Online Supplements for Bai, Uhlmann, & Berdahl,

"The Robustness of the Win-Win Effect"

Supplement 1: National wealth is a conservative control when predicting Olympic medal wins from gender equality (but should still be controlled for)

It is worth elaborating on why controlling for national wealth represents a conservative test of the "Win-Win" hypothesis when estimating the relationship between gender equality and medal wins. Let us make the plausible assumption that the causal relationship between gender equality and economic growth is bidirectional. If the effects are reciprocal, then there likely exists (1) a direct effect of gender equality on medal wins, (2) an indirect effect of gender equality on medal wins mediated by national wealth, and (3) a spurious component of the correlation between gender equality and medal wins that is really due to the third variable of national wealth. If so, then controlling for national wealth removes not only (3) but also (2), underestimating the true relationship between gender equality and medal wins, which consists of both direct (unmediated) and indirect (mediated) effects. Figure S1 below illustrates this visually.

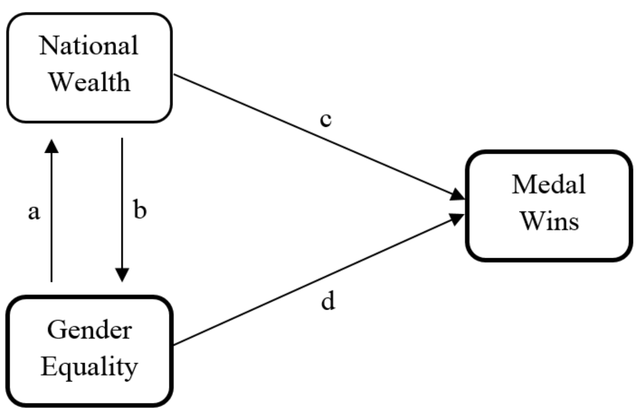

Figure S1. Theoretical path model of the relationships between gender equality, national wealth, and Olympic medal wins.

Let us assume a path model with two variables that cause each other (national wealth and gender equality, paths a and b) and a third (outcome) variable that is caused by these two variables (medals, paths c and d). The total effect of gender equality on medals is (1) the direct effect (path d) and (2) the indirect effect through national wealth (paths a*c). By statistically controlling for national wealth, we remove the influence of national wealth and thus examine only the direct effect (i.e., the part of the total effect that is not genuinely mediated by national wealth *and* that is not a spurious relationship due to the third variable of national wealth). As a consequence, the regression coefficients controlling for national wealth in Tables 2-5 of the main text underestimate the total effect of gender equality on medal wins.

At the same time, it is also the case that estimating the relationship between gender equality and medal wins without controlling for economic wealth would represent an overly liberal test of the hypothesis. This would leave not only (1) the direct effect of gender equality on medal wins and the (2) indirect effect of gender equality on medal wins that is mediated by economic wealth, but also (3) the spurious component of the correlation between gender equality and Olympic medal wins that is due to the third variable of national wealth.

Supplement 2: Effects of different analytic approaches on support for the Win-Win hypothesis

GDP per capita and national population are both positively skewed variables. Kuppens and Pollet (2015) inconsistently log transform GDP per capita to correct for this skew and use raw scores for national population. This is a simple and understandable mistake to make, but turns out to have major implications for the degree of support obtained for the Win-Win hypothesis.

The potential analytic approaches can be more fully represented as a 2 (log transform vs. use raw scores) x 2 (GDP per capita vs. national population) matrix. How do these different specifications affect the results?

There exist six significant zero-order correlations between measures of gender equality and Olympic medal wins (Table 1): specifically, between overall gender equality, educational gender equality, and economic gender equality and medal wins for both male and female athletes. Table S2 below summarizes whether these relationships remain statistically significant in regressions using each analytic approach.

		GDP per capita															
		Log transformed								Raw							
		Women's medals				Men's medals				Women's medals				Men's medals			
		Quasipoisson		Negative binominal		Quasipoisson		Negative binominal		Quasipoisson		Negative binominal		Quasipoisson		Negative binominal	
		b	*p*	*b*	*p*	*b*	*p*	*b*	*p*	*b*	*p*	*b*	*p*	*b*	*p*	*b*	*p*
Population																	
Log transformed	Overall gender equality	.33	.013	.83	<.001	.05	.643	.27	.128	.40	.009	1.09	<.001	.23	.097	.54	.006
	Educational gender equality	1.42	.004	.75	.031	1.01	.010	.72	.031	2.09	<.001	1.24	<.001	1.83	<.001	1.42	<.001
	Economic gender equality	.56	<.001	.97	<.001	.20	.082	.41	.005	.64	<.001	.97	<.001	.34	.013	.41	.020
Raw	Overall gender equality	**.06**	**.702**	*.71*	*.010*	**-.15**	**.259**	*.11*	*.556*	.29	.097	.87	.002	.15	.327	.29	.171
	Educational gender equality	**.72**	**.194**	*.27*	*.427*	**.52**	**.215**	*.32*	*.308*	1.69	.005	.78	.011	1.48	.001	1.06	<.001
	Economic gender equality	.43	.025	.76	.001	-.001	.993	.20	.180	.60	.004	.67	.007	.23	.167	.12	.485

Table S2. Regression results under each possible specification. Results based on inconsistently log transformed variables are presented with a grey background. The analyses reported in Kuppens and Pollet's (2015) main text are in **bold**. Analyses reported in their supplement are presented in *italics*. Analyses that were not part of their commentary are in regular font.

It is illustrative to examine the number of statistically significant effects in each quadrant of Table S2. As seen in the top left quadrant, when both GDP per capita and national population are log transformed (as in our analyses from the main text), 9 of 12 tests of the relationship between gender equality and medal wins are positive and statistically significant, one is positive and marginally significant, and two are positive but nonsignificant.

The bottom right quadrant of Table S2 displays the results if raw scores are used for both GDP per capita and national population despite their positively skewed distributions. In this specification, seven statistical tests for the relationship between gender equality and medal wins are positive and significant, one is positive and marginally significant, and four are positive but nonsignificant.

The top right quadrant of Table S2 displays the results if raw scores are used for GDP per capita and national population is log transformed. In this specification, 11 statistical tests for the relationship between gender equality and medal wins are significantly positive, one is positive and marginally significant, and none are nonsignificant.

The specification that produces the least support for the Win-Win hypothesis is log transforming GDP per capita but not national population, as Kuppens and Pollet (2015) did. As shown in the bottom left quadrant of Table S2, this analytic approach results in nine nonsignificant relationships between gender equality and medal wins: eight positive and one negative. The other three relationships are significant and positive. One of these, a significant positive relationship between overall gender equality and medal wins for female athletes in the negative binomial regression, is reported in their supplement.

Two of these three significant positive effects are not included in Kuppens and Pollet's (2015) commentary (either the main text or supplement). The reason is that Kuppens and Pollet analyzed overall World Economic Forum gender gap scores and only one of the four World Economic Forum gender gap subindexes, specifically educational equality between men and women. This was done because educational gender equality was the best predictor out of the four subindexes in Berdahl et al. (2015). Unfortunately, however, this approach led Kuppens and Pollet to overlook significant effects of *economic* gender equality.

In sum, log transforming GDP per capita and using raw scores for population yields less support for the Win-Win effect than the other potential analytic approaches (see Table S2). Kuppens and Pollet also overlook positive relationships between measures of gender equality and medal wins that emerge when the available measures are more fully analyzed.

Supplement 3: Further empirical issues with the use of regions in cross-national analyses

Researchers should strive to avoid not only Type 1 but also Type 2 errors. There are a limited number of nations with reliable data for many variables of interest, and further including regions in the analysis risks increasing the false negative rate too high. The loss of degrees of freedom when regional controls are included is a comparatively minor issue when the sample includes 121 countries, as in the Olympic medals and gender equality dataset. However, it is a significant problem for studies that only include twenty or thirty countries (e.g., Gelfand et al., 2011; Glick et al., 2000). Cross-national investigations based on new original data collected by the primary investigators and colleagues at other universities (e.g., Glick et al., 2000) are potentially crippled by reduced power.

At the same time, meaningful variability in both predictors and outcomes is reduced, in that prediction must occur within regions (Brauer, 2015). For instance, for national collectivism to predict an outcome of interest, it must do so within each area of the world (e.g., "Central and Eastern Europe," "Southern Europe"), rather than across all countries of the world. The field of cross-national comparisons becomes the study of variability within regions. Typically, however, the hypothesis is that *across the world*, nations with high scores on variable A (e.g., gender equality) further exhibit characteristic B (e.g., Olympic medal wins).

These empirical issues add to the more fundamental problem that regional groupings of nations are inherently subjective and arbitrary, regardless of whether the regional distinctions are made by Kuppens and Pollet (2014, 2015) or by other investigators.

References

Brauer, M. (2015). *Different ways to deal with non-independent data in scientific research: The example of world regions in cross-cultural research.* Manuscript under review.

Gelfand, M.J., Raver, J.L., Nishii, L.M., Leslie, L., Lun, J., Lim, B.C., et al. (2011). Differences between tight and loose cultures: A 33-Nation Study. *Science, 332*, 1100-1104.

Glick, P., Fiske, S. T., Mladinic, A., Saiz, J, Abrams, D., Masser, B., Adetoun, B., Osagie, J., Akande, A., Alao, A., Brunner, A., Willemsen, T. M., Chipeta, K., Dardenne, B., Dijksterhuis, A., Wigboldus, D., Eckes, T., Six-Materna, I., Expósito, F., Moya, M., Foddy, M., Kim, H-J., Lameiras, M., Sotelo, M. J., Mucchi-Faina, A., Romani, M., Sakalli, N., Udegbe, B., Yamamoto, M., Ui, M., Ferreira, M. C., & López, W. L. (2000). Beyond prejudice as simple antipathy: Hostile and benevolent sexism across cultures. *Journal of Personality and Social Psychology, 79*, 763-775.

Kuppens, T., & Pollet, T. V. (2014). Mind the level: problems with two recent nation-level analyses in psychology. *Frontiers in Psychology, 5*, 1110. doi:10.3389/fpsyg.2014.01110

Kuppens, T., & Pollet, T. V. (2015). Gender equality probably does not affect performance at the Olympic games: A comment on Berdahl, Uhlmann, and Bai (2015). *Journal of Experimental Social Psychology.*