

Brailey, Thomas; Kelly, Edmund; Odermatt, Angela; Ward, Albert

Working Paper

Reproduction and Robustness of Kao et al. (2024): "Female Representation and Legitimacy". A Report from the 2024 UC Berkeley Replication Games

I4R Discussion Paper Series, No. 146

Provided in Cooperation with:

The Institute for Replication (I4R)

Suggested Citation: Brailey, Thomas; Kelly, Edmund; Odermatt, Angela; Ward, Albert (2024) : Reproduction and Robustness of Kao et al. (2024): "Female Representation and Legitimacy". A Report from the 2024 UC Berkeley Replication Games, I4R Discussion Paper Series, No. 146, Institute for Replication (I4R), s.l.

This Version is available at:

<https://hdl.handle.net/10419/302052>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



No. 146

I4R DISCUSSION PAPER SERIES

Reproduction and Robustness of Kao et al. (2024): “Female Representation and Legitimacy”. A Report from the 2024 UC Berkeley Replication Games

Thomas Brailey

Edmund Kelly

Angela Odermatt

Albert Ward

August 2024

I4R DISCUSSION PAPER SERIES

I4R DP No. 146

Reproduction and Robustness of Kao et al. (2024): “Female Representation and Legitimacy”. A Report from the 2024 UC Berkeley Replication Games

Thomas Brailey¹, Edmund Kelly¹, Angela Odermatt², Albert Ward¹

¹University of Oxford/Great Britain

²Princeton University, Princeton/USA

AUGUST 2024

Any opinions in this paper are those of the author(s) and not those of the Institute for Replication (I4R). Research published in this series may include views on policy, but I4R takes no institutional policy positions.

I4R Discussion Papers are research papers of the Institute for Replication which are widely circulated to promote replications and meta-scientific work in the social sciences. Provided in cooperation with EconStor, a service of the [ZBW – Leibniz Information Centre for Economics](#), and [RWI – Leibniz Institute for Economic Research](#), I4R Discussion Papers are among others listed in RePEc (see IDEAS, EconPapers). Complete list of all I4R DPs - downloadable for free at the I4R website.

I4R Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

Editors

Abel Brodeur
University of Ottawa

Anna Dreber
Stockholm School of Economics

Jörg Ankel-Peters
RWI – Leibniz Institute for Economic Research

Reproduction and robustness of Kao *et al.* (2024): "Female representation and legitimacy". A report from the 2024 UC Berkeley Replication Games

Thomas Brailey[†] Edmund Kelly[‡]
Angela Odermatt[§] Albert Ward[¶]

August 8, 2024

Abstract

Kao *et al.* (2024) use phone-based survey experiments in Jordan, Tunisia and Morocco to test whether established theories about the effect of descriptive representation on perceived democratic legitimacy hold in the Middle East. They find that the presence of women in deliberative bodies legitimizes decision-making even in more socially conservative, less democratic societies. We blindly reproduced their study, and then extend their analysis with five additional robustness checks. We find that their analysis is reproducible and robust in several ways, although there were ambiguities in the original text which prolonged this process. Finally, we also extended their analysis by using iterative machine learning models to study heterogeneous treatment effects. We find that marital status as well as pre-treatment attitudes on related issues affect the response to the treatment.

[†]Department of Politics and International Relations, University of Oxford. Email: thomas.brailey@politics.ox.ac.uk

[‡]DPIR, University of Oxford. Email: edmund.kelly@politics.ox.ac.uk

[§]Princeton Politics. Email: angela.odermatt@princeton.edu

[¶]DPIR, University of Oxford. Email: albert.ward@politics.ox.ac.uk

1 Introduction

Recent findings in developed democracies have linked women's descriptive representation to perceptions of the legitimacy of democratic institutions, and of their outputs (Banducci, Donovan, and Karp 2004; Mansbridge 1999; Scherer and Curry 2010). Kao *et al.* (2024) extend this analysis to different polities by examining the effect of the gender composition of legislative bodies on citizens' evaluations of their decision-making processes in three Middle East and North African (MENA) countries: Morocco, Jordan and Tunisia. They employ vignette computer-assisted telephone interview (CATI) survey experiments to test whether citizens respond to treatments regarding both women's representation and policy outcomes favoring women. They find that, contrary to their expectations, citizens in these three countries display the same preferences as citizens in developed democracies. Women's presence on legislative bodies promoted perceptions of the legitimacy of these bodies and their outcomes, and pro-women outcomes were associated with increased legitimacy.

We found that the study was reproducible and robust to several additional tests. First, we add controls for income and marital status, which potentially confound sexist attitudes. Second, we investigate potential heterogeneity based on whether respondents found the treatment to be realistic. Third, we add some additional tests for social desirability bias. In each case, we find that the original results are robust. In the last part of our replication, we extend their original analysis by analyzing heterogeneous treatment effects using iterative machine learning models. We find that marital status and pre-treatment attitudes on related issues facet the response to the treatment. However, we encountered some difficulties during our initial blind replication of the analysis due to a few small ambiguities in the original replication package. In particular, the codebook omitted information on the coding or reference categories for some variables, and the precise way in which variables had been standardized was not immediately apparent. Nonetheless, this did not in general affect our ability to reproduce the original results.

This report proceeds as follows. First, we report the results of our blind replication. We then conduct several additional robustness checks: adding controls for income and marital status, and dropping controls altogether; subsetting by participant's perception

of the legitimacy of the treatment; and further testing concerns regarding social desirability bias by removing non-responders to questions about sexism. Finally, we discuss heterogeneous treatment effects based on an estimation strategy using Bayesian Additive Regression Trees (BART). We conclude by making some suggestions for how authors of future work can maximize the clarity of their replication materials based on our experience with this paper.

This report was completed for the UC Berkeley Replication Games on March 7, 2024, organized by the Institute for Replication ¹. Our replication package can be found at the following [link](#)².

2 Blind replication

We first attempted a blind replication in R using the text of the paper and supplementary appendices. By a blind replication, we mean attempting to generate the original results without looking at the authors' code, relying only on what is written in the manuscript or supplementary documents. This follows generally accepted standards for replication: the *replication standard* holds that "sufficient information exists with which to understand, evaluate, and building upon a prior work if a third party could replicate the results without any additional information from the author" (King 1995, p. 444).

While we were able to replicate Figures 2, 3, and 4 in full with no discrepancies from the original results, there were some small ambiguities in the text and the replication package which slightly prolonged this process. In particular, the precise way in which the dependent variables had been standardized could have been made more apparent

¹ For more information about the Institute for Replication, see: Brodeur, Mikola, and Cook (2024).

² https://osf.io/7nqrm/?view_only=c14f6dfb2f42412887fc923d0b536490.

in the text or the supplementary appendices. While the standardization is implied when the original authors discuss their results in terms of standard deviation changes in the dependent variable, this could have been made slightly clearer through direct mention in the text, labelling the axes of the figures to reflect the standardization of the dependent variable, or mentioning this in the notes below the relevant tables and figures. It could also have been clarified in the text whether the dependent variable was standardized within countries (which is the case for the individual country-level analysis), or standardized across the pooled sample. It took some trial and error to find the correct specifications: however these ambiguities were minor and did not affect our ability to reproduce the main results.

We were unable to replicate Figure 5 (Effect of Gender Balance Condition on Committee Decision, p.500) from the original manuscript. Running the same analysis we used to create the previous exhibits and interacting the two treatment variables did not produce exactly the same results seen in the manuscript. Following some tweaks to the specification, we decided to then look at the code used to produce the results. One issue was that the pooled model in Panel 1 (far left) does not include HC0 clustered standard errors. However, amending this did not completely fix the issue. While directionally the results all appear to be similar to the results in the manuscript, and the coefficients are of a similar size, we were unable to resolve this issue with the existing code and data within the timeframe of the replication games.

In addition, although this did not prevent our replication, it would have been useful if the codebook in the replication package had included slightly more information. For a few variables, the codebook and supplementary appendices did not explain the different levels of each variable. In particular, it would have been helpful if the variable 'd_issue_dv', which describes the issue treatment (domestic violence or littering), had been labelled with its different levels. While we were able to work out the correct coding based on the frequency table, we suggest that authors of future work make sure to include this information to streamline replication efforts.

One final small ambiguity was that in the original supplementary appendices in Table A4, which reports the models for Figure 2, the coding of the education variable appears different to that in the dataset. The original authors do not report the ref-

erence category they use for the education variable in the supplementary appendices, and in Table A4 the levels do not match those in the replication data. In particular, there are two missing levels 'High school' and 'Less than high school' which are present in the dataset. This did not, however, affect the results.

Regarding the data provided in the authors' replication package: all data provided had been pre-processed (cleaned, subsetting, recoded) which, while convenient for our replication of results, misses an important step in the replication process. Ideally, the raw survey data could also be included in the replication package to allow replication of the pre-processing steps. With the survey data in particular, meta-variables (such as the time taken for the respondent to complete the survey) may be useful for studying heterogeneous treatment effects and testing the overall robustness of original findings. Again, this is only a small issue, and indeed the provision of the raw survey data may be infeasible due to ethical considerations, but we recommend that authors include this information in their replication packages where possible.

We do not wish to overstate these difficulties. None of these ambiguities prevented us from replicating the original results. Even before resolving them, the differences with the original results were very slight. Nonetheless, we believe there are a few areas (particularly the codebook) in which some small amendments could be made to benefit future replicators without requiring any significant additional investment of time. We encourage authors of replication packages to make sure that their codebooks are complete and that standardized variables are clearly labelled in tables and figures. We report our finalized replication figures below in Figures 1 through 4. We report the accompanying tables in Appendix A2, alongside the original results in Appendix A1.

Figure 1: Replication of Figure 2: ATEs on Agreement that Committee made the Right Decision

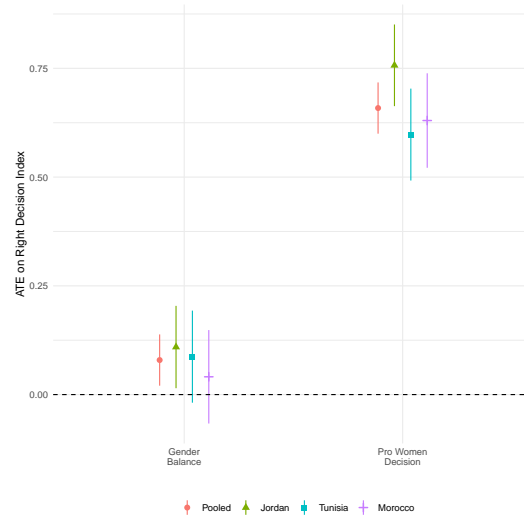


Figure 2: Replication of Figure 3: ATEs on Attitudes toward the Committee

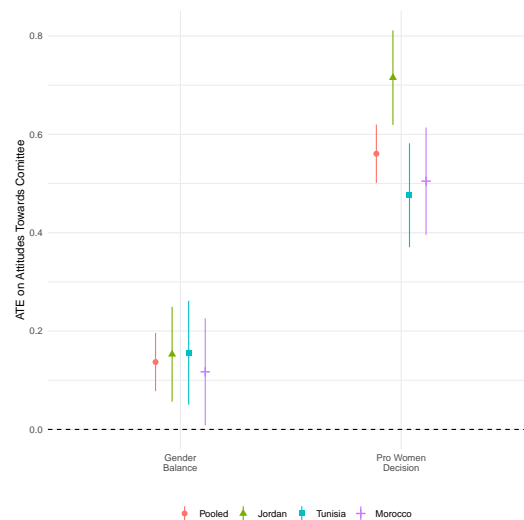


Figure 3: Replication of Figure 4: ATEs on Belief that Public Will Accept the Committee's Decision

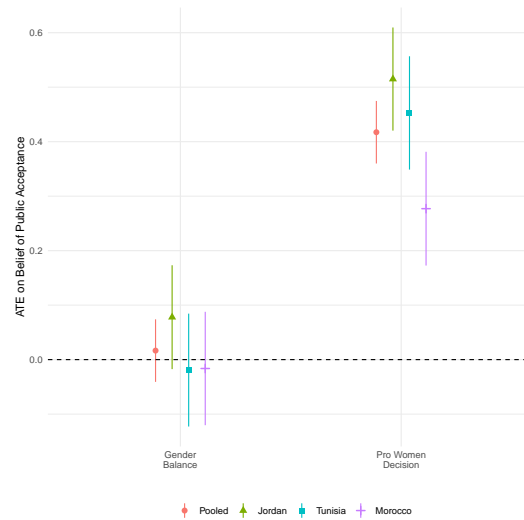
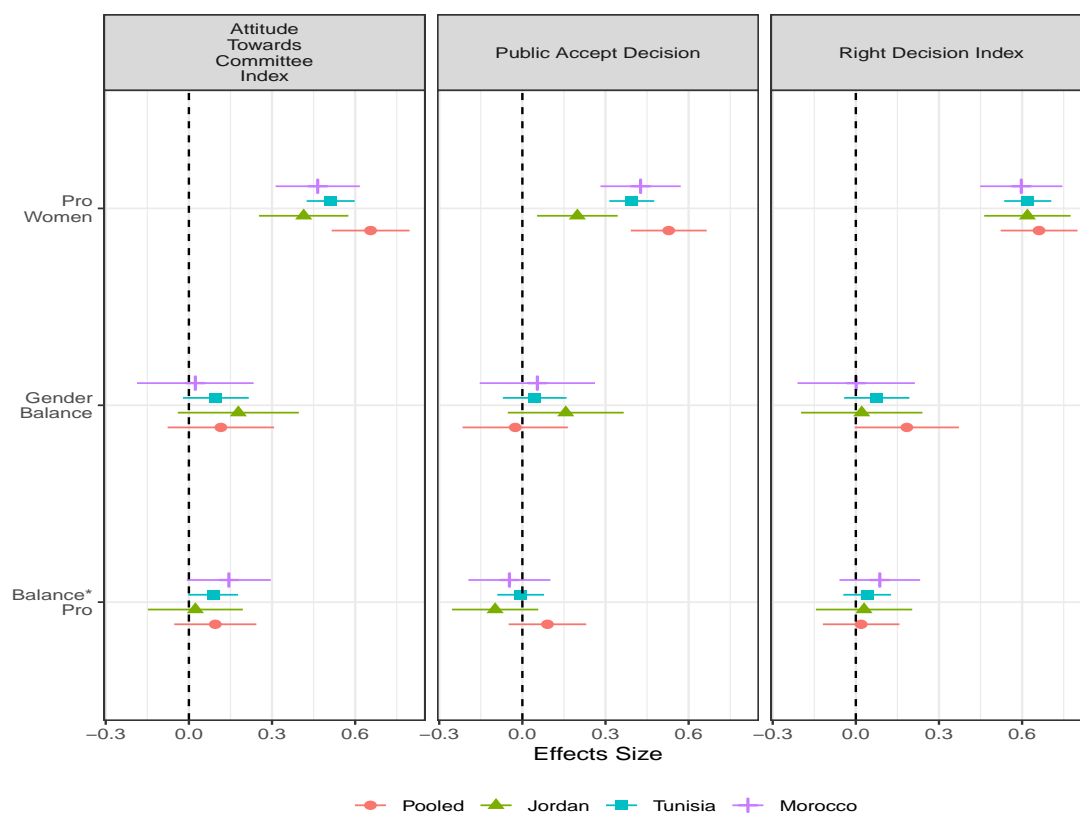


Figure 4: Replication of Figure 5: Effect of Gender Balance Conditional on Committee Decision



3 Additional robustness checks

Kao *et al.* conduct several robustness tests which they report in the supplementary appendices. In this section, we report the results of four additional robustness tests. In general, we find that the results are very similar across these different tests, and

combined with the original tests conducted by Kao *et al.*, we are confident that the results are robust in several ways.

3.1 Models with additional control variables

We additionally control for two variables which may confound sexist attitudes towards women: personal income and marital status. While much of the relationship between income and social liberalism are likely due to education, we might still expect the addition of income to capture this relationship more fully and, absent a variable for class in the data, capture some of that effect too. Marriage may confound sexist attitudes for two reasons. Firstly, because conservatives, typically measured by partisanship, are more likely to get married (Fangmeier *et al.* 2020; Wilcox 2015), and secondly because marriage may reinforce traditional gender roles, particularly in more patriarchal, socially conservative societies in the MENA. We find that the inclusion of both of these controls does not, however, substantively change any of the conclusions from the models. We report these results in Appendix A3.

3.2 Models without control variables

Controls are theoretically unnecessary for causal identification in randomized survey experiments, but they may improve the precision of the estimator and help in the case of unbalanced treatment and control groups. Kao *et al.* perform successful balance tests, but elect to include controls for age, gender and education in their final models. We replicated their models without these additional controls. Again, this did not substantively change the results; the coefficient estimates change only very slightly. We report these results in Appendix A4.

3.3 Perceived realism

One potential concern with an audio treatment is that respondents might not find the treatment realistic. Kao *et al.* account for this by asking respondents: 'Could you

imagine a real legislative committee in (country name) considering raising penalties for (domestic violence/littering)? (Yes, No)'. In each country, a majority of respondents reported perceiving the treatment to be realistic, but a non-negligible minority reported either being unsure or believed the experiment to be unrealistic.

As a robustness check, we removed those who reported being unsure and those who believed that the experiment was not realistic, to test whether the results differed substantially between these two groups. We report the results in Appendix A5. The results are similar but, in general, the effect sizes of the two main treatments are slightly stronger. This may suggest that the true effect of gender balance is slightly stronger than those estimated by Kao et al., although our evidence does not suggest that this discrepancy would be large.

3.4 Social desirability bias

As Kao et al. note in their conclusion, one substantial concern, particularly with phone-based surveys, is that respondents may feel unable to share their true opinions. They point out that many people were willing to express openly sexist views in direct questions, which limits the likelihood of social desirability bias being a significant issue in this case.

As an additional test for this, we used the variables recording whether respondents refused to answer questions on the hostile sexism and benevolent sexism scale. We suggest that those who refused to answer more of these questions may have felt more subject to social pressure against revealing sexist attitudes. We therefore constructed a new variable which is the sum of the number of questions on the hostile and benevolent sexism scales which each respondent refused to answer. We then included then included this variables as an additional control. We report these results in Appendix A6. The results are very similar to those in the paper.

4 Heterogeneous treatment effects

Lastly, we go beyond the analysis by Kao et al. by delving deeper into heterogeneous treatment effects. Kao et al. analyze theoretically motivated heterogeneous treatment effects by simply interacting hypothesized variables faceting treatment effects with the treatment in separate models. Doing so, they only find limited support for heterogeneous responses to the treatment but call for a more sophisticated analysis of these effects. We aim to provide this more extensive analysis by using Bayesian Additive Regression Trees (BART) to test for more variables causally affecting responses to the treatment. To do so, we build upon recent studies advocating for the use of machine learning in the analysis of experimental data (e.g., Green and Kern 2012; Künzel *et al.* 2019). Our approach is adapted from the study and supplementary materials by Duch *et al.* 2020. This approach has the advantages of not having to specify a functional form for the heterogeneity of treatment effects and can deal well with small sizes of subgroups.

The analysis shows that some variables indeed affect treatment outcomes. The following analyses were run for the gender balance treatment and the right decision index as the outcome variable. Figures 5 below, and 6 and 7 in the appendix show the results from the analyses for heterogeneous treatment effects using BART. The top panel of Figure 5 displays the Conditional Average Treatment Effects (CATEs) ordered by their size. The blue line denotes the mean CATE and the green line a CATE of 0. We see that there is considerable variation in the CATEs across the iterations run by the BART algorithm. 87.3% of CATEs are above 0 and 54.8% larger than the mean CATE of 0.077 (SD = 0.063). Together, this supports the effects found by Kao et al.. The lower two panels show the distribution of covariate levels across the sorted CATEs. For instance, we see that marital status affects treatment response to a large extent with married respondents having a larger, positive CATE than single respondents. Not surprisingly, pre-treatment attitudes towards increasing penalties for domestic violence also facet treatment response with people in favor of increasing the penalty reacting more positively to the treatment. These two covariates show the clearest patterns of heterogeneous treatment effects. The two figures in the appendix show the other covariate profiles used in estimating the BART models. Apart from income, age, and gender, which exhibit slight patterns, none of the other covariates exhibit signs of



Figure 5: CATEs for Marital Status and Attitudes Towards Gender Violence

heterogeneous treatment effects. Hence, we conclude that these analyses support the main findings by Kao et al. with the addition of looking at the effects of marital status as well as pre-treatment attitudes on related issues.

5 Conclusion

In this report from the 2024 UC Berkeley Replication Games, we blindly replicated Kao *et al.* (2024) and conducted several additional robustness checks. While we have highlighted a few areas in which Kao et al. could improve the clarity of their replication materials, we were able to successfully blindly replicate their study and demonstrate its robustness in a number of ways.

We recommend that authors of replication packages include the full coding of each variable in their codebooks, and that decisions to standardize variables are clearly communicated in all relevant tables, figures, and accompanying text. Nonetheless, we stress that these issues were minor and did not ultimately affect our ability to reproduce the original findings. We do not contest the substantive conclusions of the study (indeed, some of our robustness checks suggest that the true effect may be slightly stronger than that originally reported), and we found that the results are robust in a number of ways.

References

- Banducci, Susan A., Todd Donovan, and Jeffrey A. Karp (2004). "Minority Representation, Empowerment, and Participation". *The Journal of Politics* 66.2, pp. 534–556. DOI: 10.1111/j.1468-2508.2004.00163.x
- Brodeur, Abel, Derek Mikola, and Nikolai Cook (2024). "Mass Reproducibility and Replicability: A New Hope". *I4R Discussion Papers* 107
- Duch, Raymond *et al.* (2020). "Multi-modes for detecting experimental measurement error". *Political Analysis* 28.2, pp. 263–283
- Fangmeier, Troy L. *et al.* (2020). "Political Party Identification and Romantic Relationship Quality". *Couple & family psychology* 9.3, pp. 167–180. DOI: 10.1037/cfp0000141
- Green, Donald P and Holger L Kern (2012). "Modeling heterogeneous treatment effects in survey experiments with Bayesian additive regression trees". *Public opinion quarterly* 76.3, pp. 491–511
- Kao, Kristen *et al.* (2024). "Female Representation and Legitimacy: Evidence from a Harmonized Experiment in Jordan, Morocco, and Tunisia". *American Political Science Review* 118.1, pp. 495–503. DOI: 10.1017/S0003055423000357
- King, Gary (1995). "Replication, Replication". *PS: Political Science and Politics* 28.3, pp. 444–452
- Künzel, Sören R *et al.* (2019). "Metalearners for estimating heterogeneous treatment effects using machine learning". *Proceedings of the national academy of sciences* 116.10, pp. 4156–4165
- Mansbridge, Jane (1999). "Should Blacks Represent Blacks and Women Represent Women? A Contingent "Yes"". *The Journal of Politics* 61.3, pp. 628–657. DOI: 10.2307/2647821
- Scherer, Nancy and Brett Curry (2010). "Does Descriptive Race Representation Enhance Institutional Legitimacy? The Case of the U.S. Courts". *The Journal of Politics* 72.1, pp. 90–104. DOI: 10.1017/s0022381609990491
- Wilcox, Brad (2015). *More than Money: The Liberal-Conservative Divide in Marriage*. Institute for Family Studies. URL: <https://ifstudies.org/blog/more-than-money-the-liberal-conservative-divide-in-marriage> (visited on 03/07/2024)

Appendix A: Tables.

Appendix A1: Original results

Table A1: Original results for Figure 2

	term	estimate	std.error	p.value	model
1	Gender balance	0.08	0.03	0.01	Pooled
2	Pro-women decision	0.66	0.03	0.00	Pooled
3	Gender balance	0.11	0.05	0.02	Jordan
4	Pro-women decision	0.76	0.05	0.00	Jordan
5	Gender balance	0.04	0.05	0.46	Morocco
6	Pro-women decision	0.63	0.06	0.00	Morocco
7	Gender balance	0.09	0.05	0.11	Tunisia
8	Pro-women decision	0.60	0.05	0.00	Tunisia

Table A2: Original results for Figure 3

	term	estimate	std.error	p.value	model
1	Gender balance	0.14	0.03	0.00	Pooled
2	Pro-women decision	0.56	0.03	0.00	Pooled
3	Gender balance	0.15	0.05	0.00	Jordan
4	Pro-women decision	0.72	0.05	0.00	Jordan
5	Gender balance	0.12	0.06	0.03	Morocco
6	Pro-women decision	0.50	0.06	0.00	Morocco
7	Gender balance	0.16	0.05	0.00	Tunisia
8	Pro-women decision	0.48	0.05	0.00	Tunisia

Table A3: Original results for Figure 4

	term	estimate	std.error	p.value	model
1	Gender balance	0.02	0.03	0.57	Pooled
2	Pro-women decision	0.42	0.03	0.00	Pooled
3	Gender balance	0.08	0.05	0.11	Jordan
4	Pro-women decision	0.51	0.05	0.00	Jordan
5	Gender balance	-0.02	0.05	0.76	Morocco
6	Pro-women decision	0.28	0.05	0.00	Morocco
7	Gender balance	-0.02	0.05	0.72	Tunisia
8	Pro-women decision	0.45	0.05	0.00	Tunisia

Table A4: Original results for Figure 5

	term	estimate	std.error	Significance	model	outcome
1	Interaction between treatments	0.1	0.1	Not sig.	Pooled	Right decision scale
2	Decision treatment	0.6	0.0	0.001	Pooled	Right decision scale
3	Gender balance treatment	0.0	0.0	Not sig.	Pooled	Right decision scale
4	Interaction between treatments	0.2	0.1	Not sig.	Jordan	Right decision scale
5	Decision treatment	0.7	0.1	0.001	Jordan	Right decision scale
6	Gender balance treatment	0.0	0.1	Not sig.	Jordan	Right decision scale
7	Interaction between treatments	0.0	0.1	Not sig.	Tunisia	Right decision scale
8	Decision treatment	0.6	0.1	0.001	Tunisia	Right decision scale
9	Gender balance treatment	0.1	0.1	Not sig.	Tunisia	Right decision scale
10	Interaction between treatments	0.0	0.1	Not sig.	Morocco	Right decision scale
11	Decision treatment	0.6	0.1	0.001	Morocco	Right decision scale
12	Gender balance treatment	0.0	0.1	Not sig.	Morocco	Right decision scale
13	Interaction between treatments	0.1	0.1	Not sig.	Pooled	Attitudes toward committee scale
14	Decision treatment	0.5	0.0	0.001	Pooled	Attitudes toward committee scale
15	Gender balance treatment	0.1	0.0	Not sig.	Pooled	Attitudes toward committee scale
16	Interaction between treatments	0.1	0.1	Not sig.	Jordan	Attitudes toward committee scale
17	Decision treatment	0.7	0.1	0.001	Jordan	Attitudes toward committee scale
18	Gender balance treatment	0.1	0.1	Not sig.	Jordan	Attitudes toward committee scale
19	Interaction between treatments	0.0	0.1	Not sig.	Tunisia	Attitudes toward committee scale
20	Decision treatment	0.5	0.18	0.001	Tunisia	Attitudes toward committee scale
21	Gender balance treatment	0.1	0.1	Not sig.	Tunisia	Attitudes toward committee scale
22	Interaction between treatments	0.2	0.1	Not sig.	Morocco	Attitudes toward committee scale
23	Decision treatment	0.4	0.1	0.001	Morocco	Attitudes toward committee scale
24	Gender balance treatment	0.0	0.1	Not sig.	Morocco	Attitudes toward committee scale
25	Interaction between treatments	0.0	0.1	Not sig.	Pooled	Public acceptance
26	Decision treatment	0.4	0.0	0.001	Pooled	Public acceptance
27	Gender balance treatment	0.0	0.1	Not sig.	Pooled	Public acceptance
28	Interaction between treatments	-0.0	0.1	Not sig.	Jordan	Public acceptance
29	Decision treatment	0.5	0.1	0.001	Jordan	Public acceptance
30	Gender balance treatment	0.1	0.1	Not sig.	Jordan	Public acceptance
31	Interaction between treatments	0.1	0.1	Not sig.	Tunisia	Public acceptance
32	Decision treatment	0.4	0.1	0.001	Tunisia	Public acceptance
33	Gender balance treatment	-0.0	0.1	Not sig.	Tunisia	Public acceptance
34	Interaction between treatments	0.2	0.1	Not sig.	Morocco	Public acceptance
35	Decision treatment	0.2	0.1	0.01	Morocco	Public acceptance
36	Gender balance treatment	-0.1	0.1	Not sig.	Morocco	Public acceptance

Appendix A2: Tables for main replication

Table A1: Replication of Figure 2

	term	estimate	std.error	p.value	model
1	Gender balance	0.08	0.03	0.01	Pooled
2	Pro-women decision	0.66	0.03	0.00	Pooled
3	Gender balance	0.11	0.05	0.02	Jordan
4	Pro-women decision	0.76	0.05	0.00	Jordan
5	Gender balance	0.04	0.05	0.46	Morocco
6	Pro-women decision	0.63	0.06	0.00	Morocco
7	Gender balance	0.09	0.05	0.11	Tunisia
8	Pro-women decision	0.60	0.05	0.00	Tunisia

Table A2: Replication of Figure 3

	term	estimate	std.error	p.value	model
1	Gender balance	0.14	0.03	0.00	Pooled
2	Pro-women decision	0.56	0.03	0.00	Pooled
3	Gender balance	0.15	0.05	0.00	Jordan
4	Pro-women decision	0.72	0.05	0.00	Jordan
5	Gender balance	0.12	0.06	0.03	Morocco
6	Pro-women decision	0.50	0.06	0.00	Morocco
7	Gender balance	0.16	0.05	0.00	Tunisia
8	Pro-women decision	0.48	0.05	0.00	Tunisia

Table A3: Replication of Figure 4

	term	estimate	std.error	p.value	model
1	Gender balance	0.02	0.03	0.57	Pooled
2	Pro-women decision	0.42	0.03	0.00	Pooled
3	Gender balance	0.08	0.05	0.11	Jordan
4	Pro-women decision	0.51	0.05	0.00	Jordan
5	Gender balance	-0.02	0.05	0.76	Morocco
6	Pro-women decision	0.28	0.05	0.00	Morocco
7	Gender balance	-0.02	0.05	0.72	Tunisia
8	Pro-women decision	0.45	0.05	0.00	Tunisia

Table A4: Replication of Figure 5

	term	estimate	std.error	p.value	model	outcome
1	Interaction between treatments	0.04	0.04	0.35	Pooled	Right decision scale
2	Decision treatment	0.62	0.04	0.00	Pooled	Right decision scale
3	Gender balance treatment	0.08	0.06	0.21	Pooled	Right decision scale
4	Interaction between treatments	0.02	0.07	0.78	Jordan	Right decision scale
5	Decision treatment	0.66	0.07	0.00	Jordan	Right decision scale
6	Gender balance treatment	0.18	0.10	0.05	Jordan	Right decision scale
7	Interaction between treatments	0.09	0.07	0.24	Tunisia	Right decision scale
8	Decision treatment	0.60	0.08	0.00	Tunisia	Right decision scale
9	Gender balance treatment	0.00	0.11	0.99	Tunisia	Right decision scale
10	Interaction between treatments	0.03	0.09	0.74	Morocco	Right decision scale
11	Decision treatment	0.62	0.08	0.00	Morocco	Right decision scale
12	Gender balance treatment	0.02	0.11	0.85	Morocco	Right decision scale
13	Interaction between treatments	0.09	0.05	0.06	Pooled	Attitudes toward committee scale
14	Decision treatment	0.51	0.04	0.00	Pooled	Attitudes toward committee scale
15	Gender balance treatment	0.10	0.06	0.11	Pooled	Attitudes toward committee scale
16	Interaction between treatments	0.10	0.08	0.21	Jordan	Attitudes toward committee scale
17	Decision treatment	0.66	0.07	0.00	Jordan	Attitudes toward committee scale
18	Gender balance treatment	0.12	0.10	0.24	Jordan	Attitudes toward committee scale
19	Interaction between treatments	0.14	0.08	0.06	Tunisia	Attitudes toward committee scale
20	Decision treatment	0.46	0.08	0.00	Tunisia	Attitudes toward committee scale
21	Gender balance treatment	0.02	0.11	0.83	Tunisia	Attitudes toward committee scale
22	Interaction between treatments	0.02	0.09	0.79	Morocco	Attitudes toward committee scale
23	Decision treatment	0.41	0.08	0.00	Morocco	Attitudes toward committee scale
24	Gender balance treatment	0.18	0.11	0.11	Morocco	Attitudes toward committee scale
25	Interaction between treatments	-0.01	0.04	0.89	Pooled	Public acceptance
26	Decision treatment	0.39	0.04	0.00	Pooled	Public acceptance
27	Gender balance treatment	0.04	0.06	0.45	Pooled	Public acceptance
28	Interaction between treatments	0.09	0.07	0.20	Jordan	Public acceptance
29	Decision treatment	0.53	0.07	0.00	Jordan	Public acceptance
30	Gender balance treatment	-0.03	0.10	0.79	Jordan	Public acceptance
31	Interaction between treatments	-0.05	0.08	0.54	Tunisia	Public acceptance
32	Decision treatment	0.43	0.07	0.00	Tunisia	Public acceptance
33	Gender balance treatment	0.05	0.11	0.61	Tunisia	Public acceptance
34	Interaction between treatments	-0.10	0.08	0.22	Morocco	Public acceptance
35	Decision treatment	0.20	0.07	0.01	Morocco	Public acceptance
36	Gender balance treatment	0.16	0.11	0.14	Morocco	Public acceptance

Appendix A3: Tables for replication with income and marital status controls

Table A5: Replication of Figure 2

	term	estimate	std.error	p.value	model
1	Gender balance	0.08	0.03	0.01	Pooled
2	Pro-women decision	0.66	0.03	0.00	Pooled
3	Gender balance	0.11	0.05	0.02	Jordan
4	Pro-women decision	0.75	0.05	0.00	Jordan
5	Gender balance	0.04	0.05	0.43	Morocco
6	Pro-women decision	0.63	0.06	0.00	Morocco
7	Gender balance	0.08	0.05	0.12	Tunisia
8	Pro-women decision	0.60	0.05	0.00	Tunisia

Table A6: Replication of Figure 3

	term	estimate	std.error	p.value	model
1	Gender balance	0.14	0.03	0.00	Pooled
2	Pro-women decision	0.56	0.03	0.00	Pooled
3	Gender balance	0.16	0.05	0.00	Jordan
4	Pro-women decision	0.71	0.05	0.00	Jordan
5	Gender balance	0.13	0.06	0.02	Morocco
6	Pro-women decision	0.51	0.06	0.00	Morocco
7	Gender balance	0.16	0.05	0.00	Tunisia
8	Pro-women decision	0.48	0.05	0.00	Tunisia

Table A6: Replication of Figure 4

	term	estimate	std.error	p.value	model
1	Gender balance	0.02	0.03	0.60	Pooled
2	Pro-women decision	0.42	0.03	0.00	Pooled
3	Gender balance	0.08	0.05	0.10	Jordan
4	Pro-women decision	0.51	0.05	0.00	Jordan
5	Gender balance	-0.03	0.05	0.58	Morocco
6	Pro-women decision	0.28	0.05	0.00	Morocco
7	Gender balance	-0.02	0.05	0.70	Tunisia
8	Pro-women decision	0.45	0.05	0.00	Tunisia

Table A7: Replication of Figure 5

	term	estimate	std.error	p.value	model	outcome
1	Gender balance treatment	0.04	0.04	0.37	Pooled	Right decision scale
2	Decision treatment	0.62	0.04	0.00	Pooled	Right decision scale
3	Interaction between treatments	0.08	0.06	0.19	Pooled	Right decision scale
4	Gender balance treatment	0.02	0.07	0.75	Jordan	Right decision scale
5	Decision treatment	0.66	0.07	0.00	Jordan	Right decision scale
6	Interaction between treatments	0.19	0.10	0.05	Jordan	Right decision scale
7	Gender balance treatment	0.08	0.07	0.27	Tunisia	Right decision scale
8	Decision treatment	0.60	0.08	0.00	Tunisia	Right decision scale
9	Interaction between treatments	0.01	0.11	0.95	Tunisia	Right decision scale
10	Gender balance treatment	0.03	0.09	0.73	Morocco	Right decision scale
11	Decision treatment	0.62	0.08	0.00	Morocco	Right decision scale
12	Interaction between treatments	0.03	0.11	0.82	Morocco	Right decision scale
13	Gender balance treatment	0.09	0.05	0.05	Pooled	Attitudes toward committee scale
14	Decision treatment	0.51	0.04	0.00	Pooled	Attitudes toward committee scale
15	Interaction between treatments	0.10	0.06	0.12	Pooled	Attitudes toward committee scale
16	Gender balance treatment	0.09	0.08	0.23	Jordan	Attitudes toward committee scale
17	Decision treatment	0.65	0.07	0.00	Jordan	Attitudes toward committee scale
18	Interaction between treatments	0.13	0.10	0.19	Jordan	Attitudes toward committee scale
19	Gender balance treatment	0.15	0.08	0.06	Tunisia	Attitudes toward committee scale
20	Decision treatment	0.47	0.08	0.00	Tunisia	Attitudes toward committee scale
21	Interaction between treatments	0.02	0.11	0.85	Tunisia	Attitudes toward committee scale
22	Gender balance treatment	0.04	0.09	0.68	Morocco	Attitudes toward committee scale
23	Decision treatment	0.42	0.08	0.00	Morocco	Attitudes toward committee scale
24	Interaction between treatments	0.17	0.11	0.12	Morocco	Attitudes toward committee scale
25	Gender balance treatment	-0.01	0.04	0.85	Pooled	Public acceptance
26	Decision treatment	0.40	0.04	0.00	Pooled	Public acceptance
27	Interaction between treatments	0.05	0.06	0.43	Pooled	Public acceptance
28	Gender balance treatment	0.09	0.07	0.19	Jordan	Public acceptance
29	Decision treatment	0.53	0.07	0.00	Jordan	Public acceptance
30	Interaction between treatments	-0.03	0.10	0.78	Jordan	Public acceptance
31	Gender balance treatment	-0.05	0.08	0.52	Tunisia	Public acceptance
32	Decision treatment	0.43	0.07	0.00	Tunisia	Public acceptance
33	Interaction between treatments	0.06	0.11	0.59	Tunisia	Public acceptance
34	Gender balance treatment	-0.11	0.08	0.15	Morocco	Public acceptance
35	Decision treatment	0.20	0.07	0.01	Morocco	Public acceptance
36	Interaction between treatments	0.16	0.11	0.13	Morocco	Public acceptance

Appendix A4: Tables for replication without controls

Table A8: Replication of Figure 2

	term	estimate	std.error	p.value	model
1	Gender balance	0.08	0.03	0.01	Pooled
2	Pro-women decision	0.66	0.03	0.00	Pooled
3	Gender balance	0.10	0.05	0.03	Jordan
4	Pro-women decision	0.76	0.05	0.00	Jordan
5	Gender balance	0.04	0.05	0.48	Morocco
6	Pro-women decision	0.64	0.06	0.00	Morocco
7	Gender balance	0.09	0.05	0.11	Tunisia
8	Pro-women decision	0.60	0.05	0.00	Tunisia

Table A9: Replication of Figure 3

	term	estimate	std.error	p.value	model
1	Gender balance	0.14	0.03	0.00	Pooled
2	Pro-women decision	0.57	0.03	0.00	Pooled
3	Gender balance	0.16	0.05	0.00	Jordan
4	Pro-women decision	0.72	0.05	0.00	Jordan
5	Gender balance	0.13	0.06	0.03	Morocco
6	Pro-women decision	0.51	0.06	0.00	Morocco
7	Gender balance	0.15	0.06	0.01	Tunisia
8	Pro-women decision	0.46	0.06	0.00	Tunisia

Table A10: Replication of Figure 4

	term	estimate	std.error	p.value	model
1	Gender balance	0.01	0.03	0.61	Pooled
2	Pro-women decision	0.42	0.03	0.00	Pooled
3	Gender balance	0.07	0.05	0.12	Jordan
4	Pro-women decision	0.51	0.05	0.00	Jordan
5	Gender balance	-0.03	0.05	0.63	Morocco
6	Pro-women decision	0.27	0.05	0.00	Morocco
7	Gender balance	-0.02	0.05	0.70	Tunisia
8	Pro-women decision	0.45	0.05	0.00	Tunisia

Table A11: Replication of Figure 5

	term	estimate	std.error	p.value	model	outcome
1	Gender balance treatment	0.04	0.04	0.40	Pooled	Right decision scale
2	Decision treatment	0.62	0.04	0.00	Pooled	Right decision scale
3	Interaction between treatments	0.08	0.06	0.17	Pooled	Right decision scale
4	Gender balance treatment	0.02	0.07	0.74	Jordan	Right decision scale
5	Decision treatment	0.67	0.07	0.00	Jordan	Right decision scale
6	Interaction between treatments	0.17	0.10	0.09	Jordan	Right decision scale
7	Gender balance treatment	0.08	0.07	0.31	Tunisia	Right decision scale
8	Decision treatment	0.59	0.08	0.00	Tunisia	Right decision scale
9	Interaction between treatments	0.02	0.11	0.85	Tunisia	Right decision scale
10	Gender balance treatment	0.02	0.09	0.82	Morocco	Right decision scale
11	Decision treatment	0.62	0.08	0.00	Morocco	Right decision scale
12	Interaction between treatments	0.03	0.11	0.76	Morocco	Right decision scale
13	Gender balance treatment	0.08	0.05	0.09	Pooled	Attitudes toward committee scale
14	Decision treatment	0.51	0.05	0.00	Pooled	Attitudes toward committee scale
15	Interaction between treatments	0.12	0.06	0.05	Pooled	Attitudes toward committee scale
16	Gender balance treatment	0.11	0.08	0.15	Jordan	Attitudes toward committee scale
17	Decision treatment	0.67	0.07	0.00	Jordan	Attitudes toward committee scale
18	Interaction between treatments	0.10	0.10	0.30	Jordan	Attitudes toward committee scale
19	Gender balance treatment	0.11	0.08	0.17	Tunisia	Attitudes toward committee scale
20	Decision treatment	0.43	0.08	0.00	Tunisia	Attitudes toward committee scale
21	Interaction between treatments	0.07	0.11	0.50	Tunisia	Attitudes toward committee scale
22	Gender balance treatment	0.01	0.09	0.91	Morocco	Attitudes toward committee scale
23	Decision treatment	0.40	0.08	0.00	Morocco	Attitudes toward committee scale
24	Interaction between treatments	0.22	0.11	0.05	Morocco	Attitudes toward committee scale
25	Gender balance treatment	-0.01	0.04	0.81	Pooled	Public acceptance
26	Decision treatment	0.39	0.04	0.00	Pooled	Public acceptance
27	Interaction between treatments	0.05	0.06	0.40	Pooled	Public acceptance
28	Gender balance treatment	0.09	0.07	0.20	Jordan	Public acceptance
29	Decision treatment	0.53	0.07	0.00	Jordan	Public acceptance
30	Interaction between treatments	-0.03	0.10	0.73	Jordan	Public acceptance
31	Gender balance treatment	-0.04	0.08	0.57	Tunisia	Public acceptance
32	Decision treatment	0.43	0.07	0.00	Tunisia	Public acceptance
33	Interaction between treatments	0.04	0.11	0.67	Tunisia	Public acceptance
34	Gender balance treatment	-0.12	0.08	0.12	Morocco	Public acceptance
35	Decision treatment	0.18	0.08	0.02	Morocco	Public acceptance
36	Interaction between treatments	0.19	0.11	0.08	Morocco	Public acceptance

Appendix A5: Tables for replication with only those who thought the experiment was realistic

Table A12: Replication of Figure 2

	term	estimate	std.error	p.value	model
1	Gender balance	0.09	0.03	0.01	Pooled
2	Pro-women decision	0.69	0.03	0.00	Pooled
3	Gender balance	0.13	0.05	0.01	Jordan
4	Pro-women decision	0.85	0.05	0.00	Jordan
5	Gender balance	0.01	0.07	0.91	Morocco
6	Pro-women decision	0.58	0.07	0.00	Morocco
7	Gender balance	0.12	0.06	0.06	Tunisia
8	Pro-women decision	0.62	0.06	0.00	Tunisia

Table A13: Replication of Figure 3

	term	estimate	std.error	p.value	model
1	Gender balance	0.14	0.03	0.00	Pooled
2	Pro-women decision	0.56	0.03	0.00	Pooled
3	Gender balance	0.15	0.05	0.00	Jordan
4	Pro-women decision	0.72	0.05	0.00	Jordan
5	Gender balance	0.12	0.06	0.03	Morocco
6	Pro-women decision	0.50	0.06	0.00	Morocco
7	Gender balance	0.16	0.05	0.00	Tunisia
8	Pro-women decision	0.48	0.05	0.00	Tunisia

Table A14: Replication of Figure 4

	term	estimate	std.error	p.value	model
1	Gender balance	0.02	0.03	0.57	Pooled
2	Pro-women decision	0.42	0.03	0.00	Pooled
3	Gender balance	0.08	0.05	0.11	Jordan
4	Pro-women decision	0.51	0.05	0.00	Jordan
5	Gender balance	-0.02	0.05	0.76	Morocco
6	Pro-women decision	0.28	0.05	0.00	Morocco
7	Gender balance	-0.02	0.05	0.72	Tunisia
8	Pro-women decision	0.45	0.05	0.00	Tunisia

Table A15: Replication of Figure 5

	term	estimate	std.error	p.value	model	outcome
1	Interaction between treatments	-0.00	0.05	0.94	Pooled	Right decision scale
2	Decision treatment	0.60	0.05	0.00	Pooled	Right decision scale
3	Gender balance treatment	0.18	0.07	0.01	Pooled	Right decision scale
4	Interaction between treatments	0.00	0.08	0.99	Jordan	Right decision scale
5	Decision treatment	0.72	0.08	0.00	Jordan	Right decision scale
6	Gender balance treatment	0.27	0.11	0.01	Jordan	Right decision scale
7	Interaction between treatments	0.04	0.09	0.68	Tunisia	Right decision scale
8	Decision treatment	0.54	0.09	0.00	Tunisia	Right decision scale
9	Gender balance treatment	0.15	0.12	0.22	Tunisia	Right decision scale
10	Interaction between treatments	-0.06	0.11	0.61	Morocco	Right decision scale
11	Decision treatment	0.52	0.10	0.00	Morocco	Right decision scale
12	Gender balance treatment	0.12	0.14	0.40	Morocco	Right decision scale
13	Interaction between treatments	0.11	0.05	0.05	Pooled	Attitudes toward committee scale
14	Decision treatment	0.56	0.05	0.00	Pooled	Attitudes toward committee scale
15	Gender balance treatment	0.05	0.07	0.45	Pooled	Attitudes toward committee scale
16	Interaction between treatments	0.13	0.09	0.13	Jordan	Attitudes toward committee scale
17	Decision treatment	0.75	0.08	0.00	Jordan	Attitudes toward committee scale
18	Gender balance treatment	0.13	0.11	0.24	Jordan	Attitudes toward committee scale
19	Interaction between treatments	0.18	0.09	0.04	Tunisia	Attitudes toward committee scale
20	Decision treatment	0.51	0.09	0.00	Tunisia	Attitudes toward committee scale
21	Gender balance treatment	-0.12	0.12	0.34	Tunisia	Attitudes toward committee scale
22	Interaction between treatments	0.00	0.11	0.98	Morocco	Attitudes toward committee scale
23	Decision treatment	0.35	0.10	0.00	Morocco	Attitudes toward committee scale
24	Gender balance treatment	0.20	0.14	0.15	Morocco	Attitudes toward committee scale
25	Interaction between treatments	0.02	0.05	0.73	Pooled	Public acceptance
26	Decision treatment	0.46	0.05	0.00	Pooled	Public acceptance
27	Gender balance treatment	0.03	0.07	0.64	Pooled	Public acceptance
28	Interaction between treatments	0.13	0.08	0.12	Jordan	Public acceptance
29	Decision treatment	0.62	0.08	0.00	Jordan	Public acceptance
30	Gender balance treatment	-0.01	0.11	0.96	Jordan	Public acceptance
31	Interaction between treatments	-0.09	0.09	0.31	Tunisia	Public acceptance
32	Decision treatment	0.45	0.09	0.00	Tunisia	Public acceptance
33	Gender balance treatment	0.09	0.12	0.45	Tunisia	Public acceptance
34	Interaction between treatments	-0.03	0.10	0.76	Morocco	Public acceptance
35	Decision treatment	0.23	0.09	0.01	Morocco	Public acceptance
36	Gender balance treatment	0.08	0.13	0.56	Morocco	Public acceptance

Appendix A6: Tables for replication with additional guards against social desirability bias

Table A16: Replication of Figure 2

	term	estimate	std.error	p.value	model
1	Gender balance	0.08	0.03	0.01	Pooled
2	Pro-women decision	0.66	0.03	0.00	Pooled
3	Gender balance	0.09	0.05	0.06	Jordan
4	Pro-women decision	0.75	0.05	0.00	Jordan
5	Gender balance	0.04	0.06	0.51	Morocco
6	Pro-women decision	0.63	0.06	0.00	Morocco
7	Gender balance	0.11	0.05	0.04	Tunisia
8	Pro-women decision	0.61	0.05	0.00	Tunisia

Table A17: Replication of Figure 3

	term	estimate	std.error	p.value	model
1	Gender balance	0.08	0.03	0.01	Pooled
2	Pro-women decision	0.66	0.03	0.00	Pooled
3	Gender balance	0.09	0.05	0.06	Jordan
4	Pro-women decision	0.75	0.05	0.00	Jordan
5	Gender balance	0.04	0.06	0.51	Morocco
6	Pro-women decision	0.63	0.06	0.00	Morocco
7	Gender balance	0.11	0.05	0.04	Tunisia
8	Pro-women decision	0.61	0.05	0.00	Tunisia

Table A18: Replication of Figure 4

	term	estimate	std.error	p.value	model
1	Gender balance	0.02	0.03	0.50	Pooled
2	Pro-women decision	0.43	0.03	0.00	Pooled
3	Gender balance	0.07	0.05	0.14	Jordan
4	Pro-women decision	0.52	0.05	0.00	Jordan
5	Gender balance	-0.01	0.06	0.87	Morocco
6	Pro-women decision	0.29	0.06	0.00	Morocco
7	Gender balance	-0.01	0.05	0.92	Tunisia
8	Pro-women decision	0.46	0.05	0.00	Tunisia

Table A19: Replication of Figure 5

	term	estimate	std.error	p.value	model	outcome
1	Interaction between treatments	-0.00	0.05	0.94	Pooled	Right decision scale
2	Decision treatment	0.60	0.05	0.00	Pooled	Right decision scale
3	Gender balance treatment	0.18	0.07	0.01	Pooled	Right decision scale
4	Interaction between treatments	0.00	0.08	0.99	Jordan	Right decision scale
5	Decision treatment	0.72	0.08	0.00	Jordan	Right decision scale
6	Gender balance treatment	0.27	0.11	0.01	Jordan	Right decision scale
7	Interaction between treatments	0.04	0.09	0.68	Tunisia	Right decision scale
8	Decision treatment	0.54	0.09	0.00	Tunisia	Right decision scale
9	Gender balance treatment	0.15	0.12	0.22	Tunisia	Right decision scale
10	Interaction between treatments	-0.06	0.11	0.61	Morocco	Right decision scale
11	Decision treatment	0.52	0.10	0.00	Morocco	Right decision scale
12	Gender balance treatment	0.12	0.14	0.40	Morocco	Right decision scale
13	Interaction between treatments	0.11	0.05	0.05	Pooled	Attitudes toward committee scale
14	Decision treatment	0.56	0.05	0.00	Pooled	Attitudes toward committee scale
15	Gender balance treatment	0.05	0.07	0.45	Pooled	Attitudes toward committee scale
16	Interaction between treatments	0.13	0.09	0.13	Jordan	Attitudes toward committee scale
17	Decision treatment	0.75	0.08	0.00	Jordan	Attitudes toward committee scale
18	Gender balance treatment	0.13	0.11	0.24	Jordan	Attitudes toward committee scale
19	Interaction between treatments	0.18	0.09	0.04	Tunisia	Attitudes toward committee scale
20	Decision treatment	0.51	0.09	0.00	Tunisia	Attitudes toward committee scale
21	Gender balance treatment	-0.12	0.12	0.34	Tunisia	Attitudes toward committee scale
22	Interaction between treatments	0.00	0.11	0.98	Morocco	Attitudes toward committee scale
23	Decision treatment	0.35	0.10	0.00	Morocco	Attitudes toward committee scale
24	Gender balance treatment	0.20	0.14	0.15	Morocco	Attitudes toward committee scale
25	Interaction between treatments	0.02	0.05	0.73	Pooled	Public acceptance
26	Decision treatment	0.46	0.05	0.00	Pooled	Public acceptance
27	Gender balance treatment	0.03	0.07	0.64	Pooled	Public acceptance
28	Interaction between treatments	0.13	0.08	0.12	Jordan	Public acceptance
29	Decision treatment	0.62	0.08	0.00	Jordan	Public acceptance
30	Gender balance treatment	-0.01	0.11	0.96	Jordan	Public acceptance
31	Interaction between treatments	-0.09	0.09	0.31	Tunisia	Public acceptance
32	Decision treatment	0.45	0.09	0.00	Tunisia	Public acceptance
33	Gender balance treatment	0.09	0.12	0.45	Tunisia	Public acceptance
34	Interaction between treatments	-0.03	0.10	0.76	Morocco	Public acceptance
35	Decision treatment	0.23	0.09	0.01	Morocco	Public acceptance
36	Gender balance treatment	0.08	0.13	0.56	Morocco	Public acceptance

Appendix B: Figures

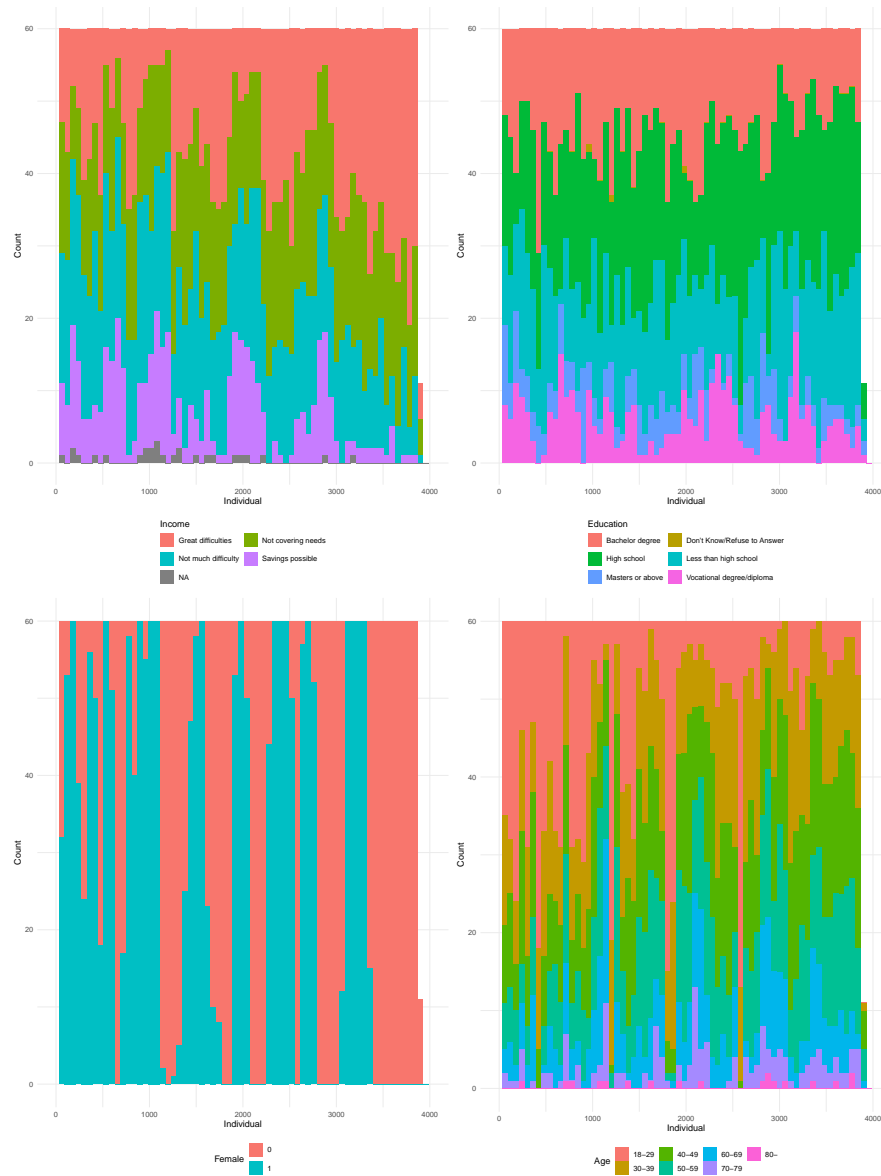


Figure 6: CATEs for Other Covariates



Figure 7: CATEs for Other Covariates (continued)