DISCUSSION PAPER SERIES

DP16740

The Gender Gap in Preferences: Evidence from 45,397 Facebook Interests

Ángel Cuevas, Ruben Cuevas, Klaus Desmet and Ignacio Ortuño-Ortín

LABOUR ECONOMICS

POLITICAL ECONOMY

CEPR

ISSN 0265-8003

The Gender Gap in Preferences: Evidence from 45,397 Facebook Interests

Ángel Cuevas, Ruben Cuevas, Klaus Desmet and Ignacio Ortuño-Ortín

Discussion Paper DP16740
Published 20 November 2021
Submitted 14 November 2021

Centre for Economic Policy Research
33 Great Sutton Street, London EC1V 0DX, UK
Tel: +44 (0)20 7183 8801
www.cepr.org

This Discussion Paper is issued under the auspices of the Centre's research programmes:

- Labour Economics
- Political Economy

The Gender Gap in Preferences: Evidence from 45,397 Facebook Interests

Abstract

This paper uses information on the frequency of 45,397 Facebook interests to study how the difference in preferences between men and women changes with a country's degree of gender equality. For preference dimensions that are systematically biased toward the same gender across the globe, differences between men and women are larger in more gender-equal countries. In contrast, for preference dimensions with a gender bias that varies across countries, the opposite holds. This finding takes an important step toward reconciling evolutionary psychology and social role theory as they relate to gender.

JEL Classification: D01, D10, D90, D91, J16, O57, Z1

Keywords: Gender Gap, preferences, Gender equality, evolutionary psychology, social role theory, gender-related preferences, non-gender-related preferences, Facebook interests, cross-country differences

Ángel Cuevas - acrumin@it.uc3m.es
Universidad Carlos III de Madrid

Ruben Cuevas - rcuevas@it.uc3m.es
University Carlos III de Madrid

Klaus Desmet - kdesmet@smu.edu
Southern Methodist University and CEPR

Ignacio Ortuño-Ortín - iortuno@eco.uc3m.es
Universidad Carlos III de Madrid

Acknowledgements
We thank Iñaki Ayensa for his help with the Facebook data, and we are grateful to Lola Collado, David Cuberes, Cynthia Desmet Villar, Ömer Özak, Javier Ruiz-Castillo, Romain Wacziarg and Beatriz Yankelevich for helpful comments. The authors furthermore acknowledge the financial support of Fundación BBVA (Ayudas a equipos de investigación for project "Aplicaciones económicas de distancias culturales"; the European Union's Horizon 2020 Innovation Action Program (PIMCITY project, grant No 871370); the Ministerio de Economía, Industria y Competitividad, Spain (grant RyC-2015-17732); the Comunidad de Madrid, Spain (EMPATIA-CM grant Y2018/TCS-5046 and MAD-ECON-POL-CM grant H2019/HUM-5891); and I+D+i Project PID2019-109157RB-IOO, MCIN/AEI/10.13039/501100011033/.

The Gender Gap in Preferences: Evidence from 45,397 Facebook Interests*

Ángel Cuevas *UC3M* | Rubén Cuevas *UC3M* | Klaus Desmet *SMU* | Ignacio Ortuño Ortín *UC3M*

October 25, 2021

Abstract

This paper uses information on the frequency of 45,397 Facebook interests to study how the difference in preferences between men and women changes with a country's degree of gender equality. For preference dimensions that are systematically biased toward the same gender across the globe, differences between men and women are larger in more gender-equal countries. In contrast, for preference dimensions with a gender bias that varies across countries, the opposite holds. This finding takes an important step toward reconciling evolutionary psychology and social role theory as they relate to gender.

1 Introduction

Do gender differences in preferences get attenuated or accentuated in more gender-equal societies? On the one hand, evolutionary psychology theory posits that gender equality accentuates differences by facilitating the expression of innate preferences that set men and women apart. On the other hand, social role theory posits that gender equality attenuates differences by eroding gender stereotypes and norms.

Using data on the prevalence of a comprehensive set of 45,397 interests by gender across most countries of the world, this paper takes an important step towards reconciling both theories. Our premise is that innately gender-specific interests should mostly conform to evolutionary psychology theory, whereas other interests should mostly conform to social role theory. We find strong evidence consistent with this premise.

*Cuevas, A.: Department of Telematic Engineering, Universidad Carlos III de Madrid, acrumin@it.uc3m.es, Cuevas, R.: Department of Telematic Engineering, Universidad Carlos III de Madrid, rcuevas@it.uc3m.es, Desmet: Department of Economics and Cox School of Business, Southern Methodist University, NBER and CEPR, kdesmet@smu.edu, Ortuño Ortín: Department of Economics, Universidad Carlos III de Madrid, iortuno@eco.uc3m.es. We thank Iñaki Ayensa for his help with the Facebook data, and we are grateful to Lola Collado, David Cuberes, Cynthia Desmet Villar, Ömer Özak, Javier Ruiz-Castillo, Romain Wacziarg and Beatriz Yankelevich for helpful comments. The authors furthermore acknowledge the financial support of Fundación BBVA (Ayudas a equipos de investigación for project "Aplicaciones económicas de distancias culturales"); the European Union's Horizon 2020 Innovation Action Program (PIMCITY project, grant No 871370); the Ministerio de Economía, Industria y Competitividad, Spain (grant RyC-2015-17732); the Comunidad de Madrid, Spain (EMPATIA-CM grant Y2018/TCS-5046 and MAD-ECON-POL-CM grant H2019/HUM-5891); and I+D+i Project PID2019-109157RB-IOO, MCIN/AEI/10.13039/501100011033/.

Our data on the prevalence of interests by gender and country come from Facebook. The social media company observes each of its almost three billion users' online activity, not just on its own platform, but also on all websites and apps where it has a presence. In addition, it tracks many of its users' offline activities by relying on GPS. Through their online and offline activities, users reveal their preferences and interests to Facebook. Using this information to assign interests to users, Facebook has unintentionally created the world's largest database on preferences.

By querying this database through Facebook's publicly available Marketing API, we collect for most countries of the world the number of male and female users interested in 45,397 different topics. Because the data are at the level of populations (e.g., Canadian men or Ghanaian women), they do not entail any individual privacy issues. Compared to other potential data sources on preferences, Facebook data have two key advantages. First, the interests are broad and comprehensive in their scope, ranging from religious beliefs and sports, to political positions and cuisine. Second, in contrast to surveys, Facebook interests constitute a bottom-up revealed measure of preferences, covering whatever users find interesting, rather than what social scientists deem important.

We start by computing for each country the cosine distance between the interest frequency vectors of men and women. This gives us a country-level metric of the overall difference in interests between genders. When regressing this metric on the degree of gender equality, we uncover a weak positive association between a country's gender equality and the interest gap between men and women. Because different interests may sometimes reflect the same underlying preferences, we use singular value decomposition of the data matrix to identify the main latent preference dimensions. When recomputing our distance metric in this lower-dimensional subspace, we find a slightly stronger positive association between a country's gender equality and its gender gap in preferences.

Next, we differentiate between gender-related and non-gender-related interests. We say that an interest is gender-related if it displays a systematic bias toward the same gender across the globe. More specifically, if in more than 90% of countries an interest is more prevalent among the same gender, then we refer to it as gender-related. For example, "cosmetics" and "motherhood" are universally more common among women, whereas "motorcycles" and "Lionel Messi" are universally more common among men. Conversely, we say that an interest is non-gender-related if its gender bias varies across countries. More specifically, if an interest is more common among men in at least 30% of countries and more common among women in at least another 30% of countries, then we refer to it as non-gender-related. For example, "world heritage site" and "physical fitness" do not display a systematic gender bias across the globe.

When exploring the relationship between a country's gender equality and the difference in interests between men and women, we uncover a sharp distinction between gender-related interests and non-gender-related interests. More gender equality is associated with greater differences between men and women for gender-related interests, whereas the opposite is true for non-gender-related interests. As an alternative way of classifying interests, we use singular value decomposition to differentiate between gender and non-gender dimensions of preferences. For a preference dimension to be gender-related, we require the relative positions of men and women along that dimension to be similar across countries. With this alternative method, we confirm the paper's central result: more gender-equal societies tend to be associated with greater differences in gender-related preferences but smaller differences in

non-gender-related preferences.

To interpret the paper's main empirical finding, we turn to two seemingly contrasting theories (Falk and Hermle, 2018). Evolutionary psychology argues that men and women differ in areas where they faced different adaptive problems in their evolutionary history (Atari, Lai and Dehghani, 2020). In societies with more equal gender rights, men and women are able to more freely express their innate predispositions, so that preference differences between men and women should widen (Buss, 1989; Schmitt, 2015; Atari, Lai and Deghani, 2020).[1] Social role theory, instead, argues that gender differences stem from gender socialization, social norms and sociocultural power structures (Schmitt et al., 2017). Since greater equality of gender rights erodes these norms, preference differences between men and women should narrow. While many papers on gender differences have been framed as a debate on the relative merits of evolutionary psychology and social role theory, these two views are not necessarily competing. Rather, their predictions apply to different preferences – evolutionary psychology to preferences that are innate and social role theory to preferences that are socially constructed.

How does the difference between innate and socially constructed preferences relate to our paper's main result? We argue that for preferences to be innate, they must display a systematic bias toward the same gender across the globe. As such, we can interpret our gender-related interests as potentially innate. In contrast, non-gender-related interests display a gender bias that varies across countries, and must hence be socially constructed. Using this interpretation, our findings are consistent with the predictions of both theories: in more gender-equal countries, differences between men and women are larger for innate (gender-related) preferences and smaller for socially constructed (non-gender-related) interests.

Our interpretation depends crucially on the way we classify interests, and hence requires caution. We refer to gender-related interests as *potentially* innate, because we cannot discard the possibility that some of these interests might be socially constructed. Of course, this would require the process of social construction to occur in the same way in all countries. While in general this seems quite unlikely, in some cases the process of globalization might have led to the homogenization of socially constructed norms across countries. In other cases nature might have given rise to universally held gender norms in the distant past that then persisted through nurture despite no longer having a biological basis.[2] For example, historically the relative physical strength of men and women was an important determinant of the division of labor between genders. As a result, universal gender norms emerged that associated some professions with men and others with women. Although technology has eroded these gendered patterns of comparative advantage, the gender norms might still survive.[3]

[1] An alternative explanation for why gender differences in preferences are larger in more gender-equal societies is that the elimination of traditional gender roles may increase the need for individuals to identify with their gender group, leading them to fall back on stereotypical gender interests and preferences (Breda et al., 2020). For example, in the workplace men may feel a stronger need to affirm their masculine identity when more women access traditionally male jobs (Akerlof and Kranton, 2000).

[2] For a discussion of the co-evolution of nature and nurture, see Boyd and Richerson (2005).

[3] For an example of the persistence of norms, Alesina, Giuliano and Nunn (2013) find that societies that depended on the plough in the distant past generated beliefs about the role of women that have survived until today. Unlike our example, these beliefs are not universal though, because the plough was not the dominant technology everywhere. For other examples on the role of culture and norms in the context of gender preferences, see Fernández, Fogli and Olivetti (2004) and Fernández and Fogli (2009). For an example where gender gaps have been reduced due to technological change, see the work by Goldin and Katz (2002) on the impact of oral contraceptives on career decisions.

While ultimately such norms still have an innate origin, they are no longer subject to biological determinism.

This paper is related to several strands of the literature on gender differences in preferences. Closest to our work is the large literature in psychology, sociology and economics that studies whether differences in values, attitudes and personality get accentuated in societies that are more gender-equal. Most empirical studies in this area have focused on gender differences in personality characteristics (Costa et al., 2001; Kaiser, 2019; Mac Giolla and Kajonius, 2019), cognitive abilities (Lippa, Collaer and Peters, 2010), education (Stoet and Geary, 2018), basic human values (Fors Connolly, Goossen and Hjerm, 2020), and specific cultural, behavioral and moral values (Falk and Hermle, 2018; Atari, Lai and Dehghani, 2020). Many of these studies find evidence of divergence between men and women in more gender-equal societies. For example, countries that are more gender-equal are found to exhibit greater sex differences in care and fairness (Atari, Lai and Dehghani, 2020), altruism, trust and risk-taking (Falk and Hermle, 2018), and the big five personality traits (Mac Giolla and Kajonius, 2018). Some other studies find the opposite or argue that this relation is not robust. For example, Guiso et al. (2008) show that in societies with greater gender equality the math gender gap narrows, and Kaiser (2019) argues that the gender divergence in personality traits disappears after controlling for ecological stress factors such as hunger and disease.

Our paper differs from this previous work in three respects. First, our data cover a broad cross-section of countries. Second, while most studies have focused on particular traits, values or abilities, we focus on 45,397 interests. Because of a lack of comprehensive data on interests and preferences, previous research has been unable to fully compare the predictions of evolutionary psychology and social role theory. Third, while these papers look at the effect of gender equality on differences in preferences, they do not address the possibility of causality running the other way. We deal with this potential endogeneity concern by taking an instrumental variable approach. Our results are suggestive of a causal interpretation of the paper's main finding.

Also related to our work is the literature that seeks to identify some of the key differences in preferences between men and women. Many experimental papers have documented systematic gender differences in risk attitudes, dislike of competition, and social preferences (see Croson and Gneezy, 2009, Bertrand, 2011, and Niederle and Vesterlund, 2011, for excellent surveys). An important, related, question is to what extent these gender differences are a consequence of nature or nurture (see Olivetti and Petrongolo, 2016, for a discussion) . Most direct evidence of the role of nature comes from studies that show that male hormones play a role in certain preferences, such as attitudes towards competition and risk-taking, as well as in career choices and activities (Archer, 2006; Dreber and Hoffman, 2007; Sapienza, Zingales and Maestripieri, 2009; Berenbaum and Beltz, 2021). More generally, the consensus points to both nature and nurture mattering. Even in the case of risk-taking, Gneezy et al. (2008) show that gender differences are society-dependent, ruling out a purely nature-based explanation.

Finally, an extensive literature in economics and political science explores how gender differences in preferences affect individual and societal choices. If women and men have different preferences, then greater female participation in political decision-making has wide-reaching consequences. Clots-Figueras (2012) demonstrates that the election of women politicians in India improves educational

attainment; Lippmann (2021) shows that in the French parliament female legislative activity focuses more on women's issues and male legislative activity more on the military; and Funk and Gathmann (2015) show that in direct democracy initiatives in Switzerland women make different choices in health, environmental protection, defense spending and welfare policy. Differences in preferences are also relevant within the household. Quisimbing and Maluccio (2000) show that giving more assets to women translates into an increase in spending on offspring in a variety of developing countries. This is an important insight for government policy that often relies on direct cash transfers to improve children's welfare. An additional effect of greater preference heterogeneity within the household is increased marital instability (Serra-Garcia, 2021). Gender differences in preferences also have important effects on career choices and other labor market outcomes (Bertrand, 2011). Hence, better understanding the evolution of gender differences in preferences is of great interest to economists.

The rest of the paper is organized as follows. Section 2 describes the data, with a special emphasis on the Facebook data on interests; Section 3 analyzes the relation between gender equality and gender differences in interests and preferences; Section 4 explores how this relation depends on whether interests and preferences are gender-related or not; and Section 5 concludes.

2 Data

This paper asks whether greater gender equality amplifies or attenuates differences in interests between men and women. Because we want to treat this question comprehensively, our biggest challenge is to get data on the prevalence of many different interests by gender for a large cross-section of countries. Below we describe how we achieve this by obtaining information on the frequency of 45,397 Facebook interests by gender and country. We also discuss, more briefly, the data on gender equality and other control variables.

2.1 Dependent Variable: Gender Differences in Interests

Data on interests by gender and country. Our data on interests by gender and country come from Facebook. The social network assigns interests to its almost 3 billion users worldwide based on their activity, both online and offline. Facebook observes its users' likes, shares, clicks and downloads, not just on its own platform but also on all other websites and apps where the company is present. Moreover, by having access to their GPS location, Facebook also observers many of its users' offline activity. By unobtrusively observing their users, Facebook has created a massive database on people's revealed preferences and interests.[4]

To construct a broad and comprehensive set of interests, we take the 1,000 most common words in English, as well as all possible combinations of one, two and three letters. For each one of these words and letter combinations, we query the Facebook Marketing API for up to 1,000 interests that match or contain these letters. This gives us a list of 308,568 interests. We keep the interests with

[4]Facebook's business model relies crucially on identifying as well as possible its users' true interests. The better it does so, the greater its ability to show its users relevant posts. If posts are less relevant, users spend less time on Facebook, causing a drop in company revenues.

a worldwide Facebook audience of more than one million but less than one billion.[5] This yields the 45,397 interests that we use in this paper.

For each one of these interests, we query Facebook's Marketing API for the corresponding number of monthly active users (MAU) by gender and country. This gives us, for example, the number of female users in France interested in "Youssou N'Dour" or the number of male users in Singapore interested in "chili crab". To automate the querying process, we developed a Facebook audience capture and analysis tool. Even when automated, this is a lengthy and time-consuming effort that spanned the entire first semester of 2019. Compared to previous work, the number of Facebook interests we use is very large and comprehensive. The paper that comes closest in number uses around 3,000 Facebook interests (Dubois et al., 2018). The only exception is our own recent work that uses 60,000 interests to measure cultural differences between countries (Obradovich et al., 2020).[6]

The Facebook data we use do not raise any privacy concerns. The Marketing API gives us information at the level of population groups, and never at the level of individuals. To further ensure anonymity, the minimum number of monthly active users (MAU) reported by the API for any demographic is 1,000. While in principle this can distort our distance measure between men and women, this is not an issue as long as the number of interests is large enough and as long as groups are not too small.[7]

To ensure that groups are sufficiently large, our baseline sample consists of 106 countries with a population above one million and a Facebook penetration rate of at least 25%. The relatively high Facebook penetration threshold also ensures that Facebook users are sufficiently representative of the population groups we are interested in. In our robustness checks, we consider alternative samples, with Facebook penetration thresholds ranging from a less restrictive 2.5% (149 countries) to a more restrictive 50% (68 countries).[8]

One issue affecting cross-country comparability is that Facebook users may be more biased towards younger populations in some countries than in others. To test the robustness of our results to this concern, we control for the ratio of young to old Facebook users. In addition, we obtain the interest frequencies by age and gender for a random subsample of 5,000 interests, and re-run our main regressions separately for the old and the young.[9]

Measuring gender differences in interests. Based on the frequency of Facebook interests of women and men, we propose a simple framework to compute gender differences. There are C countries,

[5] There are 39 interests with a Facebook audience of more than one billion. These are interests, such as "Facebook", that are very generic. To avoid an oversized effect of these large interests, we exclude them.

[6] Facebook data from the Marketing API are being increasingly used to study different socioeconomic issues, such as migrant assimilation (Dubois et al., 2018), tax policy (Lassmann et al., 2020), and political campaigns (Liberini et al., 2020). Facebook ads data have also been used to study specific aspects of the gender gap. Vieira and Vasconcelos (2021) analyze the gender balance in STEM in Brazil, Mejova et al. (2018) study the digital gender gap in India, and García et al. (2018) show how gender inequality in Facebook use is related to various aspects of gender inequality. Compared to our work, none of these papers uses a large and comprehensive set of Facebook interests.

[7] More specifically, if we choose a subset of most popular interests, our distance measure is virtually unchanged for any threshold above 20,000 interests. If, instead, we use a random subset of interests, the corresponding threshold is around 25,000 interests. When the number of interests is small, Rama et al. (2020) offer another solution to alleviate this problem.

[8] A notable country absent from our sample is China, where Facebook penetration is less than 1%.

[9] Unfortunately, even with our automated querying process, collecting data for the old and the young for all 45,397 interests would take an unreasonable amount of time.

indexed by c, and two gender groups m and w, with m referring to the group of men and w to the group of women. Individuals express their interest in all sorts of things, issues and topics. The set of possible Facebook interests has I elements, indexed by i. Each agent sends a list of signals expressing the interests that she likes most. In our data, the value of I is 45,397, and the average number of interests by user is about 300.

We do not observe the interests of individuals, but the interests of each country and gender. Let f_{ci}^w be the number of female Facebook users in country c who hold interest i, and let f_{ci}^m be the corresponding number of male users. Equivalently, these are the number of signals expressing an interest in i coming from female and male users in country c. We can then write the vector with the interest frequencies of women in country c as $f_c^w = \{f_{c1}^w, f_{c2}^w, ..., f_{cI}^w\}$. The corresponding vector for men is f_c^m. For example, an element of f_{ci}^w could be the total number of female Facebook users of country c who have sent a signal expressing an interest in "beer".

The cosine distance between the interest vector of men and women in country c measures the gender difference in interests in that country:[10,11]

$$CosDist_c = 1 - \frac{\sum_{i=1}^{I} f_{ci}^m \ f_{ci}^w}{\sqrt{\sum_{i=1}^{I} (f_{ci}^m)^2} \sqrt{\sum_{i=1}^{I} (f_{ci}^w)^2}} \tag{1}$$

In Obradovich et al. (2020) we show that this distance in Facebook interests is a good measure of cultural distances between populations. The average within-country distance between men and women is 0.08. To put this number in context, the average distance between populations of different countries is 0.25.[12] Figure 1 depicts the gender differences in Facebook interests in 149 countries with population above one million, Facebook penetration rate above 2.5% and number of Facebook users greater than 100,000. Appendix Table B.1 provides the full data.

2.2 Main Variable of Interest: Gender Equality

As the main measure of gender equality, we take the 2018 World Economic Forum's Gender Gap Index (WEF). This index is one of the best-established indices of gender equality and the only independent index published every year. It captures gender-based gaps in access to resources and opportunities in countries, rather than the actual level of those countries' available resources and opportunities. Thus, this index attempts to capture the level of gender equality separately from the level of economic development. The index is increasing in the degree of equality and has a scale from zero to one. It is made up of four subindices, related to economic opportunity, educational attainment, health outcomes and political empowerment. Examples of variables that contribute to the WEF gender equality index include female labor force participation relative to male, female earned income over male, sex ratio at birth, gender difference in healthy life expectancy, and females with seats in parliament. Figure 2

[10] This distance is proportional to the Euclidean distance if we normalize all the vectors to have modulo 1. In the Appendix we explore the robustness of our results to alternative distance measures, such as Euclidean and Manhattan.

[11] Note that the cosine distance does not change if we use interest shares. As we will explain later, a group's interest share could be defined as either the share of the group's signals that pertain to the interest under consideration or as the share of the group's users that hold the interest under consideration.

[12] Although distances between countries tend to be much larger, there are cases where the distance between men of two different countries is smaller than the distance between men and women from the same country.

Figure 1: Gender Differences in Facebook Interests

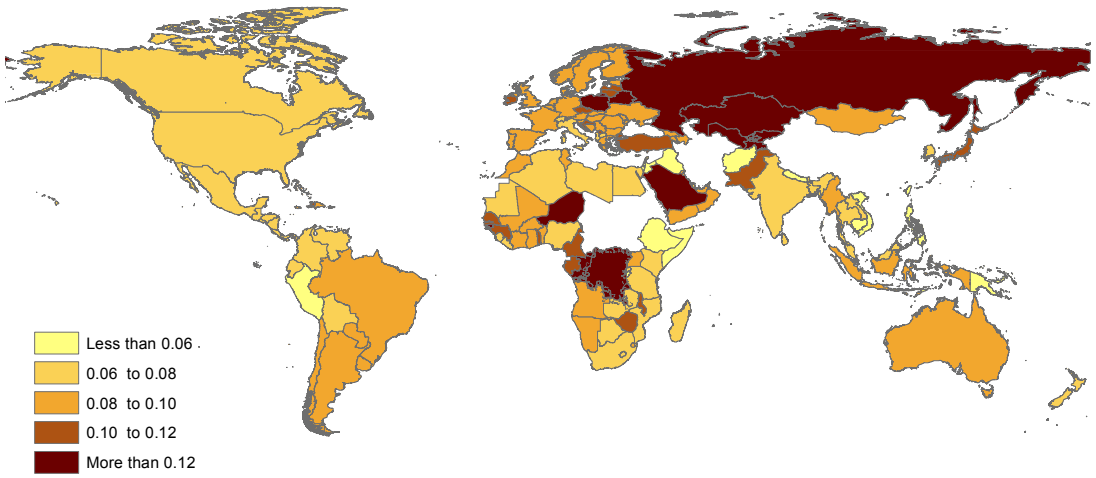


Figure shows the cosine distance between the interest frequency vector of men and women, based on 45,397 Facebook interests in countries with population above one million, Facebook penetration rate above 2.5% and number of Facebook users greater than 100,000.

shows a world map of the WEF gender equality index for the same sample of countries as Figure 1.

For robustness purposes, we also consider alternative indices, such as the UNDP's Gender Inequality Index and the OECD's Social Institutions and Gender Index. UNDP's Gender Inequality Index is similar to the World Economic Forum's GGI: it measures inequality in reproductive health, educational attainment, political empowerment and economic status. As for the OECD's Social Institutions and Gender Index, it aims to capture discrimination against women in formal and informal social institutions. More specifically, it measurers discrimination in the family, restricted physical integrity, restricted access to productive and financial resources, and restricted civil liberties. For instance, it includes measures of discrimination in divorce and inheritance laws, violence against women, genital mutilation, workplace rights, and access to justice and financial services. For reasons of comparison with the WEF index, we recode the UNDP and the OECD indices so that both are increasing in the degree of gender equality.

2.3 Other Control Variables

Other variables are likely to affect a country's gender differences in interests. In our baseline specification we include two additional control variables: the level of economic development and the overall diversity in interests. In other specifications, we add further controls.

Economic development. An increase in income reduces material constraints, allowing men and women to more freely express gender-specific desires, interests and ambitions (Falk and Hermle, 2018). As expected, economic development and gender equality are positively correlated (Fernández 2014; Cuberes and Teignier, 2014). However, that correlation is far from perfect, standing at 0.21 in our

Figure 2: Gender Equality

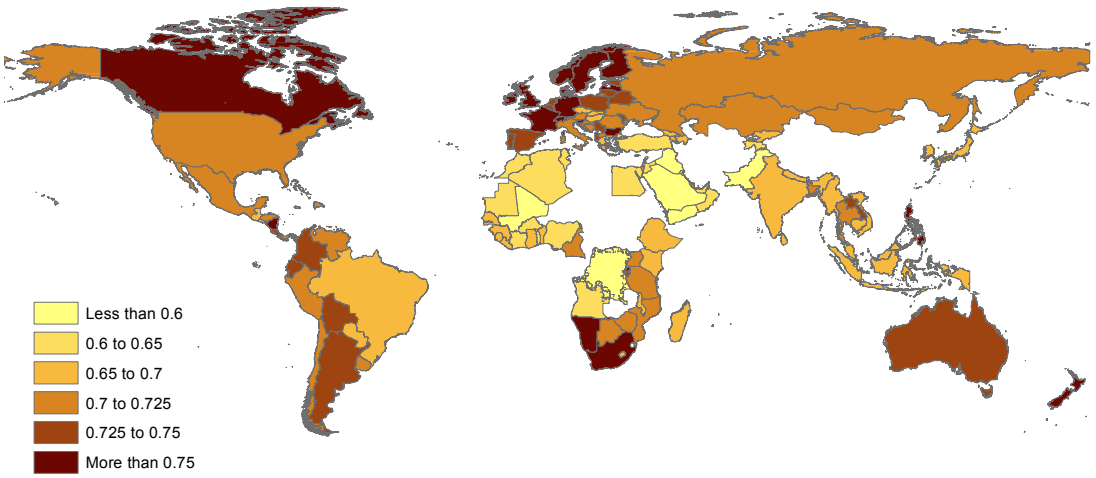


Figure depicts the 2018 Gender Gap Index of the World Economic Forum (WEF). The index is increasing in gender equality, and is based on a set of metrics related to economic opportunity, educational attainment, health outcomes and political empowerment. Sample of countries is the same as that in Figure 1.

baseline sample (and 0.09 in our expanded sample). There are poor countries with a high degree of gender equality, like Uganda, and rich countries with a low degree of gender equality, like Saudi Arabia.

Overall diversity. Bigger differences between men and women could partly reflect greater overall heterogeneity in society. Or on the contrary, more pluralistic countries in terms of interests might display smaller gender differences. To control for a country's overall diversity, we use the entropy index, given by $Ent_c = -\sum_{i=1}^{I} s_{ci} \log(s_{ci})$, where $s_{ci} = f_{ci} / \sum_i f_{ci}$ and f_{ci} is the number of individuals in country i who hold interest i . In the baseline sample of countries, the correlation between $CosDist_c$ and Ent_c is -0.25, implying that on average countries with greater gender differences in interests exhibit less overall heterogeneity in interests.

Other controls. In our robustness checks, we also consider more comprehensive specifications where we control for additional variables. First, we include regional dummies. This allows us to evaluate whether our results are mostly driven by differences between the world's large regions, or whether they also hold within regions. Second, we assess the importance of a country's religious composition. Religious beliefs affect gender norms and roles, and may hence be a confounding factor. Third, we control for the degree of Facebook penetration. Data from countries with low penetration rates may be less representative and reliable. Fourth, we check for the possible role of geographic and climatic variables in shaping gender preferences and norms.

3 Differences in Interests between Men and Women

This section explores the cross-country relation between the overall difference in interests between men and women and gender equality. In a first step, we do so by simply taking the cosine distance between men and women based on all 45,397 interests. One issue with this approach is that different interests may reflect the same latent preferences. To tackle this issue, in a second step we use singular value decomposition to reduce the dimensionality of interests. We then compute the gender difference based on the cosine distance between men and women in this lower-dimensional subspace.

3.1 All Interests

Raw correlation. Focusing on our baseline sample, Figure 3 provides a first look at the data by depicting the raw correlation between gender equality and the difference in interests between men and women, computed as the cosine distance between all 45,397 interests.[13] The correlation is slightly positive, suggesting that men and women in more gender-equal countries exhibit slightly larger differences in interests.

Figure 3: Gender Equality and Difference in Interests between Men and Women

Figure depicts the 2018 Gender Gap Index of the World Economic Forum (WEF) on the horizontal axis and the cosine distance between the vectors of 45,397 Facebook interest frequencies of men and women on the vertical axis for the baseline sample of countries (population > 1 million, Facebook penetration > 0.25 and Facebook users > 100,000).

[13] For countries with a population above one million and a Facebook penetration rate of at least 25%, we have Facebook data on 106 countries. Of those countries, 98 also have data on gender equality.

Partial correlation. To control for confounding determinants, we take a regression approach. Our baseline estimating equation is

$$CosDist_c = \beta GenderEq_c + \gamma Z_c + \varepsilon_c \tag{2}$$

where $CosDist_c$ is the cosine distance between the vectors of 45,397 interest frequencies of men and women in country i, $GenderEq_c$ is gender equality, Z_c is a vector of controls, and ε_c is an error term. Our main coefficient of interest is β, the partial correlation between gender equality and the difference in interests between women and men.

Table 1 reports the results for seven different specifications. Column (1) is our baseline specification: in addition to gender equality, it includes GDP per capita and the entropy of interests as regressors. As in Figure 3, we find a weak positive relation between gender equality and the difference in interests between men and women. However, the corresponding coefficient is not statistically significant at the 10% level. The other control variables show that economic development is associated with larger gender differences, whereas greater diversity in interests is associated with smaller gender differences.

Column (2) adds a set of regional dummies as controls. The coefficient on gender equality switches sign, but continues to be statistically insignificant. Column (3) explores the possibly confounding effects of religious composition and Soviet influence. This slightly strengthens the positive relation between gender equality and the difference in interests between men and women: the corresponding coefficient is now statistically significant at the 10% level. Having a greater percentage of Catholics or Muslims, or having been under Soviet influence, are all associated with greater differences in interests between men and women. Column (4) analyzes the effect of geography and climate on gender differences in interests. Higher agricultural land suitability is associated with larger differences in interests between men and women, whereas higher temperature is associated with smaller gender differences. When controlling for these geographic and climatic factors, the coefficient on gender equality once again switches sign and becomes negative, though its magnitude is not statistically different from zero. We observe the same absence of a statistically significant relation between a society's gender equality and its gender gap in interests when controlling for country size and Facebook penetration (column (5)). The last two regressions return to the most basic specification, but use alternative measures of gender equality from the OECD and the UNDP. There, we find a positive, statistically significant effect of gender equality on the gender interest gap (columns (6) and (7)).

Appendix Table B.2 considers the same seven specifications for different samples of countries. When expanding the sample to include countries with a population below one million, the results become slightly stronger. When changing the sample by setting the Facebook penetration threshold to 2.5%, the relation between gender equality and the gender interest gap is often negative, whereas when increasing the same threshold to 50% the results are mostly statistically insignificant. Appendix Table B.3 replaces the cosine distance between the interest vectors of men and women by either the Euclidean or the Manhattan distance. Here as well, the results tend to be statistically insignificant.

Our results so far are inconclusive. In the next subsection, we further investigate the relation between gender equality and the difference in interests between men and women by focusing on the latent structure of preferences.

Table 1: Gender Differences in Interests and Gender Equality: All Interests

Dependent Variable: Cosine Distance between Men and Women Based on 45,397 Facebook Interests							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Gender Equality (WEF)	0.050	-0.017	0.090*	-0.020	0.021		
	(0.035)	(0.042)	(0.051)	(0.045)	(0.042)		
Gender Equality (OECD)						0.001***	
						(0.000)	
Gender Equality (UNDP)							0.041**
							(0.019)
Log GDP per Capita	0.008***	0.006***	0.008***	0.005*	0.010***	0.006**	0.004
	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)	(0.002)
Entropy	-0.059***	-0.060***	-0.054***	-0.047**	-0.049**	-0.060***	-0.056***
	(0.021)	(0.017)	(0.016)	(0.019)	(0.021)	(0.019)	(0.018)
Sub-Saharan Africa		0.014					
		(0.009)					
Middle East and North Africa		0.011					
		(0.011)					
Europe and Central Asia		0.032***					
		(0.009)					
East Asia and Pacific		-0.002					
		(0.011)					
North America		0.022**					
		(0.011)					
Latin America and Caribbean		0.019**					
		(0.009)					
Share of Protestants			-0.001				
			(0.009)				
Share of Catholics			0.016***				
			(0.006)				
Share of Muslims			0.023***				
			(0.008)				
Soviet Influence			0.029***				
			(0.005)				
Log Area				0.003**			
				(0.001)			
Land Suitability				0.017*			
				(0.009)			
Terrain Roughness				-0.023			
				(0.016)			
Temperature				-0.001***			
				(0.000)			
Precipitation				-0.000			
				(0.000)			
Log Population					0.001		
					(0.002)		
Facebook Penetration					-0.034**		
					(0.015)		
Constant	0.490***	0.553***	0.401***	0.445***	0.423**	0.574***	0.549***
	(0.171)	(0.148)	(0.132)	(0.145)	(0.165)	(0.160)	(0.155)
Observations	98	98	95	91	98	84	100
R^2	0.234	0.494	0.523	0.475	0.286	0.329	0.278

Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The dependent variable is the difference between men and women based on 45,397 Facebook interests. The sample consists of countries with population > 1 million and Facebook penetration > 0.25.

3.2 Latent Preference Dimensions

Our distance measure is subject to several problems. One problem is "synonymy", the possibility that different interests reflect the same underlying preferences. For example, people interested in spaghetti and people interested in pasta should perhaps be classified as having common preferences. Failing to take this into account would tend to overestimate differences between populations. Another problem is "polysemy", the possibility that the same interest has different meanings or connotations for different populations. For example, people in favor of Trump and people opposed to Trump, though both types are interested in the same individual, should probably also be classified as having different preferences. Failing to take this into account would tend to underestimate differences between populations.

These two problems are well known from the text classification and information retrieval literature. In that field, each group is a text and each interest is a word, with each text being identified by its vector of word frequencies (Baeza-Yates and Ribeiro-Neto, 2011). Retrieval techniques that match queries to documents need to compute distances between documents and also suffer from synonymy and polysemy. The conventional methodology to deal with these issues is latent semantic indexing (LSI). It uses singular value decomposition (SVD), a method similar to principal component analysis, to create a lower-dimensional semantic space that places words that occur in similar documents close to one another (Deerwester et al., 1990).[14]

Applied to our problem, we use SVD to construct a lower-dimensional space that classifies interests held by populations with similar preferences as being closely related. Doing so allows us to address the problems of synonymy and polysemy, to get rid of noisy and redundant data, and to focus on the main associative patterns in the Facebook interest data.

Singular value decomposition. Consider the $I \times G$ interests-by-group matrix X, where the rows correspond to the I interests and the columns to the G country-gender groups.[15] Element x_{ig} of the matrix refers to the share of interest i in group g.[16] We denote the rank of matrix X by r, where $r \leq G$.[17] A well-known theorem of linear algebra says that X can be decomposed as

$$X = U\Sigma V^T \tag{3}$$

where U is an orthogonal $I \times G$ matrix, Σ is an $G \times G$ diagonal matrix, and V^T is an orthogonal $G \times G$ matrix.[18] The first r diagonal elements of Σ correspond to the square roots of the r non-zero eigenvalues of XX^T. They are referred to as the non-zero singular values and they are ordered such that $\sigma_1 \geq \sigma_2 \geq ... \geq \sigma_r$. The first r columns of U contain the orthonormal eigenvectors corresponding to the non-zero eigenvalues of XX^T. They are referred to as the left singular vectors. The first r

[14] Singular value decomposition maximizes the value of the second moment of the projections of the uncentered data, whereas principal component analysis maximizes the variance of the projected data. In our case, the two methods produce very similar results.

[15] Given that we have two genders, the number of country-gender groups is twice the number of countries, so $G = 2C$.

[16] More specifically, x_{ig} is defined as the share of signals expressed by group g that corresponds to interest i, i.e., $f_{gi}/\sum_i f_{gi}$, where f_{gi} is the number of users of group g that hold interest i. An alternative would be to define x_{ig} as the share of users in group g who are interested in i. We prefer the former measure because the number of interests per capita often differs substantially between genders within the same country.

[17] Typically, in our problem $r = G$.

[18] For an exposition, see, for example, Shores (2007).

columns of V contain the orthonormal eigenvectors corresponding to the non-zero eigenvalues of $X^T X$. They are referred to as the right singular vectors.

The goal of SVD is to discover the main latent or underlying preference dimensions. It may be useful to provide some intuition of how these dimensions are related to the matrix decomposition in (3). The columns of the $I \times G$ matrix U relate the different Facebook interests to each one of the latent preference dimensions. For example, the elements of the first column of U give the relative weights of each Facebook interest in the first preference dimension. The diagonal elements of the $G \times G$ matrix Σ then give a measure of the importance of each preference dimension. As they are declining in order, the first dimension is more important than the second, and so on. The columns of the $G \times G$ matrix V relate the different country-gender groups to each one of the latent preference dimensions. For example, the elements of the first column of V give the importance that each country-gender group attaches to the first preference dimension.

Dimensionality reduction and denoising. When computing distances between populations, it is useful to consider a reduced set of latent dimensions, rather than the full dimensionality of interests. By doing so, we focus on the main associative patterns in the Facebook interest data, while getting rid of noisy data and solving the issues of synonymy and polysemy.

To reduce the dimensionality of X to $\hat{r} < r$, we keep the first $\hat{r}$ singular values in Σ and their corresponding singular vectors in U and V. This yields

$$X_{\hat{r}} = U_{\hat{r}} \Sigma_{\hat{r}} V_{\hat{r}}^T \tag{4}$$

where $X_{\hat{r}}$ is an $I \times G$ matrix, $U_{\hat{r}}$ is an $I \times \hat{r}$ matrix, $\Sigma_{\hat{r}}$ is an $\hat{r} \times \hat{r}$ matrix, and $V_{\hat{r}}^T$ is an $\hat{r} \times G$ matrix. The matrix $X_{\hat{r}}$ is the best $\hat{r}$-rank approximation of X in the sense that it minimizes the sum of squared errors (Eckart and Young, 1936).

An important question is how to determine a reasonable value of $\hat{r}$. We use two different methods. A first consists of plotting the singular values in decreasing order, and keeping all singular values before there is a large drop in the plot. This ad-hoc approach is referred to as identifying an "elbow" in the curve of singular values. In our case, this method gives us a rank $\hat{r} = 9$. (Appendix Figure B.1 shows the plot of singular values.) A second is based on Gavish and Donoho (2014) who make certain assumptions on the noise structure of the data and then retain all singular values that are larger than those of the noise matrix. In our case, this gives us a rank $\hat{r} = 68$.[19] The matrix X_{68} has a correlation of 0.94 with the original matrix X. Thus, 68 dimensions give an excellent approximation of X. By reducing the rank to 9, we still get a very high correlation of 0.82 between X_9 and X.

Relation between gender equality and gender differences in preferences. For each one these two matrices, X_9 and X_{68}, we calculate the cosine distance between men and women by country and run the same regressions as before. Table 2 shows the results. When removing the noise and focusing on the main preference dimensions, we find a significantly stronger positive relation between gender equality and the difference in preferences between women and men. In most specifications, the coefficient on gender equality is now positive and statistically significant. In the most basic

[19] This number is based on equation (5) in Gavish and Donoho (2014).

Table 2: Gender Differences in Main Latent Preferences and Gender Equality

Panel A: Cosine Distance between Men and Women Based on First 9 Dimensions of SVD							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Gender Equality (WEF)	0.104***	0.012	0.088**	0.040	0.065*		
	(0.031)	(0.029)	(0.043)	(0.038)	(0.034)		
Gender Equality (OECD)						0.001***	
						(0.000)	
Gender Equality (UNDP)							0.078***
							(0.018)
Log GDP per Capita	0.006***	0.004***	0.007***	0.003	0.009***	0.004*	-0.001
	(0.002)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)
Entropy	-0.017	-0.015	-0.007	-0.004	-0.004	-0.015	-0.006
	(0.014)	(0.009)	(0.009)	(0.014)	(0.014)	(0.015)	(0.015)
Regional Dummies		Yes					
Religious Composition			Yes				
Geography & Climate				Yes			
FB Penetration					Yes		
Observations	98	98	95	91	98	81	97
R^2	0.254	0.666	0.612	0.519	0.335	0.387	0.305
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel B: Cosine Distance between Men and Women Based on First 68 Dimensions of SVD							
Gender Equality (WEF)	0.084**	0.027	0.126**	0.030	0.058		
	(0.038)	(0.043)	(0.059)	(0.057)	(0.046)		
Gender Equality (OECD)						0.001***	
						(0.000)	
Gender Equality (UNDP)							0.057***
							(0.020)
Log GDP per Capita	0.008***	0.006***	0.009***	0.005	0.010***	0.005	0.003
	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)
Entropy	-0.050*	-0.057***	-0.049***	-0.036	-0.041	-0.046	-0.043*
	(0.025)	(0.020)	(0.018)	(0.023)	(0.026)	(0.028)	(0.025)
Regional Dummies		Yes					
Religious Composition			Yes				
Geography & Climate				Yes			
FB Penetration					Yes		
Observations	98	98	95	91	98	81	97
R^2	0.218	0.545	0.566	0.400	0.244	0.320	0.240

Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The dependent variable is the difference between men and women based on either the first 9 dimensions (Panel A) or the first 68 dimensions (Panel B) of the singular value decomposition of 45,397 Facebook interests. The sample consists of countries with population > 1 million and Facebook penetration > 0.25. The seven specifications are identical to those in Table 1. All regressions include a constant. Religious composition refers to share of protestants, catholics and muslims, as well as a dummy for Soviet influence; geography & climate refers to log area, agricultural land suitability, terrain roughness, temperature and precipitation; continents refer to continental dummies; and FB penetration refers to Facebook penetration and log of population.

specification in column (1), the standardized β is 30% when taking 9 dimensions, meaning that a one standard deviation increase in gender equality increases the difference in preferences between men and women by 30% of its standard deviation. The corresponding standardized β when taking 68 dimensions is 19%. Overall, the evidence points to a positive relation between gender equality and the difference in interests between men and women. Next, we further analyze this relation by differentiating between gender-related and non-gender-related interests.

4 Gender and Non-Gender Interests and Preferences

In this section, we start by classifying all 45,397 interests into two groups, those that are systematically related to gender and those that are not. We then analyze the relation between gender equality and gender differences for each of these two groups of interests. Next, we use singular value decomposition as an alternative way to identify which preference dimensions are gender-related and which are not. We then analyze how the relation between gender equality and gender differences in preferences depends on that distinction.

4.1 Gender and Non-Gender Interests

Some examples. To illustrate our approach, take the interest of Facebook users in engineering or biology. There is a clear gender bias: in almost all countries, more men are interested in engineering, and more women are interested in biology. For both of these interests, the difference between men and women is larger in countries that are more gender-equal. Contrast this with the interest of Facebook users in mathematics or popular music. There is no longer a gender bias: in some countries, men are more interested in mathematics, and in others women. The same is true for popular music. More importantly, for both of these interests, the difference between men and women is now smaller in countries that are more gender-equal. These examples suggest that the relation between gender equality and the gender gap in interests and preferences depends on the type of interest. Next, we evaluate whether this insight generalizes when looking at all interests.

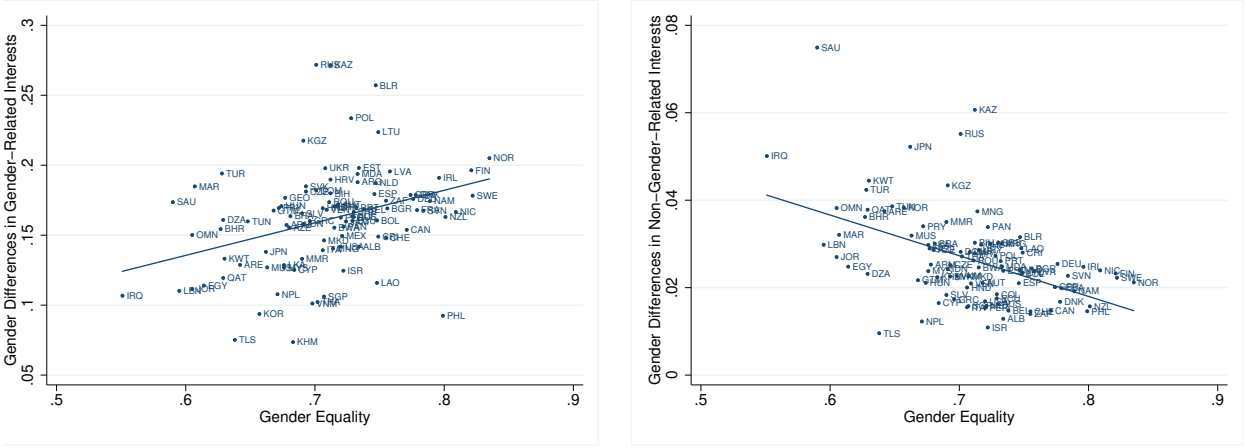
Distinguishing between gender and non-gender interests. Starting with all 45,397 interests, we define two subsets of interests, one that is gender-related and one that is non-gender-related. We call an interest gender-related if in more than 90% of countries the interest is more frequent among one of the genders. Examples include "engineering", "fatherhood", "romantic comedies", "hunting" and "baking". We call an interest non-gender-related if in at least 30% of countries the interest is more frequent among men and in at least another 30% of countries it is more frequent among women. Examples include "language school", "blood donation" and "positive attitude quotes". This procedure yields 2,685 gender-related interests and 8,755 non-gender-related interests.[20]

For each country, we compute two cosine distances between men and women, one based on the set of gender-related interests and another based on the set of non-gender-related interests. Figure

[20]To have a sufficiently broad cross-section of countries, we apply the procedure to the sample of 131 countries with a population $>$ 1 million, Facebook penetration $>$ 2.5% and Facebook users $>$ 100,000, for which we have data on gender equality from the World Economic Forum.

4 depicts the raw correlation between gender equality and each of these two distance measures. The difference is immediately apparent: greater gender equality is associated with larger differences in gender-related interests (Panel A), but smaller differences in non-gender-related interests (Panel B). Table 3 does a more in-depth analysis of these relations, based on the same seven regressions as before. When comparing our findings for gender-related interests in Panel A to those for non-gender-related interests in Panel B, we observe the same stark difference. For gender-related interests, there tends to be a strong positive association between gender equality and differences in interests between men and women, whereas for non-gender-related interests, there tends to be a strong negative association between the two. That is, gender-related interests diverge in more gender-equal societies, whereas non-gender-related interests converge in more gender-equal societies. The magnitudes of the effects are large: in the most basic specification in column (1), the standardized β corresponding to gender equality is 35% in the case of gender-related interests, and -46% in the case of non-gender-related interests.

Figure 4: Gender Equality and Differences in Gender-Related vs Non-Gender-Related Interests



a. Gender Interests

b. Non-Gender Interests

Figure depicts a scatter plot of gender equality against differences in interests between men and women. Panel A shows the differences between men and women in gender-related interests, whereas Panel B shows the differences between men and women in non-gender-related interests.

Relation to evolutionary psychology and social role theory. The debate on whether greater gender equality should enhance or mitigate preference differences between men and women has often been framed in terms of two competing theories. Evolutionary psychology argues that more gender equality allows men and women to more freely express their innate predispositions, leading to widening preference differences. In contrast, social role theory claims that more gender equality allows breaking down socially constructed barriers between men and women, leading to narrowing preference differences. This suggests that there is no reason why there should be a one-size-fit-all theory. Instead, evolutionary psychology should apply predominantly to innate preferences or interests, whereas social role theory should apply predominantly to socially constructed preferences or interests.

Table 3: Differences between Men and Women in Gender-Specific vs Non-Gender-Specific Interests

Panel A: Cosine Distance between Men and Women Based on Gender-Specific Interests							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Gender Equality (WEF)	0.232***	0.126**	0.203**	0.137*	0.168**		
	(0.061)	(0.062)	(0.097)	(0.079)	(0.072)		
Gender Equality (OECD)						0.001***	
						(0.000)	
Gender Equality (UNDP)							0.056
							(0.040)
Log GDP per capita 2000-2017	0.006	0.003	0.006*	0.000	0.010**	0.004	0.003
	(0.004)	(0.003)	(0.003)	(0.004)	(0.005)	(0.005)	(0.005)
Entropy	-0.048	-0.065**	-0.047*	-0.026	-0.028	-0.059*	-0.052
	(0.034)	(0.026)	(0.026)	(0.031)	(0.034)	(0.034)	(0.032)
Regional Dummies		Yes					
Religious Composition			Yes				
Geography & Climate				Yes			
FB Penetration					Yes		
Observations	98	98	95	91	98	84	100
R^2	0.180	0.563	0.518	0.397	0.249	0.249	0.110
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel B: Cosine Distance between Men and Women Based on Non-Gender-Specific Interests							
Gender Equality (WEF)	-0.089***	-0.079***	-0.054**	-0.112***	-0.082***		
	(0.020)	(0.024)	(0.021)	(0.019)	(0.020)		
Gender Equality (OECD)						-0.000**	
						(0.000)	
Gender Equality (UNDP)							-0.033***
							(0.010)
Log GDP per Capita	0.003***	0.003***	0.003***	0.003***	0.003***	0.003***	0.005***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Entropy	-0.030***	-0.033***	-0.035***	-0.030***	-0.032***	-0.037***	-0.039***
	(0.006)	(0.007)	(0.006)	(0.007)	(0.007)	(0.007)	(0.006)
Regional Dummies		Yes					
Religious Composition			Yes				
Geography & Climate				Yes			
FB Penetration					Yes		
Observations	98	98	95	91	98	84	100
R^2	0.436	0.474	0.548	0.557	0.452	0.489	0.403

Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The dependent variable is the difference between men and women based on either the subset of interests that are more frequent in one of the genders in at least 90% of countries (Panel A) or on the subset of interests that are more frequent in men at least 30% of countries and more frequent in women in at least 30% of countries (Panel B). The sample consists of countries with population > 1 million and Facebook penetration > 0.25. The seven specifications are identical to those in Table 1. All regressions include a constant. Religious composition refers to share of protestants, catholics and muslims, as well as a dummy for Soviet influence; geography & climate refers to log area, agricultural land suitability, terrain roughness, temperature and precipitation; continents refer to continental dummies; and FB penetration refers to Facebook penetration and log of population.

For an interest to be innate to gender, we claim that it must display a systematic bias toward the same gender across the globe. Viewing this as a necessary condition, we can interpret our gender-related interests as potentially innate. By the same token, interests that do not display such a systematic bias cannot be innate to gender. Hence, non-gender-related interests must be socially constructed. This argument provides us with a mapping from gender-related and non-gender-related interests into innate and socially constructed interests. Using this mapping, the paper's main finding is consistent with both theories: more gender-equal societies display greater differences between men and women in gender-related (innate) interests and smaller differences in non-gender-related (socially constructed) interests. In the remainder of the paper, we explore the robustness of the paper's main empirical result by considering alternative ways of identifying gender and non-gender interests and preferences.

Different ways of classifying gender and non-gender interests. For an interest to be classified as gender-related, we required it to have a common gender bias in at least 90% of countries. When rerunning the specification in column (1) of Table 3 for 25 different thresholds between 70% to 95%, the effect of gender equality is always positive and statistically significant at the 1% level. For the case of non-gender-related interests, we required the interest to be more frequent among men in at least 30% and more frequent among women in at least another 30% of countries. When varying the threshold from 10% to 45%, the effect of gender equality is always negative and statistically significant at the 1% level.[21] From this we conclude that our results are robust to less and more strict ways of classifying gender and non-gender interests.

Dimensionality reduction. In the same way that we used SVD to reduce the dimension of our overall interests-by-group matrix to focus on the main latent preference dimensions, we can use the same procedure on the subset of gender interests and on the subset of non-gender interests. Using the ad-hoc "elbow" method yields a rank $\hat{r} = 8$ for both matrices. We refer to the truncated gender-related and non-gender-related matrices as, respectively, X_8^g and X_8^{ng}. For each one of these two matrices, we recompute the cosine distances between men and women and re-run the same regressions as before. Table 4 reports our findings. We observe the same stark difference: more gender-equal societies exhibit larger differences between men and women along gender dimensions and smaller differences along non-gender dimensions.

Gender and non-gender interests based on sample of least gender-equal countries. One potential concern is that the subsets of interests that experience either a widening or a narrowing gender gap in more gender-equal societies might consist of random interests that ex post get classified as gender and non-gender interests. To illustrate this concern, consider the following hypothetical example. Suppose that in the past, when countries were less gender-equal, the gender gap for most interests was idiosyncratic across countries. In that case, almost all interests would have been classified as non-gender-related. Then, as some countries became more gender-equal, suppose that for a random

[21] It is important to note that there are many zeros in our interest matrix, implying that for many interests the difference between men and women is zero. This implies that there is almost no overlap in non-gender interests with a weak threshold and gender interests with a weak threshold.

Table 4: Gender Differences Based on Main Latent Gender and Non-Gender Dimensions

Panel A: Cosine Distance between Men and Women Based on SVD of Gender Interests							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Gender Equality (WEF)	0.252***	0.125**	0.178*	0.170**	0.180***		
	(0.057)	(0.054)	(0.090)	(0.075)	(0.066)		
Gender Equality (OECD)						0.002***	
						(0.000)	
Gender Equality (UNDP)							0.110***
							(0.035)
Log GDP per Capita	0.006*	0.003	0.006**	0.000	0.010**	0.002	-0.003
	(0.003)	(0.003)	(0.003)	(0.004)	(0.004)	(0.004)	(0.005)
Entropy	-0.006	-0.016	0.001	0.017	0.019	-0.001	0.010
	(0.027)	(0.017)	(0.019)	(0.026)	(0.027)	(0.030)	(0.028)
Regional Dummies		Yes					
Religious Composition			Yes				
Geography & Climate				Yes			
FB Penetration					Yes		
Observations	98	98	95	91	98	81	97
R^2	0.236	0.677	0.588	0.439	0.320	0.338	0.162
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel B: Cosine Distance between Men and Women Based on SVD of Non-Gender Interests							
Gender Equality (WEF)	-0.024***	-0.018**	-0.019***	-0.015**	-0.019**		
	(0.008)	(0.008)	(0.006)	(0.006)	(0.008)		
Gender Equality (OECD)						-0.000	
						(0.000)	
Gender Equality (UNDP)							-0.014***
							(0.003)
Log GDP per Capita	0.001***	0.001***	0.001***	0.001***	0.001**	0.000	0.002***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Entropy	-0.003**	-0.006***	-0.004***	-0.006***	-0.005***	-0.002	-0.006***
	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)	(0.002)
Regional Dummies		Yes					
Religious Composition			Yes				
Geography & Climate				Yes			
FB Penetration					Yes		
Observations	98	95	91	98	98	81	97
R^2	0.242	0.313	0.377	0.417	0.296	0.051	0.225

Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The dependent variable is the difference between men and women based on the first eight dimensions of SVD on either subset of interests that are more frequent in one of the genders in at least 90% of countries (Panel A) or subset of interests that are more frequent in men at least 30% of countries and more frequent in women in at least 30% of countries (Panel B). The sample consists of countries with population $>$ 1 million and Facebook penetration $>$ 0.25. The seven specifications are identical to those in Table 1. All regressions include a constant. Religious composition refers to share of protestants, catholics and muslims, as well as a dummy for Soviet influence; geography & climate refers to log area, agricultural land suitability, terrain roughness, temperature and precipitation; continents refer to continental dummies; and FB penetration refers to Facebook penetration and log of population.

subset of interests the gender gap widened in those countries. In that case, we would ex post have classified those interests as gender-related. If this were an important driver of our findings, we would expect our results to no longer hold when using a classification based on the past gender gap. While we do not have historical data, we can use as proxy a classification of gender and non-gender interests based on the subset of least gender-equal countries, since arguably gender equality has advanced less in those countries.

As a robustness check, we therefore take the interest frequencies by gender for countries with gender equality below the median, and classify them into gender-related and non-gender-related interests. For all countries, we recompute the cosine distances between men and women for those two groups of interests. Using both distance measures, we re-run the same regressions as before. Appendix Table B.4 shows that the results are unchanged: more gender-equal countries are associated with a wider gap between men and women for gender-related interests, and a narrower gap for non-gender-related interests.

Causality. So far we have mostly refrained from using causal language. A society's gender equality is potentially endogenous because of reverse causality: differences in preferences between genders may affect the degree of equality between men and women. It is not obvious in which direction this potential endogeneity would bias our coefficients. On the one hand, if men and women want different things from life, this might translate in less gender equality in certain outcomes. This would increase the coefficient on gender equality, hence strengthening our findings for gender interests and weakening them for non-gender interests. On the other hand, if men and women have different preferences, there may be more pressure for women's rights and female political empowerment, leading to greater gender equality. This would decrease the coefficient on gender equality, hence weakening our findings for gender interests and strengthening them for non-gender interests.

To address this potential endogeneity concern, we take two approaches. In a first approach. we use the earliest available version of our gender equality index. As such, in our baseline specification we replace the gender equality index of 2018 by the one of 2006. The idea is that there is less likely to be a reverse causality issue between today's differences in preferences and the gender equality index of almost 15 years ago. Columns (2) and (4) in Table 5 report our findings for gender interests and non-gender interests. When comparing to the baseline regressions reported in columns (1) and (4), there is no significant difference in the coefficients on gender equality. This somewhat allays concerns about reverse causality. Needless to say, to the extent that the unobservable factors that led to the possible identification problem in the first place are correlated over time, reverse causality is still an issue.

In a second approach, we turn to instrument variable estimation. We use the year when women gained the right to vote as an instrument for today's degree of gender equality. Since constructing gender equality through the political process takes many years, the time elapsed since female suffrage is bound to be a good predictor of today's gender equality. How long ago women gained the vote is of course likely to affect today's differences in preferences between men and women. We would expect this effect to be mediated by the degree of female political empowerment and acquired economic, social and economic rights and opportunities. Since all these mediating factors are captured by the

gender equality index we use, the exclusion restriction is likely to be satisfied.

Columns (3) and (6) in Table 5 report our findings based on IV estimation. The coefficients on gender equality are slightly larger in absolute value terms when using IV than when using OLS. In addition, the F-statistics of the first stage are larger than the Stock-Yogo critical values for 10% maximal IV size, so we can reject the hypothesis that our instrument is weak. Overall, these findings suggest that we can give a causal interpretation to our main result: more gender equality leads to larger differences between men and women in gender-specific interests and smaller differences in non-gender-specific interests. However, we must be cautious with this interpretation, because this result is based on our baseline specification. When considering more comprehensive specifications, our IV strategy ceases to pass the weak instrument test.

Table 5: Gender and Non-Gender Interests: Causality

	Cosine Distance Men - Women					
	Gender Interests			Non-Gender Interests		
	(1) OLS	(2) Lagged	(3) IV	(4) OLS	(5) Lagged	(6) IV
Gender Equality (WEF)	0.232***		0.269**	-0.089***		-0.180***
	(0.061)		(0.123)	(0.020)		(0.061)
Gender Equality (WEF, 2006)		0.202**			-0.082***	
		(0.081)			(0.030)	
Log GDP per Capita	0.006	0.004	0.006	0.003***	0.003***	0.004***
	(0.004)	(0.005)	(0.004)	(0.001)	(0.001)	(0.001)
Entropy	-0.048	-0.060*	-0.049	-0.030***	-0.039***	-0.027***
	(0.034)	(0.035)	(0.034)	(0.006)	(0.007)	(0.007)
Constant	0.366	0.520*	0.352	0.322***	0.395***	0.357***
	(0.277)	(0.305)	(0.263)	(0.055)	(0.064)	(0.062)
Observations	98	86	98	98	86	98
R^2	0.180	0.173	0.177	0.436	0.450	0.232
Cragg-Donald F			18.05			18.05
Stock-Yogo 10% max IV size			16.38			16.38

Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The dependent variable is the difference between men and women based on either the subset of interests that are more frequent in one of the genders in at least 90% of countries (columns (1)-(3)) or on the subset of interests that are more frequent in men at least 30% of countries and more frequent in women in at least 30% of countries (columns (1)-(3)). The sample consists of countries with population $>$ 1 million and Facebook penetration $>$ 0.25. Columns (1) and (4) are identical to column (1) in Table 1. Columns (2) and (5) use the gender equality index of the WEF of 2006. Columns (3) and (6) are based on IV regressions, using the year when women obtained the right to vote as instrument of the gender equality index of the WEF.

4.2 Gender and Non-Gender Preference Dimensions

Rather than classifying interests as gender-related or non-gender-related, in this subsection we take the main preference dimensions identified by SVD and classify them as gender-related or non-gender-related. We then analyze whether there are systematic differences in the relation between gender equality and gender differences in preferences depending on whether the preference dimension is gender-specific or not.

Identifying gender and non-gender preference dimensions. Along each of the latent preference dimensions identified by singular value decomposition, we can position men and women of different countries. Starting off with the $I \times G$ interest by country-gender matrix X of rank r, equation (3) shows the singular value decomposition $X = U\Sigma V^T$. The matrix ΣV^T places country-gender groups in the vector space of rank r. More specifically, the non-zero first r rows of the $G \times G$ matrix

ΣV^T give the positions of the country-gender groups along each one of the r interest dimensions. For example, the elements of the first row give the positions of men and women in different countries along the first interest dimension. The position of country-gender group g along preference dimension i can be written as $\sigma_i v_{ig}^T$, where v_{ig}^T is the element corresponding to row i and column g of matrix V^T.

To visualize the relative positions of men and women in the different countries, Figure 5 displays two-dimensional scatter plots for each one of the first nine preference dimensions, with the position of women on the horizontal axis and the position of men on the vertical axis. Consider, for example, the scatter plot that depicts the preference dimension associated with the second singular vector V_2. Each point corresponds to one country, and gives the position of women in that country on the horizontal axis and the position of men in that country on the vertical axis. Points that are above the 45° line refer to countries where the position of men along preference dimension 2 is higher than that of women.

To distinguish between gender-related preference dimensions and non-gender preference dimensions, we start with a visual inspection of the different panels of Figure 5. Of the different dimensions, the one associated with singular vector V_4 displays the strongest gender component: independently of country, women have a positive value while men have a negative value. Along that dimension women of different countries tend to be more similar to each other than to men of their own country. To further illustrate how V_4 captures a dimension along which men's and women's interests are very different, we can multiply the fourth left singular vector U_4 by the fourth singular value σ_4 to obtain the position of each one of the 45,397 different interests along the fourth preference dimension. The interests with lower values correspond to "masculine" interests, and the ones with higher values to "feminine" values. Among the most masculine interests, many relate to cars and sports, and among the most feminine interests, many relate to cooking, shopping and family.[22] The dimensions associated with singular vectors V_2 and V_5 also display a gender bias: men either have systematically higher values than women (V_2), or the other way around (V_5). However, in contrast to V_4, along V_2 and V_5 there continues to be an important country component: women of a particular country tend to be closer to men of their own country than to women in other countries, although in each country men and women are systematically different. The other dimensions V_3, V_6, V_7 and V_8 do not show a clear gender bias, and can be considered to be mostly unrelated to gender. For these dimensions, some points are above and others are below the 45° line.[23]

To more formally identify gender preferences and non-gender preferences, for each dimension we

[22]For a full list of the 500 most masculine and the 500 most feminine interests along preference dimension 4, see Appendix Tables B.5 and B.6. Examples of the most masculine interests include Automobiles, BMW, Motorcycles, Personal finance, War, Vladimir Putin, Game Consoles, Free Software, Engine, SUVs, Cameras, Outdoor recreation, UEFA Champions League, Lionel Messi, Sport cars, Wheel, Bluetooth, Martial arts, Hunting, Military, Tool, Poker, Shooter games, Computer monitors. Examples of the most feminine interests include Dresses, Cosmetics, Infant, Motherhood, Poetry, Beauty salons, Pregnancy, Boutiques, Child, Cooking, Cake, Chocolate, Jewelry, Handbags, Blouse, Hairstyle, Weddings, Recipes, Make-up artist, Skirt, Cuisine, Skin, Flower, Childbirth, Wedding dress, Weight loss, Psychology, Yoga, Breastfeeding. Male interests appear much more than female interests in the Facebook categories of Hobbies and activities, Technology, and Sports and Outdoors. Female interests appear much more than male interests in the categories of Food and Drink, Shopping and Fashion, and Family and Relationships.

[23]In our description of the different dimensions, we did not mention V_1. Along that dimension, all countries and genders present very similar values. This is the dimension that captures the mean positions. It can also be interpreted as the dimension that captures the preferences common to all groups. In principal component analysis, this dimension is absent, because of data normalization. In the rest of the analysis, we will ignore V_1.

Figure 5: Positions of Women and Men along Main Preference Dimensions

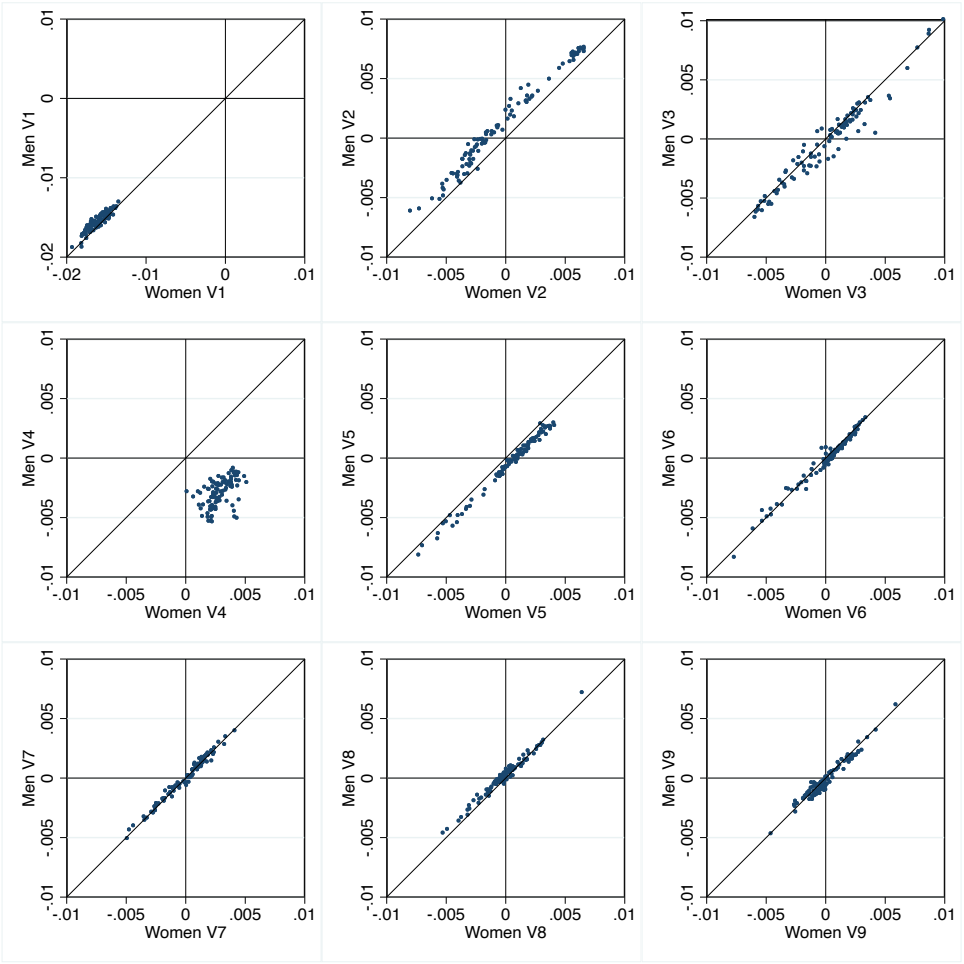


Figure shows the positions of women and men in the different countries along the nine most important preference dimensions as determined by SVD.

compute the incremental R^2 from adding gender to a regression of the positions of men and women on a full set of country dummies. The greater the explanatory power of gender, the larger the incremental R^2. This methodology confirms our visual inspection of Figure 5. Of the different dimensions, the incremental R^2 due to gender for V_4 is 88%. For V_2 and V_5, the incremental R^2 is between 2% and 4%, and for all other dimensions it is below 1%. We can therefore conclude that dimension V_4 is gender-related, dimensions V_2 and V_5 are weakly gender-related, and dimensions V_3, V_6, V_7, V_8 and V_9 are non-gender-related.

Gender differences along gender and non-gender preference dimensions. Next, we analyze whether there is a difference in the relation between gender equality and gender differences in preferences depending on whether the preference dimension is gender-specific or not. We compute the Euclidean distance between men and women based on gender and non-gender dimensions, and re-run

the same seven regressions as before.[24] Table 6 reports the results. Panels A and B focus on gender dimensions (taking either a strict definition, based on V_4, or a more lenient definition, also including V_2 and V_5), whereas Panel C focuses on non-gender dimensions. Once again, we confirm the paper's main finding. The coefficients on gender equality tend to be positive and statistically significant in Panels A and B, whereas they tend to be negative and statistically significant in Panel C. Hence, more gender-equal societies exhibit greater differences between men and women for gender-specific preferences and smaller differences between men and women for non-gender-specific preferences.

4.3 Young and Old

Given that our analysis focuses on the subset of countries with a Facebook penetration rate of at least 25%, we are fairly confident that our data are broadly representative of the population groups that we are interested in. However, some biases may persist even when reaching relatively high levels of Facebook penetration. Probably, the one that should concern us most is the age bias, since social media users tend to be younger than the overall population. This would not be too much of a concern if the age bias were the same in all countries. In that case, cross-country comparisons would still be valid, though they would disproportionately reflect the preferences of the young. However, there is substantial variation in the age bias across countries. Accounting for this bias is important because age may be a determinant of the differences in interests between men and women. For example, if older men and women are more similar than younger men and women, then the coefficient on gender equality would be biased downward if more gender-equal countries have a larger proportion of older Facebook users.

One way to address this concern is to re-run the regressions of Table 3, controlling for the ratio of older to younger Facebook users. As cutoff between the two groups, we take an age of 40 years. As can be seen in Appendix Table B.7, the results are unchanged. Controlling for the age ratio, in more gender-equal countries the difference between men and women is larger for gender-specific interests and smaller for non-gender-specific interests.

Another way to address this concern is to run separate regressions for the old and the young. This requires us to have interest frequency data by age group. Unfortunately, getting such data for all 45,397 interests would be extremely time-consuming, and goes beyond the scope of this paper. However, for 5,000 randomly chosen interests, we obtained frequency data by country for both the old (age above 40) and the young (age 40 and below). Using the same definitions as before, we identify which of these 5,000 interests are gender-specific and which are not. We then compute four distance measures: the distance between old men and old women for gender-specific interests; the distance between young men and young women for gender-specific interests; and analogous measures for the old and the young applied to non-gender-specific interests. For each one of these distance measures, we run our standard set of regressions. Appendix Tables B.8 and B.9 report the results. Two findings stand out. First, the results for the old are almost identical to the results for the young, suggesting that age composition is not material to the paper's findings. Second, we confirm our central result

[24] We use the Euclidean distance, rather than the cosine distance, because in some cases the distance is based on just one dimension. In the cases for which the distance is based on more than one dimension, using the cosine distance yields qualitatively very similar results.

Table 6: Gender and Non-Gender Dimensions Based on SVD

Panel A: Euclidean Distance between Men and Women Based on Gender Dimension V_4							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Gender Equality (WEF)	0.005***	0.002	0.005**	0.003	0.003*		
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)		
Gender Equality (OECD)						0.000***	
						(0.000)	
Gender Equality (UNDP)							0.004***
							(0.001)
Log GDP per Capita	0.000***	0.000**	0.000***	0.000	0.000***	0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Entropy	-0.003***	-0.003***	-0.003***	-0.003***	-0.002***	-0.003***	-0.003***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Regional Dummies		Yes					
Religious Composition			Yes				
Geography & Climate				Yes			
FB Penetration					Yes		
Observations	98	98	95	91	98	81	97
R^2	0.283	0.651	0.653	0.522	0.340	0.421	0.345
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel B: Euclidean Distance between Men and Women Based on Gender Dimension V_2, V_4 and V_5							
Gender Equality (WEF)	0.006***	0.002	0.006**	0.004	0.005**		
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)		
Gender Equality (OECD)						0.000***	
						(0.000)	
Gender Equality (UNDP)							0.004***
							(0.001)
Log GDP per Capita	0.000***	0.000**	0.000***	0.000	0.000***	0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Entropy	-0.003***	-0.003***	-0.003***	-0.003***	-0.003***	-0.003***	-0.003***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Regional Dummies		Yes					
Religious Composition			Yes				
Geography & Climate				Yes			
FB Penetration					Yes		
Observations	98	98	95	91	98	81	97
R^2	0.311	0.672	0.670	0.548	0.372	0.435	0.355
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel C: Euclidean Distance between Men and Women Based on Non-Gender Dimension V_3, V_6, V_7, V_8 and V_9							
Gender Equality (WEF)	-0.005***	-0.003***	-0.004***	-0.004***	-0.004***		
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)		
Gender Equality (OECD)						-0.000*	
						(0.000)	
Gender Equality (UNDP)							-0.002***
							(0.000)
Log GDP per Capita	0.000***	0.000***	0.000***	0.000***	0.000*	0.000	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Entropy	-0.001**	-0.001***	-0.001***	-0.001***	-0.001***	-0.001	-0.001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Regional Dummies		Yes					
Religious Composition			Yes				
Geography & Climate				Yes			
FB Penetration					Yes		
Observations	98	98	95	91	98	81	97
R^2	0.326	0.531	0.461	0.493	0.400	0.125	0.246

Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The dependent variable is the Euclidean distance between men and women based on V_4 (Panel A), V_2, V_4 and V_5 (Panel B) and V_3, V_6, V_7, V_8 and V_9 (Panel C). The sample consists of countries with population > 1 million and Facebook penetration > 0.25. The seven specifications are identical to those in Table 1. All regressions include a constant. Religious composition refers to share of protestants, catholics and muslims, as well as a dummy for Soviet influence; geography & climate refers to log area, agricultural land suitability, terrain roughness, temperature and precipitation; continents refer to continental dummies; and FB penetration refers to Facebook penetration and log of population.

for both the old and the young: as gender equality increases, men and women tend to diverge in gender-specific interests, and they tend to converge in non-gender-specific interests.

5 Conclusions

This paper used information on the frequency of 45,397 Facebook interests to study how the difference in preferences between men and women changes with a country's degree of gender equality. The paper's main finding is that for interests or preferences that are gender-related, we observe a larger gender gap in more gender-equal countries, whereas the opposite is true for interests or preferences that non-gender-related.

We established the paper's central finding by using many different ways of classifying interests and preferences. First, we split up all 45,397 interests into gender and non-gender interests, using a criterion that interests that are more frequent among the same gender for almost all countries get classified as innate to gender. Second, we experimented with more stringent and more lenient thresholds when classifying interests as related to gender or not. Third, we considered an alternative classification of gender and non-gender interests based on the subset of least gender-equal countries to alleviate concerns that our classification might be tautological. Fourth, we used singular value decomposition on both subsets of interests to focus on the relevant latent dimensions. Fifth, we also used singular value decomposition on all interests, to then classify the resulting latent preference dimensions as related to gender or not. We found our paper's main result to be robust to these different ways of distinguishing between gender and non-gender interests and preferences.

By interpreting gender-related preferences as being potentially innate and non-gender-related preferences as being socially constructed, we argued that the paper's main finding can help resolve a long-standing debate between evolutionary psychology and social role theory. Indeed, our evidence confirms the predictions of both theories: for gender-related (innate) preferences, our results point towards evolutionary psychology, whereas for non-gender-related (socially constructed) preferences, our results point towards social role theory.

In closing, it is important to reiterate a word of caution. While we pursued different ways of identifying preferences and interests that are gender-related, they fall short of direct proof of any interest actually being innate to gender. Rather, we explored the heterogeneity in the relation between a country's degree of gender equality and the difference in preferences between men and women, and found that it is consistent with an interpretation that reconciles evolutionary psychology and social role theory.

References

[1] Akerlof, G. and R. E. Kranton (2000). "Economics and Identity," *Quarterly Journal of Economics*, 115(3): 715-753.

[2] Alesina, A., Giuliano, P. and N. Nunn (2013). "On the Origins of Gender Roles: Women and the Plough," *Quarterly Journal of Economics*, 128(2): 469-530.

[3] Archer, J. (2006). "Testosterone and Human Behavior: An Evaluation of the Challenge Hypothesis," *Neuroscience and Biobehavioral Reviews*, 30(3): 319-345.

[4] Atari, M., Lai, M.H.C. and M. Dehghani (2020). "Sex Differences in Moral Judgements across 67 Countries," *Proceedings of the Royal Society B: Biological Sciences*, 287: 20201201.

[5] Baeza-Yates, R. and B. Ribeiro-Neto (2011). *Modern Information Retrieval. The Concepts and Technology behind Search*, Essex: Pearson Education Limited.

[6] Berenbaum, S. and A.M. Beltz (2021). "Evidence and Implications from a Natural Experiment of Prenatal Androgen Effects on Gendered Behavior," *Current Directions in Psychological Science*, 30: 202-210.

[7] Bertrand, M. (2011). "New Perspectives on Gender," in: Ashenfelter, O. and D. Card (eds.), *Handbook of Labor Economics*, edition 1, volume 4, chapter 17, 1543-1590, Elsevier.

[8] Boyd, R. and P.J. Richerson (2005). *The Origin and Evolution of Culture.* Oxford, U.K.: Oxford University Press.

[9] Breda, T., Jouini, E., Napp, C. and G. Thebault (2020). "Gender Stereotypes Can explain the Gender-Equality Paradox," *Proceedings of the National Academy of Sciences*, 117(49): 31063-31069.

[10] Buss, D. (1989). "Sex Differences in Human Mate Preferences: Evolutionary Hypotheses Tested in 37 Cultures," *Behavioral and Brain Sciences*, 12(1): 1-14.

[11] Clots-Figueras, I. (2012). "Are Female Leaders Good for Education? Evidence from India," *American Economic Journal: Applied Economics,* 4(1): 212-244.

[12] Costa, P.T., Terracciano, A., and R.R. McCrae (2001). "Gender Differences in Personality Traits across Cultures: Robust and Surprising Findings," *Journal of Personality and Social Psychology*, 81(2): 322-331.

[13] Croson, R. and U. Gneezy (2009). "Gender Differences in Preferences," *Journal of Economic Literature*, 47(2): 448-474.

[14] Cuberes, D. and M. Teignier (2014). "Gender Inequality and Economic Growth: A Critical Review," *Journal of International Development*, 26(2): 260-276.

[15] Deerwester, S., Dumais, S.T., Furnas, G.W., Landauer, T.K. and R. Harshman (1990). "Indexing by Latent Semantic Analysis," *Journal of the American Society for Information Science*, 41(6): 391-407

[16] Dreber, A. and M. Hoffman (2007). "Portfolio Selection in Utero", University of Chicago.

[17] Dubois, A., Zagheni, E., Garimella K. and I. Weber (2018). "Studying Migrant Assimilation through Facebook Interests," in: Staab, S., Koltsova O. and D. Ignatov (eds.), *Social Informatics. SocInfo 2018. Lecture Notes in Computer Science*, vol 11186. Springer.

[18] Eckart, C. and G. Young (1936). "The Approximation of One matrix by Another of Lower Rank," *Psychometrika*, 1(3): 211-218.

[19] Falk, A. and J. Hermle (2018). "Relationship of Gender Differences in Preferences to Economic Development and Gender Equality," *Science*, 362(6412), eaas9899.

[20] Fernández, R. (2014). "Women's Rights and Development," *Journal of Economic Growth*, 19(1): 37-80.

[21] Fernández, R., Fogli, A. and C. Olivetti (2004). "Mothers and Sons: Preference Formation and Female Labor Force Dynamics," *Quarterly Journal of Economics*, 119(4): 249-1299.

[22] Fernández, R. and A. Fogli (2009). "Culture: An Empirical Investigation of Beliefs, Work, and Fertility," *American Economic Journal: Macroeconomics*, 1(1): 146-177.

[23] Fors Connolly, F., Goossen, M. and M. Hjerm (2020). "Does Gender Equality Cause Gender Differences in Values? Reassessing the Gender-Equality-Personality Paradox," *Sex Roles*, 83: 101-113.

[24] Funk, P. and C. Gathmann (2015). "Gender Gaps in Policy Making: Evidence from Direct Democracy in Switzerland," *Economic Policy*, 30(81): 141-181.

[25] Gavish, M. and D.L. Donoho (2014). "The Optimal Hard Threshold for Singular Values Is $4/\sqrt{3}$," *IEEE Transactions on Information Theory*, 60(8): 5040-5053.

[26] García, D., Mitike Kassa, Y., Cuevas, A., Cebrian, M., Moro, E., Rahwan, I. and R. Cuevas, R. (2018). "Analyzing Gender Inequality through Large-Scale Facebook Advertising Data," *Proceedings of the National Academy of Sciences*, 115(27): 6958-6963.

[27] Gneezy, U., Leonard, K.L. and J.A. List (2008). "Gender Differences in Competition: Evidence from a Matrilineal and Patriarchal Society," *Econometrica*, 77(3): 909-931.

[28] Goldin, C. and L. F. Katz (2002). "The Power of the Pill: Oral Contraceptives and Women's Career and Marriage Decisions," *Journal of Political Economy*, 110(4): 730-770.

[29] Guiso, L., Monte, F., Sapienza, S. and L. Zingales (2008). "Culture, Gender, and Math," *Science*, 320(5880): 1164–65.

[30] Kaiser, T. (2019). "Nature and Evoked Culture: Sex Differences in Personality Are Uniquely Correlated with Ecological Stress," *Personality and Individual Differences*, 148: 67-72.

[31] Lassmann, A., Liberini, F., Russo, A., Cuevas, A. and R. Cuevas (2020). "Global Spillovers of Taxation in the Online Advertising Market. Theory and Evidence from Facebook," CESifo Working Paper Series 8149.

[32] Liberini, F., Redoano, M., Russo, A., Cuevas, A. and R. Cuevas (2020). "Politics in the Facebook Era - Evidence from the 2016 US Presidential Elections," CESifo Working Paper Series 8235.

[33] Lippa, R.A., Collaer, M.L. and M. Peters (2010). "Sex Differences in Mental Rotation and Line Angle Judgments Are Positively Associated with Gender Equality and Economic Development across 53 Nations," *Archives of Sexual Behavior* 39(4): 990-997.

[34] Lippmann, Q. (2021). "Gender and Lawmaking in Times of Quotas," working paper.

[35] Mac Giolla, E. and P.J. Kajonius (2019). "Sex Differences in Personality Are Larger in Gender Equal Countries: Replicating and Extending a Surprising Finding," *International Journal of Psychology*, 54(6): 705-711. (2019).

[36] Mejova, Y., Gandhi, H. R., Rafaliya, T. J., Sitapara, M. R., Kashyap, R. and I. Weber (2018). "Measuring Subnational Digital Gender Inequality in India through Gender Gaps in Facebook Use," *COMPASS '18: Proceedings of the 1st ACM SIGCAS Conference on Computing and Sustainable Societies,* 43: 1-5.

[37] Niederle, M. and L. Vesterlund (2011). "Gender and Competition," *Annual Review of Economics*, 3(1): 601-630.

[38] Obradovich, N., Özak, Ö., Martín, I., Ortuño Ortín, I., Awad, E., Cebrián, M., Cuevas, R., Desmet, K., Rahwan, I. and A. Cuevas (2020). "Expanding the Measurement of Culture with a Sample of Two Billion Humans," NBER Working Paper #27827.

[39] Olivetti, C. and B. Petrongolo (2016). "The Evolution of Gender Gaps in Industrialized Countries," *Annual Review of Economics*, 8(1): 405-434.

[40] Polachek, S. and J. Xiang (2015). "The Gender Pay Gap Across Countries: A Human Capital Approach," LIS Working papers 646.

[41] Quisimbing, A.R. and J.A. Maluccio (2020). "Intrahousehold Allocation and Gender Relations: New Empirical Evidence from Four Developing Countries," FCND Discussion Paper 84, International Food Policy Research Institute (IFPRI).

[42] Rama, D., Mejova, Y., Tizzoni, M., Kalimeri, K. and I. Weber (2020). "Facebook Ads as a Demographic Tool to Measure the Urban-Rural Divide," in: *Proceedings of the Web Conference 2020 (WWW '20)*, April 20-24, Taipei, Taiwan.

[43] Sapienza, P., Zingales, L. and D. Maestripieri (2009). "Gender Differences in Financial Risk Aversion and Career Choices are Affected by Testosterone," *Proceedings of the National Academy of Sciences*, 106(36):15268-73.

[44] Schmitt, D. P. (2015). "The Evolution of Culturally-Variable Sex Differences: Men and Women Are Not Always Different, but When They Are...It Appears Not to Result from Patriarchy or Sex Role Socialization," in: Shackelford, T. K. and R. D. Hansen (eds.), *The Evolution of Sexuality*, Springer International Publishing.

[45] Schmitt, D.P., Long, A.E., McPhearson, A., O'Brien, K., Remmert, B. and Shah, S.H. (2017). "Personality and Gender Differences in Global Perspective," *International Journal of Psychology*, 52: 45-56.

[46] Serra-Garcia, M. (2021). "Risk Attitudes and Conflict in the Household," *Economic Journal*, forthcoming.

[47] Shores, T. S. (2007). *Applied Linear Algebra and Matrix Analysis*, New Springer, New York: Springer.

[48] Stimpfle, A. and D. Stadelmann (2016). "Marriage Age Affects Educational Gender Inequality: International Evidence," CREMA Working Paper No. 2016-02.

[49] Stoet, G, and D. C. Geary (2018). "The Gender-Equality Paradox in Science, Technology, Engineering, and Mathematics Education," *Psychological Science*, 29(4): 581-593.

[50] Vieira, C. C. and M. Vasconcelos (2021). "Using Facebook Ads Data to Assess Gender Balance in STEM: Evidence from Brazil," Proceedings of the Web Conference 2021 (WWW '21 Companion).

A Data Appendix

Area. Log of total land area. Source: Ashraf and Galor (2013). Data are from World Development Indicators, World Bank. http://wdi.worldbank.org

Facebook penetration. Share of population using Facebook. Source: Garcia et al. (2018).

Facebook interests by gender and country. Number of monthly active users (MAU) associated with 45,397 interests by gender and country. Used to compute different distance measures between men and women and entropy. See Section 2 of paper.

Gender equality index (WEF). Gender Gap Index, 2018. Source: The Global Gender Gap Report 2018. https://www.weforum.org/

Gender equality index (OECD). Negative of Social Institutions and Gender Index 2019. Source: Social Institutions and Gender (indicator), OECD. https://doi.org/10.1787/7b6cfcf0-en

Gender equality index (UNDP). Negative of Gender Inequality Index 2018. Source: Human Development Reports, United Nations Development Programme. http://hdr.undp.org/en/indicators/68606

GDP per capita. GDP per capita in current US$, average 2000-2017. Source: World Development Indicators. World Bank. http://wdi.worldbank.org

Land suitability. Land suitability for agriculture. Source: Ashraf and Galor (2013). Data are from Ramankutty et al. (2002) and Michalopoulos (2012).

Population. Population (in thousands), 2015. Source: World Population Prospects: The 2017 Revision, United Nations.

Precipitation. Average monthly precipitation of a country in mm per month over the 1961? to 1990 time period. Source: Ashraf and Galor (2013). Data are based on G-Econ project (Nordhaus, 2006).

Regional dummies. Regional dummies. Source: Ashraf and Galor (2013). Data are from World Bank.

Religious composition. Share of protestants, share of catholics and share of muslims. Source: Ashraf and Galor (2013). Data are from La Porta et al. (1999).

Suffrage. Year of female suffrage defined as first year that enfranchised female adults older than the minimal voting age exceeded 90%. Source: Coppedge et al. (2021).

Temperature. Average monthly temperature of a country in degrees Celsius per month over the 1961-1990 time period. Source: Ashraf and Galor (2013). Data are based on G-Econ project (Nordhaus, 2006).

Terrain roughness. The degree of terrain roughness of a country, calculated using geospatial surface undulation data. Roughness of terrain. Source: Ashraf and Galor (2013). Data are based on G-Econ project (Nordhaus, 2006).

Additional data references.

Ashraf, Q. and Galor, O. (2013). "The Out of Africa' Hypothesis, Human Genetic Diversity, and Comparative Economic Development," *American Economic Review*, 103(1): 1-46.

Coppedge, M., Gerring, J., Knutsen, C.H., Lindberg, S.I., Teorell, J., Alizada, N., Altman, D., Bernhard, M., Cornell, A., Fish, M.S., Gastaldi, L., Gjerløw, H., Glynn, A., Hicken, A., Hindle, G., Ilchenko, N., Krusell, J., Lührmann, A., Maerz, S.F., Marquardt, K.L., McMann, K., Mechkova, V. Medzihorsky, J., Paxton, P., Pemstein, D., Pernes, J., von Römer, J., Seim, B., Sigman, R., Skaaning, S.-E., Staton, J., Sundström, A., Tzelgov, E., Wang, Y., Wig, T., Wilson, S. and D. Ziblatt (2021). "V-Dem [Country-Year/Country-Date] Dataset v11.1," Varieties of Democracy (V-Dem) Project. `https://doi.org/10.23696/vdemds21.`

La Porta, R., Lopez-de-Silanes, F., Shleifer, A. and R.W. Vishny (1999). "The Quality of Government," *Journal of Law, Economics, and Organization*, 15(1): 222-279.

Michalopoulos, S. (2012). "The Origins of Ethnolinguistic Diversity," *American Economic Review*, 102(4): 1508-39.

Nordhaus, W.D. (2006). 'Geography and Macroeconomics: New Data and New Findings," *Proceedings of the National Academy of Sciences*, 103(10): 3510?-3517.

Ramankutty, N., Foley, J.A., Norman, J. and K. McSweeney (2002). "The Global Distribution of Cultivable Lands: Current Patterns and Sensitivity to Possible Climate Change," *Global Ecology and Biogeography*, 11(5): 377-?392.

B Additional Figures and Tables

Table B.1: Gender Differences in Interests and Gender Equality: Different Samples

Country	Gender Dist.	%FB	Gender Eq.	Country	Gender Dist.	%FB	Gender Eq.	Country	Gender Dist.	%FB	Gender Eq.
Afghanistan	0.051	0.11		Honduras	0.071	0.39	0.706	Paraguay	0.086	0.53	0.672
Albania	0.068	0.50	0.734	Hong Kong	0.055	0.80		Peru	0.053	0.70	0.72
Algeria	0.073	0.52	0.629	Hungary	0.087	0.60	0.674	Philippines	0.046	0.68	0.799
Angola	0.091	0.12	0.633	India	0.073	0.24	0.665	Poland	0.130	0.44	0.728
Argentina	0.095	0.76	0.733	Indonesia	0.083	0.51	0.691	Portugal	0.093	0.63	0.732
Armenia	0.086	0.46	0.678	Iraq	0.060	0.54	0.551	Puerto Rico	0.088	0.56	
Australia	0.083	0.71	0.73	Ireland	0.102	0.64	0.796	Qatar	0.077	1.23	0.629
Austria	0.090	0.50	0.718	Israel	0.054	0.77	0.722	Rep. Congo	0.122	0.13	
Azerbaijan	0.095	0.32	0.68	Italy	0.074	0.57	0.706	Romania	0.091	0.52	0.711
Bahrain	0.095	0.99	0.627	Jamaica	0.082	0.42	0.724	Russia	0.168	0.30	0.701
Bangladesh	0.074	0.20	0.721	Japan	0.100	0.31	0.662	Rwanda	0.074	0.05	0.804
Belarus	0.152	0.26	0.747	Jordan	0.057	0.61	0.605	Saudi Arabia	0.123	0.75	0.59
Belgium	0.085	0.65	0.738	Kazakhstan	0.169	0.41	0.712	Senegal	0.107	0.22	0.682
Benin	0.092	0.12	0.654	Kenya	0.077	0.18	0.7	Serbia	0.087	0.43	0.73
Bolivia	0.070	0.58	0.748	Kosovo	0.104	0.45		Sierra Leone	0.062	0.08	0.661
Bosnia	0.094	0.47	0.712	Kuwait	0.086	1.04	0.63	Singapore	0.057	0.83	0.707
Botswana	0.079	0.41	0.715	Kyrgyzstan	0.131	0.30	0.691	Slovakia	0.096	0.50	0.693
Brazil	0.087	0.65	0.681	Laos	0.070	0.38	0.748	Slovenia	0.087	0.49	0.784
Bulgaria	0.090	0.53	0.756	Latvia	0.101	0.48	0.758	Somalia	0.046	0.10	
Burkina Faso	0.088	0.07	0.629	Lebanon	0.063	0.68	0.595	South Africa	0.079	0.36	0.755
Burundi	0.057	0.04	0.741	Lesotho	0.078	0.14	0.693	South Korea	0.065	0.40	0.657
Cambodia	0.046	0.52	0.683	Liberia	0.070	0.12	0.681	Spain	0.094	0.58	0.746
Cameroon	0.110	0.15	0.714	Libya	0.064	0.67		Sri Lanka	0.074	0.29	0.676
Canada	0.079	0.70	0.771	Lithuania	0.115	0.57	0.749	Swaziland	0.088	0.16	
Chile	0.084	0.79	0.717	Macedonia	0.078	0.54	0.707	Sweden	0.092	0.70	0.822
Colombia	0.074	0.64	0.729	Madagascar	0.073	0.09	0.691	Switzerland	0.077	0.52	0.755
Costa Rica	0.078	0.71	0.749	Malawi	0.107	0.03	0.662	Taiwan	0.059	0.84	
Cote d'Ivoire	0.095	0.20	0.627	Malaysia	0.066	0.81	0.676	Tajikistan	0.137	0.05	0.638
Croatia	0.102	0.47	0.712	Mali	0.099	0.09	0.582	Tanzania	0.071	0.09	0.704
Cyprus	0.064	0.85	0.684	Mauritania	0.067	0.18	0.607	Thailand	0.063	0.74	0.702
Czech Republic	0.100	0.51	0.693	Mauritius	0.073	0.63	0.663	The Bahamas	0.088	0.20	0.741
DRC	0.122	0.03	0.582	Mexico	0.064	0.67	0.721	The Gambia	0.097	0.18	0.642
Denmark	0.082	0.69	0.778	Moldova	0.092	0.28	0.733	Timor-Leste	0.033	0.37	0.638
Dom. Rep.	0.087	0.55	0.701	Mongolia	0.085	0.70	0.714	Togo	0.115	0.09	0.618
Ecuador	0.069	0.70	0.729	Morocco	0.089	0.49	0.607	Trinidad	0.071	0.60	
Egypt	0.061	0.41	0.614	Mozambique	0.076	0.07	0.721	Tunisia	0.087	0.63	0.648
El Salvador	0.071	0.57	0.69	Myanmar	0.084	0.39	0.69	Turkey	0.117	0.65	0.628
Estonia	0.100	0.55	0.734	Namibia	0.083	0.27	0.789	Uganda	0.088	0.06	0.724
Ethiopia	0.054	0.06	0.656	Nepal	0.044	0.33	0.671	Ukraine	0.098	0.36	0.708
Finland	0.098	0.57	0.821	Netherlands	0.099	0.64	0.747	UAE	0.081	1.10	0.642
France	0.092	0.57	0.779	New Zealand	0.078	0.74	0.801	UK	0.094	0.66	0.774
Gabon	0.114	0.36		Nicaragua	0.071	0.44	0.809	USA	0.074	0.72	0.72
Georgia	0.093	0.66	0.677	Niger	0.141	0.03		Uruguay	0.086	0.76	0.715
Germany	0.098	0.44	0.776	Nigeria	0.070	0.13	0.621	Uzbekistan	0.131	0.07	
Ghana	0.087	0.20	0.688	Norway	0.100	0.69	0.835	Venezuela	0.070	0.39	0.709
Greece	0.084	0.52	0.696	Oman	0.093	0.64	0.605	Vietnam	0.050	0.62	0.698
Guatemala	0.078	0.42	0.668	Pakistan	0.119	0.19	0.55	Yemen	0.082	0.09	0.499
Guinea	0.103	0.15	0.656	Palestine	0.049	0.41		Zambia	0.080	0.13	
Guinea-Bissau	0.137	0.08		Panama	0.080	0.55	0.722	Zimbabwe	0.107	0.07	0.721
Haiti	0.112	0.19		Papua NG	0.040	0.09			0.131	0.07	

Gender Dist.: cosine distance between men and women based on 45,397 Facebook interests; % Facebook: Facebook penetration (Garcia et al, 2018); Gender Eq.: Gender Gap Index of World Economic Forum. Bosnia refers to Bosnia and Herzegovina; Dom. Rep. to Dominican Republic; Papua NG to Papua New Guinea; UEA to United Arab Emirates; Trinidad to Trinidad and Tobago.

Table B.2: Gender Differences in Interests and Gender Equality: Different Samples

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: Expanded Sample Including Countries with Population Less Than 1 Million							
Gender Equality (WEF)	0.065*	-0.000	0.123**	0.004	0.040		
	(0.033)	(0.037)	(0.047)	(0.043)	(0.038)		
Gender Equality (OECD)						0.001***	
						(0.000)	
Gender Equality (UNDP)							0.055**
							(0.021)
Log GDP per Capita	0.006***	0.003*	0.007***	0.004**	0.009***	0.006**	0.001
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)
Entropy	-0.038**	-0.041***	-0.037***	-0.039**	-0.039**	-0.060***	-0.022
	(0.017)	(0.014)	(0.012)	(0.017)	(0.019)	(0.019)	(0.018)
Regional Dummies		Yes					
Religious Composition			Yes				
Geography & Climate				Yes			
FB Penetration					Yes		
Observations	110	109	106	97	110	85	113
R^2	0.172	0.442	0.495	0.450	0.250	0.326	0.156
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel B: Expanded Sample Including Countries with Facebook Penetration $> 2.5\%$							
Gender Equality (WEF)	-0.000	-0.070*	-0.011	-0.065*	-0.014		
	(0.034)	(0.037)	(0.047)	(0.038)	(0.034)		
Gender Equality (OECD)						0.000	
						(0.000)	
Gender Equality (UNDP)							0.040**
							(0.020)
Log GDP per Capita	0.006***	0.004**	0.005***	0.002	0.009***	0.005**	0.002
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)
Entropy	-0.042***	-0.038***	-0.035***	-0.032**	-0.032*	-0.049***	-0.047***
	(0.014)	(0.012)	(0.012)	(0.014)	(0.016)	(0.014)	(0.013)
Regional Dummies		Yes					
Religious Composition			Yes				
Geography & Climate				Yes			
FB Penetration					Yes		
Observations	131	131	128	123	131	115	134
R^2	0.134	0.379	0.371	0.371	0.202	0.203	0.190
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel C: Restricted Sample with Countries with Facebook Penetration 50%							
Gender Equality (WEF)	0.004	-0.044	0.068	-0.059	-0.025		
	(0.039)	(0.055)	(0.065)	(0.061)	(0.044)		
Gender Equality (OECD)						0.001***	
						(0.000)	
Gender Equality (UNDP)							0.013
							(0.022)
Log GDP per Capita	0.007***	0.005**	0.006**	0.007**	0.009***	0.003	0.005*
	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)	(0.002)	(0.003)
Entropy	-0.017	-0.022	-0.018	-0.019	-0.016	-0.008	-0.014
	(0.016)	(0.016)	(0.015)	(0.020)	(0.017)	(0.020)	(0.016)
Regional Dummies		Yes					
Religious Composition			Yes				
Geography & Climate				Yes			
FB Penetration					Yes		
Observations	63	63	62	59	63	52	65
R^2	0.170	0.381	0.379	0.348	0.246	0.311	0.169

Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The dependent variable is the difference between men and women based on 45,397 Facebook interests. The sample consists of countries with population $>$ 1 million, Facebook penetration > 0.25, and Facebook users $>$ 100,000, except when noted otherwise. The seven specifications are identical to those in Table 1. All regressions include a constant. Religious composition refers to share of protestants, catholics and muslims, as well as a dummy for Soviet influence; geography & climate refers to log area, agricultural land suitability, terrain roughness, temperature and precipitation; continents refer to continental dummies; and FB penetration refers to Facebook penetration and log of population.

Table B.3: Gender Differences in Interests and Gender Equality: Different Distance Measures

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: Manhattan Distance							
Gender Equality (WEF)	0.105	0.096	0.132	0.004	0.070		
	(0.079)	(0.112)	(0.110)	(0.096)	(0.096)		
Gender Equality (OECD)						0.001***	
						(0.000)	
Gender Equality (UNDP)							0.057
							(0.044)
Log GDP per Capita	0.015***	0.009**	0.013***	0.006	0.016***	0.010*	0.010
	(0.004)	(0.004)	(0.004)	(0.005)	(0.005)	(0.005)	(0.006)
Entropy	-0.059**	-0.056**	-0.034	-0.027	-0.045	-0.072**	-0.072***
	(0.027)	(0.022)	(0.024)	(0.029)	(0.030)	(0.029)	(0.027)
Regional Dummies		Yes					
Religious Composition			Yes				
Geography & Climate				Yes			
FB Penetration					Yes		
Observations	98	98	95	91	98	84	100
R^2	0.157	0.344	0.372	0.365	0.185	0.233	0.196
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: Euclidean Distance							
Gender Equality (WEF)	0.002	-0.000	0.004	-0.001	0.001		
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)		
Gender Equality (OECD)						0.000***	
						(0.000)	
Gender Equality (UNDP)							0.001
							(0.001)
Log GDP per Capita	0.000***	0.000**	0.000***	0.000**	0.000***	0.000**	0.000*
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Entropy	-0.006***	-0.006***	-0.006***	-0.006***	-0.006***	-0.006***	-0.006***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Regional Dummies		Yes					
Religious Composition			Yes				
Geography & Climate				Yes			
FB Penetration					Yes		
Observations	98	98	95	91	98	84	100
R^2	0.520	0.640	0.693	0.678	0.553	0.641	0.597

Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The dependent variable is the difference between men and women based on 45,397 Facebook interests. The sample consists of countries with population > 1 million and Facebook penetration > 0.25. The seven specifications are identical to those in Table 1. All regressions include a constant. Religious composition refers to share of protestants, catholics and muslims, as well as a dummy for Soviet influence; geography & climate refers to log area, agricultural land suitability, terrain roughness, temperature and precipitation; continents refer to continental dummies; and FB penetration refers to Facebook penetration and log of population.

Table B.4: Gender and Non-Gender Interests Based on Subset of Least Gender-Equal Countries

Panel A: Cosine Distance Men-Women Based on Gender Interests in Least Gender-Equal Countries							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Gender Equality (WEF)	0.219***	0.138**	0.203**	0.126	0.156**		
	(0.062)	(0.063)	(0.094)	(0.079)	(0.074)		
Gender Equality (OECD)						0.001***	
						(0.000)	
Gender Equality (UNDP)							0.054
							(0.040)
Log GDP per Capita	0.006*	0.003	0.005*	-0.000	0.010**	0.004	0.003
	(0.004)	(0.003)	(0.003)	(0.004)	(0.005)	(0.005)	(0.006)
Entropy	-0.053	-0.067**	-0.045*	-0.027	-0.032	-0.065*	-0.058*
	(0.034)	(0.026)	(0.025)	(0.031)	(0.035)	(0.035)	(0.033)
Regional Dummies		Yes					
Religious Composition			Yes				
Geography & Climate				Yes			
FB Penetration					Yes		
Observations	98	98	95	91	98	84	100
R^2	0.169	0.551	0.532	0.386	0.231	0.234	0.119
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel B: Cosine Distance Men-Women Based on Non-Gender Interests in Least Gender-Equal Countries							
Gender Equality (WEF)	-0.068***	-0.066***	-0.045**	-0.101***	-0.065***		
	(0.019)	(0.024)	(0.022)	(0.018)	(0.020)		
Gender Equality (OECD)						-0.000**	
						(0.000)	
Gender Equality (UNDP)							-0.021**
							(0.008)
Log GDP per Capita	0.003***	0.003***	0.004***	0.003***	0.003***	0.003***	0.005***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Entropy	-0.030***	-0.034***	-0.035***	-0.030***	-0.032***	-0.035***	-0.036***
	(0.006)	(0.006)	(0.006)	(0.006)	(0.007)	(0.006)	(0.006)
Regional Dummies		Yes					
Religious Composition			Yes				
Geography & Climate				Yes			
FB Penetration					Yes		
Observations	98	98	95	91	98	84	100
R^2	0.386	0.422	0.490	0.549	0.416	0.456	0.370

Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The dependent variable is the difference between men and women based on subset of interests that are more frequent in one of the genders in at least 90% of the 61 least gender-equal countries (Panel A) or subset of interests that are more frequent in men at least 30% of the 61 least gender-equal countries and more frequent in women in at least 30% of the 61 least gender-equal countries (Panel B). The sample consists of countries with population > 1 million and Facebook penetration > 0.25. The seven specifications are identical to those in Table 1. All regressions include a constant. Religious composition refers to share of protestants, catholics and muslims, as well as a dummy for Soviet influence; geography & climate refers to log area, agricultural land suitability, terrain roughness, temperature and precipitation; continents refer to continental dummies; and FB penetration refers to Facebook penetration and log of population.

Table B.5: 500 Most Masculine Interests according to Preference Dimension V_4

Automobiles; BMW; Motorcycles; Automotive industry; Mercedes-Benz; Cars (film); Luxury vehicle; Auto racing; Trucks; Audi; Motorsport; Engine; Smartphones; Toyota; SUVs; Outdoor recreation; Cameras; UEFA Champions League; Sports car; Association football (Soccer); Free software; Real Madrid C.F.; Lionel Messi; Application software; Motor vehicle; Cristiano Ronaldo; Wheel; Bluetooth; Martial arts; Game consoles; Internal combustion engine; Electronics; Nissan; Formula One; Drive (2011 film); Streaming media; Boxing; Porsche; Volkswagen; Fishing; Construction; Personal finance; Country; Volkswagen Group; Hunting; Premier League; Mercedes-AMG; Honda; Money; Online games; Land; Telecommunication; Mixed martial arts; American football; Transport; Tire; Tablet computers; Hybrids; UEFA; Basketball; FIFA World Cup; Cycling; Play (telecommunications); BMW M; Linux; Mobile app; Juventus F.C.; Military; Coupé ; Sedan (automobile); Ford Motor Company; Team sport; Finance; Lamborghini; Speed (1994 film); Investment; Gamer; Car tuning; Gambling; Televisions; Brand; Four-wheel drive; National Basketball Association; 24 Hours of Le Mans; Scooters; Goalkeeper (association football); First-person shooter games; Daimler AG; Sales; Asia; PlayStation 4; Nürburgring; Road; Manufacturing; Liverpool; Ultimate Fighting Championship; Aston Martin; Computer processors; Top Gear (magazine); Programming language; BMW 3 Series; BMW M3; Electric vehicle; FIFA; China; Ferrari; Chip tuning; Convertible; Football; Racing games; Telephone; Action movies; Engine tuning; Serie A; Bicycle; Liverpool F.C.; Turbocharger; Golf; United States; Combat sport; Gasoline; USB; Battery (electricity); Audi RS 6; UEFA Europa League; Russia; IPhone; La Liga; Bundesliga; Sports car racing; Product (business); Minivans; Global Television Network; BMW 3 Series (E36); Microsoft; Headphones; Polishing; Manchester City F.C.; History (European TV channel); Engineering; Eden Hazard; Norway; Bugatti; BMW 3 Series (E46); Power (physics); BMW M4; Europe; Brazilian jiu-jitsu; Women's association football; Sports games; Google; Drifting (motorsport); S.S. Lazio; A.C. Milan; Calciatori Brutti; Insurance; Diesel engine; Tool; Company; Manchester United F.C.; Football team; Google Play; Supercar; Russian language; Chelsea F.C.; Apple Inc.; Finland; Neymar; Action games; Motorcycle racing; La Gazzetta dello Sport; Huawei; Euro; Samsung; Inter Milan; Lexus; PlayStation; Grappling; Video; Information technology; Victory; Gamer (film); Champion (sportswear); Major League Soccer; Mobile game; Boats; Kickboxing; Banking; FC Barcelona; Sony; Land Rover; Skiing; New York City; President of the United States; Wi-Fi; Shooter games; Cloud computing; Electricity; Driving; LFC TV; Mazda; Multinational corporation; Volvo; Wrestling; Sound recording and reproduction; Taiwan; Audi S and RS models; Electric car; Vintage car; Touchscreen; Germany; United States women's national soccer team; Hong Kong; Copa Libertadores; Lithuania; Rugby league; Estonia; Adventure game; Beer; Yamaha Motor Company; SEAT; Desktop computers; Off-roading; Latvia; Camping; IOS; ITunes; Website; Computer monitors; College football; Role-playing games; Massively multiplayer online role-playing games; Pickup truck; Automatic transmission; YouTube; Trade; Beijing; Computer hardware; Arsenal F.C.; Adidas; Muay Thai; Machine; Tractor; Chinese language; Anfield; Heavy metal music; Jürgen Klopp; Carcare; Windows Phone; Xi Jinping; Gaming computer; Professional boxing; Politics; Xbox (console); Police; Macau; Red Bull; Types of motorcycles; Auto show; Economy; Poker; Hard drives; Motocross; United Kingdom; Portugal national football team; Bellator MMA; Light-emitting diode; Chevrolet; Compact car; Japan; Suzuki; Team; Nightclubs; International Brazilian Jiu-Jitsu Federation; Counter-Strike: Global Offensive; Computer network; Station wagon; McLaren; Yamaha Corporation; Track and field; The Cars; Grand Prix motorcycle racing; Volkswagen Golf; Top Gear; Massively multiplayer online games; Politics and social issues; Mike Tyson; GSM; Professional wrestling; League of Legends; War; Laptop; Golden State Warriors; Solar energy; Racing; Xbox One; Hyundai; Peugeot; Watch; Nvidia; Republican Party (United States); Symbian; Electronic music; Classic car; Used car; Motorcycling; PlayStation (console); World War II; Marketing; Mitsubishi Motors; ADCC Submission Wrestling World Championship; IPod; Information; Random-access memory; Stock; Volleyball; Xiaomi; Energy drinks; Digital data; Jeep; NASDAQ; Renault; Air conditioning; Graphics processing unit; PC Gamer; Heavyweight (MMA); Sport bike; Ukraine; Japanese domestic market; PHP; Muhammad Ali; Sweden; Razer Inc.; Rallying; Mass media; Video game industry; Entrepreneurship; Fuel (band); World; Personal computer; Fédération Internationale de l'Automobile; UFC 1; Amateur boxing; World Boxing Association; College basketball; GeForce; Soviet Union; App Store (iOS); Nike; Vladimir Putin; Sensor; Front-wheel drive; European Union; Jaguar Cars; RVs; IPad; Mountain biking; Epic (2013 film); Ski; Competition; Samsung Galaxy; Twitch (website); Strategy games; Garage (clothing retailer); The Ultimate Fighter; Land use; Stand-up comedy; Mark Zuckerberg; Electric motor; Judo; United Nations; Steel; The Game (rapper); Loudspeaker; National Football League; Card games; Motorcycle sport; Mountain bike; Diesel fuel; PC game; IMG Models; Snowboarding; Management; Car dealership; Mobile device; Electronic sports; Legend (1985 film); Dota 2; Ericsson; Judi; Computer servers; Platform game; Elite Model Management; Current events; Global Positioning System; Ice hockey; Car rentals; Credit cards; Career; FC Bayern Munich; Electro (music); Everything (band); Smartwatch; Loan; Custom car; Car classification; 1080; Transmission (mechanics); Ducati; Future (rapper); Price; Extra (acting); Angela Merkel; V8 engine; Toronto Raptors; DVD; Wireless; KTM; Cars; Welcome (2007 film); Ford Mustang; Network (film); Maserati; German language; Alfa Romeo; Spotify; Toronto; BMW M5; Welding; Bus; Adventure; Air pollution; Atlético Madrid; Suzuki GSX-R series; Ubisoft; Tennis; Epic Games; Dmitry Medvedev; Subaru; Modeling agency; Privately held company; 2016-17 UEFA Champions League; 2015-16 UEFA Champions League; Supermodel; Viral video; HVAC; Jujutsu; NBA Finals; Tool (band); Ambassador; Broadcasting; Swimming; Live events; Taekwondo; Multiplayer online battle arena; Brand New (band); Marathons; Top 14; Headlamp; Chrysler; Artificial intelligence; WorldStarHipHop; Africa; Kia Motors; American Civil War; Rapping; Traffic; Investor; House (TV series); LeBron James; Auto detailing; Army; South Africa; Funk; TV; Road racing; Humour; Aviation; Recreation; World Rally Championship; Music festivals; Heavy equipment; Hungary; Fuel efficiency; Thriller movies; Auction.

Table B.6: 500 Most Feminine Interests according to Preference Dimension V_4

Dresses; Cosmetics; Infant; Motherhood; Beauty salons; Hair products; Woman; Pregnancy; Boutiques; Child; Hair (film); Chocolate; Cooking; Desserts; Cake; Jewelry; Handbags; Fashion accessories; Blouse; Recipes; Hairstyle; Weddings; Make-up artist; Nail (anatomy); Baking; Skirt; Cuisine; Skin; Childbirth; Flower; Wedding dress; Spas; Aesthetics; Kids (film); Fragrances; Interior design; Men's clothing; Shoes; Fashion design; Trousers; Parent; Female; Footwear; Luxury goods; Bride; Weight loss (Fitness And wellness); Decorative arts; Toys; Pink (singer); Shopping malls; Textile; Protein; Meal; Children's clothing; Eating; Hair care; Anatomy; Psychology; Crafts; Handicraft; Manicure; Veganism; Bread; Yoga; Coupons; Mother's Day; Pastry; Love (John Lennon song); Fashion week; Medicine; Kitchen; Furniture; Chanel; Hand; Makeup brush; Breastfeeding; Love; Parenting; L'Oréal; Integumentary system; Maquiladora; Healthy diet; Fatherhood; Books; Fashion blog; Discount stores; Gift; Foodie; Pleasure; Flour; Face; Nail art; Fruit; Pedicure; Makeup Tutorials; Fashion (film); Maria B; Home and garden; Airbrush makeup; MAC Cosmetics; Birthday; Creativity; Marriage; Make Up For Ever; Philosophy; Spanish language; Sewing; Zara (retailer); Literature; Sugar; Sandal; Retail; Pizza; Personal care; Country music; Pakistani clothing; Visual arts; Coffee; Justin Bieber; Primate; H&M; Human; Latin America; Human sexuality; Poetry; Hijab; Affection; Haute couture; Confectionery; Nutrient; Eyebrow; Mama (2013 film); Baby shower; Friends; Developmental psychology; Outfit of the day; Meditation; Emotion; Vertebrate; Vegetarianism; Virtue; God; Personal development; Writing; Chef; Childhood; Cake (band); Girl; Cake decorating; Academy Award for Best Makeup and Hairstyling; Cognition; Sephora; Zainab Chottani; Amour (2012 film); Italian cuisine; Do it yourself (DIY); Cats; Cupcake; Cookie; Crochet; Design; Textile arts; Salé ; Cream (band); Tea; Do it yourself; The Walt Disney Company; Lipstick; People (magazine); Romance film; Discover Card; Nail polish; WhatsApp; Candy; Mammal; Painting; Dogs; Physician; Evening gown; Colombia; Breakfast; Eyelash; Fabindia; HIM (Finnish band); Next (TV series); Cosmetology; Milk; Organism; CoverGirl; Deco; Horoscope; Interpersonal relationship; Mind; Discounts and allowances; Wedding photography; Astrology; Icing (food); Blog; Color; Keratin; Quality of life; Parties; Religion; Mexico; Home improvement; Lip; Colors (film); Queen (band); Baby Boy (film); Latin pop; Cookbook; Health care; Slow Food; Anita Dongre; Sweetness; Embroidery; Michelin Guide; IKEA; Beverages; Hair coloring; Yarn; Gown; Physical attractiveness; Teacher; Adult; Idea; Wedding planner; Carbohydrate; Chile; Knitting; Too Faced Cosmetics; Tuxedo; Wine; Sari; Street fashion; Permanent makeup; Pasta; Cook (profession); Earring; Perception; Puberty; French cuisine; Nail salon; Spirituality; Lakme Fashion Week; Ritu Kumar; Peru; Kindergarten; Human hair color; Coffeehouses; Diaper; Country Living; Drawing; tarte cosmetics; Baker; Bakery; Tattoos; Birthday cake; Biology; Manish Malhotra; Romance novels; Father's Day; Latin music (genre); Restaurants; Cheese; Silk; Ethics; Big (film); Musical film; Chic; Amor (film); Gastronomy; Kylie Jenner; Home (2009 film); Patisserie; Organic food; Singer-songwriter; Foot; Juice; Farmhouse; Artist; Fast food; Variety (magazine); Vogue (magazine); Big Ben; Tarot; Learning; Tumblr; Zodiac; K-pop; School; Therapy; Contemporary R&B; Girls (TV series); Artificial nails; Child care; Natural product; Ready (2011 film); Ice cream; Couch; Christian Lacroix; Bed; Indo-Western clothing; Central America; Causeway Bay; Guatemala; Ring (jewellery); Magazines; Concept; Intimate relationship; Necklace; Toddler; nail; Preschool; Keeping Up with the Kardashians; Jeans; Bella (film); Consciousness; Thursday (band); Kim Kardashian; Cotton; Selena Gomez; Ecuador; Baby sling; Ontology; Abdomen; Paper; Metaphysics; Liu Jo; Needlework; Spanish Empire; Elle (magazine); Canada; Gordon Ramsay; Designer clothing; Self-love; Teen drama; World Health Organization; Vegetable; Angel; Valentine's Day; Romance (love); Etsy; romantic comedies; Coaching; Behavior; Offspring; Bolivia; Mexico City; Victoria's Secret; Adolescence; Chair; Bracelet; Ewa Chodakowska; Telenovela; Designer; Henna; Baby monitor; Ballet; Conde Nast; Nursing; Scrapbooking; Lunch; Inditex; Sana Safinaz; Alternative medicine; States of Brazil; Friday (1995 film); Pharmacy; Snack food; Wedding Planners; Dentistry; Glamour (magazine); Interior Design Ideas; Ketone; Vegetarian cuisine; Object (philosophy); Fondant icing; Wella Professionals; Lingerie; Microsoft Office; moda; Health & wellness; High-heeled footwear; Idealism; Nestle; Kendall Jenner; Organ (anatomy); Pampers; Dermatology; Work of art; Boy band; Maybelline; Blossom; The Business of Fashion; Eyelash extensions; Ageing; Nature; Gemstone; Flickr; art; Engagement; Gender; Theme parks; Collagen; Sleep; Mickey Mouse; Gourmet (magazine); Milan Fashion Week; Lima; Culinary art; Gambero Rosso; Pain (musical project); Chinese cuisine; British Royal Family; Catherine; Down (band); Butter; Another (novel); TLC (TV network); Supermarket; Swarovski; Lace; Summer; Fisher-Price; Eye; Illustration; Paris Fashion Week; Dish Network; Massage; Bride and Groom; Seafood; Mumbai; Handmade jewelry; Banarasi saris; Sunglasses; Soul; Flowering plant; Heart; Chalene Johnson; Mamas & Papas; Girl group; Ariana Grande; Newborn; Female Entrepreneur Association; Feminism; Women's rights; Jamie Oliver; Home Decor Products; Laser; Lehenga; Rihanna; Telemundo; Rock and roll; Braid; Swimsuit; Gardening; Pop rock; Hobby; Horticulture; Southern Living; Happiness; Gluten-free diet; Environmental science; Yo Amo los Zapatos; Spain; Spanish cuisine; Home; Plus-size clothing; Glitter; Catholic Church; Self-esteem; Silver; Pop music; Bridesmaid; Elle Decoration; Abstraction; Engagement ring; 1080i; Graphic design; health; Hairdressers Journal.

Table B.7: Gender and Non-Gender Interests, Controlling for Age Ratio

Panel A: Cosine Distance between Men and Women Based on Gender Interests (with Age Ratio)							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Gender Equality (WEF)	0.137*	0.140**	0.187*	0.147*	0.109		
	(0.070)	(0.062)	(0.101)	(0.079)	(0.073)		
Gender Equality (OECD)						0.001***	
						(0.000)	
Gender Equality (UNDP)							-0.023
							(0.046)
Ratio Old-Young Facebook	0.049***	-0.016	0.022	-0.009	0.037**	0.029	0.073***
	(0.016)	(0.022)	(0.020)	(0.026)	(0.018)	(0.024)	(0.018)
Log GDP per Capita	0.001	0.005	0.004	0.001	0.006	0.000	0.002
	(0.004)	(0.004)	(0.004)	(0.005)	(0.005)	(0.006)	(0.004)
Entropy	-0.060*	-0.059**	-0.056**	-0.023	-0.040	-0.065**	-0.078***
	(0.034)	(0.024)	(0.026)	(0.030)	(0.036)	(0.032)	(0.029)
Regional Dummies		Yes					
Religious Composition			Yes				
Geography & Climate				Yes			
FB Penetration					Yes		
Observations	98	98	95	91	98	84	100
R^2	0.226	0.565	0.524	0.398	0.273	0.261	0.214
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel B: Cosine Distance between Men and Women Based on Gender Interests (with Age Ratio)							
Gender Equality (WEF)	-0.081***	-0.074***	-0.053**	-0.100***	-0.078***		
	(0.018)	(0.022)	(0.021)	(0.019)	(0.018)		
Gender Equality (OECD)						-0.000**	
						(0.000)	
Gender Equality (UNDP)							-0.019
							(0.013)
Ratio Old-Young Facebook	-0.004	-0.006	-0.002	-0.010*	-0.003	-0.003	-0.013*
	(0.005)	(0.007)	(0.005)	(0.006)	(0.005)	(0.005)	(0.007)
Log GDP per Capita	0.003***	0.003***	0.003***	0.004***	0.003**	0.003**	0.006***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Entropy	-0.029***	-0.031***	-0.034***	-0.026***	-0.031***	-0.036***	-0.035***
	(0.006)	(0.007)	(0.007)	(0.007)	(0.008)	(0.007)	(0.007)
Regional Dummies		Yes					
Religious Composition			Yes				
Geography & Climate				Yes			
FB Penetration					Yes		
Observations	98	98	95	91	98	84	100
R^2	0.440	0.478	0.549	0.570	0.453	0.491	0.435

Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The dependent variable is the Euclidean distance between men and women based on V_4 (Panel A), V_2, V_4 and V_5 (Panel B) and V_3, V_6, V_7, V_8 and V_9 (Panel C). The sample consists of countries with population > 1 million and Facebook penetration > 0.25. The seven specifications are identical to those in Table 1. All regressions include a constant. Religious composition refers to share of protestants, catholics and muslims, as well as a dummy for Soviet influence; geography & climate refers to log area, agricultural land suitability, terrain roughness, temperature and precipitation; continents refer to continental dummies; and FB penetration refers to Facebook penetration and log of population.

Table B.8: Gender Interests, Old vs Young

Panel A: Cosine Distance between Old Men and Old Women Based on Subset of Gender Interests							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Gender Equality (WEF)	0.194***	0.023	0.081	0.084	0.130**		
	(0.056)	(0.073)	(0.077)	(0.065)	(0.056)		
Gender Equality (OECD)						0.001***	
						(0.000)	
Gender Equality (UNDP)							0.133***
							(0.037)
Log GDP per Capita	0.014***	0.011***	0.013***	0.009***	0.018***	0.013***	0.002
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)	(0.005)
Entropy	-0.050**	-0.038**	-0.021	-0.030	-0.029	-0.056***	-0.029
	(0.022)	(0.017)	(0.019)	(0.020)	(0.021)	(0.021)	(0.021)
Regional Dummies		Yes					
Religious Composition			Yes				
Geography & Climate				Yes			
FB Penetration					Yes		
Observations	98	98	95	91	98	81	97
R^2	0.272	0.545	0.565	0.513	0.327	0.374	0.292
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel B: Cosine Distance between Young Men and Young Women Based on Subset of Gender Interests							
Gender Equality (WEF)	0.175***	0.058	0.132	0.081	0.084		
	(0.065)	(0.068)	(0.092)	(0.077)	(0.071)		
Gender Equality (OECD)						0.001***	
						(0.000)	
Gender Equality (UNDP)							0.122***
							(0.040)
Log GDP per Capita	0.008*	0.002	0.007**	-0.002	0.013***	0.005	-0.003
	(0.004)	(0.003)	(0.003)	(0.004)	(0.005)	(0.005)	(0.006)
Entropy	-0.051	-0.053**	-0.022	-0.015	-0.019	-0.056	-0.035
	(0.033)	(0.022)	(0.023)	(0.027)	(0.033)	(0.036)	(0.033)
Regional Dummies		Yes					
Religious Composition			Yes				
Geography & Climate				Yes			
FB Penetration					Yes		
Observations	98	98	95	91	98	81	97
R^2	0.129	0.575	0.502	0.492	0.236	0.230	0.152

Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The dependent variable is the cosine distance between men and women based on the subset of interests that are more frequent in one of the genders in at least 90% of countries. This exercise is performed on a random sample of 5,000 interests, instead of the full sample of 45,397 interests. The sample consists of countries with population > 1 million and Facebook penetration > 0.25. The seven specifications are identical to those in Table 1. All regressions include a constant. Religious composition refers to share of protestants, catholics and muslims, as well as a dummy for Soviet influence; geography & climate refers to log area, agricultural land suitability, terrain roughness, temperature and precipitation; continents refer to continental dummies; and FB penetration refers to Facebook penetration and log of population.

Table B.9: Non-Gender Interests, Old vs Young

Panel A: Cosine Distance between Old Men and Old Women Based on Subset of Non-Gender Interests							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Gender Equality (WEF)	-0.167***	-0.144**	0.024	-0.148**	-0.145***		
	(0.045)	(0.058)	(0.048)	(0.063)	(0.044)		
Gender Equality (OECD)						-0.000	
						(0.000)	
Gender Equality (UNDP)							-0.045*
							(0.024)
Log GDP per Capita	0.007***	0.006**	0.007***	0.006**	0.005**	0.003	0.011***
	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)	(0.002)	(0.004)
Entropy	-0.027*	-0.023	-0.019*	-0.019	-0.034**	-0.019	-0.034**
	(0.014)	(0.015)	(0.011)	(0.017)	(0.015)	(0.013)	(0.015)
Regional Dummies		Yes					
Religious Composition			Yes				
Geography & Climate				Yes			
FB Penetration					Yes		
Observations	98	98	95	91	98	81	97
R^2	0.249	0.320	0.489	0.290	0.296	0.047	0.130
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel B: Cosine Distance between Young Men and Young Women Based on Subset of Non-Gender Interests							
Gender Equality (WEF)	-0.158**	-0.146**	-0.058	-0.203***	-0.137**		
	(0.069)	(0.062)	(0.071)	(0.066)	(0.065)		
Gender Equality (OECD)						0.000	
						(0.000)	
Gender Equality (UNDP)							-0.041
							(0.027)
Log GDP per Capita	0.007***	0.007**	0.009***	0.007**	0.006**	0.001	0.010**
	(0.003)	(0.003)	(0.002)	(0.003)	(0.003)	(0.003)	(0.004)
Entropy	-0.027*	-0.037**	-0.033**	-0.022	-0.034**	-0.017	-0.037**
	(0.015)	(0.015)	(0.016)	(0.018)	(0.017)	(0.015)	(0.016)
Regional Dummies		Yes					
Religious Composition			Yes				
Geography & Climate				Yes			
FB Penetration					Yes		
Observations	98	98	95	91	98	81	97
R^2	0.167	0.278	0.301	0.229	0.181	0.043	0.083

Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The dependent variable is the cosine distance between men and women based on the subset of interests that are more frequent in men at least 30% of countries and more frequent in women in at least 30% of countries. This exercise is performed on a random sample of 5,000 interests, instead of the full sample of 45,397 interests. The sample consists of countries with population > 1 million and Facebook penetration > 0.25. The seven specifications are identical to those in Table 1. All regressions include a constant. Religious composition refers to share of protestants, catholics and muslims, as well as a dummy for Soviet influence; geography & climate refers to log area, agricultural land suitability, terrain roughness, temperature and precipitation; continents refer to continental dummies; and FB penetration refers to Facebook penetration and log of population.

Figure B.1: Singular Values Based on SVD

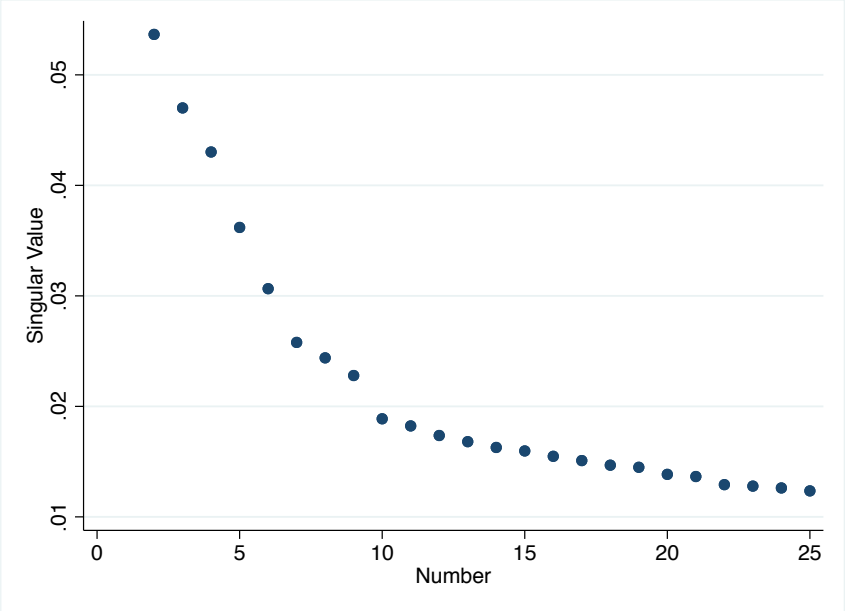


Figure depicts the singular values σ_2, σ_3, etc. To improve visualization, σ_1, which is equal to .22 is not shown. Notice the drop in value (the so-called "elbow") when going from σ_9 to σ_{10}.