Male and Female Voices in Economics

Hans Henrik Sievertsen
Sarah Smith

Discussion Paper 22/761

7 March 2022

School of Economics

University of Bristol
Priory Road Complex
Bristol
BS8 1TU
United Kingdom

University of BRISTOL
School of Economics

Male and female voices in economics

Draft: March 2022

Hans Henrik Sievertsen (University of Bristol, VIVE, IZA)

Sarah Smith (University of Bristol, IFS, CEPR)[1]

Abstract

Women's voices are likely to be even more absent from economic debates than headline figures on female under-representation suggest. Focusing on a panel of leading economists we find that men are more willing than women to express an opinion and are more certain and more confident in their opinions, including in areas where both are experts. Women make up 21 per cent of the panel but 19 per cent of the opinions expressed and 14 per cent of strong opinions. We discuss implications for the economics profession and for promoting a genuine diversity of views.

[1] Corresponding author: sarah.smith@bristol.ac.uk. The authors would like to thank Amanda Bayer, Niels-Jakob Harbo Hansen, Stephen Hansen, Erin Hengel, Michael McMahon, Mark Otoni-Wilhelm and Almudena Sevilla for helpful comments.

1. Introduction

The widespread under-representation of women in economics has been well-documented. In the US, UK, Canada, Australia and New Zealand just one in seven economics professors is a woman (CSWEP, 2021), while for Europe, the figure is one in four (Auriol et al, 2019). Outside academia, the picture is similar; in the private sector, women make up 10 – 20 per cent of chief economists in banking and finance, including central banks and finance ministries (Hansbach et al, 2021). When economic advisory councils meet to advise national governments, three out of four seats are taken by men (Hansbach et al, 2021).

In this paper, we argue that women's voices are likely to be less prominent in economic debates, even than these headline figures on representation would suggest. Focusing on a panel of expert economists, we show that men's voices are louder and stronger than women's across several dimensions; men are more willing to express an opinion than women and are more certain and confident in those opinions. Evidence also suggests that confident and strong opinions expressed by men are more likely to be heard.

One implication is that the nature of debate in economics is shaped by the fact that it is a male-dominated profession. We show that, if the gender balance among the expert panel of economists was reversed, opinions would be expressed with less certainty and greater caution. Fourcade et al (2015) argue that "confidence is perhaps the greatest achievement of the economics profession – but it is also its most vulnerable trait, its Achilles heel". This suggests that it is not clear whether less confident views would be unambiguously better or worse. However, we present some evidence that women calibrate their opinions to background uncertainty to a greater extent than men; in particular, women are more likely to be uncertain in expressing an opinion when there is greater background uncertainty.

A second implication is that attempts to increase diversity in economics need to go beyond ensuring that different groups are represented and focus also on whose voices are heard. We show that, even if there were equal gender representation, men's greater willingness to express strong opinions means that their views are likely to dominate. Increasing diversity can bring in groups with different perspectives (Levine et al, 2014) and different views on key economic issues (May et al, 2014; May et al, 2018). However, these benefits of diversity will only be realised if different voices are heard.

Our paper is related to a literature in economics and psychology showing that men are less reluctant than women to share ideas and give opinions (Thomas-Hunt and Phillips 2004; Babcock and Laschever 2004; Coffman, 2014). These studies typically focus on groups drawn from the general population. Here, we show that there are gender differences in willingness to give opinions among a selected group of men and women who are not only relatively homogeneous but who are also selected on the basis of their expertise and a willingness to give opinions (Adams and Funk, 2018). We study the IGM expert economists panel (EEP), a group of leading economists who are otherwise very similar (PhD institutions and current institutions), who have made it to the very top of the profession, who are acknowledged experts in their field and who have agreed to be on the panel precisely to give their opinions on different issues.

Our paper builds on two previous studies which have also analysed EEP survey responses. Gordon and Dahl (2013) study consensus among economists. They show that there is a high degree of consensus, which increases with the state of economics knowledge on a topic. They report that women are more likely than men to be uncertain and have a lower level of confidence. Sarsons and Xu (2021) also show that women are less confident in their opinions but argue that this is only when issues are outside women's areas of expertise.

We extend previous analysis in several ways. First, we add new measures of voice and show that men are more voluble (i.e. they are more willing to give an opinion and more likely to comment). Second, we exploit the panel nature of the survey and show that the gender differences persist over time. Third, we look by field of expertise, by types of questions (theory, positive and normative) and within more- and less-male dominated fields. Fourth, we consider whether the differences in voice are likely to carry over to real-world economic debates and decision-making. We present additional analyses of twitter showing that more confident men have more followers and of the votes of Bank of England Monetary Policy Committee (MPC) members showing that women are typically more cautious in how they vote. Finally, we discuss the implications for the economics profession.

The focus of this paper is on the presence and consequences of gender differences in voice; we do not speak directly to why men are more willing to express their opinions than women. In the psychology literature (Crosby and Nyquist 1977; Thomas-Hunt and Phillips 2004; Babcock and Laschever 2007), higher male confidence in expressing opinions is attributed to

both "over-estimation", i.e. men have a higher estimation of their ability (Lundeberg et al, 1994; Niederle and Vesterlund, 2007) and "over-precision", i.e. men have a greater level of precision with respect to their beliefs (Moore and Healy, 2008; Barber and O'Dean, 2001). In a group-setting, another possibility is that men have higher social confidence, i.e. "a type of self-confidence that goes beyond the individual belief that one can do well and interacts with social concerns" (Alan et al, 2020). Our findings – men are more voluble, more certain in their opinions and more confident in their certain opinions – are consistent with over-estimation, over-precision and higher social confidence. These behaviours may reflect innate differences between men and women or they may be learned as a rational strategy in anticipation of different reactions to confident men and women. Brescoll (2012) argues that backlash concerns explain why more powerful male politicians (but not more powerful female politicians) are more voluble. Thomas-Hunt and Phillips (2004) found that expertise increased men's influence in a group setting, but not women's. Our analysis of twitter shows that more confident male economists tend to be more influential (i.e. have more followers) but that the same is less true for women. This suggests, first, that men's greater confidence matters because it is likely to translate into greater influence and second, that the solution is not for women to "lean in", in this case to speak up, because they may not be listened to in the same way as men. Instead, processes and practices, such as seminar and meeting guidelines, are required to create an inclusive environment in which all voices are heard.

2. Data and results

2.1 The panel

We analyse data on economic opinions expressed by Economist Expert Panel (EEP) members in surveys carried out by the Initiative on Global Markets (IGM) at the University of Chicago Booth School of Business. We combine survey responses from 53 panellists in the US panel, which started in September 2011, with survey responses from 51 panellists in the European panel, which started in December 2016. In both cases, panellists are typically senior faculty members at leading research universities. These data are uniquely rich in offering insights into the opinions of a large sample of leading economists across a broad range of topics, over a sustained period. Our combined sample contains 18,990 observations, representing responses to 396 distinct questions.

The survey is intended to explore economists' views on vital policy issues. Panel members are given approximately two questions each month, asking whether they agree or disagree with a statement about economic policy, for example "The current combination of US fiscal and monetary policy poses a serious risk of prolonged higher inflation". Responses are on a five-point Likert scale: Strongly disagree, Disagree, Uncertain, Agree and Strongly agree. In addition to these responses, panellists are also asked to say how confident they are in their answer (on a scale of 1 – 10) and they are given an opportunity to provide an additional written comment.

Gordon and Dahl (2013) describe the US panel as geographically diverse with members drawn from across the political spectrum and a mix of older and younger scholars. However, what is striking about the panel is the similarity of backgrounds of panel members, particularly those based in the US. The 53 US panel members are drawn from nine institutions, and they obtained their PhDs from 13 institutions (of which the top four institutions account for 76 per cent). The European panel of 51 experts is slightly more diverse representing 29 current institutions and 23 PhD institutions, of which the top four account for 55 per cent.

The panellists are mainly men, although the representation of women – 11 out of 53 (21 per cent) US panel members and 11 out of 51 (22 per cent) European panel members – is above the overall share of women professors in economics. Summary statistics (presented in Table 1) show that the women are younger than the men (defined by years' post PhD) and have fewer citations.[2] We control for both these factors in the regression analysis below.

2.2 Measures of voice

We use five measures of voice to analyse gender differences – these capture dimensions of volubility, certainty and confidence.

Volubility

The first measure, not used by Gordon and Dahl (2013) or Sarsons and Xu (2021), is whether the expert gives an opinion or not. This is motivated by previous studies showing that men are less reluctant than women to share ideas and give opinions (Crosby and Nyquist 1977; Thomas-Hunt and Phillips 2004; Babcock and Laschever 2007; Coffman, 2014). Not all panel

[2] Citations from Google scholar profiles as of June 2020. Where an expert does not have a profile, we manually calculate the total number of Google citations.

members respond to every question and we define "Gives opinion" = 1 if the expert responded to the question and did not respond "no opinion".[3] The mean share giving an opinion, by question, is 0.825, ranging from 0.530 to 1 across the 396 questions. We also look at whether panel members leave "Any comment". By question, the average share who comment is 0.396 (range 0.142 – 0.767). Panel members are more likely to comment when they are uncertain (0.432) or strongly (dis)agree (0.407) compared to when they (dis)agree (0.376). A further measure of volubility (number of words) shows no notable variation.

Certainty

Following previous studies (Gordon and Dahl, 2013; Sarsons and Xu, 2021), we look at measures of how certain men and women are in expressing an opinion. Specifically, we consider whether panellists report they are "Uncertain" (mean share = 0.223, range across questions from 0 to 0.710) and whether they express a "Strong opinion", i.e. they strongly agree or disagree (mean share = 0.234, ranging from 0 to 0.923).

Confidence

As well as reporting agreement/ disagreement with the statement, panel members are asked to say how confident they are in their response on a scale of 1 – 10.[4] "Confidence" is correlated with the degree of certainty: Average confidence on "Uncertain" responses is 4.6 compared to 8.0 on "Strongly (dis)agree" responses. But there is also variation within opinion groups; panel members can be uncertain but confident and agree strongly but not be confident. We interpret confidence as capturing the degree of precision that people hold about their views (Moore and Healey, 2008).

2.3 Results

We test for differences between men and women in these five different measures of voice, controlling for current institution (fixed effects), years since PhD (fixed effects), citations and H-index (both continuous, linear) [5], US-sample (0/1), Nationality (0/1) and question fixed

[3] Analysis shows that respondents are more likely to choose no opinion (rather than not replying) when they have answered another question in the same batch.
[4] One panel member chose to report confidence of 11 on one question – we recoded this to be 10.
[5] Since citations may be gendered (Koffi, 2021) we also test robustness to excluding these controls.

effects.[6] We estimate linear models, using reghdfe to deal with multiple fixed effects (Correia, 2017):

$$y_{iq} = \alpha + \beta F_i + \gamma' X_i + \varphi_q + u_{iq}$$

where y_{iq} is the outcome measure (one of five) for expert *i* in question *q*. *F* takes the value 1 if the expert is female, X_i is our vector of controls, and φ_q is a vector of question fixed effects.

Baseline results

Our baseline results – gender differences across the combined sample of all US and European experts – are shown in the first "bar" of each of the five panels in Figure 1, where each panel depicts one of the five measures. The results consistently indicate that men's voices are louder and stronger than women's. Men are 9.4 percentage points more likely to give an opinion than women and, conditional on giving an opinion, are 20.1 percentage points more likely to leave a comment than women. Among those who give an opinion, men are 4.5 percentage points less likely to be uncertain and 8.5 percentage points more likely to express a strong opinion. Men also express greater confidence (0.46 points) in their opinions. In part this reflects the lower level of uncertainty. Running separate regressions for each level of certainty (uncertain, (dis)agree and strongly (dis)agree), there is no gender gap among those who are uncertain/ strongly (dis)agree. But men who (dis)agree are more confident in their opinions than women (0.39 points, p=[0.000]). The gender gap in commenting is greatest among panel members who strongly (dis)agree (27.8 percentage points).

We expand on these average differences by looking at where men and women economists lie in the distributions of experts for each of our measures. Taking the residuals from our regressions (excluding an indicator for female), we calculate expert-level averages and then determine the share of women in each quintile of the distribution of experts. The full results, shown in Figure 2, show a mix of men and women across the distribution – some women are among the most confident in expressing themselves and some men are among the least confident. But there are also clear differences. Women (21 per cent of the full sample) make

[6] For comparison, Sarsons and Xu (2021) control for Harvard/MIT PhD, current institution FE, Question FE, Years since tenure FE, number of solo publications and number of co-authored publications. Gordon and Dahl (2013) control for years since PhD (0-15,16-30,>30), a dummy indicating experience working in Washington, question FE

up less than 5 per cent of the top quintile expressing strong opinions and more than one-third of the most uncertain experts.

By question type and expertise

The remaining six bars in each panel in Figure 1 show gender differences for each of six groups defined by the type of question (T = theory, P = positive and N = normative[7]) and whether the question is in field or out of field for the panellist (X = expert, N = non).

Theory questions (18 per cent questions) ask about core aspects of economic theory. An example: "Unless they have inside information, very few investors (if any) can consistently make accurate predictions about whether the price of an individual stock will rise or fall"

Positive questions (42 per cent questions) ask for opinions where there is relevant evidence that experts can draw on in their responses. They are questions that are, in principle, empirically verifiable. An example: "Raising the federal minimum wage to $9 would make it noticeably harder for low-skilled workers to find work"

Normative questions (40 per cent questions) require a value judgement. Existing evidence is likely to be relevant but there are also trade-offs to be made. An example: "Considering both distributional effects and changes in efficiency, it is a good idea to let companies that send video content to customers pay more to internet service providers"

There are several reasons for distinguishing between these question types. First, there is evidence from May et al (2014, 2018) of fewer substantive differences between men and women on elements of core theory. Second, Gordon and Dahl (2013) find that uncertainty among the expert panel decreases with the availability of evidence; similarly, we find the highest level of uncertainty (and the lowest level of consensus) on normative questions and the highest level of certainty (and the highest level of consensus) on theory questions. It is natural to ask whether women are more likely to be uncertain in relation to normative questions, where they have to exercise judgement, compared to positive questions where they can draw on evidence. Finally, positive and normative questions identify the kinds of real-world policy issues that economists might be asked directly to advise on.

[7] The authors classified the questions independently and discussed cases where there was any discrepancy. All macro and finance questions were separately classified by two colleagues.

Distinguishing whether a question is inside/ outside a panellist's field of expertise is also done to highlight real-world situations in which economists typically advise on policy within their field. We use panel members' National Bureau of Economic Research (NBER) and Centre for Economic Policy Research (CEPR) affiliations[8] to determine their fields of expertise, where the fields are development, finance, industrial organisation, international, labor, macro and public.[9] Note that each expert can have multiple fields. Each of the questions was also classified into the same fields to determine whether it is "in field" for a panel member or "out of field". Men and women do not differ in the number of fields in which they have expertise (1.93 compared to 1.83, p=0.625) but men have more questions "in field" (see Table 1) because of differences in areas of expertise and the distribution of questions across fields.

Sarsons and Xu (2021) argue that differences in confidence and certainty are greatest in areas where the panellists are not experts. This would be consistent with what is referred to as "male answer syndrome" the idea that men have opinions about things they know nothing about. This would suggest that gender differences in voice may not matter for policy-making where economists are more likely to be asked questions within their field of expertise.

We find that the biggest gender difference in volubility (i.e. giving an opinion and leaving a comment) is among non-experts on normative questions. This is also where women are more uncertain compared to men. However, the patterns do not consistently suggest greater gender differences when men and women are not experts and on normative questions. The biggest differences in confidence and strong opinions are on core theory questions between men and women who are experts. Indeed, the main take-away from Figure 1 is that there are significant gender differences among both experts and non-experts, across all three types of questions, i.e. men typically have louder and stronger voices irrespective of question type and field expertise.

We carry out additional analysis of whether gender differences are related to the extent to which the fields are seen as traditionally (even) more male. This is motivated by findings from previous studies that gender differences tend to be more pronounced in relation to stereotypically male behaviours (Coffman, 2014). We follow Chari and Goldsmith-Pinkham

[8] Ten experts have neither (NBER/ CEPR) affiliation; we manually assign fields using information on their websites.

[9] These broad fields are chosen to map to NBER and CEPR programmes. Further information is provided in the Appendix (Table A1).

(2017) in splitting fields into "male" (macro, international and finance) and "female" (development, labor, public, IO).[10] The results, reported in Appendix Table A5, show that the gaps are reduced slightly – but are still present and statistically significant – in female fields. Stereotypically male dominated fields (macro, international and finance) cannot account for all the gender differences.

Over time

Many of the panel members have answered hundreds of questions over a period of several years, providing a unique opportunity to look at whether the strength of men's and women's voices changes over time. It is possible that gender differences attenuate, for example, because women and men observe, and learn from, others' past responses and choose to update their beliefs about their relative ability or precision. It is possible that men tone down their stronger opinions or women gain social confidence after they have performed the same tasks in public several times. If this is the case – and men and women learn to behave like each other – then gender differences in voice may be less important in real-world policy-making.

In fact, there is little evidence that this is the case. Figure 3 plots the coefficients from estimating gender differences separately by the number of questions each panellist has faced (1 – 50, 51 – 100 and 101 – 150); non-parametric polynomial plots are shown separately for men and women in Figure A1 in the Appendix.[11] There is some reduction in the gender difference in uncertainty over time, but the gaps in strong opinions, in confidence and volubility remain fairly constant over time. We take this as evidence that gender differences that are observed among EEP panel members persist.

3. Discussion and implications

Our analysis shows that there are important differences between male and female voices in economics. Even among a group of leading academics, who are a homogeneous group with similar backgrounds and experience and a highly selected group who have agreed to be part

[10] Note that this split also lines up with the areas of expertise for men and women in Dolado et al (2012).
[11] There is variation in when panel members start being asked questions, not only between US and European panels but also within the two panels. A second batch of panellists joined the US panel in Jan 2012 and the European panel in Aug 2018.

of a panel that is asked for opinions on topical issues, men's voices are typically louder and stronger than women's. Our findings imply that women's voices are even less prominent in economic debates than headline figures on their under-representation would suggest. Women make up 21 per cent of EEP members but 19 per cent of opinions expressed, 14 per cent of strong opinions and 12 per cent of comments.

Do gender differences in voice matter?

Just because men's voices are louder and stronger does not necessarily mean that women's voices are less likely to be heard. It is possible that loud men are not listened to, while quiet women are listened to. To shed light on this, we look at the relationship between panel members' voices, as measured by their EEP responses, and the number of followers that they have on twitter, taking the latter as a measure of how much different voices are heard.

Of the EEP members, 38 per cent of men and 50 per cent women are on twitter. Being on twitter is not correlated with any of the voice measures but is correlated with panel members' age. By contrast, there is evidence of correlation between voice and followers – particularly for men, but less so (or reversed) for women. We show this for two voice measures (Strong Opinion and Confidence) in Figure 4. Men who are more likely to express strong opinions and who are more confident typically have more twitter followers – but the same is not true for women. These are the measures for which the relationship is clearest, but Table A6 in the Appendix presents regression results with similar findings for all measures of voice except "Gives Opinion".

This evidence is tentative but it has two implications. First, the positive relationship between voice and influence for men suggests that louder and stronger male voices are more likely to be heard. Second, the absence of a relationship for women implies that advice to women to be less cautious in expressing their opinions (i.e. to lean in) is unlikely to increase influence. We return to this at the end.

Implications for economic debate

Our findings suggest that the nature of economic debate is likely to be shaped by the fact that it is male dominated. Increasing the share of women in economics would likely lead to greater caution among economists in expressing opinions. To demonstrate the effect of changing the gender composition of the panel on collective opinions, we re-weight the opinions of

individual EEP members and calculate the new "consensus" for each question. As an extreme example, we flip the gender ratio (i.e. assume 80 per cent women and 20 per cent men). The effect of this is to increase the share of questions where there is "no consensus" (i.e. modal response = uncertain) from 12 per cent (47/396 questions) to 19 per cent (75/396 questions).[12] It also reduces (by 10/396 questions) the number of questions where a strong opinion (strongly agree/ disagree) is the consensus. By contrast, there are only three out of the 396 questions where flipping the gender ratio results in a substantive change in the consensus view, i.e., changes in the consensus from agree to disagree or from disagree to agree. In other words, the biggest effect of increasing the share of women is not to change the substantive opinions of this group of economists, but to increase the level of caution with which the opinions are expressed.

Caution in expressing opinions can be both a good and bad thing. Fourcade et al (2015) argue that "confidence is perhaps the greatest achievement of the economics profession – but it is also its most vulnerable trait, its Achilles heel". Decision-makers often welcome the absence of uncertainty; Harry Truman famously demanded a "one-handed economist" who could give a clear opinion. Giving strong opinions may help economists to gain influence in policy-making. But confident opinions are not a good thing if the opinions turn out to be wrong. Although there is strong demand for certainty in economic advice, there is often genuine uncertainty in the policy situations that are faced. Strong opinions could be damaging if they are not calibrated to the context.

We can look at how sensitive individual panel members' responses are to background uncertainty, defined for each panel member as the share of (other) panel members who are uncertain. The results are reported in Appendix Table A7. We find that on two key measures – whether women are willing to give an opinion and whether they are uncertain – women are more sensitive to background uncertainty than men. In terms of giving a substantive opinion, women appear to calibrate more than men, although whether women give a strong opinion and their level of confidence are less sensitive to this measure of background uncertainty.

[12] 37 questions change from consensus to no consensus and nine change the other way.

Implications for policy-making

The opinions expressed by the EEP members (in the survey and on twitter) are in a personal capacity, without direct policy implications. We therefore turn to a different setting – voting by the members of the Bank of England Monetary Policy Committee (MPC)[13] who are responsible for setting the Bank rate – to provide evidence of gender differences in expressing opinions when those opinions directly affect real-world outcomes. The MPC meets twice as a group for discussion before members vote and the Governor suggests a rate on which members vote, but the Bank of England has an individualistic, one-person, one-vote philosophy and members are encouraged to express their personal policy preferences (Hansen et al, 2014) an approach which means that there is a relatively high level of dissent compared to the Federal Open Market Committee. Individual members' votes are published; our data capture votes from 260 meetings from June 6th, 1997 through January 30th, 2020. Over this period, there were 43 different members who voted an average of 53 times. 19 per cent of votes are made by women (11 per cent internal votes and 24 per cent external votes). We define two outcome measures intended to capture caution in voting. The first is whether MPC members vote to keep the status quo. The second is whether members vote against the consensus. The results are shown in Appendix Table A8. We find that women are 11.6 percentage points more likely than men to vote with the status quo. Also, women are 9.9 percentage points less likely than men to vote against the MPC consensus. Our take-away from this analysis is that women are cautious in their opinions than men and that gender differences carry over to real-world, policy-relevant settings.

Implications for improving diversity

Finally, gender differences in the way opinions are expressed have implications for current debates about achieving diversity within the economics profession. Achieving diversity means going beyond improving representation in numbers to ensuring that diverse voices are heard. Consider another thought experiment. Suppose that, in an economic debate, the voice that is heard is the one that expresses a strong opinion and (in a tie) is the most confident. In a world with equal representation of women, whose voices would be heard? For each of the

[13] At any one time, the MPC consists of five internal members (the Governor, the Chief Economist and three Deputy Governors) and four external members, chosen for their expertise in economics and monetary policy and drawn from industry and academia. The number of women on the MPC since it was formed in 1997 has ranged from one to three members.

nearly 400 questions, we randomly pick one man and one woman from the panel and determine whose voice is heard according to this rule. We repeat the process 500 times. By construction, the share of women's responses is 50 per cent but the (average) share of the 400 questions on which women's voices are heard over the 500 repetitions is only 43 per cent. Even though women are equally represented, their voices are not equally heard because men are more likely to express strong opinions and are more confident in their opinions. As we have already argued, the solution is not that women should behave like men and speak up, partly because they may not be listened to in the same way.

Instead, the solution lies in giving men and women an equal voice. As a starting point, IGM should re-consider its current policy of weighting panellists' opinions by their confidence since this gives more weight to male opinions. Women provide 19 per cent of the opinions but only 18 per cent of the confidence-weighted opinions. Weighting by confidence systematically downgrades the opinions of women compared to men. Within the discipline, policies and practices, such as MIT's guidance for making seminars more constructive[14] and AEA's guidelines for inclusive meetings,[15] both of which emphasize the importance of giving equal time to different groups, are likely to have an important role to play, not only in improving the culture in economics but also in ensuring that different voices (loud and quiet) can be heard

[14] http://blogs.bu.edu/ellisrp/2019/11/guidance-for-a-constructive-culture-of-exchange-plus-two-addenda/
[15] https://www.aeaweb.org/resources/best-practices/leading-departments

References

Adams, R. and Funk, P. (2012) “Beyond the Glass Ceiling: Does Gender Matter?” Management Science, 58(2), pp. 219-235.

Alan, S., Ertac, S., Kubilay, E. and Loranth, G. (2020) “Understanding Gender Differences in Leadership”. The Economic Journal, 130(626), pp. 263-289.

Auriol, E, G Friebel and S Wilhelm (2019), “Women in European economics”, VoxEU.org, 19 November

Babcock, L. and S. Laschever (2004) “Women don’t ask: Negotiation and the Gender Divide”, Southern Economic Journal

Barber, B. M. and Odean, T. (2001), “Boys will be Boys: Gender, Overconfidence, and Common Stock Investment”, The Quarterly Journal of Economics 116 (1), pp. 261–292

Brescoll V L (2011) “Who Takes the Floor and Why: Gender, Power, and Volubility in Oganizations”. Administrative Science Quarterly 56(4) pp. 622-641.

Coffman,K (2014) "Evidence on Self-Stereotyping and the Contribution of Ideas." Quarterly Journal of Economics 129(4) pp. 1625–1660.

Correia, S (2017). reghdfe: Stata module for linear and instrumental-variable/gmm regression absorbing multiple levels of fixed effects. Statistical Software Components s457874, Boston College Department of Economics.

Chari, Anusha, and Paul Goldsmith-Pinkham (2017) Gender representation in economics across topics and time: Evidence from the NBER summer institute. No. w23953. National Bureau of Economic Research

CSWEP (2021) FOCUS Introduction: Women in Economics, An International Perspective file (aeaweb.org)

Dolado, J.J., Felgueroso, F. & Almunia, M. (2012) Are men and women-economists evenly distributed across research fields? Some new empirical evidence. *SERIEs* **3, pp.** 367–393

Fourcade, M., Ollion, E. and Y. Algan (2015) "The Superiority of Economists." *Journal of Economic Perspectives*, 29 (1), pp. 89-114.

Gordon, Roger, and Gordon B. Dahl. 2013. "Views among Economists: Professional Consensus or Point-Counterpoint?" American Economic Review, 103 (3): pp. 629-35.

Hansen, S., McMahon, M. and C. Velasco Rivera (2014) “Preferences or private assessments on a monetary policy committee?” Journal of Monetary Economics

Hanspach, P., Sondergeld, V and J. Palka (2021) Few top positions in economics are held by women VoxEU

Levine, S. Apfelbaum, E., Bernard, M. Bartelt, V., Zajac, E. and Stark, D. (2014) Ethnic diversity deflates price bubbles Proceedings of the National Academy of Sciences Dec 2014, 111 (52) 18524-18529

Lundeberg, A., Fox, P. and J. Punc´cohaar´ (1994) “Highly Confident but Wrong: Gender Differences and Similarities in Confidence Judgments.” Journal of Educational Psychology 86 (1): 114–21.

Koffi, M (2021) Gendered Citations at Top Economic Journals” AEA Papers and Proceedings, Vol.11

May, A., McGarvey, M. and R. Whaples (2014) “Are Disagreements among Male and Female Economists Marginal at Best?: A Survey of AEA Members and Their Views on Economics.” Contemporary Economic Policy 32 (1): 111–32

May, A., McGarvey, M. and D. Kucera (2018 “Gender and European Economic Policy: A survey of the views of European Economists on Contemporary Economic Policy" Kyklos 71 (1): 162–83.

Moore and Healy (2008), “The Trouble With Overconfidence”, Psychological Review 115(2):502-17

Niederle, M. and Vesterlund, L., 2007. Do Women Shy Away From Competition? Do Men Compete Too Much?. The Quarterly Journal of Economics, 122(3), pp.1067-1101.

Sarsons, Heather, and Guo Xu. 2021. "Confidence Men? Evidence on Confidence and Gender among Top Economists." AEA Papers and Proceedings, 111: 65-68.

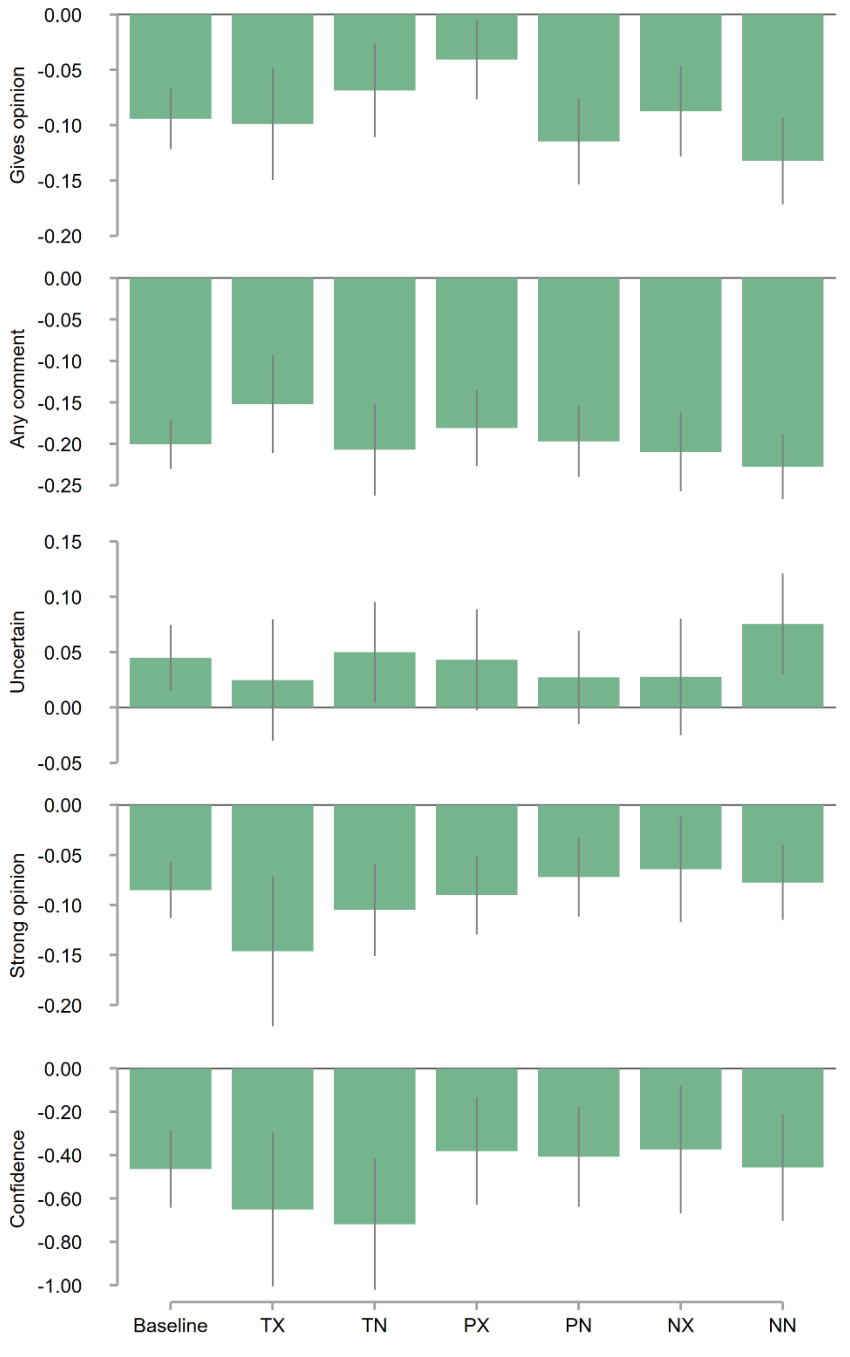
Thomas-Hunt MC, Phillips KW. When What You Know Is Not Enough: Expertise and Gender Dynamics in Task Groups. Personality and Social Psychology Bulletin. 2004;30(12):1585-1598

Table 1: Summary statistics, by gender

	Female		Male		P-value
	Mean	SD	Mean	SD	
A. Expert characteristics					
US sample	0.50	0.51	0.49	0.50	0.92
Age of expert (2020-PhD Year)	26.68	8.64	33.96	9.95	0.00
PhD from MIT, Harvard, LSE, Oxford, Stanford, Yale, Princeton	0.68	0.45	0.72	0.48	0.73
Citations (total)	16129.09	9777.54	38722.54	34168.76	0.00
Number of questions received	164.77	90.50	187.38	101.59	0.35
Number of questions answered	126.91	72.16	156.70	96.67	0.18
Observations	22		82		
B. Response characteristics					
Question is within expert's field	0.35	0.48	0.37	0.48	0.02
Gives opinion	0.77	0.42	0.84	0.37	0.00
Response: Strongly Disagree	0.04	0.20	0.05	0.23	0.01
Response: Disagree	0.13	0.33	0.15	0.36	0.00
Response: Uncertain	0.25	0.43	0.21	0.41	0.00
Response: Agree	0.42	0.49	0.39	0.49	0.00
Response: Strongly Agree	0.16	0.37	0.20	0.40	0.00
Any comment	0.26	0.44	0.43	0.49	0.00
Confidence (1=lowest, 10=highest)	6.01	2.56	5.99	2.40	0.66
Observations	3,625		15,365		

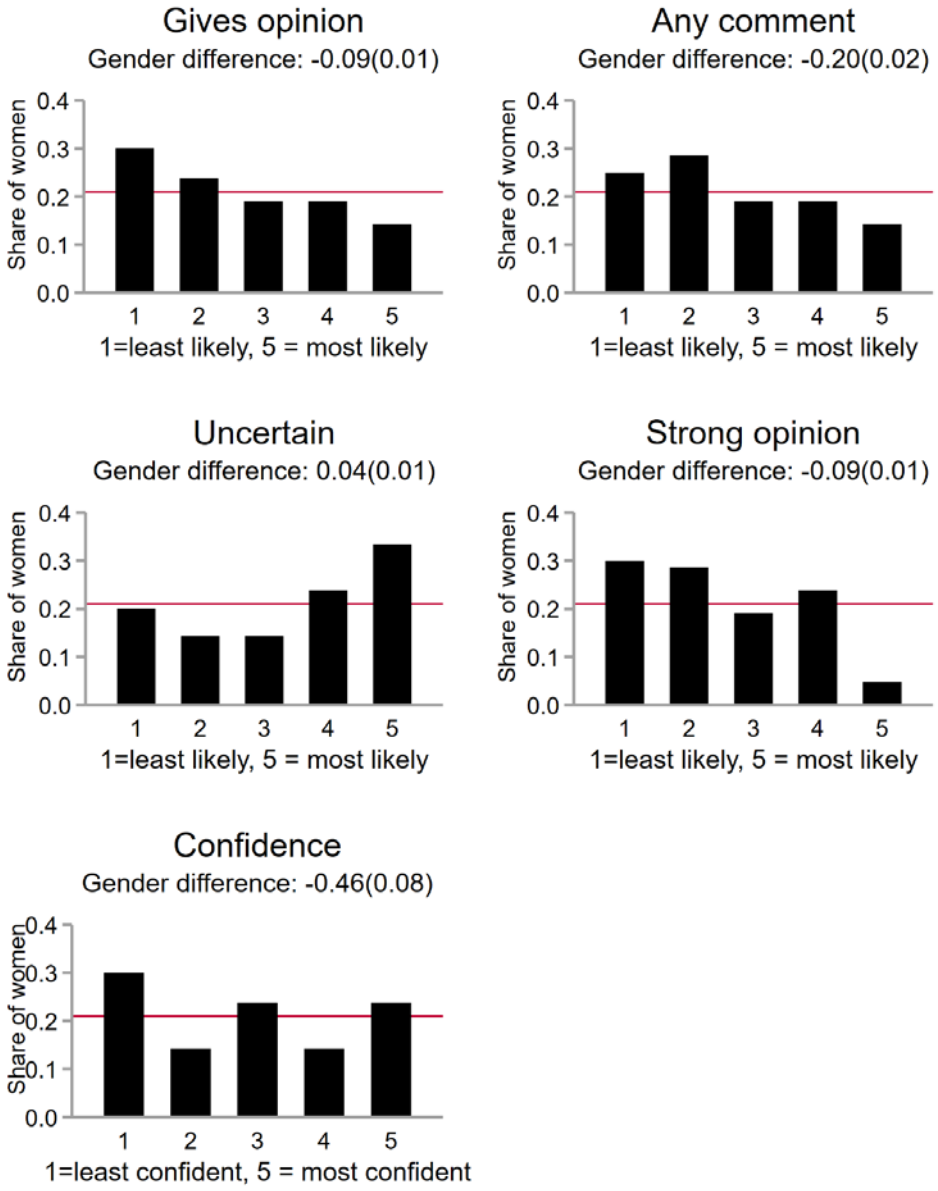
Notes: IGM Economist Expert Survey. Citations are according to Google Scholar in June 2020. Experts' fields based on NBER/ CEPR affiliations (see Appendix Table A1). The aggregate summary statistics are shown in Appendix Table A2.

Figure 1: Estimated gender gaps – Whole Sample (baseline) and by Question type/ Expertise



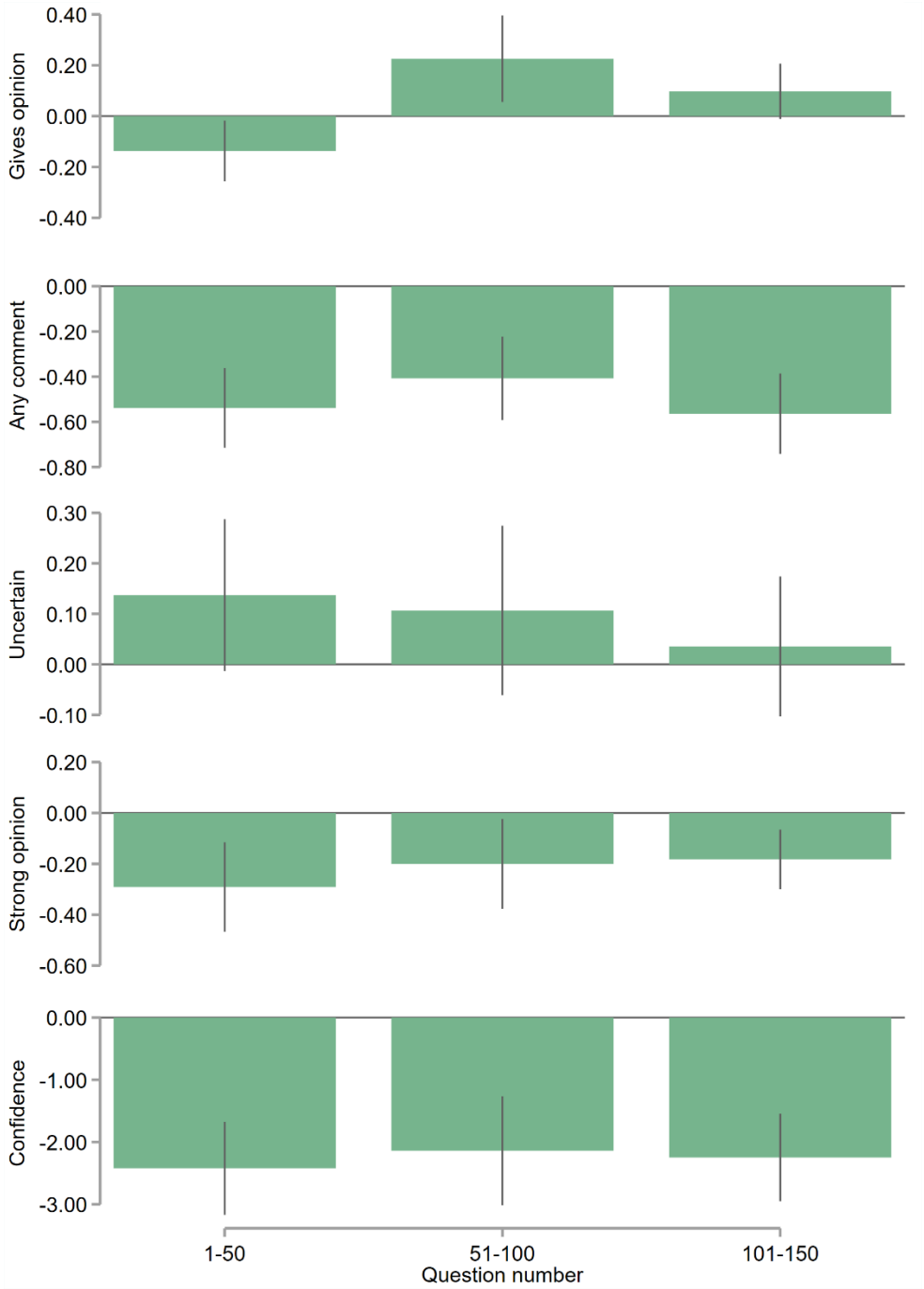
Notes: The chart is based on estimating specifications with question, institution, and PhD year fixed effects, as well as linear controls for the number of citations, the H-index, and dummy variables for the US sample, and for nationality. The bars show the coefficients on an indicator for female; spikes show the 95 percent confidence intervals based on standard errors clustered at the question level. ***Baseline*** *refers to the gender difference in the full sample.* ***TX*** *= theory questions, panellist is an expert.* ***TN*** *= theory questions, outside field.* ***PX*** *= positive questions (i.e. questions for which there is evidence), expert.* ***PN*** *= positive questions, outside field.* ***NX*** *= normative questions (i.e. questions which require a value judgement), expert.* ***NN*** *= normative questions, outside field. Appendix Table A3 shows the baseline regression coefficients and standard errors and Appendix Table A4 shows the coefficients and standard errors for interactions by question type and expert status.*

Figure 2: Female representation across the distribution



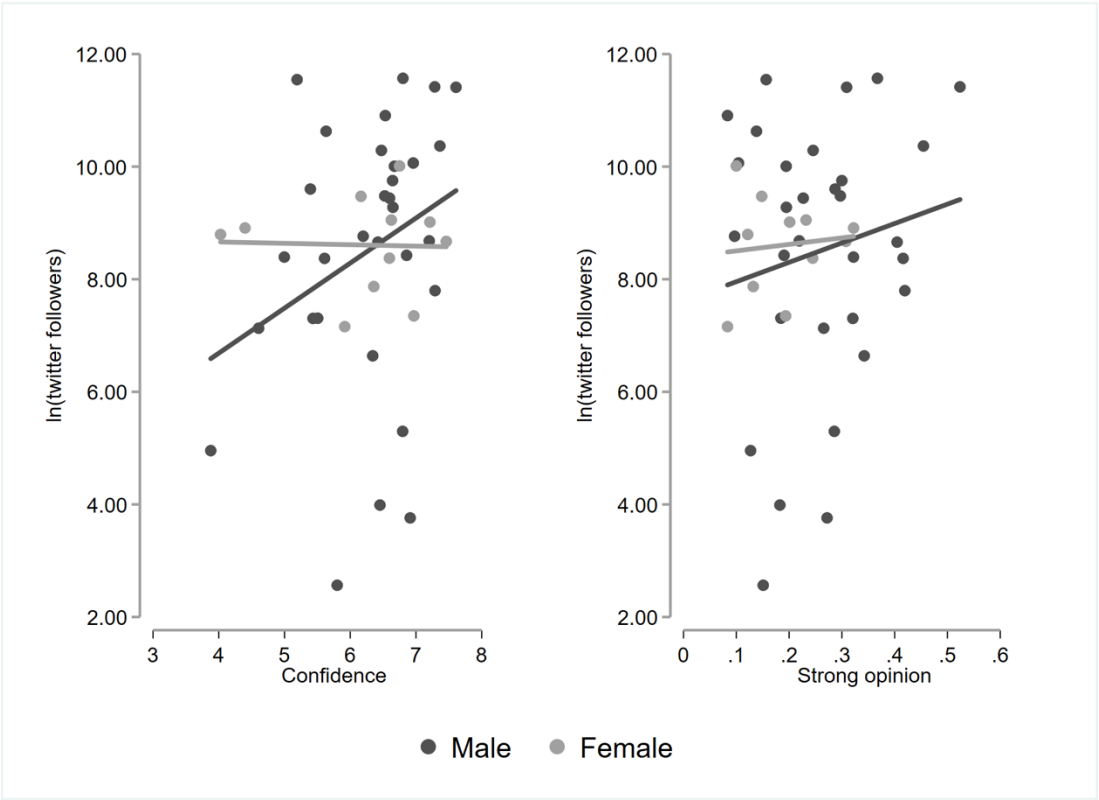
Notes: The chart is based on residuals from specifications, excluding an indicator for female, including the same controls as in Figure 1. The distribution is of average residuals at the expert-level. The bars show the shares of women in each quintile of the distribution. The horizontal line shows the overall share of female economists in the panel (21 per cent).

Figure 3: Estimated gender gaps, by question number



Notes: The chart is based on estimating the same, baseline specifications as in Figure 1, but separately for different question numbers. The bars show the coefficients on an indicator for female; spikes show the 95 percent confidence intervals based on standard errors clustered at the question level.

Figure 4: Male/ female voices and twitter followers



Notes: The charts plot the log of the number of twitter followers for each panel-member against measures of voice (average share of questions answered by the panel member with a strong opinion and panel members' average confidence in their responses), together with an estimated linear fit.

Charts and tables for Online Appendix

Table A1: Mapping of NBER and CEPR program affiliations to fields

	NBER	CEPR
Development	Development Political Economy Environment and Energy Economics	Development
International	International Finance and Macroeconomics International Trade and Investment	International Macroeconomics and Finance International Trade and Regional Economics
Finance	Asset Pricing Corporate Finance	Financial Economics
Industrial Organisation	Industrial Organisation Productivity	Industrial Organisation
Labor	Labor Education	Labor Economics
Macroeconomics	Economic Fluctuations and Growth Monetary Economics	Macroeconomics and Growth Monetary Economics and Fluctuations
Public	Aging Public Economics Environment and Energy Economics Health Care Children	Public Economics

Notes: NBER and CEPR affiliations are mapped to seven fields to define panel members' areas of expertise. Fields are not exclusive.

Table A2: Descriptive Statistics, full sample

	Mean	SD	Min	Max
A. Expert characteristics				
US sample	0.49	0.50	0	1
Age of expert (2020-PhD Year)	32.42	10.10	14	53
Citations (total)	33943.15	31993.63	2668	150,000
Number of questions received	182.60	99.36	27	323
Number of questions answered	150.39	92.52	2	323
Observations	104			
B. Question characteristics				
Type: Theory	0.18	0.38	0	1
Type: Positive	0.42	0.49	0	1
Type: Normative	0.40	0.49	0	1
Observations	396			
C. Response characteristics				
Expert's field	0.37	0.48	0	1
Response: Strongly Disagree	0.05	0.22	0	1
Response: Disagree	0.14	0.35	0	1
Response: Uncertain	0.22	0.41	0	1
Response: Agree	0.39	0.49	0	1
Response: Strongly Agree	0.19	0.39	0	1
Any comment	0.40	0.49	0	1
Confidence (1=lowest, 10=highest)	5.99	2.43	1	10
Gives opinion	0.82	0.38	0	1
Observations	18,990			

Notes: IGM Economist Expert Survey. Citations are according to Google Scholar in June 2020. Experts' fields based on NBER/ CEPR affiliations (see Table A1).

Table A3: Regression results – Gender differences, whole sample

	Gives opinion	Any comment	Uncertain	Strong opinion	Confidence
	(1)	(2)	(3)	(4)	(5)
Female	-0.094	-0.201	0.045	-0.085	-0.464
	(0.014)	(0.015)	(0.015)	(0.014)	(0.090)
Citations (total)	0.000	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
H index overall	0.001	-0.006	-0.000	-0.001	-0.019
	(0.000)	(0.000)	(0.000)	(0.000)	(0.003)
US sample	0.022	-0.148	0.007	0.036	-0.517
	(0.022)	(0.028)	(0.022)	(0.027)	(0.146)
American	-0.019	0.048	-0.046	0.032	0.245
	(0.015)	(0.022)	(0.019)	(0.017)	(0.099)
European	0.051	-0.249	-0.051	0.029	-0.029
	(0.018)	(0.023)	(0.021)	(0.019)	(0.120)
Constant	0.776	0.829	0.244	0.255	7.339
	(0.032)	(0.038)	(0.034)	(0.036)	(0.212)
Observations	18990	15641	15641	15641	15645
R-squared	0.18	0.27	0.16	0.23	0.25
Mean of dep var	0.82	0.40	0.22	0.24	5.99

Notes: All coefficients are based on the specification with question, institution, and PhD year fixed effects, as well as linear controls for the number of citations, the H-index, and a dummy variable for the US sample, a dummy variable for expert being European, and a dummy variable for the expert being American. Standard errors clustered at the question level in parentheses.

Table A4: Regression results – Gender differences by question type X expertise

	Gives opinion	Any comment	Uncertain	Strong opinion	Confidence
	(1)	(2)	(3)	(4)	(5)
Female	-0.099	-0.152	0.025	-0.146	-0.651
	(0.026)	(0.030)	(0.028)	(0.038)	(0.180)
Positive X Female	0.058	-0.029	0.018	0.056	0.269
	(0.027)	(0.034)	(0.033)	(0.040)	(0.204)
Normative X Female	0.011	-0.058	0.003	0.082	0.276
	(0.029)	(0.034)	(0.035)	(0.043)	(0.211)
Theory X Non-expert X Female	0.030	-0.055	0.025	0.041	-0.068
	(0.026)	(0.035)	(0.034)	(0.041)	(0.219)
Positive X Non-expert X Female	-0.074	-0.016	-0.016	0.018	-0.026
	(0.020)	(0.026)	(0.028)	(0.024)	(0.142)
Normative X Non-expert X Female	-0.045	-0.018	0.048	-0.013	-0.083
	(0.024)	(0.025)	(0.027)	(0.029)	(0.145)
Theory X Non-expert	-0.010	0.002	0.032	-0.082	-0.655
	(0.017)	(0.020)	(0.014)	(0.019)	(0.097)
Positive X Non-expert	-0.046	-0.039	0.017	-0.072	-0.827
	(0.011)	(0.014)	(0.012)	(0.012)	(0.076)
Normative X Non-expert	-0.041	-0.047	0.016	-0.065	-0.833
	(0.011)	(0.014)	(0.012)	(0.012)	(0.070)
Citations (total)	0.000	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
H index overall	0.000	-0.007	0.000	-0.002	-0.023
	(0.000)	(0.000)	(0.000)	(0.000)	(0.003)
US sample	0.017	-0.152	0.009	0.027	-0.614
	(0.022)	(0.028)	(0.022)	(0.026)	(0.149)
American	-0.022	0.046	-0.045	0.030	0.207
	(0.015)	(0.022)	(0.019)	(0.017)	(0.099)
European	0.045	-0.253	-0.050	0.027	-0.072
	(0.018)	(0.023)	(0.021)	(0.020)	(0.120)
Constant	0.821	0.866	0.223	0.326	8.151
	(0.033)	(0.039)	(0.034)	(0.037)	(0.215)
Observations	18990	15641	15641	15641	15645
R-squared	0.18	0.27	0.16	0.23	0.25
Mean of dep var	0.82	0.40	0.22	0.24	5.99

Notes: All coefficients are based on the specification with question, institution, and PhD year fixed effects, as well as linear controls for the number of citations, the H-index, and a dummy variable for the US sample, a dummy variable for expert being European, and a dummy variable for the expert being American. Standard errors clustered at the question level in parentheses.

Table A5: Regression results – Gender differences by male/ female fields

	Gives opinion	Any comment	Uncertain	Strong opinion	Confidence
	(1)	(2)	(3)	(4)	(5)
Female	-0.106	-0.213	0.048	-0.071	-0.579
	(0.017)	(0.017)	(0.019)	(0.017)	(0.113)
Fem X femfield	0.020	0.020	-0.006	-0.025	0.184
	(0.016)	(0.017)	(0.017)	(0.016)	(0.097)
Expert	0.045	0.038	-0.023	0.070	0.798
	(0.007)	(0.008)	(0.007)	(0.008)	(0.044)
Citations (total)	0.000	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
H index overall	0.000	-0.007	0.000	-0.002	-0.023
	(0.000)	(0.000)	(0.000)	(0.000)	(0.003)
US sample	0.015	-0.153	0.009	0.027	-0.611
	(0.022)	(0.028)	(0.022)	(0.026)	(0.149)
American	-0.020	0.047	-0.044	0.029	0.206
	(0.015)	(0.022)	(0.019)	(0.017)	(0.098)
European	0.047	-0.251	-0.050	0.025	-0.073
	(0.018)	(0.023)	(0.021)	(0.019)	(0.119)
Constant	0.780	0.829	0.243	0.257	7.355
	(0.032)	(0.038)	(0.034)	(0.036)	(0.210)
Observations	18990	15641	15641	15641	15645
R-squared	0.18	0.27	0.17	0.24	0.27

Notes: All coefficients are based on the specification with question, institution, and PhD year fixed effects, as well as linear controls for the number of citations, the H-index, and a dummy variable for the US sample, a dummy variable for expert being European, and a dummy variable for the expert being American. Female fields are Development, Labor, Public and IO. Standard errors clustered at the question level in parentheses.

Table A6: Regression results – Relationship between voice and influence (twitter followers)

Dependent variable = ln(twitter followers)

	(1)	(2)	(3)	(4)	(5)	(6)
Female	1.363	2.333	2.194	-0.207	4.301	7.485
	(0.693)	(3.409)	(1.319)	(1.830)	(1.848)	(3.317)
Twitter age (years)	0.194	0.168	0.203	0.201	0.234	0.175
	(0.095)	(0.093)	(0.103)	(0.101)	(0.101)	(0.085)
Mean(Gives opinion)		-3.547				
		(1.895)				
Female X Gives opinion		-1.479				
		(4.210)				
Mean(Any comment)			2.453			
			(1.703)			
Female X Any comment			-1.616			
			(2.496)			
Mean(Uncertain)				-5.074		
				(5.413)		
Female X Uncertain				7.268		
				(8.421)		
Mean(Strong opinion)					5.137	
					(3.486)	
Female X Strong opinion					-12.671	
					(6.503)	
Mean(Confidence)						1.069
						(0.474)
Female X Confidence						-0.973
						(0.532)
Observations	42	42	42	42	42	42
R-squared	0.249	0.331	0.302	0.276	0.316	0.364

Notes: Mean variables defined at the individual panel-member level. *Standard errors in parentheses*

Table A7: Regression results – Effect of background uncertainty

	Gives opinion	Any comment	Uncertain	Strong opinion	Confidence
	(1)	(2)	(3)	(4)	(5)
Female	-0.073	-0.204	0.020	-0.107	-0.710
	(0.015)	(0.020)	(0.017)	(0.018)	(0.103)
FemXUncertainty	-0.095	0.015	0.081	0.105	1.190
	(0.041)	(0.057)	(0.047)	(0.052)	(0.290)
Uncertainty	-0.066	-0.436	-6.681	0.807	7.483
	(0.108)	(0.143)	(0.118)	(0.129)	(0.724)
Citations (total)	0.000	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
H index overall	0.001	-0.006	-0.000	-0.001	-0.019
	(0.000)	(0.001)	(0.000)	(0.000)	(0.003)
US sample	0.024	-0.139	0.152	0.018	-0.680
	(0.020)	(0.026)	(0.022)	(0.024)	(0.133)
American	-0.019	0.050	-0.028	0.030	0.229
	(0.015)	(0.021)	(0.017)	(0.018)	(0.104)
European	0.050	-0.248	-0.035	0.027	-0.046
	(0.017)	(0.023)	(0.019)	(0.021)	(0.118)
Constant	0.790	0.917	1.594	0.092	5.829
	(0.037)	(0.050)	(0.041)	(0.045)	(0.250)
Observations	18990	15641	15641	15641	15645
R-squared	0.18	0.27	0.31	0.24	0.25

Notes: All coefficients are based on the specification with question, institution, and PhD year fixed effects, as well as linear controls for the number of citations, the H-index, and a dummy variable for the US sample, a dummy variable for expert being European, and a dummy variable for the expert being American. Uncertainty is defined by the share of other panel members who respond uncertain. Standard errors clustered at the question level in parentheses.

Table A8: Regression results – Voting by Monetary Policy Committee members

	Vote status quo (1)	Vote against majority (2)	Only one voting against (3)
Female	0.116	-0.099	-0.037
	(0.018)	(0.019)	(0.011)
Internal member	0.110	-0.129	-0.045
	(0.016)	(0.015)	(0.008)
Constant	0.682	0.182	0.589
	(0.010)	(0.010)	(0.005)
Observations	2280	2230	2230
Mean dep var	0.760	0.100	0.030

Notes: All coefficients are based on estimating a specification with meeting fixed effects and a dummy for whether the member is an internal member. Standard errors clustered at the question level in parentheses.

Figure A1: Non-parametric analysis, opinions over time