Confidence Men? Gender and Confidence: Evidence among Top Economists*

Heather Sarsons† and Guo Xu‡

December 12, 2016

Abstract

Using data from economists working in top U.S. universities, we find that women are less confident than men along three margins. When asked about their level of agreement on survey questions about the economy, women are less likely to provide a judgement than their male counterparts. Conditional on providing a judgement, women are less likely to give "extreme" answers in which they strongly agree or disagree. Women are also less confident in the accuracy of their answer. We show that the confidence gap is driven by women being less confident when asked questions outside their field of expertise.

*We thank Marianne Bertrand, Katherine Baldiga Coffman, John Coglianese, Laura Derksen, Claudia Goldin, Daniele Paserman and Dina Pomeranz for their valuable comments. All errors remain our own.

†Harvard University, Department of Economics, Littauer Center, 1805 Cambridge Street, Cambridge MA 02138, USA (e-mail: sarsons@fas.harvard.edu, phone: +1-6174952144)

‡London School of Economics, Department of Economics and STICERD, Houghton Street, London WC2A 2AE, UK (e-mail: g.xu2@lse.ac.uk, phone: +44-2074057686)

1 Introduction

Gender gaps in labour market outcomes have remained large despite continued efforts to promote equality. While the sources of these gaps have traditionally been ascribed to differences in human capital accumulation and discrimination, a growing body of literature attributes gender gaps to psychological factors such as confidence (Bertrand 2011, Samek 2015).

This paper uses survey data from a select group of tenured economists at top U.S. universities to document a confidence gap even between highly successful men and women. Women are vastly underrepresented in the economics profession. Currently, approximately 30% (15%) of all assistant (full) professors in economics departments are women (Bayer and Rouse, 2016). Women who are tenured at top universities are therefore very selected. If underconfidence hinders a woman's chance at promotion, those women who do "break the glass ceiling" should be at least as confident as men. Rising through the ranks could even heighten a woman's sense of confidence and beliefs about her ability, making her more confident than her male counterparts. Evidence for confidence gaps among men and women who make it to the top of their careers, however, remains scant.

We show that women are less confident than men along three margins. First, when asked to provide assessments to survey questions about the economy, women are 4% points less likely to provide any answer[1]. Second, conditional on answering, women are 7% points less likely to give "extreme" answers in which they strongly agree or disagree. This is known as overestimation in psychology: men have more extreme answers because they believe themselves to be better able to judge the correct answer. Finally, women are less confident in the accuracy of their answer. Women express a level of confidence that is 0.27 points (or 0.11 standard deviations) lower than a comparable man, as measured on a scale of 1 (unconfident) to 10 (very confident). This is known as overprecision: men are overly confident that their answer is the correct one regardless of how extreme it is. The results persist after controlling for the year the PhD was granted, the PhD awarding institution, the current institution, and the number of solo and co-authored publications up to tenure. An independent dataset for a select group of U.K. economists validates the results.

We also shed light on mechanisms driving the confidence gap. We find that the gap is largely driven by women being less confident when asked questions that are outside of their field of expertise. Women are as confident as men when answering questions related to their primary field but are less confident when asked a question outside of their primary field. This result is not driven by men having a greater breadth of expertise than women. This suggests that previously documented confidence gaps could be explained by women having a better sense of the bounds of their expertise than men. Finally, we show that men's confidence increases when others agree with their view while women are slow to adjust their confidence.

The remainder of the paper is organized as follows. Section 2 describes the data used in our analysis. In Section 3, we present the main results and explore mechanisms that might explain the findings. Section 4 concludes.

[1]This could reflect other factors such as women having less time to respond to the survey.

2 Data

Our main data source is the Initiative of Global Markets (IGM) survey. The survey asks a select group of 51 economists, all at top U.S. institutions, questions related to the current state of the economy as well as other policy issues. The economists are chosen to represent a range of political views, ages, and research interests. All are reported to have a "keen interest in public policy".[2]

A total of 255 questions[3] were asked between September 2011 and November 2016, leaving us with 6,986 responses. Respondents are not required to answer each question, but the response rate is high at 83%. The questions cover a range of economic topics, from education to the minimum wage to monetary policy. Participants answer individually and their answers are publicly displayed.

When respondents decide to provide a judgement, they can answer each question along two dimensions. First, they are asked the degree to which they agree with a given statement (e.g. "Some Americans who work in the production of competing goods, such as clothing and furniture, are made worse off by trade with China."), measured on a Likert scale of 1 (strongly disagree) to 5 (strongly agree). Second, they are asked how much confidence they have in their answer which is measured on a scale of 1 (lowest) to 10 (highest).

Each response captures a different aspect of confidence. First, providing a publicly visible answer is in itself a signal of confidence. Second, the extent to which respondents agree or disagree captures whether respondents are overly confident in their level of dis/agreement with the statement, known as overestimation in the psychology literature (Moore and Healy 2008). Finally, the confidence measure explicitly elicits the subjective perception over the accuracy of an individual's response, known as overprecision. We look at whether the confidence gap is primarily a result in differences in extremeness of opinions or confidence in views. As shown in Figures 1 and 2, women are less likely to hold strong views than men and are more dispersed in their self-reported confidence.

We combine the IGM data with pre-tenure characteristics for the 51 economists. The economists are at different stages of their careers, but all have received tenure. This is therefore the period for which more comparable data is obtainable from their public CVs. These data include each economist's primary field of specialization, the year the PhD was received, the number of published papers broken down into solo and coauthored papers, and whether the individual received tenure at his or her initial placement school. These data are collected and coded as described in Sarsons (2015). Due to missing values, the final sample consists of 47 economists.

Table 1 summarizes the differences among male and female economists of the IGM panel. Given the highly selected nature of a sample of top economists, male and female economists are, in terms of their average characteristics, comparable (Panel A). However, given the small sample size, some differences may be insignificant due to lack of power. Women appear, for example, to be younger than their male counterparts, as indicated by the later average year of PhD award, although the

[2]Refer to http://www.igmchicago.org/igm-economic-experts-panel/, retrieved 11 June 2015 and Gordon and Dahl (2013) for a description and alternative use of the data. The set of economists answering questions only changes in that new members have been added to the panel over time.

[3]The full list of questions can be viewed at http://www.igmchicago.org/igm-economic-experts-panel/poll-results?SurveyID=SV_72JJHkpH4FvJb9j.

difference is statistically insignificant. The share of PhDs awarded from Harvard and MIT are comparable at roughly 60%. While there are differences in fields, these are, with the exception of International/Trade, insignificant. The number of publications is also comparable across gender.

Even with the highly selected sample, we observe stark differences in the voting behavior (Panel B). Women are significantly less likely to provide an assessment altogether. While the average level of agreement with a statement is the same for those who do provide assessments, the distributions for men and women are very different. Men hold more extreme views, as measured by the share of extreme answers, and believe these views to be more accurate, as measured by the self-reported confidence level. These differences are significant at the 1% level.

3 Results

3.1 Main Results

To test whether women are less confident than men in a multivariate setting, we estimate:

$$y_{ijs} = \beta_1 fem_i + x_i'\gamma + \theta_j + \mu_s + \epsilon_{ijs} \quad (1)$$

where fem_i is a female indicator for respondent i, x_i is a vector of individual-level controls discussed in Section 2, and θ_j and μ_s are question and school fixed effects. The question fixed effect is included to confine the identifying variation to within-question comparisons. As a conservative specification and to ensure that the results are not driven by differences across schools, we also include current school fixed effects to compare gender differences only among academics at the same school.[4]

As discussed in the data section, we measure confidence, y_{ijs}, in three ways: (i) the propensity to provide any expert judgement (ii) the propensity to provide extreme judgements, as measured by a dummy for whether the respondent strongly agreed or disagreed with a statement, and (iii) the self-reported confidence level on an integer scale of 1 to 10. If a confidence gap exists, with women being less confident than men, we expect β_1 to be negative. We cluster the standard errors on the question-level.[5]

The results are reported in Table 2 and confirm that a gap exists for both measures of confidence. When asked to provide a judgement to a question pertaining the economy, women are significantly less likely to provide an answer (Column 1). This gap is not driven by gender-specific differences across questions or schools as it is robust to question, school fixed effects and pre-tenure individual controls (Columns 2)[6]. However, women could be less likely to provide public answers for reasons other than confidence. If women are more likely to sit on many committees or to take care of

[4]The same university may both have a business school and an economics department. We distinguish between both institutions within the same university by including separate school fixed effects.

[5]We also cluster at the individual level. However, this leaves us with only 46 clusters which greatly reduces the significance of our results. More data is therefore needed to determine whether the results are robust to individual-level clustering.

[6]The results are also robust to controlling for an individual's current number of citations.

children, the pulls on their time might prevent them from responding to the survey. With this caveat in mind, we move to the other more direct measures of confidence.

When providing an answer, women are 7% points less likely to provide extreme judgements (Column 3-4).[7]In terms of magnitude, the confidence gap is economically large. Compared to the mean of the dependent variable (24.5%), this corresponds to a gap of 30%. The gap is somewhat smaller for the self-reported confidence level but nonetheless non-trivial (Columns 5-6). On average, women tend to report a confidence score that is 0.25 points lower than men. This corresponds to a gap of 6% when evaluated against the mean. Interestingly, those who were awarded PhDs from Harvard or MIT are 0.47 points more confident than respondents who received their PhD elsewhere. Among these schools, the gender confidence gap is about half of the size.[8]

3.2 Mechanisms

Having confirmed that a confidence gap exists among these highly selected individuals, we now attempt to disentangle mechanisms that could be driving the results. In particular, we explore whether there is a difference in men and women's confidence in answering questions outside of their primary field of expertise. To do so, we exploit variation in the questions asked. Respondents are asked about questions related to many areas of economics, allowing us to look at differences in confidence when being asked about topics they are less familiar with. From respondents' CVs, we coded up to two listed primary and secondary fields[9]. We then estimate the following equation to test whether women's lack of confidence appears when they are asked questions about topics outside of their fields:

$$y_{ijs} = \beta_1 fem_i + \beta_2 foreign_{ij} + \beta_3 \left(fem_i \times foreign_{ij}\right) + x_i^{'}\gamma + \theta_j + \mu_s + \epsilon_{ijs} \tag{2}$$

Here, $foreign_{ij}$ equals one when the question j is outside of respondent i's primary field and all other variables are defined as in Section 3.1.

The results are presented in Table 3 for both measures of confidence. The results show that both men and women are less confident when being asked questions outside of their field, but that a confidence gap persists. For example, while men are 4.6 percentage points less likely to give an extreme answer when speaking on a topic outside of their primary field, women are 6 percentage points less likely to do so (Column 2). This result is statistically insignificant for the self-reported level of confidence (Column 3), but the point estimates indicate that women are also less confident in their answers when responding to questions outside of their field.[10] More importantly, accounting

[7]This finding is consistent with Mondak and Anderson (2004), who document that women are more likely to report "I don't know" in surveys of political knowledge than men.

[8]The reader can therefore expect the confidence gap between the co-authors of this paper to be more than closed upon completion of their PhDs.

[9]If primary fields were not listed on a respondent's CV, we inferred a primary field based on their publication record.

[10]We would expect women to be less confident than men when answering questions outside of their field if women

for the differential confidence when moving beyond one's own field "explains away" the level effect of gender. It appears that the confidence gap emerges when women are speaking on topics for which they might be less informed.

These findings resonate with Coffman (2014) who shows that both men and women are less likely to contribute ideas in an area that is traditionally associated with the opposite sex. For example, women are less likely to contribute an idea when asked a math question and men are less likely to contribute an idea when asked about the arts. Coffman also finds that even when men and women find out that they are the most knowledgeable on a topic, they are still unwilling to contribute their ideas provided the topic is stereotypically associated with the opposite sex. While we find that women who are asked a question related to their area of expertise are not less confident than men are, it is possible that they feel uncomfortable expressing their opinion because the field as a whole is steretoypically male. It could also be that successful women are not underconfident but rather are more aware of the bounds of their expertise. Given the data constraints, however, we are unable to distinguish between these two interpretations.

3.2.1 Are women responding to disagreement?

Some of the questions on the IGM panel are more controversial than others. For example, all economists agree that the benefits of free trade and NAFTA outweigh the costs. There is much disagreement, though, as to whether using tax incentives to affect a firm's location choice is beneficial. In Figure 3, we examine the relationship between women's confidence and the level of disagreement on a question. Specifically, we estimate

$$y_{ijs} = \beta_1 SDrating_j + \beta_2 fem_i + \beta_3 \left(fem_i \times SDrating_j\right) + x_i^{'}\gamma + \theta_j + \mu_s + \epsilon_{ijs} \tag{3}$$

where $SDrating_j$ is the standard deviation in responses for question j, and plot the regression slopes $\hat{\beta}_1$ and $\hat{\beta}_1 + \hat{\beta}_3$. The corresponding regression table is shown in Appendix Table 5 In Figure 3, we see that a confidence gap exists when there is consensus on a question, with men being significantly more confident than women. On questions with more disagreement, though, men and women have roughly equal levels of confidence. This result suggests that both men and women recognize when a topic is contestable and adjust their confidence accordingly. However, women do not seem to take others' agreement with their view as a signal to be confident. We cannot say, though, whether men are being overconfident in these situations or if women are being underconfident. It also could be that while there is a consensus within the economics community on a topic, there is broad disagreement in other communities and women take this into account more than men. From this analysis, we can only say that women seem to be "sticky" in their confidence as they do not adjust it as much as men do. Note that the coefficient on *female* × *foreign field* remains negative and significant, providing further evidence for the robustness of our results.

actually have a narrower range of expertise than men do. The results do not change when we control for the respondents' breadth of expertise using RePEc data and allow breadth to vary by gender (Appendix Table 6)

3.2.2 The Wallflower Effect

A complementary explanation is the existence of a "wallflower effect" in which women do not want to stand out and therefore give answers that are closer to the mean. Researchers have documented this phenomenon experimentally. Linardi and Jones (2014), for example, show that individuals whose charitable donations are made public are more likely to donate the mean amount of past donations. Women are particularly likely to condition their behaviour on what others have done. In our context, women may avoid giving extreme answers to avoid standing out. This might in part have to do with confidence but there could also be other aspects, such as being given a hard time if they are wrong, that drive women to stick to the mean. In Appendix Table 8, we look at whether women are less likely to both strongly disagree or strongly agree with a statement. Women are 6 percentage points less likely to strongly agree with an answer but are not statistically less likely to strongly disagree with an answer. The wallflower effect would imply that women shade their answers on both ends, making them less likely to both strongly agree and strongly disagree. That women are no less likely to strongly disagree than men are suggests that the results are not driven by a desire to blend in.

3.2.3 Robustness and External Validity

We provide robustness checks to rule out alternative mechanisms (Table 7). An alternative explanation for the gender gap in confidence is that women sort into fields in which people are generally less confident or less extreme in their answers. It could be, for example, that macroeconomists are especially confident. Since there are few women in macroeconomics, the effects could be picking up this sorting rather than measuring an overall lack of confidence among women. However, the confidence gap persists, and is in fact larger, when controlling for field of study (Column 2) suggesting that sorting is not at play. We also include answer fixed effects in Column 3 of Table 7. In this sense, we are holding constant one margin of confidence and asking, for example, for all individuals who strongly agree with a statement, are there differences in how confident they are about strongly agreeing? The fixed effect takes out any correlation between the extremeness of the answer and the confidence in the answer. The fact that the size of the coefficient decreases, suggests that strongly agreeing with an answer is correlated with being more confident and that women are less likely to strongly agree with answers. Further, within each answer type, women are still less confident than men. Some of the confidence gap is thus driven by men having more extreme stances in addition to being more confident in their stance.[11]

Finally, we use a similar survey from the Centre for Macroeconomics (CFM) to show that our results hold in other samples. The CFM runs a survey similar to the IGM, asking economists based in Europe questions on macroeconomics and public policy questions. We do not focus on the CFM

[11]For brevity, we report the results for the explicit confidence-level measure but the results are similar for the propensity to provide extreme judgements. We also estimate the equations controlling for tenure. Tenure is negatively correlated with confidence, however, because only 4 people in our sample did not receive tenure and 3 of them are women, we refrain from drawing any conclusions from this result.

results as the sample consists of only 5 women (out of 58 participants).[12] However, the results are similar in magnitude in both samples. The CFM survey results, reported in Table 4, show that women are less likely to hold an extreme view on a topic and less confident in their views.

4 Conclusion

Several papers have found that women are less confident than men. We test whether the confidence gap persists for women who have reached the top of their careers. While we do find that a confidence gap persists, it is primarily driven by women being less confident when asked about topics they are not an expert on. It is therefore difficult to state that women being less confident is always negative. It could be that women have an optimal level of confidence and adjust it depending on whether they are an expert whereas men are consistently overconfident. We also look at measures of confidence and find that the confidence gap has two components. Women hold less extreme views and are also less confident in their views.

Our paper helps explain why women can still be less confident than men even after breaking through the glass ceiling. Further research would help to understand why this gap exists and the implications that holding less extreme views and being less confident in their views has for women. For example, are women penalized for holding extreme views and does this contribute to the low number of women reaching upper-level positions? We leave such questions for future research.

[12]Non-response is nearly absent (1.2%), preventing us from examining whether it varies by gender. We are also unable to replicate the foreign field findings as nearly all economists are macroeconomists.

Bibliography

Bayer, Amanda and Cecilia Elena Rouse (2016). "Diversity in the Economics Profession: A New Attack on an Old Problem", Journal of Economic Perspectives, 30(4), pp.221-242.

Bertrand, Marianne (2011). "New Perspectives on Gender" Chapter 17 in Ashenfelter, Orley and Card, David "Handbook of Labor Economics", Elsevier, edition 1, vol. 4 (5)

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

Gordon, Roger and Gordon B. Dahl (2013). "Views among Economists: Professional Consensus or Point-Counterpoint?", American Economic Review: Papers & Proceedings, 103 (3), pp. 629-635.

Mondak, Jeffery and Mary Anderson (2004). "The Knowledge Gap: A Reexamination of Gender-Based Differences in Political Knowledge", Journal of Politics, 66 (2), pp. 492-512.

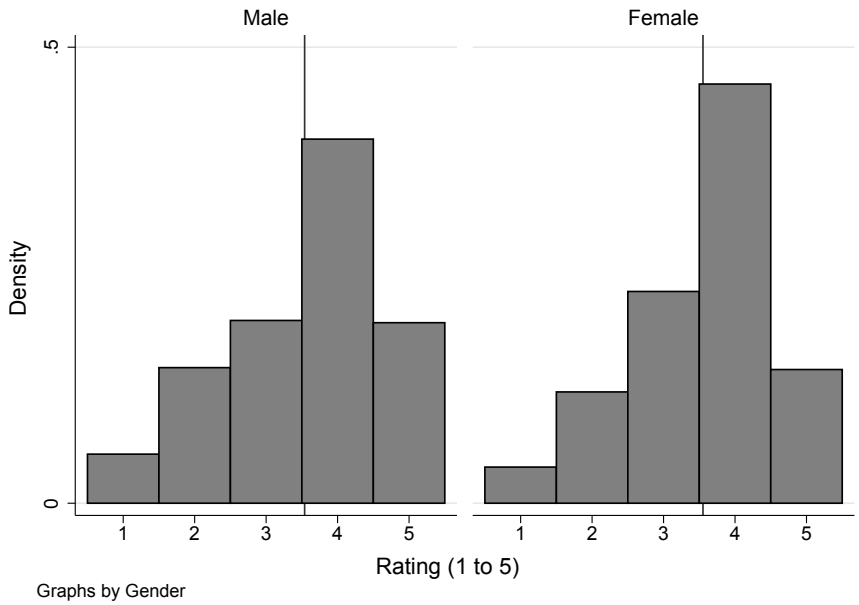
Moore, Don and Paul Healy (2008). "The Trouble with Overconfidence", Psychological Review, 115 (2), pp. 502-17.

Samek, Anya (2015). "A University-Wide Field Experiment on Gender Differences in Job Entry Decisions", mimeo

Sarsons, Heather (2015). "Gender Differences in Recognition for Group Work", mimeo.

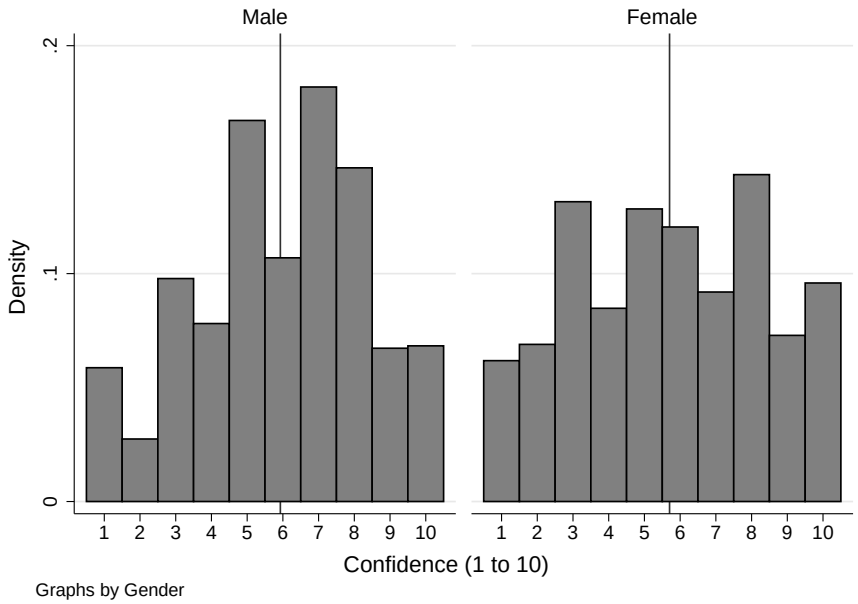
Tables and Figures

Figure 1: Distribution of responses (Likert scale)



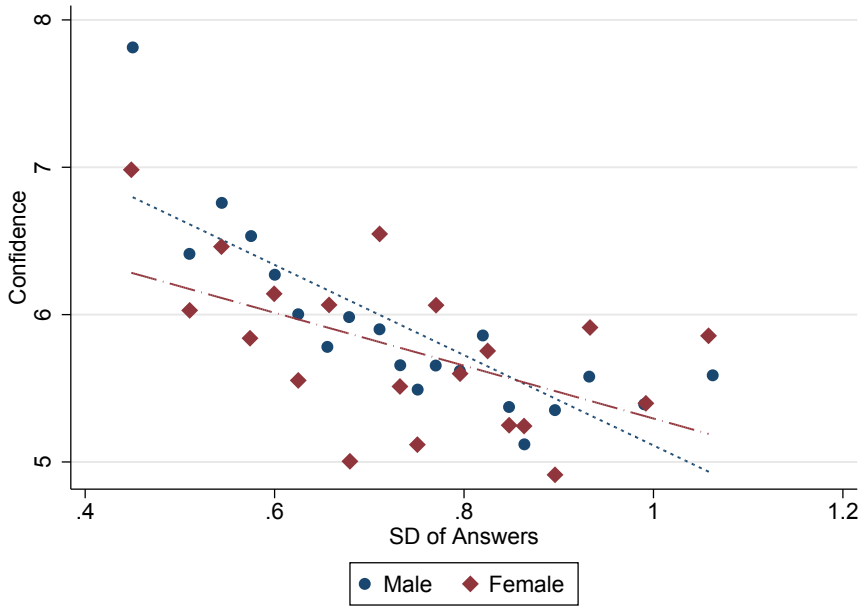
N=6986. Vertical line marks the mean

Figure 2: Distribution of subjective scores for effectiveness



N=6986. Vertical line marks the mean

Figure 3: Confidence and disagreement



Self-reported confidence (1-10) and level of disagreement (as measured by the standard deviation in ratings) for a question, broken down by gender.

Table 1: Comparison of background characteristics and confidence by gender

	(1)	(2)	(3)	(4)	(5)	(6)
	Male		Female		p-value	
	Mean	SD	Mean	SD	Diff mean	KS-test
Panel A						
Individual characteristics						
Year PhD award	1985.6	10.222	1990.6	2.887	0.174	0.135
Harvard/MIT PhD	0.567	0.502	0.600	0.516	0.857	1.000
Field: IO	0.027	0.164	0.100	0.316	0.320	1.000
Field: Labour	0.108	0.314	0.300	0.483	0.136	0.882
Field: Dev/Hist/Pol.	0.162	0.373	0	0	0.180	0.969
Field: Behavioural/Exp.	0.027	0.164	0	0	0.608	1.000
Field: Public/Health/Env	0.189	0.397	0.400	0.516	0.169	0.797
Field: Finance	0.081	0.276	0	0	0.362	1.000
Field: International	0	0	0.100	0.316	0.053	1.000
Field: Macro	0.216	0.417	0.100	0.316	0.418	1.000
Field: Microtheory	0.162	0.373	0	0	0.180	0.969
Field: Econometrics	0.027	0.164	0	0	0.608	1.000
Solo pubs. until tenure	5.540	3.870	4.600	2.319	0.469	0.448
Co-authored pubs. until tenure	6.405	4.312	6.100	1.462	0.840	0.958
Observations	37		10			
Panel B						
Voting behaviour						
Any judgement given (dummy)	0.846	0.360	0.767	0.422	0.000	0.000
Observations	6763		1645			
Judgement (Likert 1-5)	3.538	1.124	3.551	1.011	0.711	0.009
Extreme judgement (1 or 5)	0.251	0.434	0.186	0.389	0.000	0.000
Confidence (1-10)	5.929	2.393	5.703	2.659	0.003	0.000
Observations	5724		1262			

Column (5) shows the p-value of the simple t-test for equality of means between male and female. Column (6) shows the p-value for the Kolmogorov-Smirnov test for equality of distributions. *** p<0.01, ** p<0.05, * p<0.1

Table 2: Confidence and gender

	(1)	(2)	(3)	(4)	(5)	(6)
	Any judgement		Extreme judgement		Confidence (1 to 10)	
Mean of dep. var.	0.830	0.827	0.239	0.245	5.888	5.878
Female	-0.079***	-0.047***	-0.066***	-0.070***	-0.226***	-0.246***
	(0.011)	(0.016)	(0.012)	(0.014)	(0.077)	(0.081)
Year PhD award		-0.001***		-0.002***		-0.012***
		(0.000)		(0.001)		(0.003)
PhD Harvard/MIT		-0.039***		-0.006		0.468***
		(0.009)		(0.012)		(0.071)
Total solo pubs.		0.000		-0.000		0.000
		(0.001)		(0.002)		(0.010)
Total co-authored pubs.		0.000		-0.005***		-0.048***
		(0.001)		(0.001)		(0.009)
Question FE		X		X		X
School FE		X		X		X
Observations	8408	7871	6986	6511	6986	6511

The unit of observation is the economist-question pair. The dependent variable is a dummy that is 1 if the respondent provided a judgement (Columns 1-2), replied either strongly disagree or strongly agree to a question (Columns 3-4) and the self-reported measure of confidence vis-a-vis a given question's reply measured on a scale of 1 (unconfident) to 10 (confident) (Columns 5-6). *Female* is a dummy that is 1 if the respondent is female. *Year of PhD award* is the year the respondent was awarded the PhD. *PhD Harvard/MIT* is a dummy that is 1 if the respondent was awarded a PhD from either Harvard or MIT. *Total solo pubs.* is the number of total single authored publication up to tenure. *Total co-authored pubs.* is the total number of co-authored publications up to tenure. Robust standard errors in parentheses, clustered at the question level. *** p<0.01, ** p<0.05, * p<0.1

Table 3: Differential confidence in answering foreign field questions

	(1)	(2)	(3)
	Any judgement	Extreme answer	Confidence (1 to 10)
Mean of dep. var.	0.827	0.245	5.878
Female	-0.047*	-0.021	-0.024
	(0.025)	(0.033)	(0.169)
Foreign field question	-0.076***	-0.046**	-0.816***
	(0.012)	(0.018)	(0.101)
Female × Foreign field	-0.000	-0.060*	-0.282
	(0.024)	(0.035)	(0.189)
Female + Female × Foreign field	-0.047***	-0.081***	-0.305***
	(0.012)	(0.014)	(0.088)
Controls	X	X	X
Question FE	X	X	X
School FE	X	X	X
Observations	7871	6511	6511

The unit of observation is the economist-question pair. The dependent variable is a dummy that is 1 if the respondent provided a judgement (Columns 1), replied either strongly disagree or strongly agree to a question (Columns 2) and the self-reported measure of confidence vis-a-vis a given question's reply (Columns 3). *Female* is a dummy that is 1 if the respondent is female. *Foreign field* is a dummy that is 1 if the question does not cover the respondent's primary field of research. Controls are: *Year of PhD award* is the year the respondent was awarded the PhD. *PhD Harvard/MIT* is a dummy that is 1 if the respondent was awarded a PhD from either Harvard or MIT. *Total solo pubs.* is the number of total single authored publication up to tenure. *Total co-authored pubs.* is the total number of co-authored publications up to tenure. Robust standard errors in parentheses, clustered at the question level. *** p<0.01, ** p<0.05, * p<0.1

Table 4: Center for Macroeconomics (Robustness)

	(1)	(2)	(3)	(4)
	Extreme judgement		Confidence (1 to 5)	
Mean of dep. var.	0.215	0.215	0.215	3.199
Female	-0.070*	-0.076*	-0.188**	-0.169*
	(0.036)	(0.040)	(0.073)	(0.087)
Year PhD award		-0.002**		-0.008***
		(0.001)		(0.003)
PhD Harvard/MIT		0.039		0.142
		(0.028)		(0.088)
Total pubs.		0.002***		0.000
		(0.000)		(0.002)
Question FE		X		X
School FE		X		X
Observations	2,498	1,727	2,498	1,726

The unit of observation is the economist-question pair. The dependent variable is a dummy that is 1 if the respondent replied either strongly disagree or strongly agree to a question (Columns 1-2) and the self-reported measure of confidence vis-a-vis a given question's reply measured on a scale of 1 (unconfident) to 5 (confident) (Columns 3-4). *Female* is a dummy that is 1 if the respondent is female. *Year of PhD award* is the year the respondent was awarded the PhD. *PhD Harvard/MIT* is a dummy that is 1 if the respondent was awarded a PhD from either Harvard or MIT. *Total pubs.* is the number of total publications. Robust standard errors in parentheses, clustered at the question level. *** p<0.01, ** p<0.05, * p<0.1

Appendix (for online publication)

Table 5: Confidence and disagreement by gender (Online Appendix)

	(1)	(2)	(3)
	Confidence (1 to 10)		
Mean of dep. var.	5.920	5.920	5.920
Female	0.054	-0.848***	-0.964***
	(0.165)	(0.321)	(0.309)
Year PhD award	-0.012***	-0.012***	-0.013***
	(0.003)	(0.003)	(0.003)
PhD Harvard/MIT	0.399***	0.398***	0.402***
	(0.069)	(0.069)	(0.069)
Total solo pubs	0.002	0.002	0.003
	(0.009)	(0.009)	(0.009)
Total co-authored pubs	-0.048***	-0.048***	-0.047***
	(0.008)	(0.008)	(0.008)
Foreign field question	-0.731***	-0.731***	-0.817***
	(0.106)	(0.106)	(0.101)
Female × Foreign field	-0.340*	-0.334*	-0.276
	(0.185)	(0.184)	(0.187)
SD rating	-2.810***	-3.043***	
	(0.425)	(0.441)	
Female × SD rating		1.211***	1.264***
		(0.385)	(0.367)
Question FE			X
School FE	X	X	X
Observations	5,241	5,241	5,241

The unit of observation is the economist-question pair. The dependent variable is the self-reported measure of confidence vis-a-vis a given question's reply (1 lowest, 10 highest on integer scale). *Female* is a dummy that is 1 if the respondent is female. *Year of PhD award* is the year the respondent was awarded the PhD. *PhD Harvard/MIT* is a dummy that is 1 if the respondent was awarded a PhD from either Harvard or MIT. *Total solo pubs.* is the number of total single authored publication. *Total co-authored pubs.* is the total number of co-authored publications. *SD rating* measures the standard deviation of ratings submitted, capturing the level of disagreement in a given question. Robust standard errors in parentheses, clustered at the question level. *** p<0.01, ** p<0.05, * p<0.1

Table 6: Differential confidence in answering foreign field questions and breadth (Online Appendix)

	(1)	(2)	(3)
	Any judgement	Extreme answer	Confidence (1 to 10)
Mean of dep. var.	0.830	0.251	5.929
Female	-0.068**	-0.009	0.256
	(0.027)	(0.034)	(0.185)
Foreign field question	-0.085***	-0.046**	-0.788***
	(0.012)	(0.019)	(0.103)
Female × Foreign field	0.029	-0.074**	-0.440**
	(0.026)	(0.037)	(0.208)
Breadth	0.005***	0.004**	-0.018*
	(0.002)	(0.002)	(0.010)
Female ×Breadth	-0.036***	-0.024**	-0.075
	(0.010)	(0.011)	(0.064)
Female + Female × Foreign field	-0.038***	-0.082***	-0.184**
	(0.012)	(0.015)	(0.093)
Controls	X	X	X
Question FE	X	X	X
School FE	X	X	X
Observations	7871	6511	6511

The unit of observation is the economist-question pair. The dependent variable is a dummy that is 1 if the respondent provided a judgement (Columns 1), replied either strongly disagree or strongly agree to a question (Columns 2) and the self-reported measure of confidence vis-a-vis a given question's reply (Columns 3). *Female* is a dummy that is 1 if the respondent is female. *Foreign field* is a dummy that is 1 if the question does not cover the respondent's primary field of research. *Breadth* measures the number of distinct fields in which there are one or more papers citing an economist's work. Controls are: *Year of PhD award* is the year the respondent was awarded the PhD. *PhD Harvard/MIT* is a dummy that is 1 if the respondent was awarded a PhD from either Harvard or MIT. *Total solo pubs.* is the number of total single authored publication up to tenure. *Total co-authored pubs.* is the total number of co-authored publications up to tenure. Robust standard errors in parentheses, clustered at the question level. *** p<0.01, ** p<0.05, * p<0.1

Table 7: Self-reported confidence level and gender (Online Appendix)

	(1)	(2)	(3)
	Confidence (1 to 10)		
Mean of dep. var.	5.920	5.920	5.920
Female	-0.024	-0.240	0.060
	(0.169)	(0.169)	(0.144)
Year PhD award	-0.013***	-0.021***	-0.006**
	(0.003)	(0.004)	(0.003)
PhD Harvard/MIT	0.404***	0.201***	0.433***
	(0.069)	(0.089)	(0.062)
Total solo pubs	0.003	0.006	0.008
	(0.009)	(0.010)	(0.008)
Total co-authored pubs	-0.047***	-0.058***	-0.034***
	(0.008)	(0.009)	(0.008)
Foreign field question	-0.816***	-0.890***	-0.712***
	(0.101)	(0.102)	(0.082)
Female × Foreign field	-0.282	-0.382**	-0.130
	(0.189)	(0.193)	(0.164)
Female + Female × Foreign field	-0.305***	-0.622***	-0.070
	(0.088)	(0.117)	(0.077)
Question FE	X	X	X
School FE	X	X	X
Field FE		X	
Answer FE			X
Observations	6511	6511	6511

The unit of observation is the economist-question pair. The dependent variable is the self-reported measure of confidence vis-a-vis a given question's reply (1 lowest, 10 highest on integer scale). *Female* is a dummy that is 1 if the respondent is female. *Year of PhD award* is the year the respondent was awarded the PhD. *PhD Harvard/MIT* is a dummy that is 1 if the respondent was awarded a PhD from either Harvard or MIT. *Total solo pubs.* is the number of total single authored publication. *Total co-authored pubs.* is the total number of co-authored publications. Robust standard errors in parentheses, clustered at the question level. *** p<0.01, ** p<0.05, * p<0.1

Table 8: Providing extreme judgements, broken down (Online Appendix)

	(1)	(2)	(3)
	Extreme judgement	Strongly agree	Strongly disagree
Mean of dep. var.	0.245	0.193	0.052
Female	-0.070***	-0.061***	-0.008
	(0.014)	(0.012)	(0.012)
Controls	X	X	X
Question FE	X	X	X
School FE	X	X	X
Observations	6511	6511	6511

The unit of observation is the economist-question pair. The dependent variable is a dummy that is 1 if the respondent provided a judgement (Columns 1), replied either strongly disagree or strongly agree to a question (Columns 2) and the self-reported measure of confidence vis-a-vis a given question's reply measured on a scale of 1 (unconfident) to 10 (confident) (Columns 3). *Female* is a dummy that is 1 if the respondent is female. *Year of PhD award* is the year the respondent was awarded the PhD. *PhD Harvard/MIT* is a dummy that is 1 if the respondent was awarded a PhD from either Harvard or MIT. *Total solo pubs.* is the number of total single authored publication up to tenure. *Total co-authored pubs.* is the total number of co-authored publications up to tenure. Robust standard errors in parentheses, clustered at the question level. *** p<0.01, ** p<0.05, * p<0.1