The gender gap in adolescent mental health: a cross-national investigation of 566,827 adolescents across 73 countries

O L K Campbell[1], Praveetha Patalay[1*], David Bann[1*]
[1]University College London, UK
*Contributed equally
Corresponding author: O L K Campbell: Olympia.campbell.15@ucl.ac.uk

Table of Contents

Table S1: Regional groupings

EUR A	EUR B	EUR C	Eastern Mediterranean (EMR)	South East Asia/ Western Pacific (SEA WPR)	Americas (AMR)
Austria	Albania	Belarus	Jordan	Australia	Argentina
Belgium	Baku (Azerbaijan)	Estonia	Lebanon	Brunei Darussalam	Brazil
Croatia	Bosnia and Herzegovina	Hungary	Morocco	Chinese Taipei	Canada
Czech Republic	Bulgaria	Kazakhstan	Qatar	Hong Kong	Chile
Denmark	Georgia	Latvia	Saudi Arabia	Indonesia	Colombia
Finland	Kosovo	Lithuania	United Arab Emirates	Japan	Costa Rica
France	Montenegro	Moldova		Korea (Republic of)	Dominican Republic
Germany	Poland	Russian Federation		Macao	Mexico
Greece	Romania	Ukraine		Malaysia	Panama
Iceland	Serbia			Philippines	Peru
Ireland	Slovak Republic			Vietnam	United States
Italy	Turkey			Thailand	Uruguay
Luxembourg	North Macedonia				
Malta					
Netherlands					
Portugal					
Slovenia					
Spain					
Sweden					
Switzerland					
United Kingdom					

Regional groupings based on World Health Organisation categories. South East Asia and Western Pacific groups were combined, and Kosovo, Serbia and Montenegro were all classified as Europe B as the WHO does not include them in their classifications

Table S2: Measurement Invariance for Psychological Distress, Hedonia, and Eudaemonia

	Configural	Metric	Scalar
Gender (2 groups)			
RMSEA[1]	0.077	0.074	0.071
SRMR[2]	0.042	0.043	0.043
CFI[3]	0.969	0.969	0.967
TLI[4]	0.960	0.963	0.967
Region (6 groups)			
RMSEA[1]	0.081	0.078	0.079
SRMR[2]	0.045	0.045	0.048
CFI[3]	0.967	0.966	0.954
TLI[4]	0.958	0.961	0.960
Region X Gender (12 groups)			
RMSEA[1]	0.080	0.077	0.081
SRMR[2]	0.044	0.046	0.049
CFI[3]	0.968	0.965	0.950
TLI[4]	0.959	0.961	0.958

[1] Root Mean Square Error of Approximation
[2] Standardised Root Mean Square Residual
[3] Comparative Fit Index
[4] Tucker-Lewis Index

We tested measurement invariance for the 3 outcome measures that had more than one item (the life satisfaction is a single item measure), by fitting a model where the items of each of the three scales loaded onto the respective scale. We investigated measurement invariance at three levels: configural, metric and scalar[1]. For the RMSEA and SRMR, a score closest to zero indicates good model fit, whereas for CFI it is closest to one. Rules of thumb for model comparison are as follows: for RMSEA invariance is met if a change is smaller than 0.015, for SRMR if it is smaller than 0.03 and for CFI if it is smaller than 0.01[2]. We investigated three measurement invariance by three factors: gender (2 groups: male and female), region (6 groups: Europe A, Europe B, Europe C, Eastern Mediterranean, South East Asia/Western Pacific, and Americas) and genderxregion (12 groups: each region by sex). As can be seen from Table S2 above, the measurement invariance conditions were adequately met. We are therefore satisfied that the group invariance assumption is met.

[1] Putnick & Bornstein (2016) Measurement invariance conventions and reporting: The state of the art and future directions for psychological research. Developmental review. 41:71-90. doi: 10.1016/j.dr.2016.06.004

[2] Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling*, *14*(3), 464–504. https://doi.org/10.1080/10705510701301834

Table S3: Comparison of the different gender equality measures and the indicators that make up each dimension.

Index	*Dimension*			
	Political	**Educational**	**Economic**	**Health**
Gender Inequality Index	1. Female and male shares of parliamentary seats	1. Female and male population with at least secondary education	1. Female and male labour force participation rates	1. Maternal mortality ratio 2. Adolescent birth rate
Global Gender Gap Index	1. Ratio of female to male seats in parliament 2. Ratio of females to males at ministerial level 3. Ratio of number of years with a female head of state (last 50 years) to years with male head of state	1. Ratio of female literacy rate to male 2. Ratio of female net primary enrolment rate to male 3. Ratio of female net secondary enrolment rate to male 4. Ratio of female gross tertiary enrolment ratio to male	1. Ratio of female labour force participation to male 2. Wage equality between women and men for similar work 3. Ratio of female estimated earned income to male 4. Ratio of female legislators, senior officials and managers over male value 5. Ratio of female professional and technical workers to males	1. Sex ratio at birth 2. Ratio of female healthy life expectancy to males
Gender Social Norms Index	1. "men make better political leaders than women do" 2. "women have the same rights as men"	1. "University is more important for a man than for a woman"	1. "men should have more right to a job than women" 2. "men make better business executives than women do"	1. Proxy for intimate partner violence 2. Proxy for reproductive rights

The GGGI measures gaps in gender equality rather than levels. For example, countries will score higher or lower on the GGGI depending on the difference in higher education enrolment between men and women but not for the overall level of higher education in the country. This separates the GGGI from a country's level of development to a greater extent than the GII. The items in the GSNI are derived from the World Values Survey

Table S4: Correlation matrix showing the country-level correlations between the country-level variables.

	GII	GGGI	GSNI	GDP per cap	Gini
GII	1.000	-0.56	0.82	-0.79	0.60
GGGI		1.000	-0.65	0.48	-0.20
GSNI			1.000	-0.85	0.47
GDP per cap				1.000	-0.41
Gini					1.000

Abbreviations: GII = Gender Inequality Index; GGGI = Global Gender Gap Index; GSNI = Gender Social Norms Index; GDP per cap = Gross Domestic Product per capita. The GII and GSNI are more highly correlated with the economic indicators GDP per capita and Gini than the GGGI is.

Table S5: Regression coefficients with standard errors (SE) and variance components of multi-level models with controls for life satisfaction

	Model A: Baseline Model Coef (SE)	**Model B: + Sex** Coef (SE)	**Model C: + country-level indicators** Coef (SE)	**Model D: + cross-level interactions** Coef (SE)
Intercept	-0.00046 (0.029)	-0.12 (0.028) ***	-0.87 (0.37)*	-0.82 (0.37)*
Sex female		-0.16 (0.015) ***	-0.17 (0.017)***	0.32 (0.21)
Age		-0.0019 (0.0004) ***	-0.0015 (0.0004)***	-0.0015 (0.0004)***
ESCS		0.047 (0.0016) ***	0.045 (0.002) ***	0.045 (0.002)***
Second-generation immigrant		-0.077 (0.0065) ***	-0.082 (0.007) ***	-0.082 (0.007)***
First-generation immigrant		-0.11 (0.007) ***	-0.12 (0.007) ***	-0.12 (0.007)***
GGGI x 10			0.15 (0.047)**	0.14 (0.047)**
GDP per capita x 10^{-4}			-0.027 (0.01) *	-0.028 (0.010)**
Gini			-0.00049 (0.003)	-0.0006 (0.003)
GGGI x 10 X sex female				-0.072 (0.03)*
GDP per cap x 10^{-4} X sex female				-0.018 (0.006)**
Gini X sex female				0.002 (0.002)
Variance Components				
Individual	0.95 (0.97)	0.92 (0.96)	0.93 (0.96)	0.93 (0.97)
National**:** Country	0.056 (0.24)	0.051 (0.23)	0.033 (0.18)	0.033 (0.18)
Sex Female		0.015 (0.12)	0.016 (0.13)	0.011 (0.10)
VPC	5.6%	5.2%	3.4%	3.4%

Model A baseline model to illustrate the between country variance. Model B includes sex and the control variables. ESCS is the economic, social and cultural status index produced by PISA as a measure of socioeconomic status. The comparison category for second- and first-generation immigrant is native. Model C includes all country-level variables. Model D includes all cross-level interactions with sex female. Since the GGGI scale runs from 0-1 we multiply it by 10 so that the coefficient for GGGI represents a 0.1-point increase in the scale. GDP per capita is rescaled by dividing by 10,000, so that the coefficient represents the association with an increase of 1 x 10^4 GDP per capita. Variance partition coefficient (VPC) indicates the proportion of unexplained variance attributable to differences between countries. *p<0.05 **p<0.01 ***p<0.00

Table S6: Regression coefficients with standard errors (SE) and variance components of multi-level models with controls for psychological distress

	Model A: Baseline Model Coef (SE)	**Model B: + Sex** Coef (SE)	**Model C: + country-level indicators** Coef (SE)	**Model D: + cross-level interactions** Coef (SE)
Intercept	-0.019 (0.032)	-0.27 (0.033)***	-0.52 (0.50)	0.20 (0.48)
Sex female		0.47 (0.016) ***	0.47 (0.017)***	-0.081 (0.21)
Age		0.001 (0.0004)*	0.001 (0.0004)*	0.001 (0.0004)*
ESCS		0.016 (0.002)***	0.017 (0.002)***	0.017 (0.002)***
Second-generation immigrant		0.011 (0.006)	0.012 (0.006)	0.012 (0.006)
First-generation immigrant		0.048 (0.006)***	0.064 (0.007)***	0.064 (0.007)***
GGGI x 10			0.014 (0.059)	-0.033 (0.061)
GDP per capita x 10^{-4}			0.018 (0.013)	0.005 (0.013)
Gini			0.002 (0.004)	0.004 (0.004)
GGGI x 10 X sex female				0.082 (0.027) **
GDP per cap x 10^{-4} X sex female				0.024 (0.006) ***
Gini X sex female				-0.003 (0.002)
Variance Components				
Individual	0.93 (0.96)	0.87 (0.93)	0.87 (0.93)	0.87 (0.93)
National**:** Country	0.073 (0.27)	0.075 (0.27)	0.058 (0.24)	0.056 (0.24)
Sex Female		0.018 (0.13)	0.018 (0.13)	0.010 (0.10)
VPC	7.3%	7.8%	6.2%	6.0%

Model A baseline model to illustrate the between country variance. Model B includes sex female and the control variables. ESCS is the economic, social and cultural status index produced by PISA as a measure of socioeconomic status. The comparison category for second- and first-generation immigrant is native. Model C includes all country-level variables. Model D includes all cross-level interactions with sex female. Since the GGGI scale runs from 0-1 we multiply it by 10 so that the coefficient for GGGI represents a 0.1-point increase in the scale. GDP per capita is rescaled by dividing by 10,000, so that the coefficient represents the association with an increase of 1 x 10^4 GDP per capita. Variance partition coefficient (VPC) indicates the proportion of unexplained variance attributable to differences between countries. *p<0.05 **p<0.01 ***p<0.001

Table S7: Regression coefficients with standard errors (SE) and variance components of multi-level models with controls for Hedonia

	Model A: Baseline Model Coef (SE)	**Model B: + Sex** Coef (SE)	**Model C: + country-level indicators** Coef (SE)	**Model D: + cross-level interactions** Coef (SE)
Intercept	-0.012 (0.028)	0.041 (0.03)	-1.3 (0.44) **	-1.4 (0.44)**
Sex female		-0.042 (0.015) **	-0.049 (0.016)**	0.97 (0.22)***
Age		-0.002 (0.0004)***	-0.002 (0.0004)***	-0.002 (0.0004)***
ESCS		0.062 (0.002)***	0.062 (0.002)***	0.062 (0.002)***
Second-generation immigrant		-0.025 (0.006)***	-0.025 (0.006)***	-0.025 (0.006)***
First-generation immigrant		-0.054 (0.007)***	-0.064 (0.007)***	-0.064 (0.007)***
GGGI x 10			0.14 (0.055) *	0.15 (0.056)**
GDP per capita x 10^{-4}			-0.026 (0.012) *	-0.026 (0.012)*
Gini			0.011 (0.004)**	0.011 (0.004)**
GGGI x 10 X sex female				-0.14 (0.028)***
GDP per cap x 10^{-4} X sex female				0.0037 (0.006)
Gini X sex female				-0.0003 (0.002)
Variance Components				
Individual	0.94 (0.97)	0.92 (0.96)	0.93 (0.96)	0.93 (0.96)
National**:** Country	0.056 (0.24)	0.063 (0.25)	0.047 (0.22)	0.047 (0.22)
Sex Female		0.016 (0.13)	0.016 (0.13)	0.011 (0.11)
VPC	5.6%	6.2%	4.7%	4.8%

Model A baseline model to illustrate the between country variance. Model B includes sex female and the control variables. ESCS is the economic, social and cultural status index produced by PISA as a measure of socioeconomic status. The comparison category for second- and first-generation immigrant is native. Model C includes all country-level variables. Model D includes all cross-level interactions with sex female. Since the GGGI scale runs from 0-1 we multiply it by 10 so that the coefficient for GGGI represents a 0.1-point increase in the scale. GDP per capita is rescaled by dividing by 10,000, so that the coefficient represents the association with an increase of 1 x 10^4 GDP per capita. Variance partition coefficient (VPC) indicates the proportion of unexplained variance attributable to differences between countries. *p<0.05 **p<0.01 ***p<0.001

Table S8: Regression coefficients with standard errors (SE) and variance components of multi-level models with controls for Eudaemonia

	Model A: Baseline Model Coef (SE)	**Model B: + Sex** Coef (SE)	**Model C: + country-level indicators** Coef (SE)	**Model D: + cross-level interactions** Coef (SE)
Intercept	-0.0024 (0.028)	0.042 (0.025)	0.040 (0.33)	0.096 (0.33)
Sex female		-0.071 (0.015) ***	-0.071 (0.016)***	0.096 (0.21)
Age		0.0001 (0.0004)	0.0002 (0.0004)	0.0002 (0.0004)
ESCS		0.015 (0.002)***	0.014 (0.002)***	0.014 (0.002)***
Second-generation immigrant		0.030 (0.006)***	0.040 (0.007)***	0.040 (0.007)***
First-generation immigrant		-0.003 (0.007)	-0.005 (0.007)	-0.005 (0.007)
GGGI x 10			-0.0006 (0.041)	-0.021 (0.042)
GDP per capita x 10^{-4}			-0.032 (0.009) ***	-0.034 (0.009)***
Gini			0.003 (0.003)	0.006 (0.003)
GGGI x 10 X sex female				-0.061 (0.026)*
GDP per cap x 10^{-4} X sex female				-0.006 (0.006)
Gini X sex female				0.008 (0.002)***
Variance Components				
Individual	0.94 (0.97)	0.93 (0.96)	0.94 (0.97)	0.94 (0.97)
National**:** Country	0.056 (0.24)	0.041 (0.20)	0.030 (0.17)	0.027 (0.16)
Sex Female		0.041 (0.20)	0.015 (0.12)	0.01 (0.1)
VPC	5.6%	4.2%	2.8%	2.8%

Model A baseline model to illustrate the between country variance. Model B includes sex female and the control variables. ESCS is the economic, social and cultural status index produced by PISA as its measure of socioeconomic status. The comparison category for second- and first-generation immigrant is native. Model C includes all country-level variables. Model D includes all cross-level interactions with sex female. Since the GGGI scale runs from 0-1 we multiply it by 10 so that the coefficient for GGGI represents a 0.1-point increase in the scale. GDP per capita is rescaled by dividing by 10,000, so that the coefficient represents the association with an increase of 1 x 10^4 GDP per capita. Variance partition coefficient (VPC) indicates the proportion of unexplained variance attributable to differences between countries. *p<0.05 **p<0.01 ***p<0.001

Table S9: Robust regression coefficients with standard errors (SE) and variance partition coefficient of final multi-level model with top 3 outlier countries removed.

	Life Satisfaction Coef (SE)	**Psychological Distress** Coef (SE)	**Hedonia** Coef (SE)	**Eudaemonia** Coef (SE)
Sex female	0.36	-0.41	0.86***	-0.05
	(0.20)	(0.21)	(0.22)	(0.22)
GGGI x 10	0.15**	-0.074	0.15*	-0.035
	(0.049)	(0.063)	(0.064)	(0.049)
GDP per capita x 10^{-4}	-0.032**	0.013	-0.026*	-0.034***
	(0.010)	(0.013)	(0.013)	(0.009)
Gini	-0.001	0.001	0.011*	0.005
	(0.004)	(0.004)	(0.004)	(0.003)
GGGI x 10 X sex female	-0.082**	0.126***	-0.12***	-0.044
	(0.02)	(0.027)	(0.028)	(0.0270
GDP per cap x 10^{-4} X sex female	-0.012*	0.024***	0.005	-0.003
	(0.005)	(0.006)	(0.006)	(0.005)
Gini X sex female	0.002	-0.002	-0.001	0.008***
	(0.002)	(0.002)	(0.002)	(0.002)
VPC	3.2%	5.4%	5%	2.8%

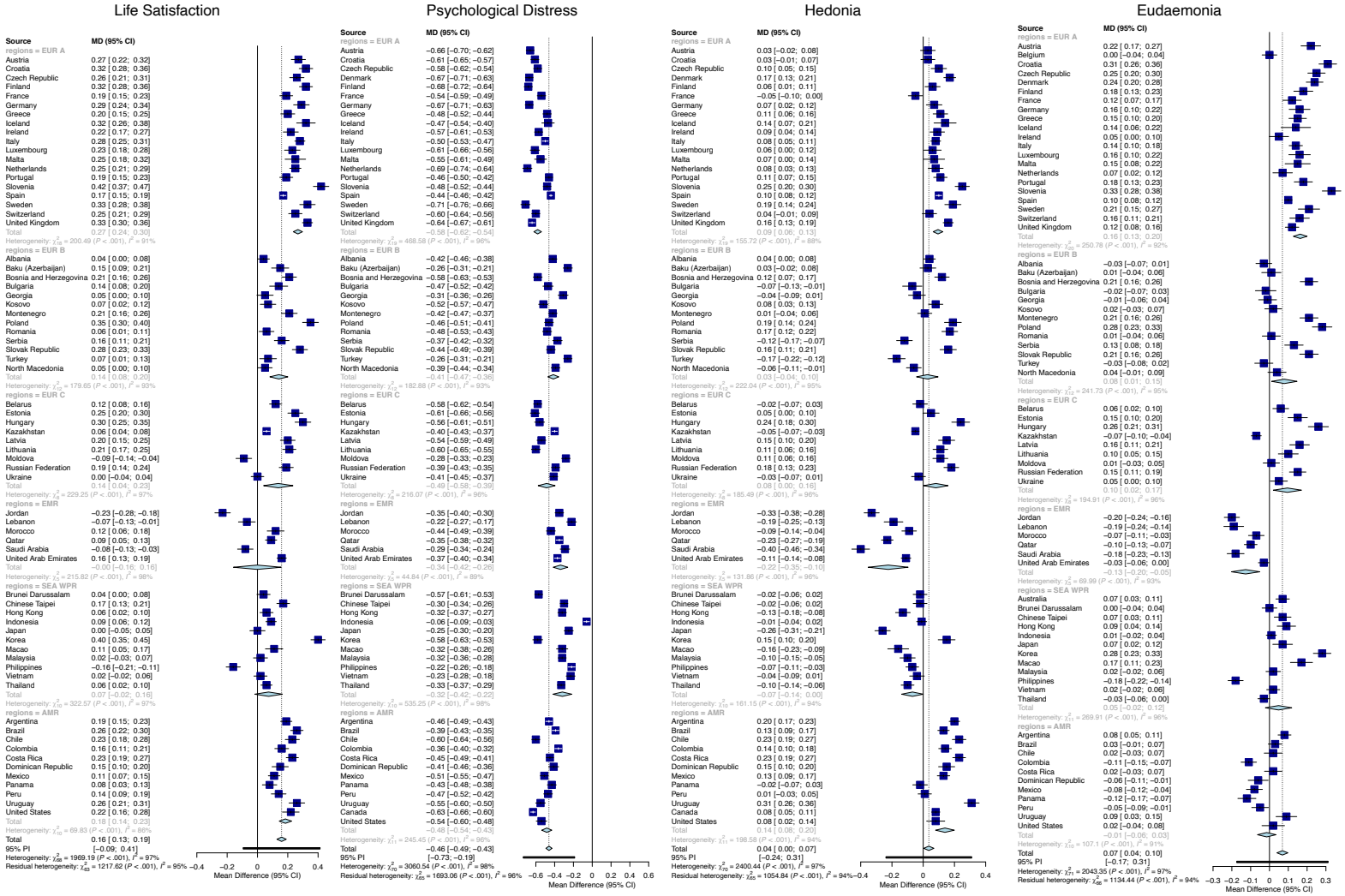
Final model with all cross-level interactions and individual level controls (ESCS, age, and immigration status). The top 3 countries with the highest cook's distance were removed to ensure the models are robust to outliers. Cook's distance of countries were identified in a single-level model with the average difference in each mental health variable as the outcome variable and input variables of the three country-level indicators (GGGI, GDP per capita and Gini). For LS the 3 countries removed were: Jordan; Moldova; Philippines. For PD: Ireland; Iceland; Philippines. For hedonia: Jordan; Philippines; Saudi Arabia. For eudaemonia: Jordan; Lebanon; Philippines. Models are robust to removal of these countries aside from eudaemonia where the coefficient for the interaction between GGGI and sex female loses significant.

Table S10: Multi-level models for each mental health outcome showing regression coefficients and between-country variance (VPC) using the GSNI

	Life Satisfaction Coef (SE)	Psychological Distress Coef (SE)	Hedonia Coef (SE)	Eudaemonia Coef (SE)
Model A: Baseline model				
Country VPC	5.6%	7.3%	5.6%	5.6%
Model B: including gender + controls				
Female	-0.16 ***	0.47***	-0.042**	-0.071***
	(0.015)	(0.016)	(0.015)	(0.015)
Country VPC	5.2%	7.8%	6.2%	4.2%
Model C: including country indicators + controls				
Female	-0.15***	0.45***	-0.04	-0.05*
	(0.022)	(0.02)	(0.02)	(0.02)
GSNI	-0.002	-0.003	0.001	0.006**
	(0.002)	(0.002)	(0.003)	(0.002)
GDP per cap x 10^{-4}	-0.03	0.032	-0.015	-0.005
	(0.02)	(0.021)	(0.023)	(0.014)
Gini	-0.004	0.007	0.009	-0.0005
	(0.005)	(0.005)	(0.005)	(0.003)
Country VPC	3.8%	6.8%	5.0%	1.9%
Model D : cross level interactions + controls				
Female	0.47**	-0.89***	-0.50***	0.56***
	(0.14)	(0.12)	(0.14)	(0.12)
GSNI	-0.002	0.003	-0.002	0.006***
	(0.002)	(0.003)	(0.003)	(0.002)
GDP per capita x 10^{-4}	-0.03	0.023	-0.029	-0.004
	(0.02)	(0.024)	(0.024)	(0.01)
Gini	-0.004	0.007	0.011*	0.0001
	(0.005)	(0.005)	(0.005)	(0.003)
GSNI X Female	0.005***	-0.006***	0.006***	0.004**
	(0.001)	(0.001)	(0.001)	(0.001)
GDP per cap x 10^{-4} X Female	-0.007	0.009	0.028*	0.007
	(0.01)	(0.009)	(0.012)	(0.01)
Gini X Female	-0.001	-0.0005	-0.003	0.005*
	(0.002)	(0.002)	(0.003)	(0.002)
Country VPC	3.8%	5.4%	4.9%	1.9%

Regression coefficients with standard errors (SE) from multilevel models. Model A presents the baseline model to calculate the country variance partition coefficient (VPC). Model B includes only sex, Model C contains all country-level factors and Model D contains all cross-level interactions between sex female and country-level factors. Models B-D are controlled for age, socioeconomic background and immigration status. GDP per capita is divided by 10,000, so that the coefficient represents the association with an increase of 1×10^4 GDP per capita. Note that higher values on the GSNI and Gini indicate greater inequality and that a positive coefficient for psychological distress indicates worse mental health in contrast to the other outcomes. Only the GGGI as a measure of gender equality is used due to the high correlations between the GII and GSNI and the economic variables (Table S3). *p<0.05 **p<0.01 ***p<0.001.

Figure S1: Forest plots of meta-analyses for each mental health outcome by regional sub-group



I^2 is the percentage of variation across nations due to heterogeneity rather than chance. MD = mean difference. High I^2 >85% indicates that there is considerable variation both within and between regions.

Figure S2: Distributions of life satisfaction by gender and country

Figure S3. Distributions of psychological distress for males and females by country and region.

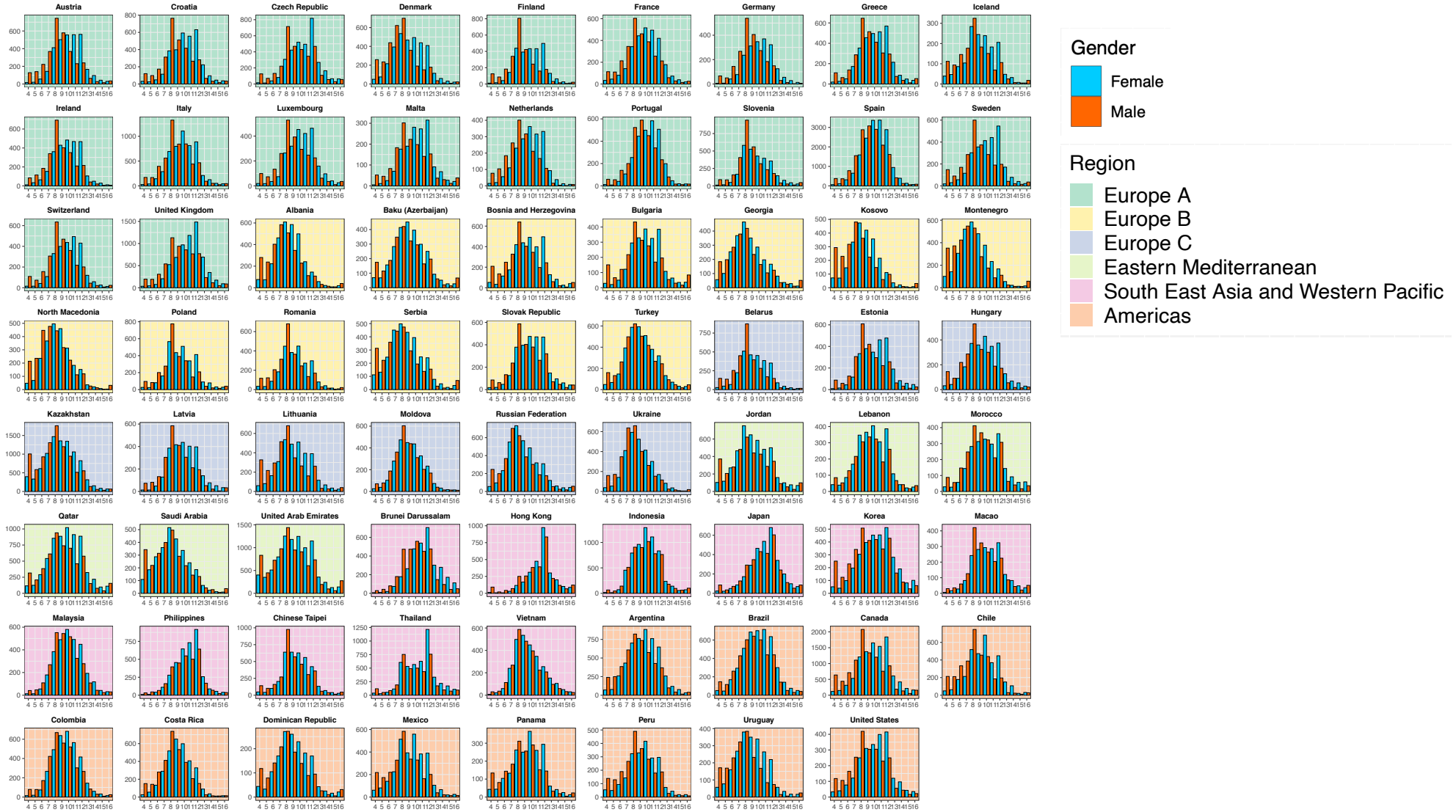


Figure S4: Distributions of hedonia by gender and country

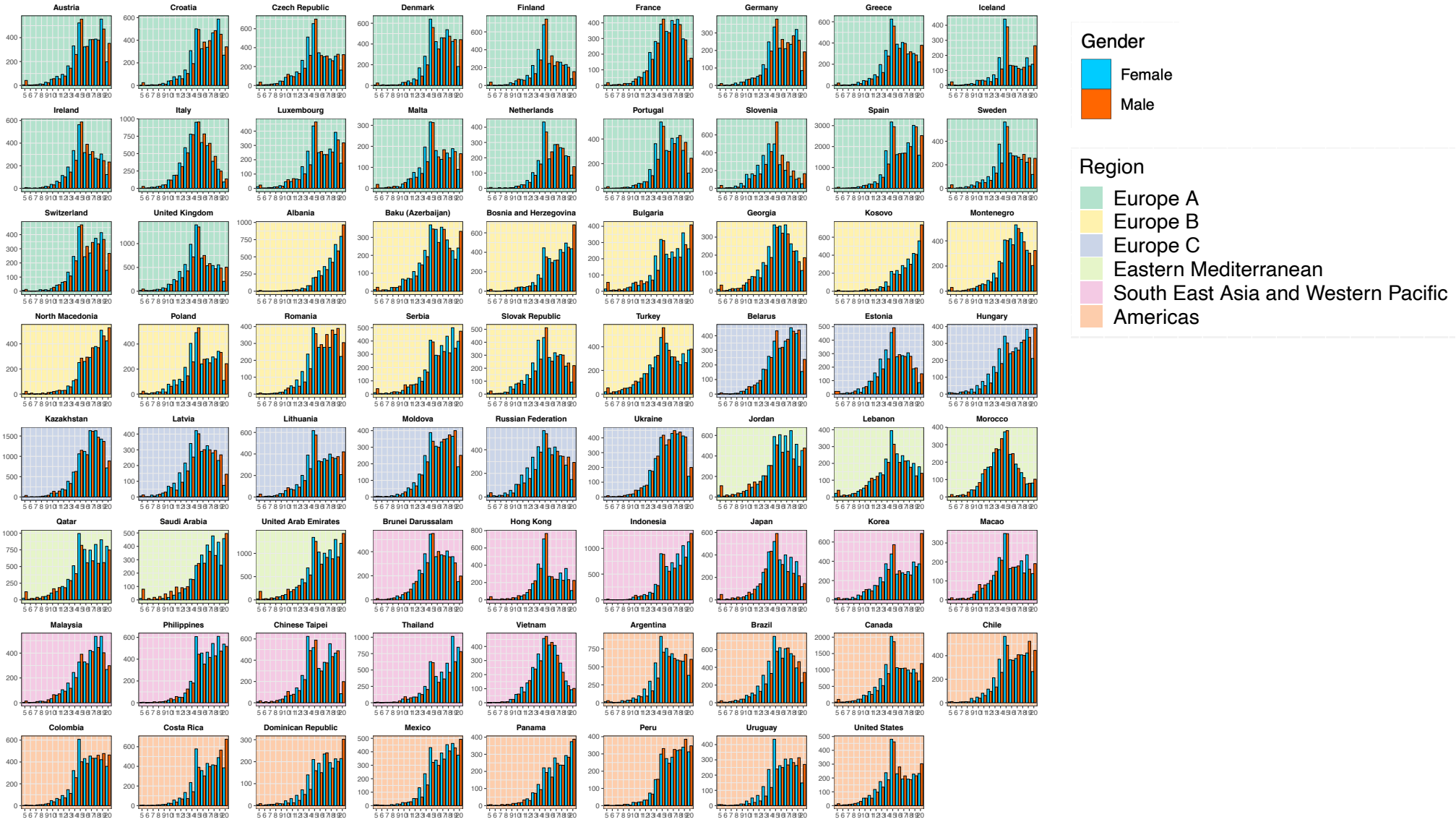


Figure S5: Distributions of Eudaemonia by gender and country

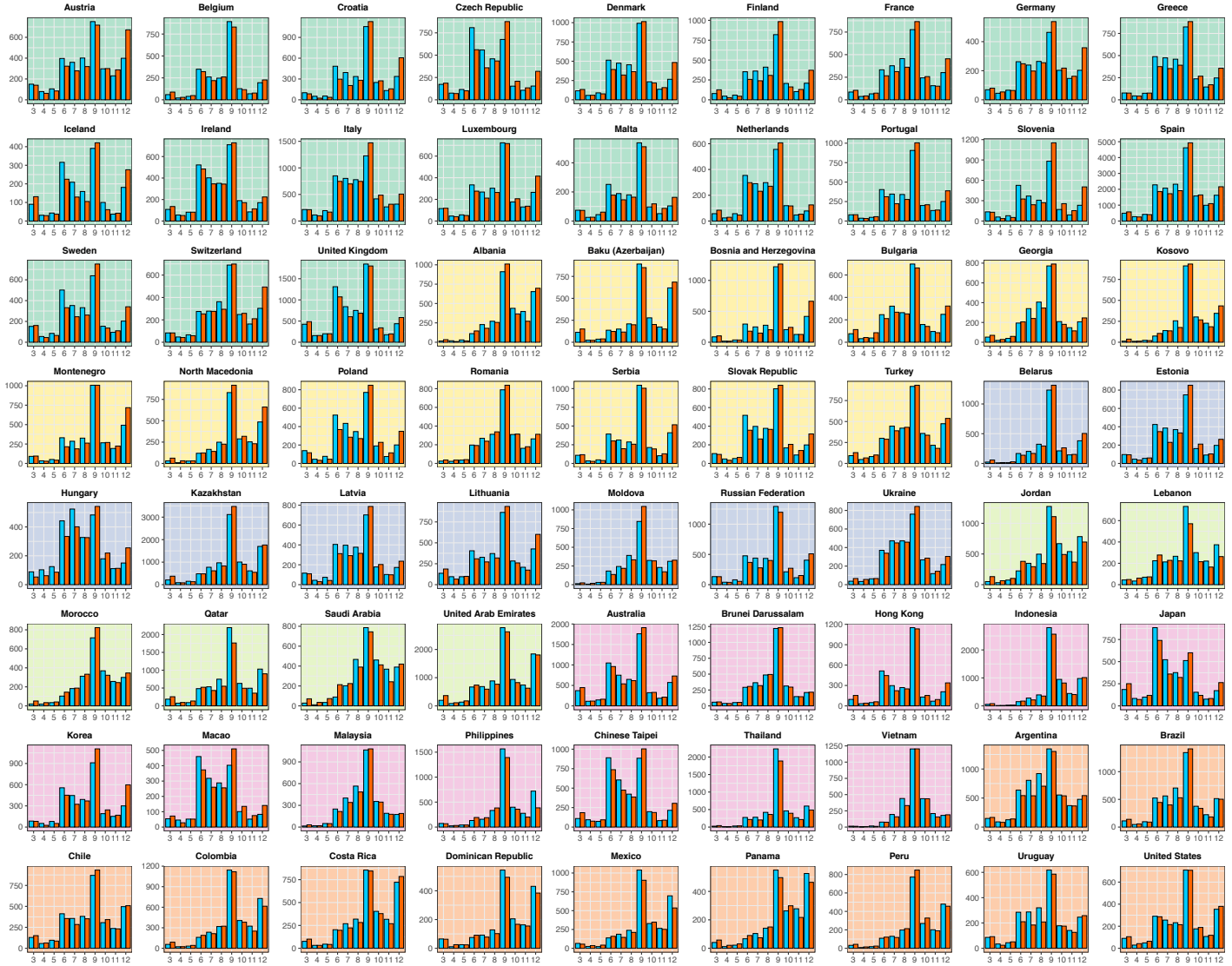
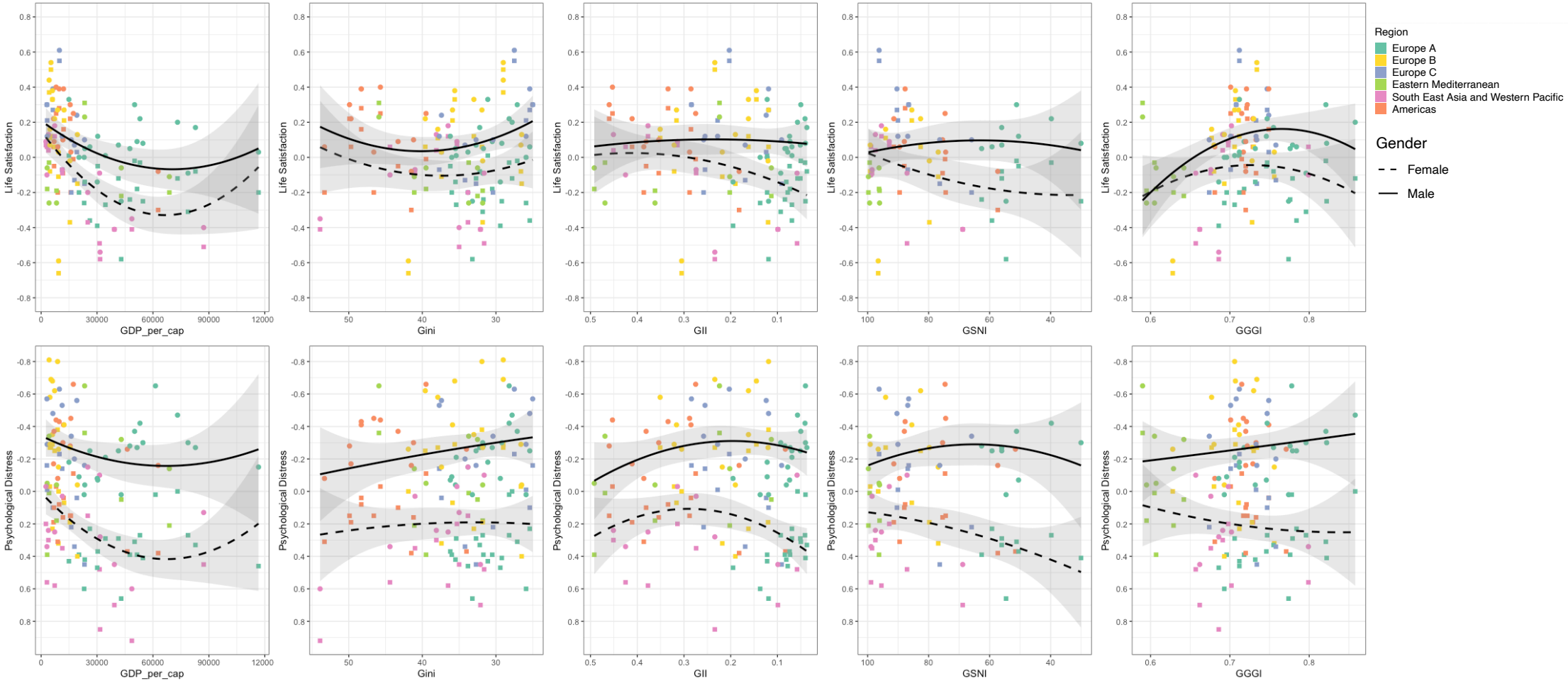
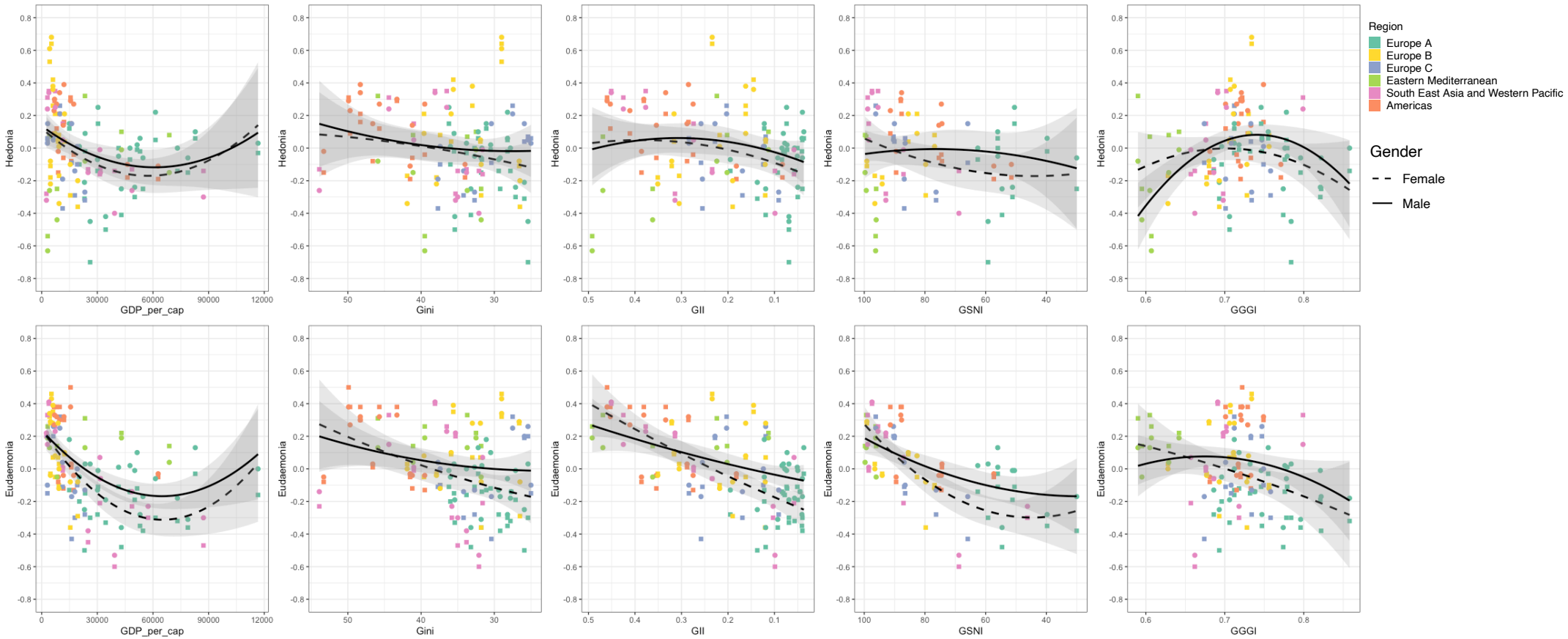


Figure S6: Quadratic associations of country-level economic and gender equality indicators with average life satisfaction and psychological distress.



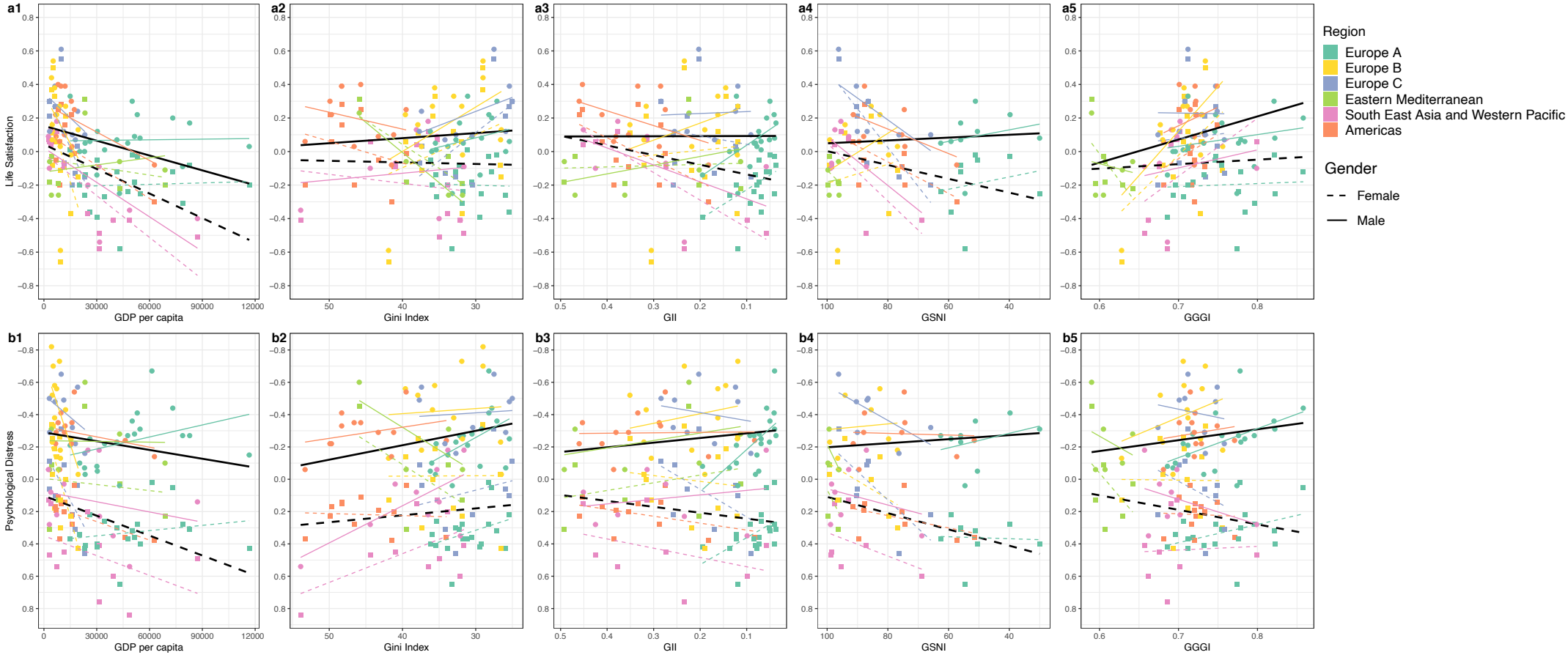
Quadratic country-level associations of economic indicators (GDP per capita and Gini) and gender equality indicators (GII, GSNI, and GGGI) with average standardised life satisfaction (a1-5) and psychological distress (b1-5) for females and males and coloured by region. The GII, GSNI and Gini scales are reversed so that they run from less equal to more equal permitting consistent interpretation with GGGI. A negative relationship indicates worse mental health across all outcomes. A larger distance between the regression lines indicates a larger gender gap. Abbreviations: Gini = income inequality, GII = gender inequality index, GSNI = gender social norms index, GGGI = global gender gap index.

Figure S7: Quadratic associations of country-level economic and gender equality indicators with average life satisfaction and psychological distress.



Quadratic country-level associations of economic indicators (GDP per capita and Gini) and gender equality indicators (GII, GSNI, and GGGI) with average standardised hedonia (c1-5) and eudaemonia (d1-5) for females and males and coloured by region. The GII, GSNI and Gini scales are reversed so that they run from less equal to more equal permitting consistent interpretation with GGGI. A negative relationship indicates worse mental health across all outcomes. A larger distance between the regression lines indicates a larger gender gap. Abbreviations: Gini = income inequality, GII = gender inequality index, GSNI = gender social norms index, GGGI = global gender gap index.

Figure S8: Associations of country-level economic and gender equality indicators with average life satisfaction and psychological distress by gender and with regional regression lines.



Country-level associations of economic indicators (GDP per capita and Gini) and gender equality indicators (GII, GSNI, and GGGI) with average standardised life satisfaction (a1-5) psychological distress (b1-5) for females and males and coloured by region. The GII, GSNI and Gini scales are reversed so that they run from less equal to more equal permitting consistent interpretation with GGGI. A negative relationship indicates worse mental health across all outcomes. A larger distance between the regression lines indicates a larger gender gap. Abbreviations: Gini = income inequality, GII = gender inequality index, GSNI = gender social norms index, GGGI = global gender gap index.