor market policies.

2. Design and hypotheses of the main study

This section provides a simple model of identity and risk-taking behavior. We then describe the design, procedures and hypotheses of our online field experiment in our main Study 1.

2.1. A theoretical framework

To facilitate a better understanding of what may drive the gender gap in risk-taking we consider, building on Akerlof and Kranton (2000) and Benjamin et al. (2010), a model incorporating identity considerations. Identity models are based on the idea that several social categories—like being a woman, a man, a parent, a teacher—are available and that each social category comes with a stereotypical or prescribed behavior. Individuals identify with social categories and derive disutility when they deviate from prescribed behavior, as this means a loss of identification.

Let there be C different social categories, indexed by $j = 1, \dots, C$ and let $\hat{A}_j$ denote prescribed behavior in category j. Individual i assigns weight $w_i(s_{ij})$ to complying with prescribed behavior of a category based on the strength of identification with that category s_{ij}. An increase in the identification—such as caused by priming—usually increases the weight assigned to that category, $\partial w_i / \partial s_{ij} \geq 0$. If the identification with a social category is zero, $s_{ij} = 0$ the weight assigned to the social category is also zero, $w_i(0) = 0$, rendering it behaviorally irrelevant.

We assume a two-stage process: first, individuals choose their identification s_{ij} with different social categories to maximize some long-run utility function. In other words, individuals have a baseline or steady state identification with the different social categories that determines their general behavior. Second, taking this baseline identification as given, individuals choose their short-run action to comply with their general identity as well as the environmental context. A specific context or prime may then induce individuals to deviate from this baseline behavior to shift towards the relevant prescribed behavior. Our focus is on this second stage, i.e., on the short-run reaction to environmental cues.

Given the choices of baseline identification with the different social categories, let the individual i choose (short-run) action a_i in order to maximize the following instantaneous utility function

$$U(a_i) = -\sum_{j=1}^{C} w_i(s_{ij})(a_i - \hat{A}_j)^2. \tag{1}$$

When several social categories are important and prescribed behaviors in these social categories deviate from one another, a short-run identity conflict arises. The optimal action a_i^* is then a weighted average of the prescribed behaviors in the different social categories,

$$a_i^* = \sum_{j=1}^{C} w_i(s_{ij})\hat{A}_j. \tag{2}$$

Three remarks are in order. First, although an identity conflict may reduce short-run utility, the identification choices are—per assumption—optimal in the long-run perspective. Second, as individuals assign different weights on complying with prescribed behaviors, the actions of individuals also differ. Third, a strong reaction of the weighting to the prevalent context may induce that temporarily only one social category is of importance, which increases instantaneous utility relative to a situation with an identity conflict.

We now turn to our experiment on whether or not the saliency of a certain identity of a person, especially with respect to their professional and their private environment, influences risk-taking. In terms of the model, risk-taking relates to the observed action a_i, with a higher a_i denoting more risk-taking. As the identification with other social categories should not be impacted and as we compare the effect of one identity relative to another identity, we reduce the model to only include two possible identities, *Private [PRIV]* and *Professional [PROF]*. One can interpret each identity as a bundle of social categories and prescribed behaviors. Accordingly, the overall prescribed behavior may be gender-specific. For gender g, we also simplify and only consider either male, denoted m, or female, denoted f.[8] Eq. (2) then becomes

$$a_i^{g*} = w_i(s_{i,PROF})\hat{A}^g_{PROF} + \underbrace{(1 - w_i(s_{i,PROF}))}_{w_i(s_{i,PRIV})}\hat{A}^g_{PRIV}. \tag{3}$$

Priming a certain identity means that the identification strength s_{ij} with the corresponding social category is temporarily increased. In other words, we are interested in the marginal impact on risk-taking of making the professional identity more (or less) salient:

$$\frac{\partial a_i^{g*}}{\partial s_{i,PROF}} = \frac{\partial w_i(s_{i,PROF})}{\partial s_{i,PROF}}\left(\hat{A}^g_{PROF} - \hat{A}^g_{PRIV}\right). \tag{4}$$

To use our model for clear predictions on the gender gap in risk-taking, we make three assumptions on prescribed behavior based on existing evidence. First, Bursztyn et al. (2017) suggest that possible differences in professional behavior between males and females are driven by private (i.e. marriage) considerations. We thus assume that there is a gender-invariant idea of 'professional behavior': $\hat{A}^f_{PROF} = \hat{A}^m_{PROF} = \hat{A}_{PROF}$. Behavioral norms related to risk-taking in the private context in turn may be gender-specific. Females are perceived as being less risk-taking (Ball et al., 2010), which may be related to females being expected to act less risk-taking. We therefore, second, assume that $\hat{A}^f_{PRIV} \neq \hat{A}^m_{PROF}$ and especially that prescribed risk-taking behavior in the private context is such that females are deemed to take fewer risks than males: $\hat{A}^f_{PRIV} < \hat{A}^m_{PRIV}$. As science has been and still is a male-dominated field of occupation (Knights and Richards, 2003; West et al., 2013), males are likely to have shaped professional norms, among others with regard to risk-taking behavior. We thus, finally, assume that $\hat{A}_{PROF} \approx \hat{A}^m_{PRIV}$, which implies via the second assumption that $\hat{A}_{PROF} > \hat{A}^f_{PRIV}$.[9]

Aggregating over all $i = 1, \ldots, N$ individuals of a population yields the mean risk choice $\bar{A} = \frac{1}{N}\sum_{i=1}^{N} a_i$, which can be disaggregated for different groups within a population. For instance, we denote the mean risk choice of females in the *Professional* identity treatment as $\bar{A}^f_{PROF}$. Optimal risk-taking as implied by our model and given our assumptions on prescribed behaviors $\hat{A}^f_{PRIV}$, $\hat{A}^f_{PROF}$, $\hat{A}^m_{PRIV}$, and $\hat{A}^m_{PROF}$ predicts that, on average, females are more risk-averse than males, i.e., $\bar{A}^m > \bar{A}^f$ (Croson and Gneezy, 2009; Charness and Gneezy, 2012). Our model also predicts that females make more risky decisions when their professional identity is salient compared to the private identity, i.e. $\bar{A}^f_{PROF} > \bar{A}^f_{PRIV}$. Also, Eq. (4) and our assumption that $\bar{A}^f_{PRIV} < \bar{A}^m_{PRIV}$ imply that the impact of professional identity priming, i.e. increasing the identification with the professional environment, is more pronounced for females than for males. Professional identity priming should therefore diminish the gender gap in risk-taking.

If results are in line with our calibrated model, we can use the model's logic to hypothesize on the 'type' of females that remain in academia.[10] In other words, what does our model imply about the different channels—sorting, adaptation and identity switching—affecting the (dis)appearance of the gender gap in risk-taking in professional versus private contexts? The difference between $\hat{A}^f_{PRIV}$ and $\hat{A}_{PROF}$ means a utility loss for females. Therefore, leaving science to work in a different sector in which the prescribed behavior is more similar to the prescribed behavior in *Private* may increase utility. According to such possible self-selection, the following 'types' of females would stay in academia:

First, we could observe females in academia who attach a lot of weight to the professional identity (high w_i) and thus stick to one form of risk-taking behavior without experiencing a large utility loss. In this case, risk choices should be similar

[8] In Study 2, we differentiate the gender-related *Private* identity into its two components, i.e. *gender* identity and non-gender *private* identity.

[9] Furthermore, as engaging in research is inherently risky as results are not known a priori, academia may be regarded as a field in which risk-taking norms can be deemed prevalent. This would provide another justification for the assumption $\hat{A}_{PROF} \approx \hat{A}^m_{PRIV} > \hat{A}^f_{PRIV}$.

[10] Certainly, many other drivers influence which 'types' of females remain in academia besides risk-taking. Our analysis thus only applies to an 'all else equal' environment.

Table 1
Identity priming questions.

Professional identity treatment	*Private* identity treatment
Who is your current employer?	What is your current city of residence?
How many years have you worked for this institution?	How many years have you lived in your current accommodation?
Do you have a tenured position?	Are you married?
How large is your direct working team (yourself included)?	How large is your direct family (yourself included)?
Where did you last go to for a conference/workshop?	Where did you last go on holiday?
In which year did you start your Ph.D.?	In which year did you kiss the first boy/girl?
At what time do you usually arrive at the office?	At what time do you usually arrive at home?
What activity in your work do you enjoy the most?	What activity in your leisure time do you enjoy the most?
How satisfied are you with your work in general?	How satisfied are you with your life in general?

across our two treatments. A high weight on the professional identity and high risk-taking of female scientists would correspond to the standard selection channel. One would observe a gender gap in risk-taking for the general population, but not for scientists (neither when the private nor the professional identity is salient).

Second, we could observe females in academia whose preferences are shaped by the academic environment and whose weights, $w_i(s_{i,\ PROF})$, on complying with the professional norm might increase with their academic age. In terms of the model, this relates to long-run changes, i.e. to changes in the first stage, to adapt to a changing environment. If this environmental adaptation effect would be the primary driver, risk choices of female scientists would develop with their exposure to the work environment, i.e. academic age, such that their overall risk-taking—in *both* the private and in the professional environment—would increase over time. This adaptation to the work environment would observationally go against the general finding that risk-taking behavior reduces with age. Risk-taking of female scientists might therefore remain constant—as the age and environmental adaptation effects cancel each other out—while males' risk-taking would decline with age. The gender gap in risk-taking would be especially pronounced for young scientists. Controlling for age, risk choices should be similar across treatments.

Finally, we could observe females in academia whose weights, $w_i(s_{i,\ PROF})$, on complying with the professional norm are (strongly) impacted by a change in the strength of identifying with the professional environment, $s_{i,PROF}$, i.e. $w_i'(s_{i,PROF})$ would be large. In this case, we expect the behavior to differ between the two identity treatments. In particular, we expect $\bar{A}^f_{PROF} > \bar{A}^f_{PRIV}$, and accordingly a diminished gender gap in risk-taking when the professional identity is made salient. If we assume learning about norms, the gender gap in risk-taking in the professional context should be smaller for older scientists.

2.2. Data collection

To test the hypotheses derived from our model and to learn more about the drivers of the gender gap in risk-taking, we use data from an online field experiment conducted with members of an international scientific organization in mid 2016. The members are mostly natural scientists, with a focus on the marine environment.[11]

The administrative office of the scientific organization provided an e-mail list of their 1930 members. We contacted all members by e-mail and invited them to participate in a short online survey that consisted of ten pages and took about 15 minutes to complete. We stated that participation would be compensated with 25 Euro on average (equivalent to US$27 at the time of the experiment) and that individual responses would be kept confidential. Upon clicking the link to the online survey in the invitation e-mail, subjects were assigned to one of two treatments by the computer: Either the professional identity treatment (abbreviated *Professional*) or the private identity treatment (*Private*). Thus, our study relies on 'between-subjects' comparisons. For random assignments to the two treatments, differences between the members cancel out. A preamble page provided further details on the experiment and the mode of payment (Amazon vouchers).

The survey began with simple descriptive questions on age, gender and nationality. This was followed by our manipulation that consisted of nine questions either relating to the professional identity (*Professional* treatment) or relating to the private identity (*Private* treatment). The purpose of these questions was to make the subjects' professional identity, and associated prescribed behavior, more salient in *Professional* as compared to *Private*. The questions for the *Professional* treatment are adapted to and extend upon Cohn et al. (2014). We designed the questions for the *Private* treatment to be as equivalent as possible to the *Professional* treatment in terms of content or activity (see Table 1 for a list of all priming questions; cf. Figs. A.2 and A.3 in Supplementary for screenshots from the online survey).[12] The identity manipulation was followed by three experimental tasks that were always presented in the same order.[13]

In this paper we focus on the first task, an established incentivized risk preference elicitation task based on Binswanger (1981) and Eckel and Grossman (2002). This task presents subjects with six different choice options in the form of lotteries.

[11] We do not report the name of the scientific organization to assure respondents' anonymity.

[12] Besides the priming questions, the only other difference across treatments was that the preamble stated that the study was on "Work [Life] satisfaction, including individual attitudes and behavior" in *Professional* [*Private*].

[13] Once a subject had completed a task, it was not possible to go back. The risk task was followed by a truth-telling task, which we analyze in a companion paper Drupp et al. (2017). Finally, we posed a hypothetical social time preference task.

Table 2
Description of the risk choices in the experimental task.

Choice options	Payment A (in Euro)	Payment B (in Euro)	Expected payout (in Euro)	Std. Dev.	CRRA range, $U(x)=\frac{x^{1-r}}{1-r}$
1	7	7	7	0	$3.460<r$
2	6	9	7.5	1.5	$1.161<r<3.460$
3	5	11	8	3	$0.706<r<1.161$
4	4	13	8.5	4.5	$0.499<r<0.706$
5	3	15	9	6	$0<r<0.499$
6	2	16	9	7	$r<0$

Subjects had to decide on their most preferred lottery (see Table 2, subjects only saw the information in the first three columns; cf. Supplementary, Fig. A.4, for a screenshot from the online survey). Each option is related to two possible payoffs, either the amount stated in columns 2 (Payment A) or 3 (Payment B) of Table 2, each occurring with 50% probability. The next column of the table indicates the expected payoff for each option. For options 1–5, the expected payoff increases with the standard deviation of the gamble, depicted in the fifth column. Options 5 and 6 have the same expected payoff, but option 5 has a smaller standard deviation compared to option 6. In particular, we see that there is a qualitative difference between options 1 to 5 and 6. While choosing options 1–5 indicates risk-averse behavior, a choice of option 6 indicates risk-seeking behavior (or at least risk-neutrality). The risk choices can also be related to a range of Constant Relative Risk Aversion (CRRA)-parameters (see last column in Table 2).[14]

Following the experimental tasks, participants were asked to complete a short follow-up survey that included a word-completion task designed to provide an implicit measure of how well the identity priming manipulation had worked (Cohn et al., 2014; Kahneman, 2011). Subjects were presented with eight word fragments and they were asked to fill in the gaps with letters to form existing words. The idea is that when the professional identity is salient other words come to the participantsâ mind as compared to when the private identity is salient. For example, they were shown the word fragment "j o u r___", which they could complete with the word "journal" that scientists would frequently encounter in their professional lives, or the word "journey", which might be more salient to those in the *Private* treatment.[15] We classified all completed words and either assigned the number 1 to words related to the professional work identity or 0 to words classified as related to a private life. Words that could not be classified as relating to either context or without actual meaning were coded as missing. For each subject we aggregate the numbers assigned to completed words (1 for words associated with professional life, 0 for words associated with private life).

Together with the payoff from the second task (a truth-telling task), which ranged from 0 to 20 Euro, and a 5 Euro compensation for completing the short follow-up survey, each subject could earn up to 41 Euro. The payoff from the risk task was revealed only after subjects had completed the short follow-up survey.[16]

3. Results

We have received 599 responses to the survey, amounting to a response rate of more than 30%.[17] Our results are based on 474 scientists who have completed the risk task.[18] Fig. 1 depicts a world map, in which the red balloons indicate the locations of the participants. Participating scientists come from most continents, predominantly from Europe but also North America.

Before we turn to analyzing risk-taking behavior, we test whether the implicit measure of identity priming—the word completion task—indicates that priming has been successful. We find that the mean number of 'professional' words, such as "journal", "paper" or "session", is higher in *Professional* (2.87 words) as compared to the 'professional' words in *Private* (2.65 words, t-test: $p=0.058$).[19] While this standard priming check can only provide indicative evidence, it suggests that the *Professional* treatment was able to make the professional identity more salient compared to the *Private* treatment.

We now turn to examining the risk choices of scientists. For this purpose, we assign the scale in the first column of Table 2 to the risk choices, ranging from 1 (no-risk choice) to 6 (risk-seeking choice). As there is a qualitative difference

[14] In line with the previous literature (Eckel and Grossman, 2008), the table shows strict inequalities and excludes CRRA-parameters for which two neighboring gambles yield the same expected utility. This allows relating preferences to non-overlapping CRRA-parameter ranges.

[15] The first two of the eight word fragments ("_ a l k" and "_ o o k") had no unambiguous professional science interpretation. These two were meant as an easy start for participants and served, following Cohn et al. (2014, 2017), the purpose of disguising the purpose of the task. The other word fragments were: " _ i s _", "_ _ s s i o n", "c o _", "_ _ o c k" as well as " _ _ p e r".

[16] We also offered the possibility to donate fractions (in 10% steps) of their earnings to the charity 'Doctors Without Borders'. This option was not pre-announced and the donation decision could not have influenced risk-taking behavior.

[17] Overall, 946 individuals clicked on the link to our study. Supplementary investigates potential response bias and balance across treatments. Males appear to drop out of *Professional* more frequently than out of *Private*. This attrition by males does not seem to be problematic for our subsequent findings. For example, hypothetically adding males to the *Professional* treatment such that balance is achieved does not alter our Result 1 on the gender gap in risk-seeking even if we increase the share of risk-seekers in the added group from the current level of 28% to up to 48%.

[18] We dropped 10 observations because they responded more than once as well as one observation because we could identify her as still being a master student.

[19] All p-values reported in this paper are based on two-sided tests.

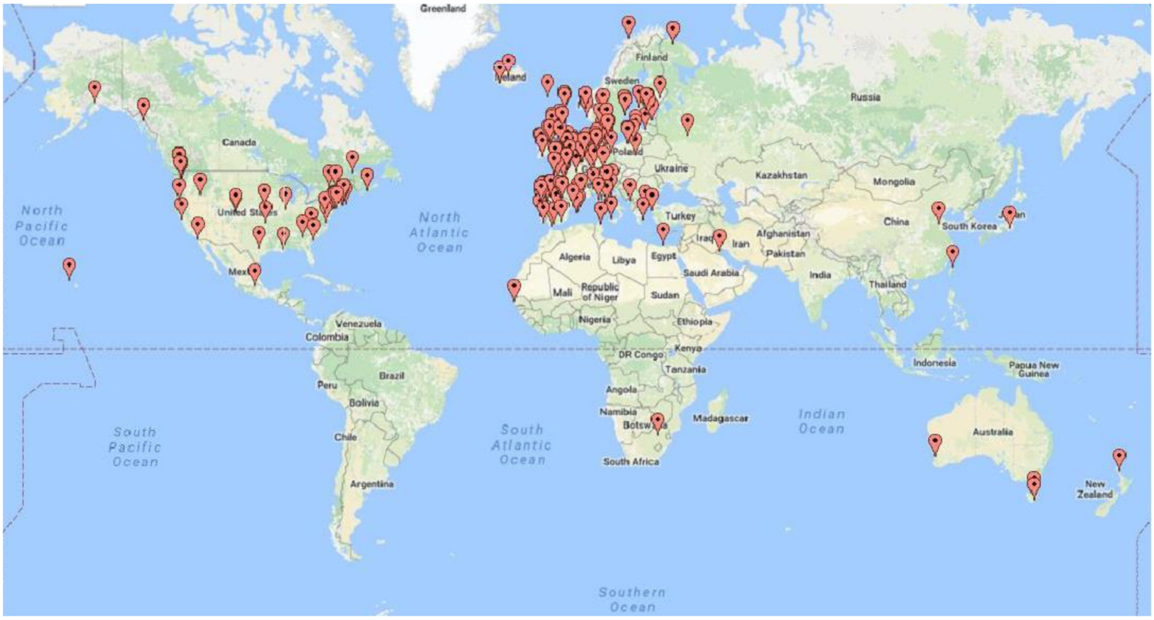

Fig. 1. Map of the world with the locations of our subjects depicted as (red) balloons.

between risk choices 1–5 (indicating risk-averse behavior) and choice 6 (indicating risk-seeking), and since risk-seeking behavior may be an important trait of scientists, we also examine risk-seeking behavior more closely.[20] We observe in particular that there is, with 29.96%, a substantially higher frequency of risk-seeking choices of scientists as compared to other subject pools (Dave et al., 2010; Khadjavi, 2018). This unusual proportion of risk-seeking individuals could be taken as an indication that academia involves taking considerable risks.

To provide a comprehensive picture while accounting for observed co-variates, we examine the interaction of gender and identities using regression analyses. The equation to be estimated is

Risk choices = Constant + Female (dummy, female=1) + *Private* treatment (dummy, *Private* =1) + Female * *Private* [+ controls] + error term,

where the constant represents the impact from males in *Professional*.

Table 3 presents the results of Tobit (from risk choices 1–6)[21] and Logit (the likelihood of a risk-seeking decision by choosing option 6) estimations for decision-making under risk, respectively, in terms of average marginal effects for the categories of interest. We complement our simple regression analysis by including controls for other observable characteristics when estimating the impact of our treatments on risk choices (columns 2 and 4). Besides the risk choices, the gender and our identity priming treatments, we collected data on age, whether the subject has tenure and on the location of the subject for both treatments.

First, unlike Cadsby et al. (2013), we find that, on average, females are more risk-averse (Columns 1 and 2, row indicated "Females") and less risk-seeking (Columns 3 and 4, row indicated "Females") as compared to males. These findings are in line with our model and confirm the first stylized finding from the survey of Croson and Gneezy (2009) that females appear to be more risk-averse than males.

Second, we do not find a general treatment effect for risk choices between *Professional* and *Private* (see row indicated *Private* treatment). Our calibrated model would suggest that only females change their behavior to exhibit more risk-taking behavior in *Professional* as compared to *Private*. The relatively smaller number of females in our sample may prevent a general treatment effect for *Professional* to be detectable. Overall, scientists thus do not seem to take greater risks in the professional environment.

Third, regarding cross-treatment differences, we only find a difference for females in risk-seeking. However, this effect vanishes if controls are included (second block of results in Table 3).

Third, we find that the difference in risk-taking or risk-seeking between males and females is larger in *Private* than in *Professional* (third block of results in Table 3). In *Private*, females act less risk-seeking or risk-taking compared to males (sec-

[20] The only other study on professional identity priming and risk-taking behavior, Cohn et al. (2017), elicited risk-taking behavior using the Gneezy and Potters (1997) investment task without the possibility to distinguish risk-averse from risk-seeking behavior.

[21] As we only have information on the risk choices in terms of six categories, using an ordered logit model is an obvious alternative to a tobit model. We present result from a tobit model as main specification for ease of interpretation. Table C.2 in the Supplementary reports estimated coefficients for a tobit, an ordered logit as well as an OLS specification. Direction and significance of effects are similar across the different specifications.

Table 3
Average marginal effects based on regression analysis of risk choices without and with controls.

	Tobit regression Dependent variable: Risk choice (from 1 to 6)		Logit regression Dependent variable: Risk-seeking choice (=1 if risk-seeking, else 0)	
	(1)	(2)	(3)	(4)
Female	−1.430***	−1.569***	−0.193***	−0.186***
(dummy, female=1)	(0.000)	(0.000)	(0.000)	(0.000)
Private treatment	0.0662	−0.0252	0.00154	0.000909
(dummy, *Private*= 1)	(0.826)	(0.935)	(0.970)	(0.983)
Private treatment				
Male	0.524	0.415	0.0737	0.0646
	(0.183)	(0.300)	(0.204)	(0.280)
Female	−0.582	−0.649	−0.101*	−0.0883
	(0.217)	(0.183)	(0.063)	(0.117)
Female				
Professional treatment	−0.936**	−1.115***	−0.115**	−0.120**
	(0.019)	(0.006)	(0.037)	(0.032)
Private treatment	−2.042***	−2.178***	−0.289***	−0.273***
	(0.000)	(0.000)	(0.000)	(0.000)
Controls	No	Yes	No	Yes
Observations	474	457	474	457

Note: Professional is the baseline of the estimations. The lower [upper] limit of the Tobit is 1 [6]. p-values in parentheses, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Controls include age (continuous), a dummy that equals one if the person is tenured, and a dummy that equals one if the person lives in Europe.

ond block of results in Table 3). While this is also the case in *Professional*, the gender gap is significantly smaller. Specifically, the differences in gender gaps in risk-taking and risk-seeking between *Private* and *Professional* are significant at the 5 or 10% level.[22]

We summarize

Result 1. The gender gap in risk-taking and risk-seeking is reduced when the professional identity is salient compared to the private identity.

This result is in line with our calibrated model and the second stylized finding on the reduced gender gap in risk-taking in professional contexts summarized by Croson and Gneezy (2009). It is also in line with the model's prediction that the gender gap in risk-taking reduces when the professional identity is salient. It confirms and extends the findings of Cadsby et al. (2013).

Furthermore, our Tobit regression in Table C.2 in the Supplementary reports a negative correlation between age and risk-taking.[23] In light of this effect, we now examine whether our treatment effects are stronger for older, tenured scientists compared to younger scientists. Fig. 2 provides a graphical overview of how mean risk-taking behavior of males (blue lines) and females (red lines) changes with age across the two treatments for the age range 25–65. In *Private* we observe the common result that older individuals are less risk-seeking than younger ones (Dohmen et al., 2017; Grubb et al., 2016; Mather et al., 2012). We observe this direction almost in parallel for females and males in *Private* without any mentionable overlap. Interestingly, for females in *Professional* we find that choices become more risky with age, in contrast to the general decline with age. Spliting the sample at median age (42 years), we find that junior and senior male scientists take greater risks than junior and senior female scientists respectively in *Private* (Mann–Whitney tests, $p = 0.000$ and $p = 0.029$). While we also find the usual gender gap in risk-taking for junior scientists in *Professional* (Mann-Whitney test, $p = 0.001$), the gender gap is completely closed for senior scientists ($p > 0.70$).[24] Likewise, as age and tenured are highly correlated, we find no gender gap in risk-taking for tenured scientists in *Professional* (Mann-Whitney test, $p > 0.40$).

We additionally test our results on age and the gender gap by including a dummy for young scientists (=1 if younger than 42) into the regression model. Table 4 presents results in terms of marginal effects for the different groups. It shows that the gender gap is largest for young scientists in *Private*, i.e. females act significantly more risk-averse than males, while there is no significant difference between males and females for the old(er) age group in *Professional*.

[22] The p-values based on a Z-statistic for the differences in coefficients reported in Table 3 for Females (relative to Males) for *Professional* vs. *Private* for the different models are 0.072, 0.091, 0.028, and 0.063, respectively.

[23] Controlling for age in a more flexible way does not affect our main result. Of course, it should be noted that we only observe an association between age and risk-taking for different individuals at a single point in time. We therefore cannot disentangle whether the observed effect in terms of risk-taking is due to age, cohort effects, or academic seniority, which are all highly correlated.

[24] The closing of the gender gap in risk-taking for senior scientist if the professional identity is salient holds for risk-taking, but also for risk-seeking behavior.

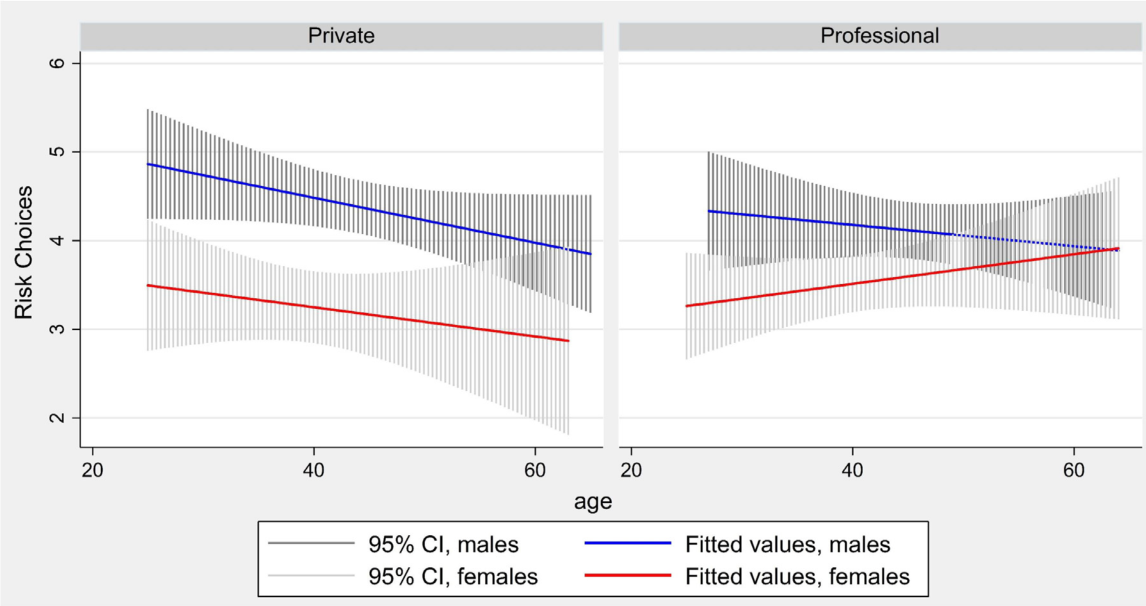

Fig. 2. Risk Choices by treatment, gender and age. *Note*: The lower-red (upper-blue) lines depict the linear fits for females (males).

Table 4
Regression results, marginal effects.

Dependent variable:	Tobit regression Risk choice (from 1 to 6)
Female	
old, *Professional*	−0.207
	(0.724)
young, *Professional*	−1.656***
	(0.003)
old, *Private*	−1.666**
	(0.025)
young, *Private*	−2.560***
	(0.000)
Observations	474
Adjusted R^2	

Note: Professional is the baseline of the estimations. The lower [upper] limit of the Tobit is 1 [6], no controls included. *p*-values in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

These comparisons yield
Result 2. The gender gap in risk-taking is closed for senior scientists when the professional identity is salient.

4. Replication and extension study

Following our own curiosity and the comments of the editorial team, we decided to extend our main study to further delve into the private identity of scientists. To this end, we ran an extension and replication online field experiment with three treatments and more than 1000 scientists from around the world and across very different scientific subdisciplines.

4.1. Design

Since our *Private* treatment may consist of priming that relates to both gender and private life more generally, we designed a pre-registered second experiment to disentangle these two effects.[25] Specifically, in this replication and extension

[25] RCT ID: AEARCTR-0004023, available at https://doi.org/10.1257/rct.4023–2.0. Note that this pre-analysis plan also contains two companion experiments.

Table 5
Identity priming questions for three treatments.

Professional	*Private*	*Gender*
Who is your current employer?	What is your current city of residence?	What is your gender?
How many years have you worked for this employer?	How many years have you lived in your current accommodation?	How many weeks ago did you last meet with an all female or all male group?
How large is your direct working team (yourself included)?	How large is your circle of close friends (yourself included)?	How large is your direct family (yourself included)?
Where did you last go to for a conference/workshop?	Where did you last go on holiday?	Where did you last meet with your extended family?
Do you coordinate your work hours with your colleagues?	Do you coordinate your work hours with your close friends?	Do you need to coordinate your work hours with your family obligations?
How satisfied are you with your professional life in general? (1–9)	How satisfied are you with your private life in general? (1–9)	How satisfied are you with your family role in general? (1 to 9)
What part of your work do you enjoy the most? (bullet points are sufficient)	What part of your leisure time do you enjoy the most? (bullet points are sufficient)	What do you like most about your appearance? (bullet points are sufficient)
How many hours per week do you usually spend in the office?	How many hours per week do you usually sleep?	How many hours per week do you usually spend on housework?
What is your favorite academic journal?	What is your favorite newspaper?	What is your favorite lifestyle or sports magazine?

study, we allocated participants to three different treatments: *Professional, Private* and *Gender*. For *Professional* and *Private* we retained previous questions that we deemed unrelated to gender. We otherwise devised the questions again such that they would be as similar as possible across treatments while picking up treatment-specific associations. Priming questions for the three treatments are provided in Table 5. Priming was followed again by three experimental tasks, where the first two—risk-taking and truth-telling—were the same as in the main study, and the third task was a donation task. The expected pay-off for participating was $25.

After the three experimental tasks, we used the following words for the word-completing task to test whether the priming was successful: '_alk', '__nday', 'journ__', 'gr__t', '__ssion', '__per', and '_ook', with '_alk' and '_ook' serving to disguise the purpose of the task as before. Given the much broader science population of this second study (described below), some of these words are somewhat less specific compared to our main, first study. Since we could not devise a format that would test priming across all treatments, this priming check was designed to pick up difference between the *Professional* and *Private* treatments for five word completions, yet we also use it to examine potential differences between *Professional* and *Gender*.[26] From these five word completions, we sum up words associated with professional life.

As we had previously already targeted all members of the international scientific organization focusing on the marine environment in our main study and to see whether our results may generalize, we decided to target a more general science population in our second, extension study. To this end, we recruited subjects from the pool of corresponding authors on the science literature platform SCOPUS. Specifically, we sent out invitation e-mails for participation in our study to a random sample of corresponding authors of published papers in the year 2017 from eight different scientific subjects, with two subjects from each of the four major science categories life sciences, social sciences, health science and physical sciences, as categorized by SCOPUS.[27] In addition, we targeted a random sample of corresponding authors of 2017 publications in Science, Nature and PNAS.

Given our budget constraint and expected costs, we aimed for 1080 participants, with 432 in *Professional*, 432 in *Private* and 216 in *Gender*.[28] We collected twice the number of observations in *Professional* and *Private* to examine how well the results from the main study replicate with a larger, geographically and scientifically much more diverse sample.[29] Within each science area, we randomly allocated e-mail addresses to the three treatments and invited subjects until the slots were filled, that is 48 slots in *Professional* and *Private*, and 24 in *Gender* per field. We sent out invitations at different times of the day to acquire a more balanced representation across time zones around the world. Given that one of our word completion tasks contains âSundayâ and âMondayâ as solutions, we only sent out invitations to our study on Wednesdays and Thursdays. We initially sent out invitations on March 20, 2019 and closed the study on April 10, 2019.

We now state our hypotheses and refer the reader for more details on how we specified the analysis including statistical testing to the pre-analysis plan.

[26] Exemplary solutions to the five incomplete words are: 'Monday' vs. 'Sunday', 'journal' vs. 'journey', 'grant' vs. 'great', 'session' vs. 'passion/mission', 'paper' vs. 'super/upper'.

[27] These scientific subjects are (1) Biochemistry, Genetics, and Molecular Biology (2) Pharmacology, Toxicology, and Pharmaceutics, (3) Economics, Econometrics, and Finance, (4) Psychology, (5) Medicine, (6) Nursing, (7) Environmental Sciences and (8) Physics and Astronomy.

[28] We ended up collecting 1083 observations with payouts, as simultaneous answers led to a "too late" closure of the survey.

[29] A further reasons is that the two companion studies were targeted towards this comparison.

#	Formulation of the hypothesis
0a	The mean number of word completions in the *Professional* column is greater in the *Professional* treatment than in the *Private* treatment
0b	The mean number of word completions in the *Professional* column is greater in the *Professional* treatment than in the *Gender* treatment
1a	Females act more risk-averse than males
1b	Females act less risk-seeking than males
2	The gender gap in risk-taking is larger in *Gender* than in *Private*
3	The gender gap in risk-taking is smaller in *Professional* than in *Private*
4a	Females in *Professional* act less risk-averse relative to *Private* if the percentage of females working in the department is relatively high
4b	Males in *Professional* act more risk-averse relative to *Private* if the percentage of females working in the department is relatively high
5	The gender gap in risk-taking in *Professional* vanishes for older or tenured scientists

4.2. Results

Analogous to Study 1, we first report the results for hypotheses 0a and 0b on the success of the priming as measured by the word-completion task. For the five incomplete words described above, we find mostly insignificant results. In four out of five cases, i.e. for Monday vs. Sunday, grant vs. grant, session vs. passion/mission and paper vs. super/upper, we find no significant difference between the prevalence of professional word completion in *Professional* and the other two treatments. The only exception is that there are more completions of journal vs. journey in *Professional* compared to both *Private* ($p = 0.000$) and *Gender* ($p = 0.019$). There is no difference in testing the mean number of professional words in *Professional* against *Private* (2.59 vs. 2.51, $p = 0.320$), but there is a difference in the mean number of professional words between *Professional* and *Gender* (2.59 vs. 2.34, $p = 0.014$). While our set-up was not designed to test for priming differences between *Professional* and *Gender*, this findings suggests that *Gender* identity plays a distinct role in the non-professional identity setting. However, we overall find mixed evidence that, unfortunately, does not consistently show that our priming has worked to change the salience of the different identities in Study 2. Given that the priming check indicates that the priming did not work as well as in Study 1, the results in Study 2 have, accordingly, to be interpreted with caution.

Table 6 presents the results.[30] We confirm Hypotheses 1a and 1b: Female scientists act more risk-averse and less risk-seeking than males (Table 6, row indicated by “Female”). However, we do not find a significant difference in the gender gap between the different scenarios (based on a Z-test statistic), thus failing to confirm Hypotheses 2 and 3. The differences in risk-taking or risk-seeking of females relative to males are reported in the second block of results in Table 6. Across treatments, i.e. *Private* and *Gender* relative to *Professional*, we generally do not find significant differences within gender groups (third and fourth block of results in Table 6). One exception is the result that males act significantly more risk-averse in *Gender* relative to *Professional*. This result is opposite to our expectations based on previous results. Also, for Hypotheses 4 and 5 we do not find clear results.[31]

As our priming exercise has not worked well and since professional norms may differ between different scientific fields and cultures, the results are difficult to interpret. As we had not planned for the case that priming would not work, we decided to run an exploratory analysis in Section 4.3 by focussing on a subset of responses that are more similar to those in Study 1.

4.3. Comparison of the two studies

While Study 2 reproduces the result that female scientists act more risk-averse and less risk-seeking than males, it does not detect a difference in the gender gap in risk-taking between the different treatments. The results from the word-completion task suggest that the priming did not work across the *Professional* and the *Private* identity treatments in Study 2. It may be that the participants from the more diverse subject pool of scientists in Study 2 did not react in the same way or did not relate to the priming questions as much as the participants from the more European and Northern American subject pool in Study 1. Overall, Study 2 does not support the treatment effects of Study 1, yet Study 2 also does not provide evidence to reject our model and theory as the priming did not work.

To better understand what may drive the different outcomes resulting from the two studies, we compare both samples in terms of the scientists' characteristics that we have for both samples, i.e., whether scientists are male or female, tenured or not, living in Europe, North America or elsewhere, and how old they are. Table 7 shows that the two samples are relatively similar in terms of the age range covered and the share of tenured vs. non-tenured scientists. The table also shows that the samples differ in terms of the share of females and Europeans.

For our explorative analysis, we define a sub-sample from the sample of Study 2 with the aim to reproduce the sample of Study 1 in terms of the share of females, Europeans and Northern Americans as well as in terms of successful priming.

[30] Estimated coefficients from the tobit model as well as from an ordered logit and and OLS are presented in Table C.3.

[31] We present some results in Supplementary.

Table 6
Marginal effects based on regression analysis of risk choices without and with controls.

	Tobit regression Dependent variable: Risk choice (from 1 to 6)		Logit regression Dependent variable: Risk-seeking choice (=1 if risk-seeking, else 0)	
	(1)	(2)	(3)	(4)
Female	−0.113***	−0.115***	−0.740***	−0.780***
(dummy, female=1)	(0.000)	(0.000)	(0.001)	(0.001)
Private treatment	−0.0318	−0.0321	−0.153	−0.171
(categorical)	(0.300)	(0.295)	(0.512)	(0.462)
Gender treatment	−0.0444	−0.0458	−0.424	−0.458
(categorical)	(0.233)	(0.217)	(0.137)	(0.107)
Female				
Professional treatment	−0.163***	−0.167***	−1.050***	−1.120***
	(0.000)	(0.000)	(0.005)	(0.002)
Private treatment	−0.0829*	−0.0814*	−0.582	−0.576
	(0.066)	(0.071)	(0.104)	(0.106)
Gender treatment	−0.0745	−0.0758	−0.433	−0.505
	(0.232)	(0.223)	(0.390)	(0.315)
Private treatment				
Male	−0.0546	−0.0564	−0.286	−0.325
	(0.145)	(0.133)	(0.300)	(0.236)
Female	0.0258	0.0291	0.182	0.219
	(0.620)	(0.574)	(0.675)	(0.614)
Gender treatment				
Male	−0.0696	−0.0717	−0.599*	−0.633*
	(0.127)	(0.115)	(0.078)	(0.062)
Female	0.0191	0.0194	0.0179	-0.0181
	(0.759)	(0.754)	(0.973)	(0.972)
Controls	No	Yes	No	Yes
Observations	1083	1083	1083	1083

Note: Professional is the baseline of the estimations. The lower [upper] limit of the Tobit is 1 [6]. p-values in parentheses, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The treatment variable consists of three categories: Professional, Private and Gender. Controls include age (continuous), a dummy that equals one if the person is tenured, and a dummy that equals one if the person lives in Europe.

Table 7
Descriptive statistics.

	Study 1		Study 2	
	mean	sd	mean	sd
Female	.414	.492	.284	.451
Age	43.401	11.133	42.943	11.321
Tenured	.521	.500	.547	.498
Europe	.789	.408	.479	.450
North America	.190	.393	.166	.372

To do so, we define the sub-sample to only include Europeans and Northern Americans, scientists from the fields that have a relatively high share of females, i.e., from ‘nursing’ or ‘psychology’, and those individuals for which the word completion task suggests that priming has worked. We define successful priming if scientists in the *Professional* treatment came up with at least 3 words related to the professional context, and if scientists in the *Gender* or *Private* treatment came up with at most 3 words related to the professional context. With this procedure, we end up with 111 observations in *Professional*, 105 in *Private*, and 68 in *Gender*.

We report the results in Table 8. We find that there is no gender gap in risk-taking in *Professional* and none in *Private*, but that there is a gender gap in risk-taking in *Gender* (second block of results). This suggests that *Gender* related identity aspects entangled in the *Private* identity treatment in Study 1 may be the driver of the different gender gaps across the *Professional* and *Private* identity treatments in Study 1. Results are similar between *Professional* and *Private* in this exploratory analysis and only the gender gaps in risk-taking and risk-seeking between *Private* and *Gender* are significantly different from each other in most cases (with p-values 0.127, 0.046, 0.098, 0.047 for the four different specifications reported in Table 8). Since the sample size is now considerably smaller, we do not examine e.g., age effects on the gender gap in risk-taking.

Table 8
Marginal effects based on regression analysis of risk choices without and with controls, reduced sample.

	Tobit regression Dependent variable: Risk choice (from 1 to 6)		Logit regression Dependent variable: Risk-seeking choice (=1 if risk-seeking, else 0)	
	(1)	(2)	(3)	(4)
Female	−0.0892	−0.0917*	−0.751*	−0.836*
(dummy, female=1)	(0.103)	(0.091)	(0.084)	(0.054)
Private treatment	0.00720	−0.0180	0.0204	−0.246
(categorical)	(0.906)	(0.767)	(0.964)	(0.585)
Gender treatment	0.0240	0.00944	−0.376	−0.537
(categorical)	(0.723)	(0.888)	(0.470)	(0.295)
Female				
Professional treatment	−0.0724	−0.0975	−0.725	−0.979
	(0.417)	(0.274)	(0.296)	(0.153)
Private treatment	−0.00914	0.0166	−0.0637	0.170
	(0.920)	(0.855)	(0.925)	(0.800)
Gender treatment	−0.240**	−0.253**	−1.854*	−2.155**
	(0.024)	(0.014)	(0.053)	(0.025)
Private treatment				
Male	−0.0117	−0.0523	−0.177	−0.590
	(0.875)	(0.484)	(0.747)	(0.284)
Female	0.0515	0.0617	0.483	0.559
	(0.618)	(0.550)	(0.544)	(0.478)
Gender treatment				
Male	0.0742	0.0564	−0.0375	−0.185
	(0.377)	(0.510)	(0.950)	(0.757)
Female	−0.0937	−0.0992	−1.167	−1.361
	(0.396)	(0.344)	(0.252)	(0.176)
Controls	No	Yes	No	Yes
Observations	284	284	284	284

Note: Professional is the baseline of the estimations. The lower [upper] limit of the Tobit is 1 [6]. p-values in parentheses, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The treatment variable consists of three categories: Professional, Private and Gender. Controls include age (continuous), a dummy that equals one if the person is tenured, and a dummy that equals one if the person lives in Europe.

5. Discussion and conclusion

Our research has investigated drivers of the gender gap in risk-taking by experimentally varying the salience of the professional identity in two online field experiments. To this end, we have focused on a particular labor market: Academia. Our results, based on risk task decisions of more than 1500 scientists around the world, confirm a gender gap in risk-taking for scientists: On average, females are more risk-averse than males. The identity-priming intervention in our main study reveals that more females make risk-taking and risk-seeking decisions when their professional identity is salient as compared to the private identity, while this is not the case for males. Thus, risk choices of females and males are more similar when the professional identity is salient, implying that the general gender gap in risk-taking is reduced. This confirms and extends previous findings by Cadsby et al. (2013). We further show that the gender gap in risk-taking is closed for senior scientists when the professional identity is salient.

Study 2 was aimed at replicating the results from Study 1 and to extend our understanding of what exact private identity, a gender-based private identity or a non-gender-based private identity contributes to this treatment effect. Unlike in Study 1 where the word-completion task showed significant and successful priming, the word-completion task results in Study 2 did not show that priming had worked successfully between the private and professional identity treatments. Reasons may include characteristics of the subject pool and changes in the priming questions (to fit the new experimental design). Consequently, our treatment effects remain insignificant in Study 2 and the results of Study 2 are inconclusive.

In an exploratory analysis, we restrict our sample in Study 2 to those scientists for which priming seems to have worked and further mirror the high share of Europeans and the relatively high share of females that participated in Study 1. This exploratory analysis suggests that the non-professional gender related identity aspects may be the key driver of the observed gender gap in risk-taking that is reduced in a professional context.

Our findings paint a more nuanced picture regarding the drivers of the gender gap in risk-taking. Our results can neither be explained by a simple story of sorting into risky occupations based on general risk preferences, nor by the hypothesis that the work environment shapes the general risk-preferences. A simple story of sorting into risky occupations would—contrary to what we observe—imply similar risk-taking of female scientists across treatments that is close to male scientists. Instead, female scientists remain relatively risk-averse in the private identity treatment in Study 1, and in the gender identity treatment in Study 2. The hypothesis that the work environment shapes general risk-preferences would also imply similar risk-taking of females across treatments. In particular, one would expect that older females would also show more risk-taking behavior in the private identity treatment. Still, we observe that risk-taking behavior differs between treatments in Study 1, and older female scientists in the private identity treatment remain relatively risk-averse. Instead, our findings can be viewed as being consistent with the literature on sorting (Buser et al., 2014; Flory et al., 2015) in a more subtle way: Our evidence suggests a sorting that depends on identity considerations. Indeed, our results indicate that those females stay in academia who can adapt better to the prevailing professional norms and behavioral modes of higher risk-taking in a professional environment, while still complying with âfemaleâ norms in the private context. This also relates in a nuanced way to recent findings that agents adapt their behavior to different environmental contexts related to the hypothesis that the work environment shapes preferences (Booth and Nolen, 2012; Cohn et al., 2017; Gneezy et al., 2009; Leibbrandt et al., 2013).

While Study 1 provides some causal evidence regarding our priming treatments and related gender effects, our cross-sectional datasets are not able to clearly identify the exact mechanism leading to the gender gap for non-tenured scientists and its disappearance for older or tenured scientists. One would ideally trace subjects over a long time horizon to measure how their risk-taking changes and how this impacts the gender gap while (female) scientists select in or out of academia. We leave the question on gender differences in assigning weights to the different identities, and how this may change with age, to future research. For instance, Cinamon and Rich (2002) identify three different profiles for the family-work importance: Work is more important, family is more important, both are very important. Bénabou and Tirole (2011) develop a dynamic model of identity and one could use their idea of investing into certain identities when analyzing long-term panel data. We leave the inter-linkages of family planning, identity considerations and career planning for future research. Both aspects may hold important additional insights that our data is not able to speak to.

Our results have important implications for policies aimed at closing gender gaps in labor markets.[32] First, our analysis suggests that one policy approach for attenuating or closing the gender gap in risk-taking would be to change prevailing norms and expected behaviors. One could try to change either the professional science context or the private context. Changing prevailing norms, however, may be a rather slow process. Furthermore, risk-taking may also be warranted in science since uncertainty is inherent in the research process. Also from a social plannerâs perspective, it may not be optimal to induce norms of lower risk-taking in academia, as groundbreaking research is very valuable and often entails taking a considerable amount of risk.[33] Thus, besides changing prevailing norms in science, other mechanisms could be developed that would allow for better coping strategies, for example based on risk-diversification or risk-sharing that allow a more risk-averse person to undertake risky projects.

Second, our results indicate that those females tend to stay in science that are good at adapting to the prevailing norms of a given context and that may thus be adept at ‘switching’ their identities from a private to the professional context, all else equal. While part of this ability may be related to an individual’s personality, it may still be possible to learn or improve such a skill.

Finally, and relatedly, while our identity model assumed that the prevailing professional norm is known, this may not be the case in particular for younger scientists. The prescribed behavior thus has to be discovered before the issue of compliance with the norm arises. This may matter especially for females, as the prescribed behavior in science seems to differ from the prescribed behavior for females in private settings more strongly than for males. These considerations suggest that programs focusing on facilitating a better understanding of the prevailing norms and expected behaviors in the professional context, and how to find and shape your own way of acting in a professional context, might be very helpful. Policy approaches to address these issues include, for example, mentoring programs, which have been shown to have a significant impact on the academic performance of participating females (Blau et al., 2010).[34]

Declaration of Competing Interest

I hereby confirm that neither I nor any of my co-authors have a conflict of interest related to this manuscript.

Acknowledgments

We are grateful to Billur Aksoy, Thomas Buser, Ilyana Kuziemko, Andreas Leibbrandt, Lukas Menkhoff, Martin Quaas, Eva Ranehill, Ulrich Schmidt, Renate Schubert, Roel van Veldhuizen, two anonymous reviewers and Co-Editor Laura Schechter as well as to seminar audiences at EEA 2017, ETH Zurich, and the University of Duisburg-Essen for helpful comments. We thank the participating scientists, the administrative office of the science organization for providing e-mail address data, Joern Schmidt for handling e-mail correspondence, Olaf Bock and his team at the Experimental Lab in Hamburg for help in administering the experiment, as well as Pia Foerster, Nico Ohlenmacher, Clara Paczkowski and Lasse Schmidt for research

[32] See Ranehill and Weber (2017) for a recent study on how gender preference gaps impact policy outcomes.

[33] For example, Rzhetsky et al. (2015) analyze millions of papers and patents on chemical relationships in biomedicine and conclude that increased risk-taking by scientists would considerably speed up the generation of new scientific discoveries.

[34] Examples in the economics context on such initiatives include, for example, the Standing Committee on Women in Economics of the European Economic Association, or the Committee on the Status of Women in the Economics Profession of the American Economic Association.

assistance. This work was supported by the German Ministry of Education and Research [grant 01UT1410] and by the Cluster of Excellence 80, which is funded by the German Research Organization (DFG) on behalf of the German federal and state governments.

Supplementary material

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.jebo.2019.12.020.

References

Adams, R.B., Funk, P., 2012. Beyond the glass ceiling: does gender matter? Manag. Sci. 58 (2), 219–235.
Akerlof, G.A., Kranton, R.E., 2000. Economics and identity. Q. J. Econ. 115 (3), 715–753.
Atkinson, S.M., Baird, S.B., Frye, M.B., 2003. Do female mutual fund managers manage differently? J. Financ. Res. 26 (1), 1–18.
Azmat, G., Petrongolo, B., 2014. Gender and the labor market: what have we learned from field and lab experiments? Labour Econ. 30 (C), 32–40.
Ball, S., Eckel, C., Heracleous, M., 2010. Risk aversion and physical prowess: prediction, choice and bias. J. Risk Uncertain. 41 (3), 167–193.
Beckmann, D., Menkhoff, L., 2008. Will women be women? Analyzing the gender difference among financial experts. Kyklos 61 (3), 364–384.
Bellstrom, K., 2015. GM's Mary Barra sets a Fortune 500 record for female CEOs. fortune.com/2015/06/29/female-ceos-fortune-500-barra/ [Accessed: 31.01.2017].
Bénabou, R., Tirole, J., 2011. Identity, morals, and taboos: beliefs as assets. Q. J. Econ. 126 (2), 805–855.
Benjamin, D.J., Choi, J.J., Strickland, A.J., 2010. Social identity and preferences. Am. Econ. Rev. 100 (4), 1913–1928.
Bertrand, M., 2011. New perspectives on gender. In: Volume 4 of Handbook of Labor Economics, chapter 17. Elsevier, pp. 1543–1590.
Binswanger, H.P., 1981. Attitudes toward risk: theoretical implications of an experiment in rural India. Econ. J. 91 (364), 867–890.
Blau, F.D., Currie, J.M., Croson, R.T., Ginther, D.K., 2010. Can mentoring help female assistant professors? Interim results from a randomized trial. Am. Econ. Rev. 100 (2), 348–352.
Blau, F.D., Kahn, L.M., 2017. The gender wage gap: extent, trends, and explanations. J. Econ. Lit. 55 (3), 789–865.
Boll, C., Rossen, A., Wolf, A., 2016. The EU gender earnings gap: job segregation and working time as driving factors. HWWI Research Papers 176, Hamburg Institute of International Economics (HWWI).
Booth, A.L., Nolen, P., 2012. Gender differences in risk behaviour: does nurture matter? Econ. J. 122 (558), 56–78.
Bordalo, P., Coffman, K., Gennaioli, N., Shleifer, A., 2016a. Stereotypes. Q. J. Econ. 131 (4), 1753–1794.
Bordalo, P., Coffman, K.B., Gennaioli, N., Shleifer, A., 2016b. Beliefs about gender. NBER Working Papers 22972, National Bureau of Economic Research, Inc.
Bursztyn, L., Fujiwara, T., Pallais, A., 2017. 'Acting wife': marriage market incentives and labor market investments. Am. Econ. Rev. 107 (11), 3288–3319.
Buser, T., Niederle, M., Oosterbeek, H., 2014. Gender, competitiveness, and career choices. Q. J. Econ. 129 (3), 1409–1447.
Cadsby, C.B., Servátka, M., Song, F., 2013. How competitive are female professionals? A tale of identity conflict. J. Econ. Behav. Org. 92 (C), 284–303.
Charness, G., Gneezy, U., 2012. Strong evidence for gender differences in risk taking. J. Econ. Behav. Org. 83 (1), 50–58.
Cinamon, R.G., Rich, Y., 2002. Gender differences in the importance of work and family roles: implications for work–family conflict. Sex Roles 47 (11), 531–541.
Cohn, A., Fehr, E., Maréchal, M.A., 2014. Business culture and dishonesty in the banking industry. Nature 516 (7529), 86–89.
Cohn, A., Fehr, E., Maréchal, M.A., 2017. Do professional norms in the banking industry favor risk-taking? Rev. Financ. Stud. 30 (11), 3801–3823.
Cohn, A., Maréchal, M.A., 2016. Priming in economics. Curr. Opin. Psychol. 12, 17–21.
Croson, R., Gneezy, U., 2009. Gender differences in preferences. J. Econ. Lit. 47 (2), 448–474.
Dave, C., Eckel, C., Johnson, C., Rojas, C., 2010. Eliciting risk preferences: when is simple better? J. Risk Uncertain. 41 (3), 219–243.
Dohmen, T., Falk, A., Golsteyn, B.H.H., Huffman, D., Sunde, U., 2017. Risk attitudes across the life course. Econ. J. 127 (605), F95–F116.
Drupp, M. A., Khadjavi, M., Voss, R., 2017. Do scientists tell the truth? Evidence from a field experiment. Mimeo, Kiel University, Kiel.
Dwyer, P.D., Gilkeson, J.H., List, J.A., 2002. Gender differences in revealed risk taking: evidence from mutual fund investors. Econ. Lett. 76 (2), 151–158.
Eckel, C.C., Grossman, P.J., 2002. Sex differences and statistical stereotyping in attitudes toward financial risk. Evol. Hum. Behav. 23 (4), 281–295.
Eckel, C.C., Grossman, P.J., 2008. Forecasting risk attitudes: an experimental study using actual and forecast gamble choices. J. Econ. Behav. Org. 68 (1), 1–17.
Filippin, A., Crosetto, P., 2016. A reconsideration of gender differences in risk attitudes. Manag. Sci. 62 (11), 3138–3160.
Flory, J.A., Leibbrandt, A., List, J.A., 2015. Do competitive workplaces deter female workers? A large-scale natural field experiment on job entry decisions. Rev. Econ. Stud. 82 (1), 122–155.
Gneezy, U., Leonard, K.L., List, J.A., 2009. Gender differences in competition: evidence from a matrilineal and a patriarchal society. Econometrica 77 (5), 1637–1664.
Gneezy, U., Potters, J., 1997. An experiment on risk taking and evaluation periods. Q. J. Econ. 112 (2), 631–645.
Grubb, M.A., Tymula, A., Gilaie-Dotan, S., Glimcher, P.W., Levy, I., 2016. Neuroanatomy accounts for age-related changes in risk preferences. Nat. Commun. 7, 13822.
Johnson, J., Powell, P., 1994. Decision making, risk and gender: are managers different? Br. J. Manag. 5 (2), 123–138.
Kahneman, D., 2011. Thinking, Fast and Slow. Farrar, Straus and Giroux, New York.
Khadjavi, M., 2018. Deterrence works for criminals. Eur. J. Law Econ. 46, 165–178.
Knights, D., Richards, W., 2003. Sex discrimination in UK academia. Gender Work Org. 10 (2), 213–238.
Leibbrandt, A., Gneezy, U., List, J., 2013. Rise and fall of competitiveness in individualistic and collectivistic societies. Proc. Natl. Acad. Sci. USA 110 (23), 1067–1101.
Mather, M., Mazar, N., Gorlick, M.A., Lighthall, N.R., Burgeno, J., Schoeke, A., Ariely, D., 2012. Risk preferences and aging: the certainty effect in older adults decision making. Psychol. Aging 27 (4), 801–816.
National center of education statistics, 2016. Characteristics of postsecondary faculty. https://nces.ed.gov/programs/coe/indicator_csc.asp [Accessed:31.01.2017].
Nelson, J.A., 2015. Are women really more risk-averse than men? A re-analysis of the literature using expanded methods. J. Econ. Surv. 29 (3), 566–585.
Nelson, J.A., 2016. Not-so-strong evidence for gender differences in risk taking. Femin. Econ. 22 (2), 114–142.
Niederle, M., Vesterlund, L., 2007. Do women shy away from competition? Do men compete too much? Quarterly Journal of Economics 122 (3), 1067–1101.
OECD, 2016. Gender wage gap. https://www.oecd.org/gender/data/genderwagegap.htm [Accessed: 31.01.2017].
Ranehill, E., Weber, R., 2017. Do gender gaps affect policy outcomes? Mimeo.
Rzhetsky, A., Foster, J.G., Foster, I.T., Evans, J.A., 2015. Choosing experiments to accelerate collective discovery. Proc. Natl. Acad. Sci. USA 112 (47), 14569–14574.
Schubert, R., Brown, M., Gysler, M., Brachinger, H.W., 1999. Financial decision-making: are women really more risk-averse? Am. Econ. Rev. 89 (2), 381–385.
van Veldhuizen, R., 2016. Gender differences in tournament choices: risk preferences, overconfidence or competitiveness? Discussion Papers, Research Unit: Market Behavior SP II 2016-207, Social Science Research Center Berlin (WZB).
West, J.D., Jacquet, J., King, M.M., Correll, S.J., Bergstrom, C.T., 2013. The role of gender in scholarly authorship. PloS One 8 (7), e66212.