Converging evidence for greater male variability in time, risk, and social preferences

Christian Thöni[a,1] and Stefan Volk[b]

[a]Faculty of Law, Criminal Justice and Public Administration, University of Lausanne, 1015 Lausanne, Switzerland; and [b]University of Sydney Business School, University of Sydney, Camperdown NSW 2006, Australia

Edited by Elke U. Weber, Princeton University, Princeton, NJ, and approved April 22, 2021 (received for review December 27, 2020)

Gender differences in time, risk, and social preferences are important determinants of differential choices of men and women, with broad implications for gender-specific social and economic outcomes. To better understand the shape and form of gender differences in preferences, researchers have traditionally examined the mean differences between the two genders. We present an alternative perspective of greater male variability in preferences. In a meta-analysis of experimental economics studies with more than 50,000 individuals in 97 samples, we find converging evidence for greater male variability in time, risk, and social preferences. In some cases, we find greater male variability in addition to mean differences; in some cases, we only find greater male variability. Our findings suggest that theories of gender differences are incomplete if they fail to consider how the complex interaction of between-gender differences and within-gender variability determines differential choices and outcomes between women and men.

gender | greater male variability | preferences | meta-analysis

Individual preferences such as preferences for risk taking, patience, altruism, and trust are important determinants of human behavior. Much research has studied gender differences in preferences and their implications for differential social and economic outcomes of men and women (1–4). The extant literature has demonstrated that such preference differences constitute an important basis for differential choices and outcomes between the two genders in many domains including financial decision making, education, consumption, and the labor market (5). The foundations of gender differences in preferences are studied by a large body of theoretical and empirical work that has compared evolutionary origins (6, 7) with explanations emphasizing the role of the social environment (8).

Much of the recent debate on the origins of psychological gender differences has focused on sociocultural and environmental factors. Two prominent theories dominate this debate. Sociocultural theory (8) argues that gender differences result from a historic division of labor by gender because of men's greater strength and size and women's capacity to bear children. The higher status and wealth men achieved through their roles gave them power over women who accommodated to the restrictions imposed on the female gender role in society. Cognitive social learning theory (9) holds that gender differences emerge and are sustained by children imitating gender-appropriate behaviors and internalizing gender norms in society and then conforming to these behaviors and norms throughout their lives. According to both theories, psychological gender differences are largely the result of contextual factors and social roles men and women are assigned to in society. In contrast, evolutionary theorizing holds that, underneath social context and role-specific gender differences, more fundamental differences exist that have emerged from evolutionary selection and that are context independent (7). These more fundamental gender differences are rarely recognized, as they are often hidden in the tails of the distribution of many characteristics, traits, and skills, while most research on gender differences focuses on mean differences (10).

The aim of our meta-analysis is to uncover such hidden differences by studying gender differences in the variability of three fundamental economic preferences: time, risk, and social preference. The greater male variability (GMV) hypothesis posits that men are characterized by greater variability in many attributes because differentiation had survival value for men but not for women (11–13). According to this perspective, males and females of many species faced different adaptive problems that resulted in sex-differentiated psychological adaptations (14). Due to higher parental investments and fewer chances to reproduce, females of most species are more selective in mate choices and engage in less competition over mates. Males, on the other hand, invest less in parenting, are less selective in mate choice, and engage more in intrasexual competition for access to mates (15). Due to their higher selectivity and lower intrasexual competition, most females are able to attract a mate and reproduce. For males, on the other hand, those who deviate from the average on traits that are desired by females or that are otherwise important for intrasex competition will be likely to attract a mate and reproduce (11, 16).

On the other hand, deviating from the average often comes at the cost of fitness during harsh times. For example, male height is related to reproductive success in many species, but a tall body also requires more resources and is therefore less likely to survive during resource-poor episodes such as droughts or long winters (11, 17). Costly and resource intensive mating signals (e.g., the antlers of stags, the tails of peacocks) are widespread across species but can be a burden under demanding environmental conditions. During lean times, the stag with the small rack will beat the stag with the big rack because he will find it easier to move from one food source to another or to survive on fewer resources.

Significance

There is continuing interest in the study of gender differences in economic and social outcomes. An important factor underlying gender differences in outcomes are gender differences in fundamental economic preferences, which are at the core of many differential choices of women and men. We provide strong evidence for greater male variability in preferences. We find that men are more likely to have extreme time, risk, and social preferences, while women are more likely to have moderate preferences. With the focus on mean differences, the current literature underestimates the importance of gender differences and their effects on differential choices and outcomes between women and men.

Author contributions: C.T. and S.V. designed research; C.T. and S.V. performed research; C.T. analyzed data; and C.T. and S.V. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

[1]To whom correspondence may be addressed. Email: christian.thoeni@unil.ch.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2026112118/-/DCSupplemental.

Published June 4, 2021.

ECONOMIC SCIENCES

The adaptive problems faced by males therefore encouraged the development of more variability in traits and abilities, such as attractiveness and reproductive success versus robustness and individual longevity, than did the adaptive problems faced by females (18, 19). An alternative explanation for GMV assumes that there are risk-return trade-offs in the production of offspring. Because males can potentially sire many more offspring than females, the benefits of having a high return in the form of an exceptional male offspring outweigh the increased risk of having an inferior male offspring. According to this perspective, evolution would select a less risky process for the selective sex (females), while a more risky process is favored for the sex that is subject to strong intrasex competition for mating opportunities (20, 21).

Indeed, GMV has been found for a number of attributes including physical characteristics such as weight, height and blood parameters (12), brain structure (22), physical aggression (23), verbal and spatial performance (24), math performance (25), and intelligence (26). The fact that GMV can be observed already at birth suggests the existence of more fundamental gender differences that are independent of social role and context (12). In contrast, the existing literature on gender differences in preferences has exclusively focused on mean differences (1, 4). However, since GMV has been observed in behaviors such as cooperation (27) an important question is whether the preferences underlying these behaviors demonstrate similar variability. In line with our above arguments, preferences for extreme behaviors can represent costly mating signals. For example, preferences for high trust, fairness, altruistic, or risk-taking behaviors can communicate attractive male traits but can also have costly consequences, providing survival benefits for those with opposite preferences (28–31). This perspective suggests that the preferences of men demonstrate more evolved variability than the preferences of women who did not face similar adaptive problems. Similar predictions are made by perspectives based on risk-return trade-offs in the production of offspring (20, 21).

The existence of GMV in preferences would have significant implications given the importance of preference differences for differential choices and differential social and economic outcomes of the two genders. To test for GMV in individual preferences, we conducted a meta-analysis of individual participant data, which is a pooled analysis of multiple individual data sets (32). A traditional meta-analysis using summary statistics from published articles was not feasible, as most studies do not report measures of variability separately for the two genders. Our meta-analysis focused on decision-making experiments eliciting individual preferences which are published in leading economics journals. In 97 samples with more than 50,000 subjects, we find consistent evidence for GMV in time, risk, and social preferences.

Methods

Preference Measures. The field of experimental economics has developed a distinct set of methodological principles to elicit preferences. In contrast to stated preferences, often measured through hypothetical questions such as whether individuals are willing to take hypothetical risks, economic experiments elicit preferences through monetarily incentivized decisions. For example, preferences for risk taking are assessed through lotteries in which subjects can invest real money provided by the experimenter. By observing real behavioral choices in various incentivized decision situations, researchers can draw inferences about individuals' preferences. Although often conducted in highly stylized and anonymous environments, the literature suggests that these measures generalize to richer settings (33), and they have been shown to reliably reflect cross-cultural differences in preferences (34–36).

The last decades have seen a surge of experimental studies measuring preferences in relation to economically relevant decisions. Arguably, the three main dimensions of preferences that experimental and behavioral economists are concerned with are the following:

- Time preferences: How do people trade off costs and benefits over time?
- Risk preferences: How do people act in the face of risk?
- Social preferences: How do people evaluate their own payoffs in relation to the payoffs of others?

In our meta-analysis, we focus on studies using well-established and standardized elicitation methods for each preference category. Specifically, since we are interested in assessing gender differences in both central tendencies and variability, we focus on elicitation methods that offer subjects a broad spectrum of responses (i.e., more than binary strategy space). Furthermore, we restrict our selection to one-shot interactions that are not repeated with the same partner over time. This does not rule out that we consider multiple decisions by a subject. For example, for studies of time and risk preferences in which subjects make a series of allocation decisions (of which typically only one is relevant for their payoff), we consider all decisions. Our core inclusion criterion is that subjects received no feedback from previous outcomes of the particular game (e.g., about the realization of a random draw for risk preferences or another subject's decision for social preference measures).

In the following, we provide a brief description of each preference measure (for more details, reference *SI Appendix*, section S1*A*). In all tasks, respondents receive a monetary endowment, which they have to allocate between two options. For all studies, we normalize the endowment to one and denote the allocation by $y \in [0,1]$. For time preferences, we focus on the elicitation method introduced by Andreoni and Sprenger (37). Subjects decide how much of their endowment they prefer to receive early and how much they dedicate to be paid out at a later point in time with interest. For risk preferences, subjects decide how much of their endowment they want to invest in a risky lottery (38). For social preferences, we focus on three canonical situations involving two players. First, in the dictator game (39), subjects can share their endowment with another subject in the experiment (or, in some cases, with a charity). Second, in the ultimatum game (40), the first mover offers the second mover a split of the endowment. If the second mover accepts, the two players receive the money as proposed. If the second mover rejects, none of the players receives any money. Third, in the trust game (41), the first mover decides what fraction of the endowment to transfer to the second mover. The second mover receives three times the amount transferred and can decide how much of the money to send back to the first mover. For the ultimatum and trust game, we only consider first mover decisions.

Search Strategy. A review of the literature revealed that the vast majority of relevant studies does not report variability measures separately for men and women. Consequently, we had to restrict our analysis to studies for which we had access to raw data, permitting us to calculate variability measures directly. As a first step, we performed a power analysis to determine the number of observations required for testing our hypothesis. Effect sizes for variability are typically assessed by the variance ratio, which is the ratio between male and female variance in the respective measure. Previous studies of GMV, for example, in cooperation, mathematical performance, verbal skills, and spatial performance, report variance ratios between 1.03 and 1.30 (10, 27). In the current study, we aimed to detect variance ratios in the middle range of those reported in the literature and determined the required power to detect a variance ratio of 1.15. Our power analysis revealed that the estimated sample size for a two-sample variances test with a power of 0.80 and a type I error rate of 0.05 is ~3,220 observations.

As a next step, we identified a list of leading journals that publish studies relevant for our meta-analysis and performed a keyword search for the five preference measures (for details on the search strategy, reference *SI Appendix*, section S1*B*). We followed standard procedures (32) by randomly sampling from a homogeneous pool of studies applying the same standardized methodological principles of experimental economics. We included articles that met our selection criteria and for which we were able to obtain raw data through public repositories or via contacting the authors directly. The resulting data set constitutes our "primary sample" (37, 42–76). With the exception of the ultimatum game, our search comfortably exceeded the required sample size of our power analysis for all preference measures. In order to reach the required sample size for the ultimatum game, we collected additional data sets from a broader range of scientific journals. In this process, we also expanded our sample for the other preference measures. We refer to this extended data set as the "full sample" (77–101).

Results

Table 1 presents the meta-analytic coefficient estimates for the five preference measures. We report results for four indicators of gender differences in means and variability. To aggregate across samples, we estimate restricted maximum likelihood random effects models (102). Our main analysis is based on the full sample. All inferential results hold if we rely on the primary sample only (*SI Appendix*, Table S2). We first present results for time and risk preferences, followed by the findings for the three measures for social preferences.

Time Preferences. The first line of Table 1 shows the results for time preferences. Our data covers observations from nine samples with more than 8,500 individuals. The coefficients in the columns labeled means and variability are sample size–weighted coefficients, with 95% CIs in brackets. The first indicator, Prob(♂ > ♀), shows the probability that a male subject exhibits a more patient response than a female subject does (with ties resolved at random, reference *SI Appendix*, section S1*C* for details). The estimated sample size–weighted probability for time preferences is 0.50, suggesting the absence of any gender differences. This is confirmed by the second indicator for mean differences, Cohen's *d*, which is practically zero.

The two rightmost columns report two indicators for variability. The first is the risk ratio of extreme strategies. We define extreme strategies as either 1) allocating nothing of the endowment or 2) all of it to the later option (i.e., to be paid out later in time with interest). Throughout the text, we will be using the term extreme strategy to refer to strategies at the boundaries of the respective strategy space. It is important to point out that these strategies are extreme in the sense that they represent maximum or minimum available choices but not in the sense that they necessarily represent rare or irrational choices. The risk ratio is the relative frequency of extreme strategies observed among males divided by the relative frequency of extreme strategies among females. A risk ratio of one would indicate the absence of gender differences. The estimated risk ratio is 1.25, indicating that for every 100 women there are 125 men opting for either the most patient or most impatient option. The fourth indicator is the standard effect size measure in the variability literature, the ratio between male and female variance in the dependent variable. With an estimated variance ratio of 1.15, it ranks in the intermediate range of effect sizes found in other research on GMV such as cooperation, mathematical performance, and verbal skills as well as spatial performance (10). While the two effect size measures for mean differences are far from significant, both variability indicators show highly significant gender differences in support of the GMV hypothesis ($P < .001$). In *SI Appendix*, we report results separately for each of the nine samples (*SI Appendix*, Table S3) and provide forest plots for the meta-analyses for Cohen's *d* (*SI Appendix*, Fig. S1) as well as for the log variance ratio (*SI Appendix*, Fig. S2). The standard measures for between-sample heterogeneity (τ^2, I^2, and H^2) indicate substantial heterogeneity in means and variance ratios. However, despite the heterogeneity, we find variance ratios above one in eight of our nine samples. We will address the between-sample heterogeneity in more detail in *SI Appendix*, section S2*H*.

Fig. 1, *Left* visualizes the findings. Fig. 1, *Top* shows gender differences in means (full sample). Dotted lines indicate the admissible range from zero to one; horizontal error bars indicate 95% CIs. The absence of mean differences is confirmed by the almost identical CIs. The bars indicate the risk ratio of male to female relative frequencies across the strategy space. We merge the responses to five bins for ease of exposition. The two outermost bins contain only the responses zero or one. The three intermediate bins are sized such that they attract roughly similar mass. The percentage number below the bin indicates the number of observations in the respective category (male and female combined). For example, 14% of all decisions indicate full impatience (i.e., subjects prefer to get all the money as early as possible). For every 100 females, there are 150 males exhibiting full impatience. The three intermediate categories are more frequently chosen by female subjects, while the most patient option is again highly significantly more popular among male subjects, with 125 male subjects for every 100 female subjects.

Risk Preferences. The second line in Table 1 shows results for risk preferences from 27 samples with more than 19,000 individual decisions. In line with previous research, we find that men invest on average more money into the risky option, according to both our probabilistic indicator (57%) as well as Cohen's *d* (0.27). In addition to gender differences in central tendencies, we also find strong evidence for GMV. With a sample size–weighted mean variance ratio of 1.25, we find a level of GMV that is at the upper end of the effect size range reported by previous research (10).

Risk ratio findings imply that for every 100 female subjects, there are 172 male subjects opting for one of the two extremes of the strategy space. For all four indicators (means and variability), we can reject the null hypothesis at $P < .001$. Fig. 1, *Right* illustrates that these findings do not only reflect the male tendency for accepting risk. We also find clear evidence that males are overrepresented among those who are unwilling to take risks and hence invest zero into the risky option. *SI Appendix*, Table S2 shows that the inferential results are confirmed if we consider only the primary sample, and *SI Appendix*, Table S4 and Figs. S3 and S4 provide the details of the meta-analysis.

ECONOMIC SCIENCES

Table 1. Meta-analysis, main results

Measure	Samples	Subjects	Decisions	Prob (♂ > ♀)	Cohen's *d*	Risk ratio	Variance ratio
Time preferences	9	8,575	87,276	0.50	−0.02	1.25***	1.15***
				[0.48, 0.52]	[−0.08, 0.04]	[1.13, 1.38]	[1.08, 1.22]
Risk preferences	27	15,833	19,101	0.57***	0.27***	1.72***	1.25***
				[0.55, 0.59]	[0.19, 0.35]	[1.48, 2.00]	[1.13, 1.37]
Dictator game	31	18,765	21,564	0.48*	−0.05	1.30***	1.18***
				[0.46, 0.50]	[−0.11, 0.01]	[1.19, 1.43]	[1.12, 1.25]
Ultimatum game	13	4,624	4,624	0.50	0.00	1.14	1.12
				[0.49, 0.52]	[−0.06, 0.06]	[0.92, 1.41]	[0.90, 1.39]
Trust game	17	5,194	5,329	0.54*	0.13**	1.39***	1.28***
				[0.51, 0.56]	[0.04, 0.22]	[1.20, 1.61]	[1.18, 1.39]

Meta-analytic results of restricted maximum likelihood random effects models. Coefficients show the estimated effect sizes with 95% CIs in brackets below. Prob (♂ > ♀) is the probability for a randomly drawn male subject to choose a higher *y* than a randomly drawn female subject (with ties resolved at random). The risk ratio is the male relative frequency of extreme strategies divided by the female relative frequency. Extreme strategies are $y \in \{0,1\}$, except for the ultimatum game, in which we define extreme strategies as $y \in [0,.3) \cup (0.5,1]$. For the data source, reference *SI Appendix*, Table S1. Asterisks indicate significance levels for gender differences in the respective measure, $*P < 0.05$, $**P < 0.01$, $***P < 0.001$.

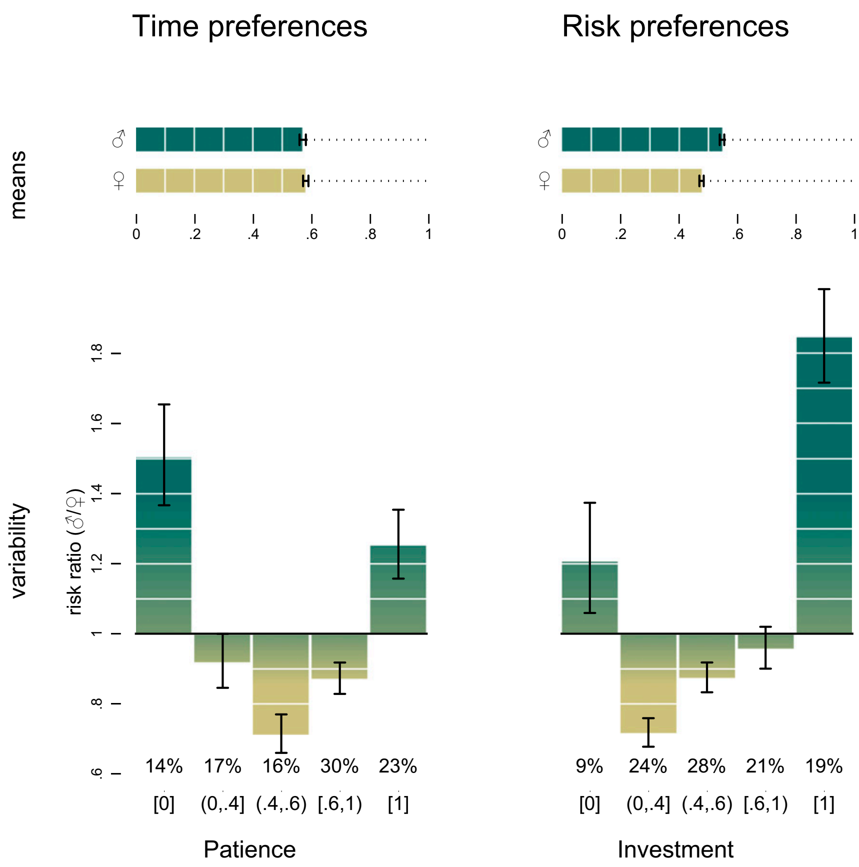

Fig. 1. Gender differences in (*Top*) means and (*Bottom*) variability for time and risk preferences. Spikes indicate bootstrapped 95% CIs. Percentages below bars indicate the percentage of cases in the particular bin (male and female combined).

Social Preferences. We first report findings for the dictator game. In our data containing more than 18,000 subjects from 31 samples, we find again highly significant evidence for GMV with a sample size–weighted mean risk ratio of 1.30 and a mean variance ratio of 1.18. Fig. 2, *Left* illustrates that males are (maybe unsurprisingly) more likely to donate zero, but they are also substantially more likely to be fully generous and donate everything. For mean differences, there is some indication that females are slightly more generous than males. When we consider the probability of drawing a more generous female (52%), the difference just reaches five percent significance. Yet, when we consider Cohen's *d*, the difference is not significant. *SI Appendix*, Table S5 and Figs. S5 and S6 provide details of the meta-analysis.

Our data on the ultimatum game comes from 13 samples with 4,624 observations. The fourth line in Table 1 reports the sample size–weighted coefficients with no indication of gender differences in mean ultimatum game offers. Our point estimate for the variance ratio is 1.12, with a CI ranging from 0.90 to 1.39. As opposed to the other preference measures, strategies $y = 0$ and $y = 1$ are only rarely observed in the ultimatum game. We therefore adjust the notion of "extreme strategy" and calculate the risk ratio for ultimatum offers of either $y < .3$ or $y > .5$. The point estimate for the risk ratio of extreme strategies is 1.14, again not reaching significance. One reason for the lack of overall significance in the variability measures is the considerable variation across samples. Whereas we rarely observe variance ratios below one in individual samples for the other preferences measures, we do so in four out of 13 samples for the ultimatum game (*SI Appendix*, Table S6 and Figs. S7 and S8). While not reaching significance, the point estimates lean toward GMV. Fig. 2, *Middle* illustrates the findings. Strategies below 0.4 and above 0.5 tend to be more common among males, while the equal split is observed more frequently among female subjects.

Finally, for the trust game, we observe the strategies of more than 5,000 subjects from 17 samples. Akin to risk preferences, we observe significant gender differences in means as well as variability. The mean differences are roughly half of those observed for risk preferences, while the sample size–weighted mean variance ratio of 1.28 is the highest reported in this meta-analysis. The risk ratio indicates that for every 100 females, there are 139 male subjects who send either zero or everything to the trustee. Fig. 2, *Right* shows again that this effect clearly originates from both ends of the spectrum. Males are significantly more likely to transfer nothing at all, but they are also more likely to transfer 60 or more percent of their endowment. *SI Appendix*, Table S7, as well as *SI Appendix*, Figs. S9 and S10, provide the meta-analytic results.

Discussion

The reported results provide converging evidence for GMV in measures of time, risk, and social preferences. Although our findings do not allow us to draw inferences about the origins of gender differences in preferences, they are in line with an evolutionary perspective according to which men are characterized by greater variability in phenotypes due to greater variability in heritable traits (11, 16, 26).

Our findings have important implications in multiple areas. For example, much research efforts have been focused on understanding gender differences in preferences and their implications for gender differences in economic and social outcomes (1–4). However, the prevailing focus on mean differences may have masked important gender differences in variability. Areas

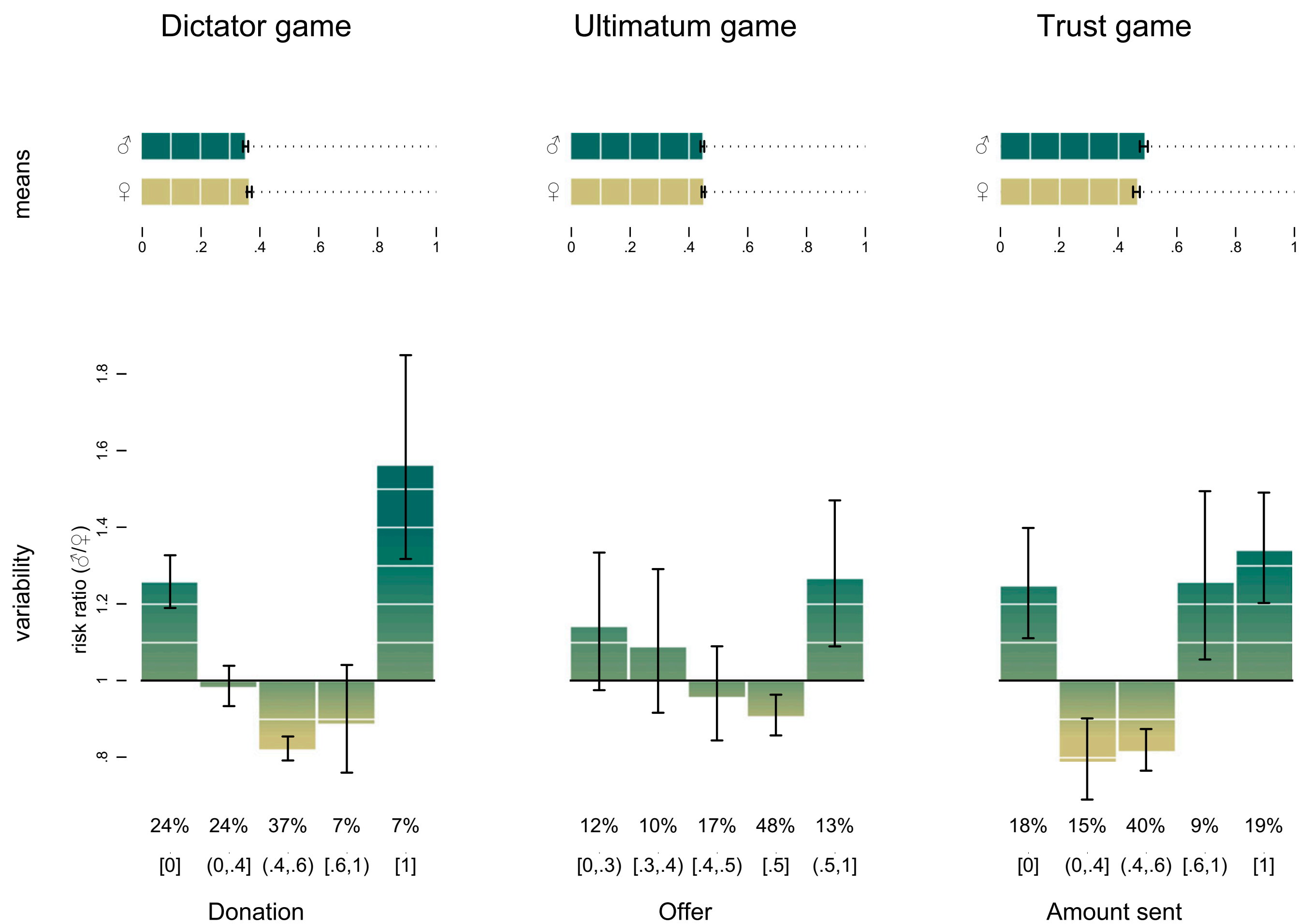

Fig. 2. Gender differences in (*Top*) means and (*Bottom*) variability for three measures of social preferences. Spikes indicate bootstrapped 95% CIs. Percentages below bars indicate the percentage of cases in the particular bin (male and female combined). For the ultimatum game, bins at the two extremes of the strategy space contain values below 0.3 and above 0.5 because 0 and 1 are only rarely observed in the ultimatum game.

in which gender differences were assumed to be absent because existing research failed to find gender differences in central tendencies may, in fact, be characterized by significant gender differences in variability (27). These so far overlooked differences are likely to have important implications for the interpretation of observed, but so far unexplained, differences in social and economic outcomes between men and women. Our findings also have potential implications for policies aimed at regulating extreme behaviors. For example, in response to what has been described as extreme trading behaviors of Game Stop stock at Wall Street, regulators are discussing new policies that prevent investors from making investment mistakes and curb extreme trading behaviors (103). Our findings suggest that preferences for extreme risk taking are more frequent among men, and policies designed to curb such behaviors are more likely to be effective when designed to appeal to men. A further area in which our results might have interesting implications is group functioning. Due to their larger variability of preferences, male groups are less likely to reach consensus and more likely to experience intragroup conflict than female groups, which can be detrimental to group functioning or beneficial for group performance depending on the nature of the tasks (104).

In sum, we find converging evidence for GMV in preferences. These findings highlight that theories are incomplete if they fail to recognize that underneath more readily detectable mean differences, also more fundamental gender differences in variability exist. In this regard, our findings point toward the critical importance of considering within-gender variability in addition to between-gender differences. With its focus on mean differences, the existing literature has underestimated the extent to which males and females differ in their preferences.

Data Availability. Data cannot be shared. (Our study is a meta-analysis combining the data of more than 90 academic articles. The majority of the data sets are available online from the journal websites. Part of the data was acquired by contacting the authors. We do not have the right to disseminate the data, but researchers who want to reproduce our work should be able to get the data from the same sources as we did.)

ACKNOWLEDGMENTS. We thank all authors of the studies included in this meta-analysis for generously sharing their data with us. For very helpful comments and suggestions, we thank the editor, two anonymous referees, Larry Cahill, Jose M. Cortina, Matthew O. Jackson, Steven Pinker, Jonathan Schulz, and the participants of the French Experimental Talks seminar.

ECONOMIC SCIENCES

1. R. Croson, U. Gneezy, Gender differences in preferences. *J. Econ. Lit.* **47**, 448–474 (2009).
2. A. Falk, J. Hermle, Relationship of gender differences in preferences to economic development and gender equality. *Science* **362**, eaas9899 (2018).
3. M. Bertrand, "New perspectives on gender" in *Handbook of Labor Economics, Volume 4b*, D. Card, O. Ashenfelter, Eds. (Elsevier B.V., 2011), vol. 4, pp. 1545–1592.
4. M. Niederle, "Gender" in *The Handbook of Experimental Economics*, J. H. Kagel, A. E. Roth, Eds. (Princeton University Press, 2017), vol. 2, pp. 481–562.

5. T. Buser, M. Niederle, H. Oosterbeek, Gender, competitiveness and career choices. *Q. J. Econ.* **129**, 1409–1447 (2014).
6. D. C. Geary, *Male, Female: The Evolution of Human Sex Differences* (American Psychological Association, 2010), p. 397.
7. D. M. Buss, D. P. Schmitt, Sexual strategies theory: An evolutionary perspective on human mating. *Psychol. Rev.* **100**, 204–232 (1993).
8. A. H. Eagly, W. Wood, The origins of sex differences in human behavior: Evolved dispositions versus social roles. *Am. Psychol.* **54**, 408–423 (1999).
9. K. Bussey, A. Bandura, Social cognitive theory of gender development and differentiation. *Psychol. Rev.* **106**, 676–713 (1999).
10. J. S. Hyde, Gender similarities and differences. *Annu. Rev. Psychol.* **65**, 373–398 (2014).
11. J. Archer, M. Mehdikhani, Variability among males in sexually selected attributes. *Rev. Gen. Psychol.* **7**, 219–236 (2003).
12. A. C. Lehre, K. P. Lehre, P. Laake, N. C. Danbolt, Greater intrasex phenotype variability in males than in females is a fundamental aspect of the gender differences in humans. *Dev. Psychobiol.* **51**, 198–206 (2009).
13. W. Johnson, A. Carothers, I. J. Deary, Sex differences in variability in general intelligence: A new look at the old question. *Perspect. Psychol. Sci.* **3**, 518–531 (2008).
14. D. M. Buss, Psychological sex differences. Origins through sexual selection. *Am. Psychol.* **50**, 164–168, discussion 169–171 (1995).
15. R. L. Trivers, "Parental investment and sexual selection" in *Sexual Selection and the Descent of Man*, B. Campbell, Ed. (Aldine Publishing Company, Chicago, 1972), pp. 136–179.
16. R. W. Wilcockson, C. S. Crean, T. H. Day, Heritability of a sexually selected character expressed in both sexes. *Nature* **374**, 158–159 (1995).
17. B. Pawlowski, R. I. M. Dunbar, A. Lipowicz, Tall men have more reproductive success. *Nature* **403**, 156 (2000).
18. J. Archer, Sex differences in aggression between heterosexual partners: A meta-analytic review. *Psychol. Bull.* **126**, 651–680 (2000).
19. E. M. Miller, Could nonshared environmental variance have evolved to assure diversification through randomness? *Evol. Hum. Behav.* **18**, 195–221 (1997).
20. S. Pinker, Sex Ed: The science of difference. The New Republic, 14 February 2005. https://newrepublic.com/article/68044/sex-ed. Accessed 27 May 2021.
21. T. P. Hill, Modeling the evolution of differences in variability between sexes. *J. Interdiscip. Math.* **23**, 1009–1031 (2020).
22. L. M. Wierenga *et al.*; Karolinska Schizophrenia Project (KaSP) Consortium, Greater male than female variability in regional brain structure across the lifespan. *Hum. Brain Mapp.*, 10.1002/hbm.25204 (2020).
23. A. Pomiankowski, A. P. Møller, A resolution of the lek paradox. *Proc. Biol. Sci.* **260**, 21–29 (1995).
24. L. V. Hedges, A. Nowell, Sex differences in mental test scores, variability, and numbers of high-scoring individuals. *Science* **269**, 41–45 (1995).
25. J. S. Hyde, S. M. Lindberg, M. C. Linn, A. B. Ellis, C. C. Williams, Diversity. Gender similarities characterize math performance. *Science* **321**, 494–495 (2008).
26. R. Arden, R. Plomin, Sex differences in variance of intelligence across childhood. *Pers. Individ. Dif.* **41**, 39–48 (2006).
27. C. Thöni, S. Volk, J. M. Cortina, Greater male variability in cooperation: Meta-analytic evidence for an evolutionary perspective. *Psychol. Sci.* **32**, 50–63 (2021).
28. M. Van Vugt, W. Iredale, Men behaving nicely: Public goods as peacock tails. *Br. J. Psychol.* **104**, 3–13 (2013).
29. M. D. Baker, J. K. Maner, Risk-taking as a situationally sensitive male mating strategy. *Evol. Hum. Behav.* **29**, 391–395 (2008).
30. D. Farrelly, P. Clemson, M. Guthrie, Are women's mate preferences for altruism also influenced by physical attractiveness? *Evol. Psychol.* **14**, 1–6 (2016).
31. D. Ehlebracht, O. Stavrova, D. Fetchenhauer, D. Farrelly, The synergistic effect of prosociality and physical attractiveness on mate desirability. *Br. J. Psychol.* **109**, 517–537 (2018).
32. P. J. Curran, A. M. Hussong, Integrative data analysis: The simultaneous analysis of multiple data sets. *Psychol. Methods* **14**, 81–100 (2009).
33. C. F. Camerer, "The promise and success of lab-field generalizability in experimental economics: A critical reply to Levitt and List" in *Handbook of Experimental Economic Methodology*, G. R. Fréchette, A. Schotter, Eds. (Oxford University Press, 2015), pp. 249–295.
34. B. Herrmann, C. Thöni, S. Gächter, Antisocial punishment across societies. *Science* **319**, 1362–1367 (2008).
35. J. Henrich *et al.*, Markets, religion, community size, and the evolution of fairness and punishment. *Science* **327**, 1480–1484 (2010).
36. C. Thöni, "Cross-cultural behavioral experiments: Potential and challenges" in *Handbook of Research Methods and Applications in Experimental Economics*, A. Schram, A. Ule, Eds. (Edward Elgar Publishing, 2019), chap. 18, pp. 349–367.
37. J. Andreoni, C. Sprenger, Estimating time preferences from convex budgets. *Am. Econ. Rev.* **102**, 3333–3356 (2012).
38. U. Gneezy, J. Potters, An experiment on risk taking and evaluation periods. *Q. J. Econ.* **112**, 631–645 (1997).
39. R. Forsythe, J. L. Horowitz, N. E. Savin, M. Sefton, Fairness in simple bargaining experiments. *Games Econ. Behav.* **6**, 347–369 (1994).
40. W. Güth, R. Schmittberger, B. Schwarze, An experimental analysis of ultimatum bargaining. *J. Econ. Behav. Organ.* **3**, 367–388 (1982).
41. J. Berg, J. Dickhaut, K. McCabe, Trust, reciprocity, and social history. *Games Econ. Behav.* **10**, 122–142 (1995).
42. S. Alan, T. Boneva, S. Ertac, Ever failed, try again, succeed better: Results from a randomized educational intervention on grit. *Q. J. Econ.* **134**, 1121–1162 (2019).
43. S. Alan, S. Ertac, Fostering patience in the classroom: Results from randomized educational intervention. *J. Polit. Econ.* **126**, 1865–1911 (2018).
44. R. Banerjee, N. D. Gupta, M. C. Villeval, The spillover effects of affirmative action on competitiveness and unethical behavior. *Eur. Econ. Rev.* **101**, 567–604 (2018).
45. S. Banuri, P. Keefer, Pro-social motivation, effort and the call to public service. *Eur. Econ. Rev.* **83**, 139–164 (2016).
46. M. Bauer, N. Fiala, I. Levely, Trusting former rebels: An experimental approach to understanding reintegration after civil war. *Econ. J. (Lond.)* **128**, 1786–1819 (2018).
47. C. Bellemare, S. Kröger, A. van Soest, Measuring inequity aversion in a heterogeneous population using experimental decisions and subjective probabilities. *Econometrica* **76**, 815–839 (2008).
48. J. Butler, P. Giuliano, L. Guiso, Trust and cheating. *Econ. J. (Lond.)* **126**, 1703–1738 (2016).
49. A. W. Cappelen, T. Halvorsen, E. Ø. Sørensen, B. Tungodden, Face-saving or fair-minded: What motivates moral behavior? *J. Eur. Econ. Assoc.* **15**, 540–557 (2017).
50. L. S. Carvalho, S. Meier, S. W. Wang, Poverty and economic decision-making: Evidence from changes in financial resources at payday. *Am. Econ. Rev.* **106**, 260–284 (2016).
51. T. N. Cason, L. Friesen, L. Gangadharan, Regulatory performance of audit tournaments and compliance observability. *Eur. Econ. Rev.* **85**, 288–306 (2016).
52. G. Charness, M.-C. Villeval, Cooperation and competition in intergenerational experiments in the field and the laboratory. *Am. Econ. Rev.* **99**, 956–978 (2009).
53. A. Cohn, J. Engelmann, E. Fehr, M. A. Maréchal, Evidence for countercyclical risk aversion: An experiment with financial professionals. *Am. Econ. Rev.* **105**, 860–885 (2015).
54. B. D'Exelle, A. Verschoor, Investment behaviour, risk sharing and social distance. *Econ. J. (Lond.)* **125**, 777–802 (2015).
55. J. de Quidt, J. Haushofer, C. Roth, Measuring and bounding experimenter demand. *Am. Econ. Rev.* **108**, 3266–3302 (2018).
56. D. L. Dickinson, T. McElroy, Sleep restriction and circadian effects on social decisions. *Eur. Econ. Rev.* **97**, 57–71 (2017).
57. O. Dijk, M. Holmen, M. Kirchler, Rank matters—The impact of social competition on portfolio choice. *Eur. Econ. Rev.* **66**, 97–110 (2014).
58. T. Dohmen, A. Falk, Performance pay and multidimensional sorting: Productivity, preferences, and gender. *Am. Econ. Rev.* **101**, 556–590 (2011).
59. C. C. Eckel, R. Petrie, Face value. *Am. Econ. Rev.* **101**, 1497–1513 (2011).
60. N. Erkal, L. Gangadharan, B. H. Koh, Monetary and non-monetary incentives in real-effort tournaments. *Eur. Econ. Rev.* **101**, 528–545 (2018).
61. G. Friebel, M. Kosfeld, G. Thielmann, Trust the police? Self-selection of motivated agents into the German police force. *Am. Econ. J. Microecon.* **11**, 59–78 (2019).
62. B. Gillen, E. Snowberg, L. Yariv, Experimenting with measurement error: Techniques with applications to the Caltech cohort study. *J. Polit. Econ.* **127**, 1826–1863 (2019).
63. X. Giné, J. Goldberg, D. Silverman, D. Yang, Revising commitments: Field evidence on the adjustment of prior choices. *Econ. J. (Lond.)* **128**, 159–188 (2018).
64. U. Gneezy, K. L. Leonard, J. A. List, Gender differences in competition: Evidence from a matrilineal and a patriarchal society. *Econometrica* **77**, 1637–1664 (2009).
65. G. Graziani, W. van der Klaauw, B. Zafar, Workers' spending response to the 2011 payroll tax cuts. *Am. Econ. J. Econ. Policy* **8**, 124–159 (2016).
66. S. P. Hargreaves Heap, D. J. Zizzo, The value of groups. *Am. Econ. Rev.* **99**, 295–323 (2009).
67. A. Heyes, J. A. List, Supply and demand for discrimination: Strategic revelation of own characteristics in a trust game. *Am. Econ. Rev.* **106**, 319–323 (2016).
68. H. J. Holm, A. Danielson, Tropic trust versus Nordic trust: Experimental evidence from Tanzania and Sweden. *Econ. J. (Lond.)* **115**, 505–532 (2005).
69. D. Houser, S. Vetter, J. Winter, Fairness and cheating. *Eur. Econ. Rev.* **56**, 1645–1655 (2012).
70. J. Huber, M. Kirchler, D. Kleinlercher, M. Sutter, Market versus residence principle: Experimental evidence on the effects of a financial transaction tax. *Econ. J. (Lond.)* **127**, F610–F631 (2017).
71. A. Imas, The realization effect: Risk-taking after realized versus paper losses. *Am. Econ. Rev.* **106**, 2086–2109 (2016).
72. P. Jakiela, O. Ozier, Does Africa need a rotten kin theorem? Experimental evidence from village economies. *Rev. Econ. Stud.* **83**, 231–268 (2016).
73. M. Lührmann, M. Serra-Garcia, J. Winter, The impact of financial education on adolescents' intertemporal choices. *Am. Econ. J. Econ. Policy* **10**, 309–332 (2018).
74. M. Sutter, S. Angerer, D. Glätzle-Rützler, P. Lergetporer, Language group differences in time preferences: Evidence from primary school children in a bilingual city. *Eur. Econ. Rev.* **106**, 21–34 (2018).
75. J. van der Weele, J. Kulisa, M. Kosfeld, G. Friebel, Resisting moral wiggle room: How robust is reciprocity? *Am. Econ. J. Microecon.* **6**, 256–264 (2014).
76. F. Zimmermann, The dynamics of motivated beliefs. *Am. Econ. Rev.* **110**, 337–363 (2020).
77. J. Andreoni, L. Vesterlund, Which is the fair sex? Gender differences in altruism. *Q. J. Econ.* **116**, 293–312 (2001).
78. U. Balakrishnan, J. Haushofer, P. Jakiela, How soon is now? Evidence of present bias from convex time budget experiments. *Exp. Econ.* **23**, 294–321 (2020).
79. A. Boschini, A. Dreber, E. von Essen, A. Muren, E. Ranehill, Gender and altruism in a random sample. *J. Behav. Exp. Econ.* **77**, 72–77 (2018).
80. P. Brañas-Garza, V. Capraro, E. Rascón-Ramírez, Gender differences in altruism on Mechanical Turk: Expectations and actual behaviour. *Econ. Lett.* **170**, 19–23 (2018).
81. P. Brañas-Garza, A. M. Espín, F. Exadaktylos, B. Herrmann, Fair and unfair punishers coexist in the ultimatum game. *Sci. Rep.* **4**, 6025 (2014).
82. T. Buser, Digit ratios, the menstrual cycle and social preferences. *Games Econ. Behav.* **76**, 457–470 (2012).
83. A. W. Cappelen, U. H. Nielsen, B. Tungodden, J.-R. Tyran, E. Wengström, Fairness is intuitive. *Exp. Econ.* **19**, 727–740 (2016).

84. M. Castillo, G. Leo, R. Petrie, Room composition effects on risk taking by gender. *Exp. Econ.* **23**, 895–911 (2020).
85. Y. Chen, M. Jiang, E. L. Krupka, Hunger and the gender gap. *Exp. Econ.* **22**, 885–917 (2019).
86. Y. Chuang, L. Schechter, Stability of experimental and survey measures of risk, time, and social preferences: A review and some new results. *J. Dev. Econ.* **117**, 151–170 (2015).
87. B. L. Cleave, N. Nikiforakis, R. Slonim, Is there selection bias in laboratory experiments? The case of social and risk preferences. *Exp. Econ.* **16**, 372–382 (2013).
88. B. Deffains, R. Espinosa, C. Thöni, Political self-serving bias and redistribution. *J. Public Econ.* **134**, 67–74 (2016).
89. E. E. Demiral, J. Mollerstrom, The entitlement effect in the ultimatum game – Does it even exist? *J. Econ. Behav. Organ.* **175**, 341–352 (2020).
90. R. Espinosa, B. Deffains, C. Thöni, Debiasing preferences over redistribution: An experiment. *Soc. Choice Welfare* **55**, 823–843 (2020).
91. F. Exadaktylos, A. M. Espín, P. Brañas-Garza, Experimental subjects are not different. *Sci. Rep.* **3**, 1213 (2013).
92. B. Gong, H. Yan, C. L. Yang, Gender differences in the dictator experiment: Evidence from the matrilineal Mosuo and the patriarchal Yi. *Exp. Econ.* **18**, 302–313 (2015).
93. G. Grimalda, A. Kar, E. Proto, Procedural fairness in lotteries assigning initial roles in a dynamic setting. *Exp. Econ.* **19**, 819–841 (2016).
94. K. E. Hauge, K. A. Brekke, L. O. Johansson, O. Johansson-Stenman, H. Svedsäter, Keeping others in our mind or in our heart? Distribution games under cognitive load. *Exp. Econ.* **19**, 562–576 (2016).
95. J. Hergueux, N. Jacquemet, Social preferences in the online laboratory: A randomized experiment. *Exp. Econ.* **18**, 251–283 (2015).
96. J. Ifcher, H. Zarghamee, The rapid evolution of homo economicus: Brief exposure to neoclassical assumptions increases self-interested behavior. *J. Behav. Exp. Econ.* **75**, 55–65 (2018).
97. D. Kistler, C. Thöni, C. Welzel, Survey response and observed behavior: Emancipative and secular values predict prosocial behaviors. *J. Cross Cult. Psychol.* **48**, 461–489 (2017).
98. V. Kurz, A. Orland, K. Posadzy, Fairness versus efficiency: How procedural fairness concerns affect coordination. *Exp. Econ.* **21**, 601–626 (2018).
99. G. Saad, T. Gill, Sex differences in the ultimatum game: An evolutionary psychology perspective. *J. Bioeconomics* **3**, 171–193 (2001).
100. J. Schmitz, Is charitable giving a zero-sum game? The effect of competition between charities on giving behavior. *Manage. Sci.* in press.
101. S. Zhong *et al*., Dopamine D4 receptor gene associated with fairness preference in ultimatum game. *PLoS One* **5**, e13765 (2010).
102. S. W. Raudenbush, "Analyzing effect sizes: Random-effects models" in *The Handbook of Research Synthesis and Meta-Analysis*, H. Cooper, L. V. Hedges, J. C. Valentine, Eds. (Russell Sage Foundation, New York, 2009), pp. 295–315.
103. B. Pisani, Wall Street worries about regulatory fallout from the GameStop saga. *CNBC*, 4 February 2021. https://www.cnbc.com/2021/02/04/wall-street-worries-about-regulatory-fallout-from-the-gamestop-saga.html. Accessed 19 May 2021.
104. K. A. Jehn, A multimethod examination of the benefits and detriments of intragroup conflict. *Adm. Sci. Q.* **40**, 256–282 (1995).

https://doi.org/10.1073/pnas.2026112118