Gender Identity, Norms, and Happiness

Natalia Danzer
Rachel E. Kranton
Piotr Paweł Larysz
Claudia Senik*

February 2026

Abstract: How do gender identity and norms relate to happiness? We study feelings of femininity and masculinity, (non-)conformity, and life satisfaction using new data from the European Social Survey. We find a robust asymmetry between men and women. For men, feeling more masculine, behaving in ways more typical of men, and life satisfaction are all positively cross-correlated. But for women, while feeling more feminine and life satisfaction are positively associated, behaving in ways more typical of women is associated with *lower* life satisfaction. These findings support theories of gender identity and reveal possible trade-offs implied by gender norms for women.

Keywords: gender identity, femininity, masculinity, measures of norms, life satisfaction, subjective well-being.

JEL Codes: I31, J16, Z10

*Danzer: Freie Universität Berlin, Berlin School of Economics, ifo Institute, CESifo, IZA, RF Berlin, natalia.danzer@fu-berlin.de; Kranton: Duke University, rachel.kranton@duke.edu; Larysz: Freie Universität Berlin, WZB, piotr.larysz@fu-berlin.de; Senik: Sorbonne University, Paris School of Economics, and IUF, claudia.senik@psemail.eu.

Acknowledgments: This research was funded by the French National Research Agency through the Investissements d'avenir framework ANR-17-EURE-0001, IUF, CEPREMAP, and the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Project numbers 465126436; 518302089. We thank participants in the GLO JOPE 2025 conference; LSE Wellbeing seminar; EHESS Sciences&Bonheur Conference; BeWell Meeting 2026 on Happiness Research for their comments. All errors remain ours.

1 Introduction

Are gender identity and norms related to people's happiness? Extensive research has documented persistent gender gaps in many realms, including occupations and time use[1], but very little evidence has been provided of the relationship between subjective well-being and the differences in the behaviors of men and women (Mencarini and Sironi 2012; Booth and Van Ours 2009). This paper examines identity and gender-related norms through the lens of subjective well-being.

To do so, we exploit the new module of the cross-national European Social Survey (2024), hereafter ESS, which asks respondents how much they feel feminine and masculine, and we construct measures of the gender typicality of respondents' behaviors; i.e. the extent to which their behavior and traits match those which distinctively characterize men and women in Europe. We relate these measures to individuals' reports of life satisfaction.

The results show a robust asymmetry between men and women. For men, feeling masculine, behavior and traits that are more typical of men, and life satisfaction are all positively cross-correlated. For women, however, while feeling feminine and life satisfaction are positively correlated, it is behavior and traits that are more typical of men (instead of women) that are associated with higher life satisfaction. This pattern is particularly strong when such behavior and traits relate to labor market behavior, personality, and political attitudes.

In a detailed regional analysis, we find that feeling more feminine for women, and more masculine for men, is consistently associated with higher life satisfaction. However, the effect of amplifying or violating gender norms varies across geographies, possibly reflecting

[1]Observed gender differences have alternatively been attributed to innate preferences and natural comparative advantages (Becker (1981)), discrimination (Becker (1957); Arrow (1973)), and, more recently, to social norms, culture, and notions of gender identity (Akerlof and Kranton (2000); Goldin (2002)). Experimental and empirical research suggests that women "shy away from competition" while men embrace it; women are more risk-averse, less patient, and express less desire to negotiate than men (see, respectively, Niederle and Vesterlund (2007); Croson and Gneezy (2009); Eckel and Grossman (2008); Falk et al. (2018); Shurchkov and Eckel (2018)). Women also show a preference for working with "people rather than things" (Su et al. (2009); Fortin (2008)) and for finding meaning at work (Burbano et al. (2024)), and perceive different trade-offs between certain working conditions, including commuting time and pay (Goldin (2014); Bertrand et al. (2015); Petrongolo and Ronchi (2020); Le Barbanchon et al. (2021)

different gender cultures that prevail in European regions.

This study advances the economic literature on gender identity and norms. Previous theoretical and empirical studies have illustrated the importance of gender norms for differentiated behavior (Cortés et al. 2025; Coffman and Klinowski 2025; Bursztyn et al. 2023) and how they persist over time through inter-generational transmission and social contact (see, for example, Bisin and Verdier 2001; Fernández 2011; Cortes and Pan 2018; Bertrand 2011). Other work has shown that institutions and policies can affect gender norms and behavior (Lippmann et al. (2020)). However, these studies do not have access to direct measures of the subjective intensity of gender identity, which is usually presumed to align with reported sex, or is captured via the set of self-reported attitudes that are generally supposed to come with it, as in (Bem (1974), Ayyar et al. (2024), Banan et al. (2025), or De Haas et al. (2024). Only a small and recent set of papers has endeavored to measure self-perceived gender identity on a continuous scale – instead of a binary one – guided by an interest in fluid gender identity (Magliozzi et al. 2016; Kachel et al. 2016; Brenøe et al. 2022, 2024; Trachman 2022). Second, these studies consider the choices and behavior of men and women, but do not have information on their individual utility.

In this study, we use direct information on these two concepts to address a new set of research questions. First, we examine the relationship between the self-assessed intensity of gender identity and life satisfaction. Second, we investigate the relationship between gender-differentiated behavior, relative to gender-specific norms, and life satisfaction, and ask whether such behavior is associated with differential utility (subjective well-being) consequences for men and women. Do gender norms simply imply horizontal differentiation - abiding by gender roles for men and for women is equally rewarded in terms of utility- or does it imply vertical differentiation, in the sense that the norm's prescriptions for one gender are drivers of higher life satisfaction? As mentioned above, we find the latter: women who behave more like men have higher life satisfaction than other women, but the reverse is not true for men.

We derive our measures as follows from the ESS. In a novel module, respondents answer two questions: "how feminine do you feel" and "how masculine do you feel," by choosing a step on a 0-6 scale, labeled "not at all" to "very ." The answers provide our measures of subjective gender intensity. Furthermore, we agnostically, i.e., without any prior, construct measures of norms for men and women based on 119 observed choices and traits variables in the sample.[2] In what follows we use the word *behavior* as short-hand to refer to a respondents' choices and traits. We construct the typical behavior of women and the typical behavior of men, and we refer to these typical behaviors as gender norms. We then consider each individual respondent's behavior as *amplifying the norm* for their gender if their behavior is in the opposite direction of the other gender (e.g., men behaving in a way more typical of men, but less typical of women), versus *violating the norm*, if their behavior is in the direction of the other gender (e.g., men behaving in a way less typical of men, but more typical of women). Finally, we measure utility through the usual self-declared life satisfaction ("Overall, how satisfied are you with your life", 0-10 scale).[3] We use the terms life satisfaction, subjective well-being, and happiness interchangeably, as is usual in the happiness literature (OECD (2024)).

We are interested in the potential trade-off in norm compliance. On the one hand, conforming to a norm *per se* should be associated with higher life satisfaction. On the other hand, the precise content of the norm could correlate in different ways – positively or negatively– with happiness. For example, gender norms could entail different labor market behaviors, which, in turn, could affect utility directly through income and status. Our estimation necessarily measures the net effect of these two forces.

[2]We consider responses to all survey questions which relate to respondents' choices and traits and do not relate to opinions or beliefs about men and women *per se*.

[3]This measure is standard in the happiness literature; it contains both a cognitive and a hedonic dimension (OECD (2013)). Following Kahneman and Deaton (2010), we interpret self-declared life satisfaction as ex-post: "experienced utility" (i.e., what people actually feel once they experience the outcome of their decision) as opposed to ex-ante "decision-utility" (i.e., what people expect when they make their decision). Life satisfaction is currently used for public policy evaluation as well as a part of government official objectives (Stiglitz et al. (2009)). The yearly World Happiness Report (2025) is dedicated to the evolution of this metric across countries and over time. Here, we examine whether gender norms and the adherence or violation thereof are potential contributors to happiness.

The analysis provides the first empirical evidence directly connecting self-assessed identity and well-being. We first examine the relationship between subjective gender intensity and life satisfaction and find a systematic positive correlation for both men and women. We also find that a stronger self-reported gender identity is associated with a higher probability of amplifying gender norms for both men and women. This result mirrors that in Brenøe et al. (2024, 2022), who introduced a continuous measure of subjective gender intensity from masculine to feminine in surveys of different populations and found that this measure adds explanatory power to the gender gaps in behaviors and preferences beyond the usual binary indicator.[4]

Second, we find robust relationships between life satisfaction and amplification or violation of gender norms for men and women. For men, amplifying the gender norms has a well-being advantage, whereas for women, amplifying the gender norm is negatively correlated with life satisfaction. Conversely, violation of gender norms is positively correlated with life satisfaction for women, but not for men.

The results indicate a certain degree of verticality in gender norms: norms for women, in contrast to those for men, seem to inherently involve behaviors and choices that, on average, lower life satisfaction.[5] We investigate the domains for which this pattern holds more or less strongly. We find a clear asymmetric pattern in the labor market domain, political attitudes, behavior relevant to health outcomes, and personality profiles (as captured by questions on personal values and habits). The pattern does not hold for other domains, such as religion, social behavior, and trust.

To consider whether the patterns hold in different geographic regions, we take the construct of the typical behaviors of men and women in the full sample, which we call the European norm, and consider sub-samples of the four main regions in Europe (Nordic, Western, Southern, and Central and Eastern). We find that, while the magnitudes differ, feeling more

[4]Trachman (2022) also uses self-reported measures of gender identity in a cross-section of French respondents.

[5]This pattern is consistent with other evidence that transgressing traditional gender roles for women can improve individual outcomes (Bursztyn et al. (2020); Fernández (2013)).

masculine for men and more feminine for women is associated with higher life satisfaction in each region. The association is particularly strong in Central and Eastern Europe. By contrast, amplifying or violating the European gender norms differs by region. Overall, Western Europe shows the clearest gender divergence, with men facing penalties while women gain utility from violating the European gender-specific norm. In Nordic countries, however, for women, violating the norm carries a large and positive correlation with life satisfaction, possibly suggesting that norm violation involves little or no well-being penalty.

We then consider the possibility that the gender norms themselves might differ by region. We construct the typical behavior of men and the typical behavior of women within each region, and *violation* or *amplification* is measured with respect to this regional standard. The patterns are unchanged for men. By contrast, in Nordic countries, the coefficient on norm *violation* for women loses its statistical significance.

These findings, highlighting the specificity of Nordic countries, are consistent with the aforementioned dual influence of gender norms on subjective well-being. In Nordic regions, where the gap between behavioral scripts for men and women is narrow, women who deviate from traditional European roles report greater life satisfaction, likely due to low social penalties for non-conformity. However, when norms are measured regionally, the already progressive context diminishes the utility gains of norm violation for women.

Finally, the difference in the association of violating European versus local gender norms likely reflects the diverse gender cultures across European regions. Nordic countries are widely recognized for their egalitarian values in both the workplace and the household (Blau and Kahn (2013); Algan et al. (2013)), whereas Central and Eastern European countries tend to exhibit more traditional domestic norms despite relatively gender-equal labor market participation (Fortin (2005); Campa and Serafinelli (2019)). These gender norms are embedded, shaped by historical and institutional trajectories, and largely exogenous to individual behavior (Alesina et al. (2013)), supporting the interpretation that deviations from or adherence to these standards can influence subjective well-being.

The paper proceeds as follows: Section 2 outlines the theoretical framework of identity-based utility. Section 3 describes the data, including the construction of measures which capture behavior that *amplifies* or *violates* gender norms. Section 4 presents the estimation model and results, examining the relationship between gender identity, *violation* or *amplification* of gender-differentiated standards, and life satisfaction, including regional heterogeneity and robustness checks. Section 7 discusses the broader implications of the findings and concludes.

2 Theoretical Framework

We consider AkerlofKranton2000's theoretical framework to guide our study. Their model introduces social identity in a utility function and emphasizes identity-based pay-offs from people's actions. Identity refers to an individual's sense of self, given prevailing social categories and the associated prescriptive characteristics and behaviors, e.g., how people in different social categories should look and behave. The identity framework posits that the divergence or convergence of individual choices and traits to these prescriptive norms has implications for a person's "identity utility," through both self-image and social image.

We present here a version of the model which focuses on gender and individual choices and traits; we then relate the model's ingredients to the data at hand.

Consider a population of individuals in a society where each person i has a set of given individual characteristics, which we denote ϵ_i, and i can choose a set of actions a_i from some set A_i. The characteristics cannot be changed by the individual and include physical attributes, such as height, and social attributes, such as country of origin. The actions, given the attributes, affect a person i's utility, denoted W_i, through standard benefits and costs such as the wages from taking on a certain job and the effort exerted at work.

To add identity to the model, we consider that society consists of different social categories of people, the set of which we denote as C. Associated with each social category are prescribed

characteristics and actions; that is, how people should look and act. We let the notation N summarize these norms. An individual i's utility then also depends on the extent to which their individual attributes, denoted ϵ_i, and actions, denoted a_i, match to the norms for their category, which we denote $c_i \in C$:

$$W_i(a_i; \epsilon_i, c_i, N)$$

We consider a simple case with the following specification of this overall utility:

$$W_i(a_i; \epsilon_i, c_i, N) = \pi_i(a_i; \epsilon_i) + u_i(a_i; \epsilon_i, c_i, N) + s_i(\psi_i(a_i; \epsilon_i, c_i, N)) \tag{1}$$

Where:

π_i is standard utility

u_i utility from the match of choices and traits to prescriptions

$s_i(\psi_i)$ utility from subjective gender intensity ψ_i.

The first term is utility which accrues to i from i's choice of action given ϵ_i as in a standard model; utility is from consumption of goods and services and general costs and benefits such as status,which are not related to social categories and norms. The second term is the utility i receives from conformity to norms, i.e. the extent to which i's actions and characteristics match the norms for their category. The final term is the utility from the subjective feelings of belonging to their category and not to another category, which could also partly come from the match between the person's actions and characteristics and the norms for their category as opposed to other categories.

In the simplest case, individual i chooses their actions a_i to maximize W_i, given their category c_i, their own attributes ϵ_i and the prescriptive norms N. Undertaking an action then has potentially three different effects on utility - through π_i, through u_i, and through $s_i(\psi_i)$. Amplifying or violating a norm then entails possible trade-offs. For example, in the case of a woman, violating the norm for women by working at a job which is more typical for

men, say, may increase utility in terms of wages but decrease utility through the violation of the norms. Actions might also affect a person's subjective belonging to their category $s_i(\psi_i)$, since acting against the norms may make one feel less part of one's category.

We connect the model to the ESS data as follows. The relevant social categories C in our setting are men and women[6].

As mentioned earlier, we consider respondents' reports of life satisfaction as a representation of overall utility, W_i. For subjective gender identity s_i, for men and for women we consider the responses to the question about how masculine they feel and how feminine they feel.

Since prescriptive norms, N, as such are not contained in the data, we impute these gender norms by considering the typical behavior of men and women in the sample. This method follows the idea that ultimately people follow prescriptive norms, so average behavior reflects the prescriptions N. Specifically, we agnostically consider the universe of surveyed individual behaviors and traits available and identify those which are distinctive of men and women. We then define an individual woman (man) as *violating* (resp. *amplifying*) a norm when their action is in the opposite (resp. same) direction of the typical action of their category.

3 Data

3.1 Survey Questions and Summary Statistics

We use micro data from the 11th round of the ess2024, which includes 45,312 observations from 29 European countries, listed in Table A.2. This dataset is among the first internationally representative general population surveys that contains information not only on respondents' gender category but also on their self-perceived intensity of how masculine and

[6]The ESS survey allowed other responses, but the vast majority of respondents (99.85%) identified as one of these two categories.

feminine they feel.

Individuals first self-classify their gender by answering the ESS question "Which of the options on this card best describes you?" with one of the categories: "a man", "a woman", or "other". In our final regression sample 46.52% respondents identify as men and 53.48% as women. In the original data, 72 individuals (0.15% of the sample) answered "other", "don't know", or "prefer not to answer" and are excluded from the estimations.

Respondents' subjective gender intensity is then assessed with two continuous rating scales, asking separately how masculine and how feminine they feel:

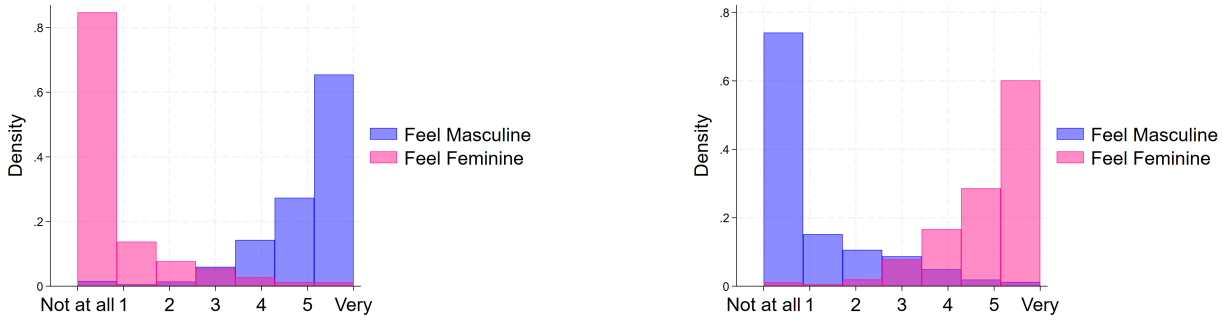
> "Regardless of whether they were born male or female, people sometimes differ in how masculine or feminine they feel.
> - Overall, how masculine would you say you feel?
> - Overall, how feminine would you say you feel?"

Responses are provided on a 7-point scale, with 0 indicating "Not at all masculine/feminine" and 6 indicating "Very masculine/feminine". The data display a clear pattern: women report higher femininity and lower masculinity, while men show the opposite pattern. Among women the average femininity rating is 5.20 (std.dev. = 1.10) and the average masculinity rating is 0.78 (std.dev. = 1.28), whereas among men the average masculinity rating is 5.28 (std.dev. = 1.08) and the average femininity rating is 0.57 (std.dev. = 1.11). Figure 1 visualizes the distributions of the feeling of femininity and masculinity: for both men and women, the answers on the two scales are skewed in opposite directions, yet they span the full range and the distributions overlap. Within each gender group, the Kolmogorov–Smirnov test confirms that the distributions of feeling feminine and feeling masculine are significantly different: *Feel Feminine*, $D = 0.88$, $p < .001$; *Feel Masculine*, $D = 0.85$, $p < .001$.

For behavior relative to gender norms, we use the extensive data on respondents' behavior and traits, detailed in the following section.

Our main proxy for utility is the life satisfaction question: "All things considered, how satisfied are you with your life as a whole nowadays? Please answer using this card, where 0

Figure 1: Distribution of Feeling Feminine and Feeling Masculine by Gender



(a) Men (b) Women

Notes: Figures display kernel density histograms derived from our own calculations on the baseline sample.

means extremely dissatisfied and 10 means extremely satisfied." On average, women report a life satisfaction of 7.18 (std.dev. = 2.05) and men report 7.24 (std.dev. = 2.01).

Summary statistics of all these variables are presented in Table 1.

Table 1: Summary Statistics

MEN (N=22,440)	**Mean**	**Std. Dev.**	**Min**	**Max**
Life Satisfaction	7.03	2.12	0	10
Feel Feminine	0.60	1.20	0	6
Feel Masculine	5.20	1.18	0	6
WOMEN (N=26,501)				
Life Satisfaction	6.96	2.17	0	10
Feel Feminine	5.11	1.17	0	6
Feel Masculine	1.01	1.55	0	6

Notes: Survey weights are applied. Numbers are derived from our own calculations on the baseline sample and rounded to two decimals.

In the robustness checks, we use additional control variables from the survey (household income, marital status) as well as matched aggregate measures of shares of women within industries and occupations[7] in the countries examined[8], which are retrieved from Eurostat and are based on the *EU Labor Force Surveys* for 2022 (Eurostat (2023)). To approximate individual income levels, we draw on the *EU Statistics on Income and Living Conditions* for 2022 from eusilc2023. We assign

[7]Eurostat provides shares of women for 1-digit occupation (ISCO 0 - 9) times 1-digit industry code (NACE 38 industries), overall 380 cells per country.

[8]For respondents from Israel, we impute the sample-wide average, as Eurostat does not collect data for Israel. For the United Kingdom, we rely on the most recent available data, from 2019.

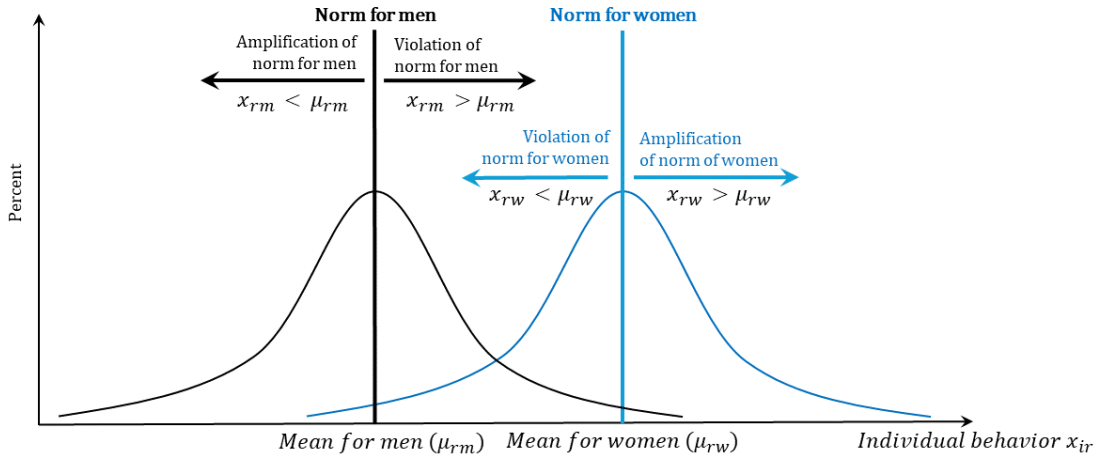
occupational prestige scores at the 4-digit ISCO level using the scores provided by G1E4BF$_2$022 $for respondents who have an ISCO code. For respondents without an ISCO code, we apply the stan$
$assigning a value of zero and accounting for missing prestige scores with a dummy indicator.$

All descriptive statistics and estimates that follow are weighted using population design weights and country weights, as recommended by the ESS.[9]

3.2 Construction of Gender Norm *Amplification* and *Violation*

Here we describe how we construct from the ESS data individual measures of the extent to which each person's behavior is more or less similar to those of their own and opposite gender.

Figure 2: Illustration of Gender Norm Violation and Amplification
(for a single gender marker r)



Before explaining the construction in detail, we first provide a brief overview of the main idea. We begin by identifying individual traits and behaviors that are distinctively typical of men or woman, which we call *gender markers*. We then compute the typical (mean) behavior of each gender, and construct a gender-typicality score for each respondent, t_i,

[9]We use analysis weights provided by ESS which adjust for selection probabilities, non-response, non-coverage, sampling error (based on four post-stratification variables), and population size differences across countries. They are created by calculating the design weight, applying post-stratification and population adjustments.

based on their own score on the gender markers. We then say an individual respondent *violates* or *amplifies* the norms by considering their behavior relative to the mean, where *violation* is in the direction towards the other gender and *amplification* is the direction away from the other gender.

Figure 2 illustrates violation and amplification for a single gender marker r. Let μ_{rm} denote the mean for men, let μ_{rw} denote the mean for women, and suppose $\mu_{rm} < \mu_{rw}$, so higher levels of this behavior are more typical for women. We say a woman i whose behavior, x_{ir}, is greater than μ_{rw} is amplifying the norm for women, and a woman whose behavior is less than μ_{rw} is violating the norm for women. A parallel vocabulary describes the behavior of men.

Step 1. Identifying gender markers

We begin by identifying gender markers, i.e., individual traits and behaviors that significantly predict whether a respondent states they are a man or a woman. Specifically, we estimate a Linear Probability Model (LPM), where the dependent variable man_i equals 1 if individual i reports being a man and equals 0 otherwise. Let B be the set of 119 individual-level potential explanatory traits and behaviors,[10] and let x_{ib} be individual i's report of $b \in B$. The probability model is specified as:

$$\text{man}_i = \alpha + \sum_{b \in B} (\beta_b x_{ib}) + \varepsilon_i \tag{2}$$

where α is a constant and ε_i denotes the error term.[11]

[10] We include variables that reflect behaviors, traits, and choices of the respondent (e.g. sector of employment, but not health problems). We exclude variables that do not point to an individual (for example the respondent's partner education level, or the composition of the household). We also exclude country-specific variables, questions that are not asked to all respondents, double-coded items, and technical variables (e.g. interview date). We further exclude variables that serve as proxies for our dependent variable, such as mental health outcomes. Lastly, we exclude questions on gender role opinions that we see as an outcome of gender identity and norm compliance rather than a component of the latter. A complete list of included and excluded variables is provided in the Online Appendix B.

[11] In cases of item non-response, when individual i has a missing value on a variable $b \in B$, we substitute the missing entry with a zero along with a variable-specific dummy to account for this imputation.

Standard errors are clustered at the country level[12] to account for possible within-country correlation in how respondents report behaviors. From this model, we retain those variables in B for which the estimated coefficient $\hat{\beta}_b$ is a statistically significant predictor of man_i at the 1% level. The resulting subset, denoted by $\mathcal{R}$, consists of 53 variables which we call the set of *gender markers*.

Next, we estimate the relative importance of each gender marker in predicting a respondent's gender. We then transform these estimates into marker-specific relative weights, which will be used to construct an individual's typicality score in Step 2. With slight adjustment of notation, let x_{ir} now denote individual i's behavior for gender marker $r \in \mathcal{R}$. To ensure comparability across markers measured on different scales, we first standardize each variable for each individual to a z-score with mean zero and standard deviation one, as follows:

$$z_{ir} = \frac{x_{ir} - \mu_r}{\sigma_r}$$

where μ_r and σ_r are the mean and standard deviation of marker r, computed across all individuals in the sample.

Next, we estimate a LPM as in (1) restricted to these standardized gender markers (z-scores):

$$\text{man}_i = a + \sum_{r \in \mathcal{R}} (\rho_r z_{ir}) + e_i \tag{3}$$

where a is constant and e_i denotes the error term. An estimated coefficient $\hat{\rho}_r > 0$ indicates that marker r is positively predictive that the respondent is a man, while a coefficient $\hat{\rho}_r < 0$ indicates that marker r is negatively predictive for being a man. For women, the same reasoning applies, but with opposite sign.

We use the estimated coefficients $\hat{\rho}_r$ for all $r \in \mathcal{R}$ to construct a relative weight denoted $\gamma_r(g)$ for each marker r, separately for each gender $g \in \{\text{man}, \text{woman}\}$. To do so, we divide

[12] We also skip clustering at the country level and obtain similar results.

each estimated coefficient $\hat{\rho}_r$ by the sum of the absolute values of all estimated coefficients for all gender markers.

For men, a marker-specific weight is, hence, constructed as:

$$\gamma_r(\text{man}) \equiv \frac{\hat{\rho}_r}{\sum_{r \in \mathcal{R}} |\hat{\rho}_r|}$$

For the weights and markers to be predictive of being a woman, we need to reverse the estimated coefficients $\hat{\rho}_r$ accordingly by multiplying the weights by (-1) :

$$\gamma_r(\text{woman}) \equiv (-1) \cdot \gamma_r(\text{man})$$

Step 2. Typicality scores

Next, we construct a typicality score for each individual, which provides a single summary measure of a person's behavior and traits across the 53 gender markers. To do so, we first standardize each marker, expressing an individual's behavior relative to others of the same gender for that specific marker. For each individual i, gender $g \in \{\text{man}, \text{woman}\}$, the standardized value for each marker r, τ_{irg}, is computed as:

$$\tau_{irg} = \frac{x_{irg} - \mu_{rg}}{\sigma_{rg}}$$

where μ_{rg} and σ_{rg} are the mean and standard deviation of marker variable r among gender g. This transformation ensures that each marker has a mean of 0 and a standard deviation of 1 within each gender. We then compute the weighted sum of the standardized marker values over all markers $\mathcal{R}$, using the respective gender- and marker-specific weights, $\gamma_r(\text{g})$, to obtain each individual i's overall *typicality score*, denoted $t_{i \in g}$:

$$t_{i \in g} \equiv \sum_r \gamma_r(\text{g}) \cdot \tau_{irg}$$

We say an individual's behavior *violates* their gender norm when their typicality score is below zero, since their behavior is more like the opposite gender, and we say an individual's behavior *amplifies* the gender norm when their typicality score is above zero. (We place individuals whose typicality score is exactly equal to zero as amplifying the norm, although empirically there are few such respondents in the sample.) The *violation* and *amplification* measures for individual i in gender g are then:

$\text{Violation}_{i \in g} \equiv |t_{ig}|$ for $t_{ig} < 0$

$\text{Violation}_{i \in g} \equiv 0$ for $t_{ig} \geq 0$

and the *amplification* measure for individual i in gender g is

$\text{Amplification}_{i \in g} \equiv t_{ig}$ for $t_{ig} \geq 0$.

$\text{Amplification}_{i \in g} \equiv 0$ for $t_{ig} < 0$.

These final scores provide individual-level measures that capture both the direction and magnitude of alignment with gender-typical traits and behavior. Note that the violation measure is expressed in absolute terms to facilitate interpretation: for both scores, higher values indicate greater violation and greater amplification, respectively.

Following this approach we can construct typicality scores for individuals in any subset and for any subset of the gender markers $\mathcal{R}$ as we will do below in the regional analysis and the analysis of particular domains of behavior.

All variables used in the analysis are presented in Appendix A, Table A.1, while Table A.2 lists all countries included in our analysis.

4 Main Estimation Model and Results

Our main empirical analysis of the relation between life satisfaction, subjective gender intensity, and individual *amplification* or *violation* scores consists of two parts. First, we assess the unconditional correlations between our main variables. Second, we explore the relationship in an econometric framework, in which we also consider the role of domain-specific violation

and amplification behavior.

4.1 Correlations among Main Variables

The basic correlations among our main variables of interest are displayed separately for men and women in Table 2. Several patterns emerge, which we find to be robust throughout the entire analysis. First, as suggested by the identity framework, for women feeling more feminine is positively correlated with the *amplification* measure and negatively with the *violation* measure. This indicates that women who feel more feminine are characterized by traits and behavior that are higher on the woman typicality score. Symmetric findings hold for men. Second, stronger subjective feelings of femininity among women and stronger subjective feelings of masculinity among men are positively associated with life satisfaction, as well as with other measures of subjective well-being and mental health (see Table A.3). However, third, the correlation coefficients between subjective well-being on the one hand, and *amplification* and *violation* on the other hand, have opposite signs for men and women. Women who violate the norms for women are on average happier than their peers, whereas men who violate men's gender norms are less happy compared to other men.

Table 2: Pairwise Correlations of Gender Identity, Violation, and Amplification

MEN	**Life Sat.**	**Feel Fem.**	**Feel Masc.**	**Violation**	**Amplification**
Life Sat.	1.00				
Feel Fem.	-0.03***	1.00			
Feel Masc.	0.08***	-0.49***	1.00		
Violation	-0.06***	0.11***	-0.10***	1.00	
Amplification	0.03***	-0.09***	0.07***	-0.46***	1.00
WOMEN	**Life Sat.**	**Feel Fem.**	**Feel Masc.**	**Violation**	**Amplification**
Life Sat.	1.00				
Feel Fem.	0.12***	1.00			
Feel Masc.	-0.14***	-0.35***	1.00		
Violation	0.04***	-0.08***	0.08***	1.00	
Amplification	-0.05***	0.07***	-0.08***	-0.46***	1.00

Notes: Survey weights are applied. Numbers are derived from our own calculations on the baseline sample and rounded to two decimals. * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

where LifeSat_i is the life satisfaction score for individual i, corresponding to the utility term

U_i in the theoretical framework of Section 2. The variables Feel feminine$_i$ and Feel masculine$_i$ are the individual's self-assessed intensity of femininity and masculinity, respectively. *Violation* and *Amplification* represent the (absolute) distance to the gender-specific norm.

The term X_i is a vector of control variables. In our baseline specification, it only contains age categories in years (18-25, 26-40, 41-55, 56-65, 66+), because they are the only exogenous characteristics. In robustness checks, we additionally control for sociodemographic (potentially endogenous) characteristics of the respondent, such as marital status, household income, occupational prestige, and predicted individual income.

Country fixed effects, κ_c, are included to account for unobserved heterogeneity across countries. Standard errors are clustered at the country level.[13]

4.2 Estimation Results for Europe

Main specification

We begin by estimating equation **??** using the full European sample, sequentially introducing the individual variables: (i) feeling masculine and feeling feminine entered separately, (ii) violation and amplification measures added, and (iii) all variables included jointly. The results are presented in Table A.6 and Figure 3.

The findings show a highly significant positive association between life satisfaction and the intensity of feeling masculine for men (col. 2) and feeling feminine for women (col. 5), respectively. Consistent with predictions from identity models, this subjective dimension of gender identity thus appears to constitute a source of utility.[14]

Second, in contrast, the association between life satisfaction and the degree to which a person's *behavior* is gender-typical differs for men and women (col. 3 and col. 7). More precisely, *violation* (below-average typicality score) is associated with a lower life satisfaction

[13] As in the regressions above deriving the typicality scores, if there is a missing value for individual i in one of the variables, we substitute this missing with a zero and include a variable-specific dummy to account for this imputation.

[14] Accordingly, subjective femininity among men and subjective masculinity among women is associated with significantly *lower* life satisfaction.

Table 3: Life Satisfaction and Gender Identity, Violation, and Amplification

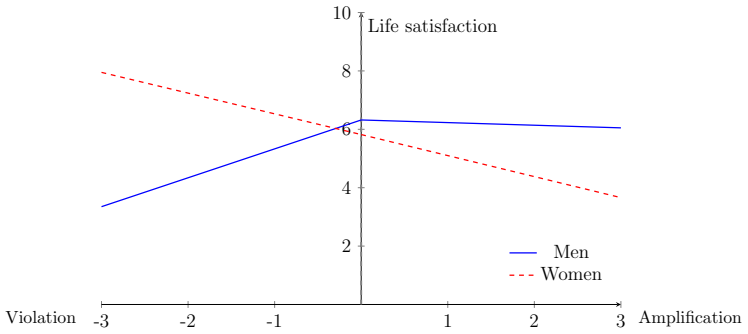
	MEN				WOMEN			
Dep. Variable: Life Satisfaction	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Feel Feminine	-0.06***			0.02	0.21***			0.22***
	(0.02)			(0.02)	(0.02)			(0.02)
Feel Masculine		0.16***		0.16***		-0.04*		0.02
		(0.03)		(0.03)		(0.02)		(0.01)
Violation of Gender Norms			-1.17***	-0.99***			0.56**	0.71***
			(0.34)	(0.33)			(0.23)	(0.23)
Amplification of Gender Norms			-0.05	-0.09			-0.64***	-0.72***
			(0.36)	(0.36)			(0.18)	(0.17)
Constant	7.15***	6.28***	7.18***	6.32***	5.91***	7.04***	7.00***	5.82***
	(0.01)	(0.14)	(0.03)	(0.14)	(0.10)	(0.02)	(0.02)	(0.10)
Obs.	22589	22589	22589	22589	26739	26739	26739	26739
R-sq.	0.10	0.11	0.10	0.11	0.13	0.12	0.12	0.13

Notes: Ordinary Ordinary Least Squares (OLS) estimates. All models include controls for age group (18–25, 26–40, 41–55, 56-65, 66+) and country fixed effects. Standard errors clustered at the country level are reported in parentheses. Survey weights are applied. Numbers are derived from our own calculations on the baseline sample and rounded to two decimals.

among men but a higher life satisfaction among women. That is, men who act more like women are less happy than other men, whereas women who act more like men are happier than other women. Conversely, *amplification* (an above-average typicality score) is negatively associated with life satisfaction among women, while for men, *amplification* shows no significant association with subjective well-being. For women (col. 7), the estimates on *violation* and *amplification* are almost identical in magnitude, which indicates that women's life satisfaction almost linearly increases as their gender typicality score moves from positive to negative values. In other words, women's life satisfaction is higher the less typical their behavior (the lower their typicality score). By contrast, for men, the coefficients are very different in magnitudes and statistical significance: men's life satisfaction decreases as their behavior gets closer to that of women, but it does not increase when their behavior exceeds the norm for men. This dichotomy is illustrated by the kink in the life satisfaction graph for men in Figure 3, where the slope of each spline is given by the estimated coefficient associated respectively with the *violation* and the *amplification* variable.

Table A.4 in the Appendix shows that the patterns of the results do not depend on whether the life satisfaction scale is treated as a cardinal measure (using OLS estimates),

Figure 3: Life Satisfaction as a Function of *Violation* and *Amplification*



Notes: Survey weights are applied. Numbers are derived from our own calculations on the baseline sample.

or is dichotomized (with a cut-off at 7 on the $0-10$ scale), or standardized at the country level. It suggests that, for men, a one standard deviation increase in *violation* is associated with a lower life satisfaction by 55% of a std. deviation, and reduces the likelihood to be highly satisfied with life (above 7 on the scale) by 17 percentage points). For women, a one std-deviation increase in violation is associated with an increase of life satisfaction by 38% std. deviation, and a rise in the likelihood of being very satisfied with life by 17 percentage points).

It is important to note that, as shown in Table A.6, when including all variables simultaneously, the associations between life satisfaction and subjective identity (feeling feminine for women; feeling masculine for men) remain unchanged. Similarly, the magnitude and significance of the coefficients on *violation* and *amplification* are not affected by including our measures of subjective gender intensity. This pattern suggests that the correlation between these two sets of variables is not measurably relevant for life satisfaction and that they each have a distinct, separate influence on a person's subjective well-being.

Alternative Measures of Subjective Well-Being

Table 4 presents the estimates of equation 3 using the other classical indicators of subjective well-being and mental health available in the survey as proxies of utility, instead of life satisfaction.

The questions are the following: "Taking all things together, how happy would you say you are?" $(0-10$ scale); "I will now read out a list of the ways you might have felt or behaved during the past week. Using this card, please tell me how much of the time during the past week you felt: depressed/ everything you did was an effort/ your sleep was restless/ you were happy/ you felt lonely/ you enjoyed life/ you felt sad/ you could not get going". (Answers from 1: "none of the time" to 4: all the time"). [15]

With each of these five alternative indicators of subjective well-being our main findings from Table A.6 are preserved and confirmed.

[15]These variables are significantly correlated with our main measure of life satisfaction (see Appendix Table A.3).

Table 4: Robustness. Alternative Subjective Well-Being and Mental Health Outcomes

	MEN						WOMEN					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	Life Sat.	Happy	Flt. Depr	Flt. Eff.	Flt. Sad.	Flt. Happy	Life Sat.	Happy	Flt. Depr	Flt. Eff.	Flt. Sad.	Flt. Happy
Feel Feminine	0.02	0.03*	0.03***	0.03**	0.03***	0.02***	0.22***	0.29***	-0.06***	-0.07***	-0.07***	0.11***
	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.03)	(0.01)	(0.01)	(0.01)	(0.01)
Feel Masculine	0.16***	0.23***	-0.06***	-0.05***	-0.05***	0.10***	0.02	0.03	0.02**	0.02	0.01	0.02**
	(0.03)	(0.03)	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)	(0.04)	(0.01)	(0.01)	(0.01)	(0.01)
Violation of Gender Norms	-0.99***	-0.86**	0.61***	0.57***	0.45***	-0.24*	0.71***	0.40	-0.10	-0.19*	-0.17	0.11
	(0.33)	(0.42)	(0.08)	(0.10)	(0.08)	(0.12)	(0.23)	(0.24)	(0.09)	(0.11)	(0.10)	(0.10)
Amplification of Gender Norms	-0.09	-0.98**	0.07	0.04	-0.04	-0.23	-0.72***	-0.35	0.59***	0.67***	0.39***	-0.22
	(0.36)	(0.42)	(0.09)	(0.14)	(0.08)	(0.16)	(0.17)	(0.26)	(0.10)	(0.10)	(0.10)	(0.14)
Constant	6.32***	6.29***	1.61***	1.80***	1.70***	2.45***	5.82***	5.87***	1.80***	2.08***	2.02***	2.32***
	(0.14)	(0.13)	(0.05)	(0.08)	(0.06)	(0.05)	(0.10)	(0.18)	(0.06)	(0.08)	(0.08)	(0.06)
Obs.	22589	22435	22493	22500	22473	22426	26739	26473	26628	26627	26624	26549
R-sq.	0.11	0.08	0.05	0.05	0.05	0.06	0.13	0.10	0.05	0.05	0.06	0.07

Notes: Ordinary Least Squares (OLS) estimates. All models include controls for age group (18–25, 26–40, 41–55, 56-65, 66+) and country fixed effects. Standard errors clustered at the country level are reported in parentheses. Survey weights are applied. Numbers are derived from our own calculations on the baseline sample and rounded to two decimals. * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$. *Flt. Depr.* reflects responses to the question: "How much of the time during the past week have you felt depressed?" Responses range from 1 (None or almost none of the time) to 4 (All or almost all of the time). *Flt. Eff.* captures responses to the question: "How much of the time during the past week have you felt that everything you did was an effort?" Responses range from 1 (None or almost none of the time) to 4 (All or almost all of the time). *Flt. Sad* refers to answers to the question: "How much of the time during the past week have you felt sad?" Responses range from 1 (None or almost none of the time) to 4 (All or almost all of the time). *Were Happy* captures responses to the question: "How much of the time during the past week were you happy?" Responses range from 1 (None or almost none of the time) to 4 (All or almost all of the time). *Happy* is based on the question: "Taking all things together, how happy would you say you are?" Responses range from 0 (Extremely unhappy) to 10 (Extremely happy).

Domains of Gender Difference

The above results reveal a strong asymmetry in how norm violation (or amplification) relates to the happiness of women versus men. Most strikingly, norm violation relates to greater happiness for women, but lower happiness for men. To investigate this pattern in more detail, we repeat our analysis considering the gender typicality of a person's behavior and traits with respect to different life domains. For this purpose, we group the set of gender markers by topic, and generate six new, domain-specific typicality scores for men and women, yielding domain-specific amplification and violation scores: (i) health-related behavior (7 markers), (ii) religion (2 markers), (iii) social behavior and trust (4 markers), (iv) personality (11 markers), (v) work (15 markers), and (vi) political attitudes and behavior (13 markers). The list of gender markers by domain appears in the online Appendix B.

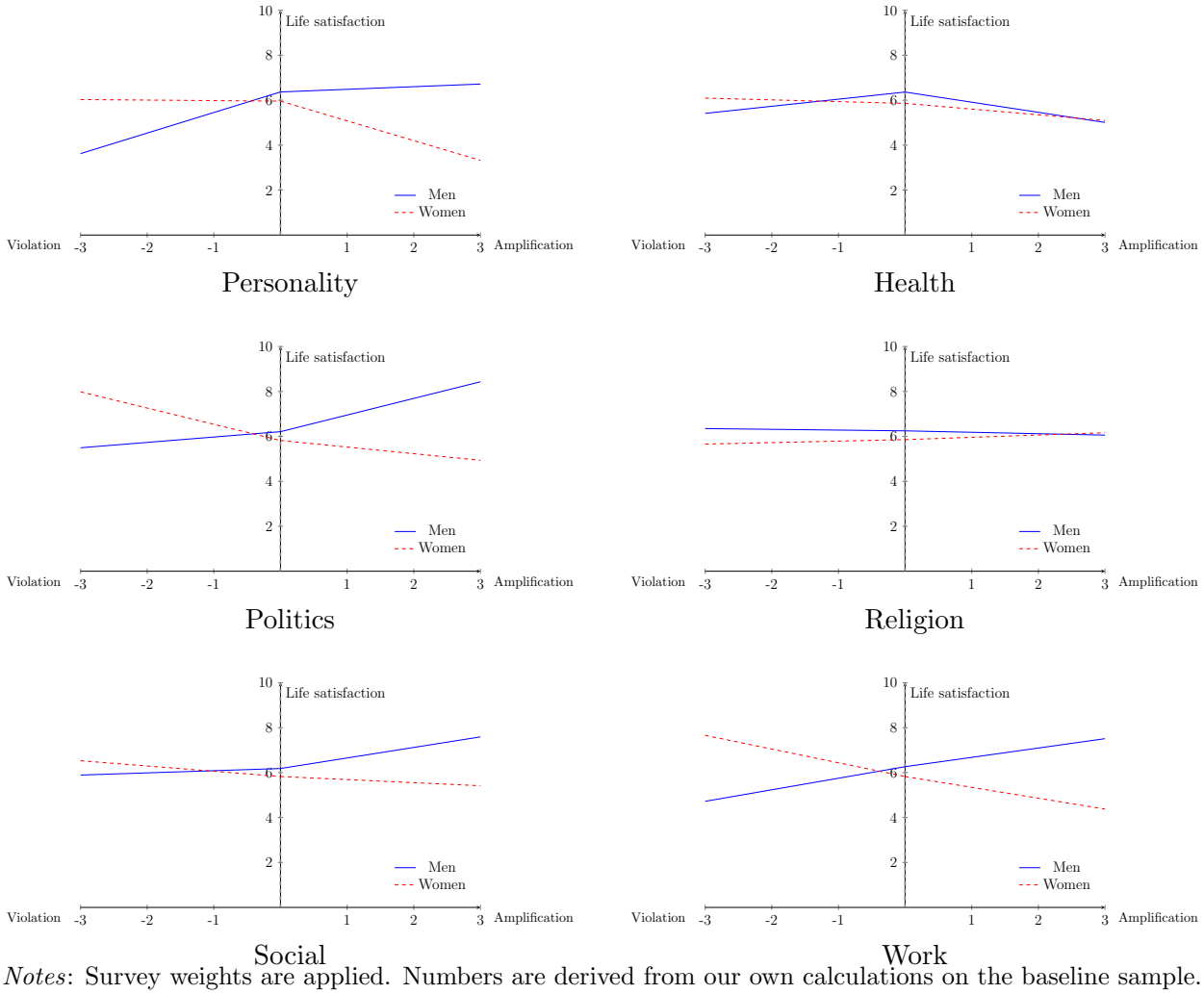
It appears (Figure 4) that it is in the labor market domain, political attitudes, behavior relevant to health outcomes, and personality profiles (as captured by questions on personal values and habits) that this asymmetric pattern is most clear. The pattern does not appear for other domains, such as religion, social behavior, and trust.

More on the Labor Market Domain

Do our estimates isolate the relation to gender norm violation or amplification as such, or do they rather capture correlated labor market characteristics, such as earnings? To address this question, we add to our main regressions some controls that could be related to norms for men and for women but are not included in the variables used to define the gender markers. These additional controls are: household income, marital status, occupational prestige, and the predicted income of the respondent. We include these variables in the estimates alternatively, one by one.

We derive these controls as follows: Because the ESS does not contain individuals' incomes, we predict them using the *EU Statistics on Income and Living Conditions* for 2022 from eusilc2023, based on the estimates of a Mincer earnings function, which

Figure 4: Life Satisfaction as a Function of *Violation* and *Amplification* in Different Domains



Notes: Survey weights are applied. Numbers are derived from our own calculations on the baseline sample.

incorporates the respondent's gender, presence of children, age, marital status, country of residence, industry of employment, and education level as predictors.[16] We assign an occupational prestige score to each individual in the sample by matching their occupational code (ISCO code) with the prestige ranking established by G1E4BF$_2$022.$Onceagain, wereplacemissingvalueswithazeroandcontrolforthisimputationusingadummy.$

The estimations, reported in Table 5 below, indicate that including measures of household income, predicted personal income, or occupational prestige, does not change the general patterns of the main results. For men, the coefficient on *violation* remains negative and statistically significant in all cases. However, when household income is included in the regression, the magnitude of the (negative) coefficient on *violation* is reduced, suggesting that part of the negative association between *violation* and life satisfaction is due to lower status. For women, the coefficients on *violation* show the same pattern, but with the opposite sign: only when controlling for household income is the positive coefficient on *violation* reduced, indicating that part of the positive association between *violation* and life satisfaction for women is due to higher household income. The magnitude of the negative association between life satisfaction and *amplification* is slightly reduced once we control for the income measures (especially household income), but becomes slightly larger (more negative) when controlling for marital status.

[16]Because this imputed individual income is estimated, we bootstrap the estimates that include it as a control. The bootstrap is used here to obtain standard errors by resampling the data. We ran 1,000 replications with a fixed random seed (12345) to ensure replicability.

[17]This applies to 11.22% of the observations in our baseline sample. Specifically 1.02% are missing 8.18% are not applicable, 0.73% are refusals, 0.21% did not know, and 1.08% did not answer

Table 5: Robustness. Additional Controls

	MEN					WOMEN				
Dep. Variable: Life Satisfaction	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Feel Feminine	0.02	-0.01	0.02	0.01	0.01	0.22***	0.21***	0.21***	0.22***	0.21***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Feel Masculine	0.16***	0.16***	0.15***	0.17***	0.15***	0.01	0.02	0.02	0.01	0.01
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Violation of Gender Norms	-0.99***	-0.57*	-0.91***	-0.85**	-0.80**	0.72***	0.38**	0.73***	0.65***	0.68***
	(0.32)	(0.29)	(0.30)	(0.32)	(0.29)	(0.23)	(0.17)	(0.23)	(0.22)	(0.23)
Amplification of Gender Norms	-0.14	-0.32	-0.06	-0.08	-0.04	-0.74***	-0.45**	-0.92***	-0.61***	-0.56***
	(0.39)	(0.38)	(0.36)	(0.39)	(0.34)	(0.17)	(0.20)	(0.16)	(0.18)	(0.16)
Constant	6.31***	6.34***	6.36***	5.63***	6.24***	5.83***	5.88***	5.91***	5.37***	5.84***
	(0.15)	(0.13)	(0.14)	(0.15)	(0.15)	(0.09)	(0.10)	(0.09)	(0.10)	(0.10)
Obs.	22258	22258	22258	22258	22258	26397	26397	26397	26397	26397
R-sq.	0.11	0.15	0.13	0.12	0.13	0.13	0.15	0.15	0.13	0.14

Notes: Ordinary Least Squares (OLS) estimates. All models include controls for age group (18–25, 26–40, 41–55, 56-65, 66+) and country fixed effects. Standard errors clustered at the country level are reported in parentheses. Survey weights are applied. Numbers are derived from our own calculations and rounded to two decimals. * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$. *Household Income* refers to the total net income from all sources, reported in deciles. *Married* is a binary indicator equal to 1 if the respondent is legally married or in a registered civil union, and 0 otherwise. *Occupational Prestige* is assigned based on the prestige rankings from $G1E4BF_2022, matchedtorespondentsviatheiroccupationcodes.Ifprestigedataismissing, itisimputedwithavalueof0, andaseparateindicatorvaria$

Altogether, these results indicate that the intensity of gender identity is a factor of higher utility (life satisfaction) as such. However, the typical traits and behavior for each gender are not equally rewarding in terms of life satisfaction. Gender norms carry a sort of verticality, with norms for men generating higher life satisfaction, in particular in the paid work domain.

5 Regional Heterogeneity

The above results are based on measures of gender norms constructed at the European level; the estimates give an association between life satisfaction and gender norm *violation* or *amplification* using the entire European sample. We now consider potential regional heterogeneity.

Following the standard geographical classification according to EuroVoc,[18] we divide the set of European countries into four large regions, which are likely to exhibit different behaviors of men and women. We first investigate whether the main results hold in each of these four regions. That is, as before, we use a respondent's typicality score vis-à-vis the European average and estimate the model in each subsample. We then consider the possibility that the gender norms themselves could be different across regions and repeat the full exercise within each region: identifying the gender markers and constructing the typicality scores for respondents with respect to the region-specific average behavior of men and women.

We conduct these analyses to study possible regional heterogeneity in the relation between life satisfaction and gendered behavior, given possibly different norms. For example, it has been shown that several countries in Western Europe have achieved their "first gender revolution" in reducing labor market-related gender gaps, but not the "second gender revolution" that consists in reducing gender differences in terms of household tasks and time-use (Goldin (2014)). In contrast, Nordic countries are more advanced on both fronts. Central

[18]We follow EuroVoc with the exception of Lithuania, Latvia and Estonia, which we group as part of Central and Eastern Europe due to their shared experience of historical state socialism.

and Eastern European countries are distinguished by their socialist history that strongly reduced gender gaps in labor-market-related domains more than anywhere else, while leaving relatively intact the traditional gender differentiation in the private sphere (Knight and Brinton (2017)).

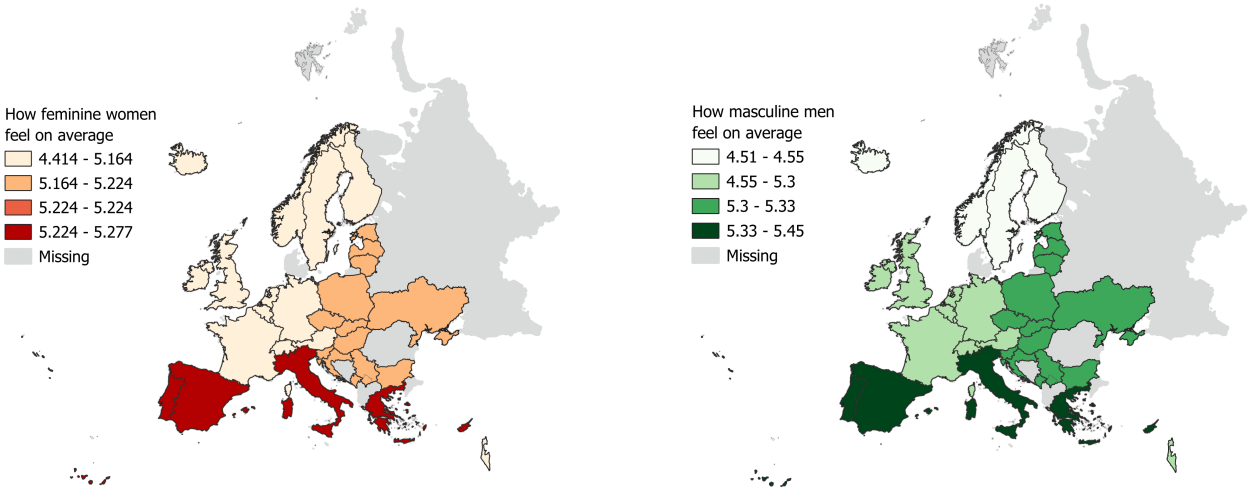
Exploiting not only individual differences, but also such regional heterogeneity helps disentangling the channels that relate gender norms and utility: matching the norms *per se*, which confer positive utility, versus the content of the norm as a source of life satisfaction. For example, women whose typicality score falls short of the average European score should be happier in countries where the role for women is less accentuated, like Norway for instance, than in more traditional countries where it is more stringent (like Spain, for instance) and where, *ipso facto*, the cost of *violation* is greater. We thus expect that in the estimates of life satisfaction, the coefficient on the gender *violation* variable —which captures the net effect of identity and *violation*— should be larger (and the benefit of *amplification* smaller) in more progressive Nordic countries than in more traditional Western European countries. This is because the toll of *violation* is lower in Nordic countries.

Before proceeding to the regressions, we first highlight the considerable regional heterogeneity in subjective gender intensity within the ESS data. Figure 5 displays the geographical maps of subjective gender intensity in Europe, based on these variables, separately for women and men.[19] The figures indicate that women and men in Southern and Central and Eastern European countries tend to report higher scores for subjective femininity and masculinity, respectively, while respondents in Northern and some Western European countries report lower scores.[20] Figures 6 and 7 show that the distribution of subjective gender intensity is quite different in Western and Southern Europe as opposed to Nordic countries, where it is much less polarized, or Central and Eastern Europe, where it is more polarized.

[19]Own calculations based on ESS. Figures display country averages of respective variables, calculated using survey weights.

[20]ANOVA with Bonferroni tests shows a clear tendency: North Europe reports the lowest mean (4.46) compared to higher but similar scores in Central (5.17), Western (5.22), and Southern Europe (5.22). Similarly, men in Northern Europe feel least masculine (M=4.52), while Central (M=5.33), Western (M=5.22), and Southern Europe (M=5.38) report significantly higher scores.

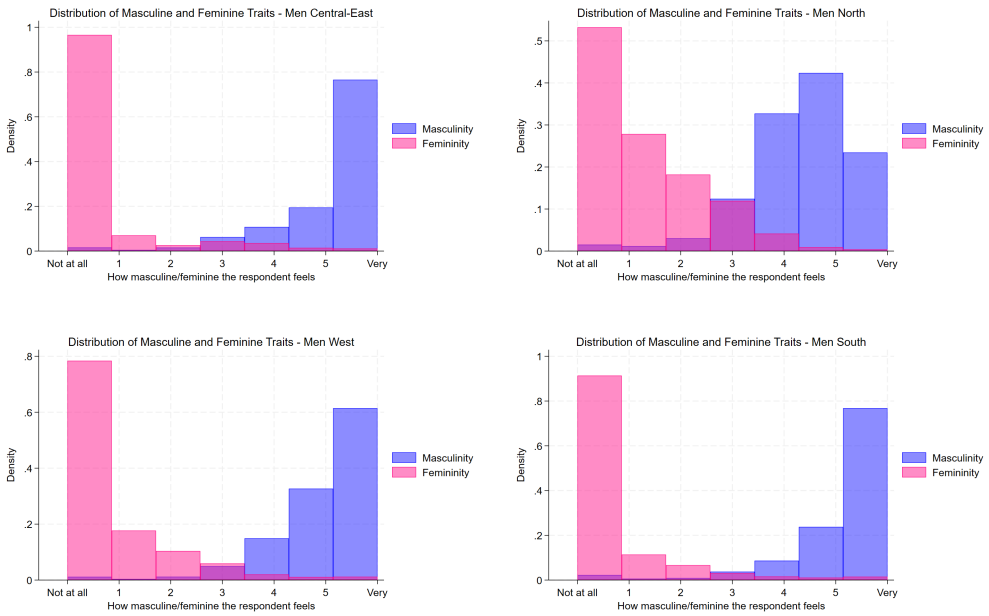
Figure 5: Feeling of femininity among women and masculinity among men in Europe



(a) Feeling feminine - women (b) Feeling masculine - men

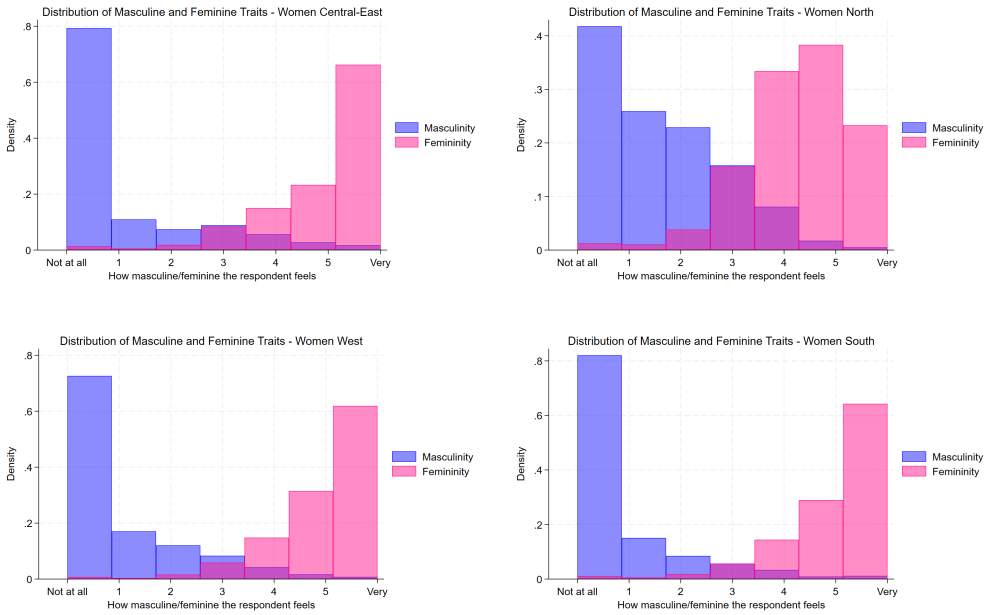
Notes: Survey weights are applied. The figures are based on our own calculations and represent averages across four European regions: Northern Europe (FI, IS, NO, SE), Central and Eastern Europe (BG, CZ, HR, HU, LT, LV, ME, PL, RS, SI, SK), Western Europe (AT, BE, CH, DE, FR, GB, IE, IL, NL), and Southern Europe (CY, ES, GR, IT, PT). The shapefiles were obtained from Eurostat$_GISCO_NUTSandmpidr2013andprocessedusingQGIS$.

Figure 6: Distribution of Subjective Identity - Men



Notes: Numbers are derived from our own calculations on the baseline sample. Northern Europe (FI, IS, NO, SE), Central and Eastern Europe (BG, CZ, HR, HU, LT, LV, ME, PL, RS, SI, SK), Western Europe (AT, BE, CH, DE, FR, GB, IE, IL, NL), and Southern Europe (CY, ES, GR, IT, PT).

Figure 7: Distribution of Subjective Identity - Women



Notes: Numbers are derived from our own calculations on the baseline sample. Northern Europe (FI, IS, NO, SE), Central and Eastern Europe (BG, CZ, HR, HU, LT, LV, ME, PL, RS, SI, SK), Western Europe (AT, BE, CH, DE, FR, GB, IE, IL, NL), and Southern Europe (CY, ES, GR, IT, PT).

Tables 6 and 7 report the results for the two sets of regional heterogeneity analyses for men and women: The estimates of the associations between life satisfaction and *violation* or *amplification* of the European norm are reported in the columns "Euro" (Columns 1, 3, 5, and 7), while the associations with the violation or *amplification* of the region specific gender norms are reported in the columns "Reg" (Columns 2, 4, 6, and 8).

The first two rows of the tables detail the relationship between life satisfaction and subjective gender intensity by region. The patterns uncovered in Table A.6 are essentially preserved as the subjective gender intensity is significantly related to life satisfaction in all regions, but the magnitude of the coefficients varies. Subjective gender intensity appears to be particularly relevant in Central and Eastern European countries (CEECs hereafter), where feeling more masculine (for men) and feminine (for women) is more strongly associated with life satisfaction. In the Nordic countries, by contrast, the subjective gender intensity matters the least for women's and men's well-being. (Note that the coefficients reported in

Table 6: Regional Heterogeneity. Men Sub-Sample

	North		**CEEC**		**West**		**South**	
	Euro	Reg	Euro	Reg	Euro	Reg	Euro	Reg
Dep. Variable: Life Satisfaction	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Feel Feminine	0.02	0.01	-0.06	-0.06	0.05*	0.04*	0.03	0.03
	(0.02)	(0.02)	(0.04)	(0.04)	(0.02)	(0.02)	(0.06)	(0.06)
Feel Masculine	0.10**	0.11***	0.24**	0.24**	0.13**	0.13**	0.15*	0.14*
	(0.02)	(0.02)	(0.08)	(0.08)	(0.05)	(0.05)	(0.06)	(0.06)
Violation of Gender Norms	-0.39	-0.00	0.24	-0.77***	-1.21**	-1.25**	-1.52**	-0.80*
	(1.13)	(0.25)	(0.30)	(0.23)	(0.42)	(0.38)	(0.47)	(0.33)
Amplification of Gender Norms	0.56	-0.10	-0.24	-0.60	0.09	-0.26	-0.57	0.14
	(0.31)	(0.21)	(0.43)	(0.44)	(0.67)	(0.38)	(0.34)	(0.23)
Constant	7.32***	7.34***	5.22***	5.34***	6.69***	6.74***	6.33***	6.27***
	(0.14)	(0.14)	(0.45)	(0.47)	(0.23)	(0.23)	(0.33)	(0.36)
Obs.	2396	2396	8761	8761	7246	7246	4186	4186
R-sq.	0.01	0.01	0.18	0.18	0.05	0.05	0.05	0.05

Notes: Ordinary Ordinary Least Squares (OLS) estimates. All models include controls for age group (18–25, 26–40, 41–55, 56-65, 66+) and country fixed effects. Standard errors clustered at the country level are reported in parentheses. Survey weights are applied. Numbers are derived from our own calculations on the baseline sample and rounded to two decimals. * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$. *Northern Europe* includes FI, IS, NO, and SE; *Central and Eastern Europe* includes BG, CZ, HR, HU, LT, LV, ME, PL, RS, SI, and SK; *Western Europe* includes AT, BE, CH, DE, FR, GB, IE, IL, and NL; and *Southern Europe* includes CY, ES, GR, IT, and PT.

Table 7: Regional Heterogeneity. Women Sub-Sample

	North		**CEEC**		**West**		**South**	
	Euro	Reg	Euro	Reg	Euro	Reg	Euro	Reg
Dep. Variable: Life Satisfaction	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Feel Feminine	0.09**	0.09**	0.31***	0.31***	0.18***	0.17***	0.22***	0.22***
	(0.02)	(0.02)	(0.06)	(0.06)	(0.01)	(0.01)	(0.02)	(0.02)
Feel Masculine	-0.02	-0.01	0.01	0.01	0.02	0.02	0.00	0.00
	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.05)	(0.05)
Violation of Gender Norms	1.14**	0.21	0.34	0.27	1.10***	0.33	0.22	0.36**
	(0.27)	(0.36)	(0.51)	(0.38)	(0.30)	(0.22)	(0.25)	(0.11)
Amplification of Gender Norms	-0.71	-0.23	-0.22	-0.92***	-0.63***	-0.55***	-0.91**	-0.63**
	(0.36)	(0.14)	(0.33)	(0.21)	(0.14)	(0.16)	(0.24)	(0.14)
Constant	7.43***	7.49***	4.74***	4.80***	6.35***	6.41***	5.81***	5.80***
	(0.08)	(0.14)	(0.36)	(0.35)	(0.04)	(0.06)	(0.12)	(0.13)
Obs.	2355	2355	11505	11505	7879	7879	5000	5000
R-sq.	0.03	0.02	0.17	0.17	0.06	0.06	0.08	0.08

Notes: Ordinary Ordinary Least Squares (OLS) estimates. All models include controls for age group (18–25, 26–40, 41–55, 56-65, 66+) and country fixed effects. Standard errors clustered at the country level are reported in parentheses. Survey weights are applied. Numbers are derived from our own calculations on the baseline sample and rounded to two decimals. * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$. *Northern Europe* includes FI, IS, NO, and SE; *Central and Eastern Europe* includes BG, CZ, HR, HU, LT, LV, ME, PL, RS, SI, and SK; *Western Europe* includes AT, BE, CH, DE, FR, GB, IE, IL, and NL; and *Southern Europe* includes CY, ES, GR, IT, and PT.

these rows are not affected by whether gender norms are calculated at the European or at the local level).

Turning to the relation between gender norm *amplification* and *violation* of the *European* norm and life satisfaction across regions, we find that, for men, the coefficient on *violation* is statistically significant and negative in Southern and Western European countries. For men, the coefficients on *amplification* remain statistically insignificant in all regions, except in the South, where it is positive. For women, the coefficient on *violation* is large and positive in all regions, but is statistically insignificant in Southern countries. In all regions except the CEECs, the coefficient on *amplification* is negative for women. Hence, overall, in most regions, women who act more like the typical European men and less like the typical European woman are more satisfied with their lives.

The finding that, in the case of CEECs, the coefficient on *amplification* is not statistically significant for women (whereas it is negative in the other regions), might be related to the region's socialist legacy. As noted, the particularity of the post-socialist countries is that gender differentiation, although small in labor market dimensions, is high in the private sphere. Hence, amplifying the norms for women in some dimensions may be a source of life satisfaction. This is consistent with the particularly large coefficient associated with subjective feeling of femininity (and masculinity) in this region.

Columns 2, 4, 6, and 8 of Tables 6 and 7 display the estimations with regionally constructed gender norms and *amplification/violation* thereof. There is some evidence that gender norms are locally different. For example, *de facto*, in Nordic countries, where gender differentiation is much more shallow, we see that violating or amplifying gender norms is not associated with any surplus in life satisfaction for women[21].

[21]In a way, there is no benefit in transgressing the local Nordic norm for women, which is already very progressive)

6 Robustness. Alternative Constructions of Typicality Measures

To test the sensitivity of our main findings, we perform several tests concerning the construction and measurement of our *violation*/*amplification* scores.

Table 8 presents estimates based on different definitions and constructions of the "typical behavior" of men and women, respectively. The *violation* and *amplification* measures are, as before, constructed as the oriented distance to this typical behavior.

To ease comparison, the first column of Table 8 recalls the results of our main specification with the baseline measures of *violation* and *amplification*. Our first alternative approach uses *violation* and *amplification* scores that are measured as the oriented distance of the behavior and traits of each respondent with respect to the average behavior of the most gender-polar men and women in the sample, i.e., men who subjectively feel entirely masculine and not at all feminine, and women who subjectively feel entirely feminine and not at all masculine (respectively 6 and 0 on the masculinity and femininity scales). Results are reported in Column 2.

Second (Column 3), we employ LASSO instead of a linear probability model to select gender markers. We set a seed (12345) to ensure reproducibility and randomly split the sample into a training set (70%) and a test set (30%). We split data into training and test to train the model on one portion and then evaluate its performance on unseen data, ensuring it generalizes well and avoids overfitting. On the training data, we then estimate a LASSO linear regression with a wide set of predictors, using the Akaike Information Criterion (AIC) for variable selection, and report the selected coefficients which we consider to be gender markers.

Third (Column 4), we construct the weights and select gender markers only on respondents with no missing values, which implies that the regressions are run on a smaller sample size compared to our baseline.

Fourth, in Column 5, we report our baseline continuous typicality score (τ_{irg}, as defined in Section 3.2, step 2). In column 6, we compare this with the same continuous typicality score, but where a probit regression is used instead of an OLS regression on the gender markers, in step one of the process. The results are highly comparable: for the baseline typicality score, the coefficients are 0.52^{**} for men and -0.83^{***} for women (s.e. 0.24 and 0.12), while for the predicted typicality score they are 0.21^{*} for men and -0.46^{***} for women (s.e. 0.12 and 0.08).

Finally, Column 7 builds the violation measure as a binary variable equal to one if the respondent violates and 0 if they conform.

All these alternative specifications of the gender norms and the corresponding *violation* and *amplification* measures produce results that are consistent with our baseline findings.

Table 8: Robustness. Alternative Measures of Gender Typicality

Dep. Variable: Life Satisfaction	(1)	(2)	(3)	(4)	(5)	(6)	(7)
MEN	Baseline	Gender-Polar	LASSO	No Missing	Baseline	Predicted	Dummy
Feel Feminine	0.02	0.02	0.02	0.02	0.02	0.02	0.02
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Feel Masculine	0.16***	0.16***	0.16***	0.16***	0.16***	0.16***	0.16***
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Violation of Gender Norms	-0.99***	-0.90***	-1.11***	-0.72***			
	(0.33)	(0.31)	(0.39)	(0.23)			
Amplification of Gender Norms	-0.09	-0.08	-0.25	0.06			
	(0.36)	(0.38)	(0.36)	(0.09)			
Typicality					0.46*	0.24*	
					(0.23)	(0.13)	
Violation (Binary)							-0.09
							(0.06)
Constant	6.32***	6.31***	6.32***	6.31***	6.25***	6.09***	6.28***
	(0.14)	(0.14)	(0.15)	(0.15)	(0.16)	(0.18)	(0.17)
Obs.	22589	22589	22589	22589	22589	22589	22589
R-sq.	0.11	0.11	0.11	0.11	0.11	0.11	0.11
Feel Feminine	0.22***	0.22***	0.22***	0.22***	0.22***	0.22***	0.22***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Feel Masculine	0.02	0.02	0.02	0.02	0.02	0.02	0.02
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Violation of Gender Norms	0.71***	0.67***	0.65*	0.39**			
	(0.23)	(0.24)	(0.35)	(0.18)			
Amplification of Gender Norms	-0.72***	-0.76***	-0.81***	-0.62***			
	(0.17)	(0.19)	(0.22)	(0.16)			
Typicality					-0.72***	-0.40***	
					(0.15)	(0.10)	
Violation (Binary)							0.15***
							(0.05)
Constant	5.82***	5.82***	5.83***	5.84***	5.82***	6.12***	5.76***
	(0.10)	(0.09)	(0.10)	(0.10)	(0.10)	(0.13)	(0.09)
Obs.	26739	26739	26739	26739	26739	26739	26739
R-sq.	0.13	0.13	0.13	0.13	0.13	0.13	0.13

Notes: Ordinary Least Squares (OLS) estimates with controls for age group (18–25, 26–40, 41–55, 56-65, 66+) and country fixed effects. Column 1 repeats the baseline measure (OLS-selected markers standardized at the European level relative to average men/women behavior). Column 2 (Gender-Polar) compares a person's typicality score to that of polar masculine men and feminine women. Column 3 uses LASSO for marker selection. Column 4 (No Missing) applies the baseline measure without imputing missing values with zero when selecting gender markers. Column 5 uses predicted probabilities from an OLS regression on gender markers. Column 6 constructs the typicality score using predicted probabilities. Column 7 builds the violation measure as a binary indicator equal to 1 if the respondent violates and 0 if they conform. Standard errors clustered at the country level; survey weights applied. * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

7 Conclusion

This paper provides the first study of links between gender norms, self-assessed femininity and masculinity, and utility. Three central findings emerge. First, self-reported gender intensity is positively correlated with life satisfaction: men who feel more masculine and women who feel more feminine are happier. Second, the implications of the gender norms for behavior are asymmetric. For men, amplifying typical behaviors supports life satisfaction, while violations are penalized. For women, however, amplifying typical behaviors is associated with lower life satisfaction, while adopting typical behaviors of men boosts well-being. Third, the main patterns hold when breaking down the larger sample into European sub-regions.

Overall, the findings point to a vertical structure of gender norms: societal organization assigns higher utility returns to the roles and traits of men. This does not imply that the behavior of men is universally superior; it only suggests that, in current institutional settings, it is more highly rewarded. This result regarding the verticality of gender norms finds an echo in some recent papers. Dean et al. (2025), for example, show that men who conform to stereotypical male behavior are expected to be "praised" by society whereas, for women, there is no behavior that is unanimously expected to be praised (hence the title: "Gender views: A restricted path for men and a mission impossible for women"). Jiang (2025) shows that the general prestige of occupations, as measured through linguistic representation, declines as they become stereotyped as female. Banan et al. (2025) show that gender non-conformity is associated with better later life outcomes for women but worse life outcomes for men compared with gender typical individuals of their own gender.

While these patterns are robust to alternative measures of typical behavior of men and women and individual divergence from such norms, there are several limitations. First, the cross-sectional design of the survey precludes causal interpretation of our findings. Second, all of our measures of norms are statistical constructs that may not fully capture cultural and societal prescriptions for gender roles and behavior. That is, there is no independent

measure of what people consider appropriate and inappropriate behavior for men and for women. Third, the self-assessments of femininity and masculinity remain ambiguous, as we do not know what baseline or criteria respondents have in mind when they answer these questions.

The study suggests at least two avenues for future research. First, researchers could strive to find exogenous variations in gender norms. This would allow for a direct test of causal relationships between gender identity and happiness. Second, researchers could elucidate the content of people's notions of masculinity and femininity. People's self-assessments could reflect their self-image or suppositions about how others perceive them, or intimate feelings that may be related to sexuality, parenting, or other imperatives related to gender. With such finer understandings of subjective masculinity and femininity, we can better investigate their relationship to both behavior and to overall well-being.

References

Akerlof, G. and Kranton, R. (2000). Economics and identity. *The Quarterly Journal of Economics*, 115(3):715–753.

Alesina, A., Giuliano, P., and Nunn, N. (2013). On the origins of gender roles: Women and the plough. *Quarterly Journal of Economics*, 128(2):469–530.

Algan, Y., Cahuc, P., and Shleifer, A. (2013). Teaching practices and social capital. *American Economic Journal: Applied Economics*, 5(3):189–210.

Arrow, K. J. (1973). The theory of discrimination. In Ashenfelter, O. and Rees, A., editors, *Discrimination in Labor Markets*. Princeton University Press.

Ayyar, S., Bolt, U., French, E., and O'Dea, C. (2024). Imagine your life at 25: Gender conformity and later-life outcomes. NBER Working Paper 32789, National Bureau of Economic Research.

Banan, A. R., Santavirta, T., and Sarzosa, M. (2025). Childhood gender nonconformity and gender gaps in life outcomes. Discussion Paper 39, Helsinki GSE.

Becker, G. (1957). *The Economics of Discrimination*. University of Chicago Press.

Becker, G. (1981). *A Treatise on the Family*. Harvard University Press.

Bem, S. L. (1974). The measurement of psychological androgyny. *Journal of Consulting and Clinical Psychology*, 42(2):155–162.

Bertrand, M. (2011). New perspectives on gender. In Card, D. and Ashenfelter, O., editors, *Handbook of Labor Economics*, volume 4, pages 1543–1590. Elsevier.

Bertrand, M., Kamenica, E., and Pan, J. (2015). Gender identity and relative income within households. *The Quarterly Journal of Economics*, 130(2):571–614.

Bisin, A. and Verdier, T. (2001). The economics of cultural transmission and the dynamics of preferences. *Journal of Economic Theory*, 97(2):298–319.

Blau, F. D. and Kahn, L. M. (2013). Female labor supply: Why is the us falling behind? *American Economic Review*, 103(3):251–256.

Booth, A. L. and Van Ours, J. C. (2009). Hours of work and gender identity: Does part-time work make the family happier? *Economica*, 76:176–196.

Brenøe, A. A., Eyibak, Z., Heursen, L., Ranehill, E., and Weber, R. (2024). Gender identity and economic decision making. *Available at SSRN*. SSRN ID: 4889851.

Brenøe, A. A., Heursen, L., Ranehill, E., and Weber, R. A. (2022). Continuous gender identity and economics. *AEA Papers and Proceedings*, 112:573–577.

Burbano, V., Padilla, N., and Meier, S. (2024). Gender differences in preferences for meaning at work. *American Economic Journal: Economic Policy*, 16(3):61–94.

Bursztyn, L., Cappelen, A. W., Tungodden, B., Voena, A., and Yanagizawa-Drott, D. H. (2023). How are gender norms perceived? NBER Working Paper 31049, National Bureau of Economic Research. Revised version 2025.

Bursztyn, L., Egorov, G., and Fiorin, S. (2020). From extreme to mainstream: How social norms unravel. *Quarterly Journal of Economics*, 135(3):1311–1367.

Campa, P. and Serafinelli, M. (2019). Politico-economic regimes and attitudes: Female workers under state socialism. *Review of Economics and Statistics*, 101(2):233–248.

Coffman, K. B. and Klinowski, D. (2025). Gender and preferences for performance feedback. *Management Science*, 71(4):3497–3516.

Cortés, P., Hwang, J., Pan, J., and Schönberg, U. (2025). Gender norms and the labor market. NBER Working Paper 34716, National Bureau of Economic Research.

Cortes, P. and Pan, J. (2018). Occupation and gender. In *The Oxford Handbook of Women and the Economy*. Oxford University Press.

Croson, R. and Gneezy, U. (2009). Gender differences in preferences. *Journal of Economic Literature*, 47(2):448–474.

De Haas, R., Baranov, V., Matavelli, I., and Grosjean, P. (2024). Masculinity around the world. Working paper, September 10, 2024.

Dean, J. T., Exley, C. L., Klinowski, D., Niederle, M., and Sarsons, H. (2025). Gender views: A restricted path for men and a mission impossible for women. Cepr discussion paper, Centre for Economic Policy Research.

Eckel, C. C. and Grossman, P. J. (2008). Men, women and risk aversion: Experimental evidence. In Plott, C. R. and Smith, V. L., editors, *Handbook of Experimental Economics Results*, volume 1, pages 1061–1073. Elsevier.

European Social Survey (2024). Ess round 11 - 2023. social inequalities in health, gender in contemporary europe. `https://doi.org/10.21338/ess11-2023`. Sikt - Norwegian Agency for Shared Services in Education and Research.

Eurostat (2023). European union labour force survey (lfs) 2023. Access to microdata provided by Eurostat under strict conditions.

Falk, A., Becker, A., Dohmen, T., Enke, B., Huffman, D., and Sunde, U. (2018). Global evidence on economic preferences. *The Quarterly Journal of Economics*, 133(4):1645–1692.

Fernández, R. (2011). Does culture matter? In Benhabib, J., Bisin, A., and Jackson, M. O., editors, *Handbook of Social Economics*, volume 1, pages 481–510. North-Holland.

Fernández, R. (2013). Cultural change as learning: The evolution of female labor force participation over a century. *American Economic Review*, 103(1):472–500.

Fortin, N. M. (2005). Gender role attitudes and the labour-market outcomes of women across oecd countries. *Oxford Review of Economic Policy*, 21(3):416–438.

Fortin, N. M. (2008). The gender wage gap among young adults in the united states: The importance of money versus people. *Journal of Human Resources*, 43(4):884–918.

Goldin, C. (2002). The rising (and then declining) significance of gender.

Goldin, C. (2014). A grand gender convergence: Its last chapter. *American Economic Review*, 104(4):1091–1119.

Jiang, W. (2025). The cultural devaluation of feminized work: The evolution of u.s. occupational prestige and gender typing in linguistic representations, 1900 to 2019. *American Sociological Review*, 90(5):755–787.

Kachel, S., Steffens, M. C., and Niedlich, C. (2016). Traditional masculinity and femininity: Validation of a new scale assessing gender roles. *Frontiers in Psychology*, 7:956.

Kahneman, D. and Deaton, A. (2010). High income improves evaluation of life but not emotional well-being. *Proceedings of the National Academy of Sciences*, 107(38):16489–16493.

Knight, C. and Brinton, M. (2017). One egalitarianism or several? two decades of gender-role attitude change in europe. *American Journal of Sociology*, 122(5):1485–1532.

Le Barbanchon, T., Rathelot, R., and Roulet, A. (2021). Gender differences in job search: Trading off commute against wage. *The Quarterly Journal of Economics*, 136(1):381–426.

Lippmann, Q., Georgieff, A., and Senik, C. (2020). Undoing gender with institutions: Lessons from the german division and reunification. *The Economic Journal*, 130(629):1445–1470.

Magliozzi, D., Saperstein, A., and Westbrook, L. (2016). Scaling up: Representing gender diversity in survey research. *Socius: Sociological Research for a Dynamic World*, 2:1–11.

Mencarini, L. and Sironi, M. (2012). Happiness, housework and gender inequality in Europe. *European Sociological Review*, 28(2):203–219. First published online 2010.

Niederle, M. and Vesterlund, L. (2007). Do women shy away from competition? do men compete too much? *The Quarterly Journal of Economics*, 122(3):1067–1101.

OECD (2013). Guidelines on measuring subjective well-being.

OECD (2024). Measuring subjective well-being across oecd countries. Technical report, OECD WISE Centre. Policy Insights, OECD.

Petrongolo, B. and Ronchi, M. (2020). Gender gaps and the structure of local labor markets. *Labour Economics*, 64:101819.

Shurchkov, O. and Eckel, C. (2018). Gender differences in behavioral traits and labor market outcomes. In *The Oxford Handbook of Women and the Economy*. Oxford University Press.

Stiglitz, J., Sen, A., and Fitoussi, J.-P. (2009). Report by the commission on the measurement of economic performance and social progress.

Su, R., Rounds, J., and Armstrong, P. I. (2009). Men and things, women and people: A meta-analysis of sex differences in interests. *Psychological Bulletin*, 135(6):859–884.

Trachman, M. (2022). Très masculin, pas très féminine. les variations sociales du genre. *Population & Sociétés*, (605).

World Happiness Report (2025). World Happiness Report 2025. Accessed October 14, 2025.

A Appendix

Table A.1: Summary Statistics

Variable	Mean	Std. Dev.	Min	Max
Man (=1)	0.47	0.50	0	1
Feel Feminine	3.01	2.56	0	6
Feel Masculine	3.01	2.52	0	6
Age Groups				
18–25	0.10	0.30	0	1
26–40	0.21	0.41	0	1
41–55	0.25	0.43	0	1
56–65	0.18	0.38	0	1
66+	0.25	0.44	0	1
Migrant (=1)	0.10	0.30	0	1
Years of Education	13.65	4.09	0	69
Number of Children	0.59	0.94	0	10
Married (=1)	0.55	0.50	0	1
Household Income (deciles)	5.58	2.79	1	10
Predicted Income (in thousands of Euros)	12.45	19.07	0	93.90
Occupational Prestige Score	35.55	24.05	0	81.49
Life Satisfaction	7.06	2.16	0	10
Happy	7.42	1.86	0	10
Were Happy	2.92	0.81	1	4
Felt Depressed	1.46	0.69	1	4
Felt Everything was Effort	1.68	0.79	1	4
Felt Sad	1.59	0.70	1	4
Violation (LASSO)	0.06	0.09	0	0.75
Amplification (LASSO)	0.06	0.08	0	0.61
Violation (OLS)	0.05	0.07	0	0.76
Amplification (OLS)	0.05	0.07	0	0.56
Violation (No Missing)	0.08	0.12	0	2.64
Amplification (No Missing)	0.08	0.11	0	0.93
Violation (Region-Specific)	0.08	0.12	0	1.39
Amplification (Region-Specific)	0.08	0.12	0	1.05
Violation of Gender Norms (Character)	0.13	0.19	0	1.52
Amplification of Gender Norms (Character)	0.13	0.18	0	1.33
Violation of Gender Norms (Health)	0.18	0.26	0	5.49
Amplification of Gender Norms (Health)	0.18	0.24	0	4.53
Violation of Gender Norms (Politics)	0.11	0.17	0	1.28
Amplification of Gender Norms (Politics)	0.12	0.17	0	1.40
Violation of Gender Norms (Religion)	0.30	0.40	0	1.75
Amplification of Gender Norms (Religion)	0.30	0.34	0	1.39
Violation of Gender Norms (Social)	0.11	0.17	0	1.70
Amplification of Gender Norms (Social)	0.11	0.17	0	1.24
Violation of Gender Norms (Work)	0.07	0.11	0	0.98
Amplification of Gender Norms (Work)	0.07	0.10	0	0.69
Violation (Binary)	0.49	0.50	0	1
Amplification (Binary)	0.05	0.08	0	0.60
Typicality (OLS)	0.00	0.14	-0.75	0.61
Typicality (Predicted)	0.70	0.24	0.00	1.00

Notes: Survey weights are applied. Numbers are derived from our own calculations on the baseline sample and rounded to two decimals.

Table A.2: List of Countries in the Sample

Code	Country	Freq.	Percent	Cum.
AT	Austria	2,285	4.67	4.67
BE	Belgium	1,518	3.10	7.77
BG	Bulgaria	2,137	4.37	12.14
CH	Switzerland	1,314	2.68	14.82
CY	Cyprus	648	1.32	16.15
CZ	Czech Republic	1,496	3.06	19.20
DE	Germany	2,319	4.74	23.94
EE	Estonia	1,217	2.49	26.43
ES	Spain	1,774	3.62	30.05
FI	Finland	1,486	3.04	33.09
FR	France	1,661	3.39	36.48
GB	United Kingdom	1,588	3.24	39.73
GR	Greece	2,661	5.44	45.16
HR	Croatia	1,473	3.01	48.17
HU	Hungary	2,008	4.10	52.28
IE	Ireland	1,921	3.93	56.20
IL	Israel	820	1.68	57.88
IS	Iceland	809	1.65	59.53
IT	Italy	2,693	5.50	65.03
LT	Lithuania	1,276	2.61	67.64
LV	Latvia	1,153	2.36	70.00
ME	Montenegro	1,530	3.13	73.12
NL	Netherlands	1,606	3.28	76.40
NO	Norway	1,270	2.59	79.00
PL	Poland	1,340	2.74	81.74
PT	Portugal	1,358	2.77	84.51
RS	Serbia	1,428	2.92	87.43
SE	Sweden	1,183	2.42	89.85
SI	Slovenia	1,176	2.40	92.25
SK	Slovakia	1,380	2.82	95.07
UA	Ukraine	2,413	4.93	100.00
	Total	**48,941**	**100.00**	

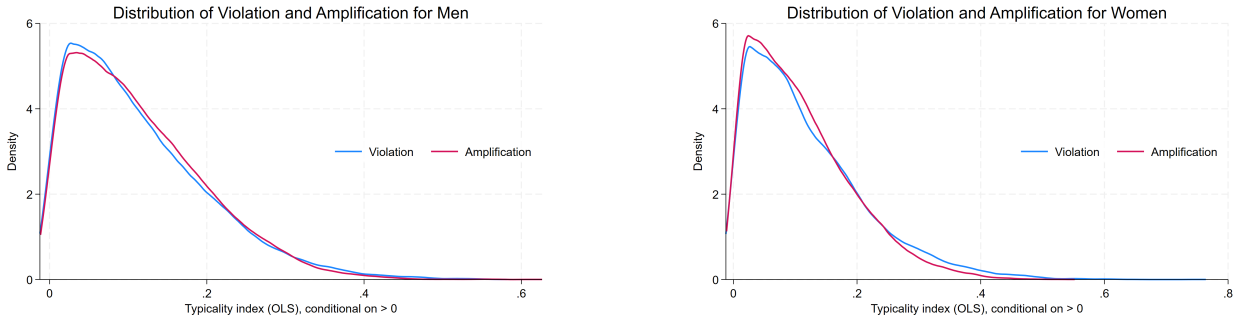
Notes: Numbers are derived from our own calculations on the baseline sample.

Table A.3: Pairwise Correlations of Well-Being Measures

MEN	**Life Sat.**	**Flt. Depr.**	**Flt. Eff.**	**Flt. Sad**	**Flt. Happy**	**Happy**
Life Sat.	1.00					
Flt. Depr.	-0.34***	1.00				
Flt. Eff.	-0.26***	0.49***	1.00			
Flt. Sad	-0.34***	0.54***	0.39***	1.00		
Flt. Happy	0.39***	-0.37***	-0.29***	-0.37***	1.00	
Happy	0.67***	-0.37***	-0.29***	-0.38***	0.49***	1.00
WOMEN	**Life Sat.**	**Flt. Depr.**	**Flt. Eff.**	**Flt. Sad**	**Flt. Happy**	**Happy**
Life Sat.	1.00					
Flt. Depr.	-0.33***	1.00				
Flt. Eff.	-0.28***	0.54***	1.00			
Flt. Sad	-0.35***	0.57***	0.45***	1.00		
Flt. Happy	0.38***	-0.42***	-0.34***	-0.42***	1.00	
Happy	0.65***	-0.37***	-0.32***	-0.39***	0.49***	1.00

Notes: Numbers are derived from our own calculations on the baseline sample. Entries are Pearson correlation coefficients. * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

Figure 8: Distribution of Amplification and Violation by Gender



(a) Men (b) Women

Notes: Figures display kernel density histograms derived from our own calculations on the baseline sample.

Table A.4: Robustness. Various Categorizations of Life Satisfaction

	MEN			WOMEN		
	(1)	(2)	(3)	(4)	(5)	(6)
	Life Sat.	Standardized Life Sat.	Binarized Life Sat.	Life Sat.	Standardized Life Sat.	Binarized Life Sat.
Feel Feminine	0.02			0.22***		
	(0.02)			(0.02)		
Feel Masculine	0.16***			0.02		
	(0.03)			(0.01)		
Violation of Gender Norms	-0.99***	-0.52***	-0.17**	0.71***	0.36***	0.15***
	(0.33)	(0.16)	(0.07)	(0.23)	(0.10)	(0.05)
Amplification of Gender Norms	-0.09	-0.05	0.11	-0.72***	-0.38***	-0.11*
	(0.36)	(0.16)	(0.11)	(0.17)	(0.09)	(0.06)
St. Feel Feminine		0.02	-0.00		0.26***	0.11***
		(0.03)	(0.01)		(0.03)	(0.01)
St. Feel Masculine		0.18***	0.09***		0.01	-0.00
		(0.04)	(0.02)		(0.01)	(0.01)
Constant	6.32***	-0.09***	0.44***	5.82***	-0.21***	0.40***
	(0.14)	(0.03)	(0.01)	(0.10)	(0.02)	(0.02)
Obs.	22589	22589	22589	26739	26739	26739
R-sq.	0.11	0.01	0.08	0.13	0.02	0.09

Notes: Ordinary Least Squares (OLS) estimates. All models include controls for age group (18–25, 26–40, 41–55, 56-65, 66+) and country fixed effects. Standard errors clustered at the country level are reported in parentheses. Survey weights are applied. Numbers are derived from our own calculations on the baseline sample and rounded to two decimals. * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$. The variables *Standardized Life Sat., St. Feel Feminine, and St. Feel Masculine* are standardized at country level to have a mean of 0 and a standard deviation of 1. The binary variable *Binarized Life Sat.* equals 1 if the original life satisfaction score is 8 or higher, and 0 otherwise.

Table A.5: Happy and Gender Identity, Violation, Amplification, and Typicality

	MEN					WOMEN				
Dep. Variable: Happy	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Feel Feminine	-0.08***			0.03*	0.03*	0.27***			0.29***	0.29***
	(0.02)			(0.02)	(0.02)	(0.02)			(0.03)	(0.03)
Feel Masculine		0.22***		0.23***	0.23***		-0.04		0.04	0.04
		(0.03)		(0.03)	(0.03)		(0.05)		(0.04)	(0.04)
Violation of Gender Norms			-1.08**	-0.83*				0.22	0.40	
			(0.46)	(0.41)				(0.22)	(0.24)	
Amplification of Gender Norms			-0.94**	-1.01**				-0.24	-0.34	
			(0.41)	(0.41)				(0.22)	(0.24)	
Typicality (OLS)					-0.07					-0.37**
					(0.16)					(0.14)
Constant	7.46***	6.28***	7.53***	6.30***	6.18***	6.00***	7.45***	7.41***	5.87***	5.87***
	(0.01)	(0.15)	(0.05)	(0.13)	(0.16)	(0.11)	(0.05)	(0.02)	(0.18)	(0.18)
Obs.	22547	22547	22547	22547	22547	26644	26644	26644	26644	26644
R-sq.	0.07	0.08	0.07	0.08	0.08	0.10	0.08	0.08	0.10	0.10

Notes: Ordinary Ordinary Least Squares (OLS) estimates. All models include controls for age group (18–25, 26–40, 41–55, 56-65, 66+) and country fixed effects. Standard errors clustered at the country level are reported in parentheses. Survey weights are applied. Numbers are derived from our own calculations on the baseline sample and rounded to two decimals.

Table A.6: Age Heterogeneity

	MEN				WOMEN			
	After 1970		Before 1970		After 1970		Before 1970	
Dep. Variable: Happy	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Feel Feminine	-0.00	-0.00	0.06**	0.06**	0.21***	0.20***	0.22***	0.22***
	(0.03)	(0.03)	(0.02)	(0.02)	(0.03)	(0.03)	(0.02)	(0.02)
Feel Masculine	0.17***	0.17***	0.14***	0.14***	-0.02*	-0.03**	0.07***	0.07***
	(0.03)	(0.03)	(0.04)	(0.04)	(0.01)	(0.01)	(0.02)	(0.02)
Violation of Gender Norms	-0.88**		-1.10**		0.65**		0.88**	
	(0.43)		(0.45)		(0.25)		(0.37)	
Amplification of Gender Norms	-0.04		-0.20		-0.34*		-0.96***	
	(0.31)		(0.54)		(0.18)		(0.18)	
Violation of Gender Norms (After 1970)		-0.96**				0.63		
		(0.45)				(0.37)		
Amplification of Gender Norms (After 1970)		0.17				-0.76**		
		(0.34)				(0.36)		
Violation of Gender Norms (Before 1970)				-1.05**				0.60
				(0.42)				(0.45)
Amplification of Gender Norms (Before 1970)				-0.44				-1.05***
				(0.39)				(0.28)
Constant	6.27***	6.27***	6.43***	6.45***	6.01***	6.05***	5.73***	5.75***
	(0.18)	(0.19)	(0.19)	(0.20)	(0.15)	(0.16)	(0.10)	(0.10)
Obs.	12059	12047	10530	10530	13822	13789	12917	12917
R-sq.	0.10	0.10	0.13	0.13	0.11	0.11	0.17	0.17

Notes: Ordinary Ordinary Least Squares (OLS) estimates. All models include controls for age group (18–25, 26–40, 41–55, 56-65, 66+) and country fixed effects. Standard errors clustered at the country level are reported in parentheses. Survey weights are applied. Numbers are derived from our own calculations on the baseline sample and rounded to two decimals.

B Online Appendix: List of Gender Markers in the Baseline Estimation

Table B.1: Health Behavior

Variable	Description	Value of the variable
cgtsmok	Cigarette smoking behaviour	1 (I smoke daily, 10 or more cigarettes) – 6 (I have never smoked)
dshltgp	Discussed health, last 12 months: general practitioner	1 = yes, 0 = no
dshltms	Discussed health, last 12 months: medical specialist	1 = yes, 0 = no
eatveg	How often eat vegetables or salad, excluding potatoes	1 (Three times or more a day) – 7 (Never)
alcwkdy	Grams alcohol, last time drinking on a weekday, Monday to Thursday	numeric
alcfreq	How often drink alcohol	1(Every day)-7(Never)
alcbnge	Frequency of binge drinking for men and women, last 12 months	1(Daily or almost daily)-5(Never)

Table B.2: Religion

Variable	Description	Value of the variable
rlgdgr	How religious are you	0 (Not at all religious) – 10 (Very religious)
pray	How often pray apart from at religious services	1 (Every day) – 7 (Never)

Table B.3: Social Behavior and Trust

Variable	**Description**	**Value of the variable**
ctrlife	How much control over life in general nowadays	0(No control at all)-10(Complete control)
trstplt	Trust in politicians	0 (No trust at all) - 10 (Complete trust)
trstlgl	Trust in the legal system	0 (No trust at all) – 10 (Complete trust)
netustm	Internet use, how much time on typical day	numeric (in minutes)

Table B.4: Personality

Variable	**Description**	**Value of the variable**
lrnobed	Obedience and respect for authority most important virtues children should learn	1(Agree strongly)-5(Disagree strongly)
sothnds	I am sensitive to others' needs	0 (Not at all) – (Completely)
actcomp	I act compassionately towards others, to what extent	0 (Not at all) – (Completely)
ipmodsta	Important to be humble and modest, not draw attention	1 (Very much like me) – 6 (Not like me at all)
ipbhprpa	Important to behave properly	1 (Very much like me) – 6 (Not like me at all)
aesfdrk	Feeling of safety of walking alone in local area after dark	1 (Very safe) – 4 (Very unsafe)
impfreea	Important to make own decisions and be free	1 (Very much like me) - 6 (Not like me at all)
likrisk	I like to take risks, to what extent	0 (Not at all) – (Completely)
liklead	I like to be a leader, to what extent	0 (Not at all) – (Completely)
impdiffa	Important to try new and different things in life	1 (Very much like me) – 6 (Not like me at all)
impsafea	Important to live in secure and safe surroundings	1 (Very much like me) - 6 (Not like me at all)
ipshabta	Important to show abilities and be admired	1 (Very much like me) - 6 (Not like me at all)
ipfrulea	Important to do what is told and follow rules	1 (Very much like me) - 6 (Not like me at all)

Table B.5: Work

Variable	**Description**	**Value of the variable**
hswrk	Doing last 7 days: housework, looking after children, others	1 = yes, 0 = no
fem_oc_r2	Share of women in the respondent's occupation and industry	numeric
pdwrk	Doing last 7 days: paid work	1 = yes, 0 = no
edctn	Doing last 7 days: education	1 = yes, 0 = no
dsbld	Doing last 7 days: permanently sick or disabled	1 = yes, 0 = no
rtrd	Doing last 7 days: retired	1 = yes, 0 = no
cmsrv	Doing last 7 days: community or military service	1 = yes, 0 = no
estsz	Establishment size	1 (Under 10) – 5 (500 or more)
jbspv	Responsible for supervising other employees	1 = yes, 0 = no
private_sec	Private sector of employment	1 = yes, 0 = no
$emplst_1$	Employee	1 = yes, 0 = no
$emplst_2$	Self-Employed	1 = yes, 0 = no
$emplst_3$	Working for own family business	1 = yes, 0 = no
wkhtot	Total hours normally worked per week in main job overtime included	numeric

Table B.6: Political Attitudes

Variable	**Description**	**Value of the variable**
polintr	How interested in politics	1 (Very interested) – 4 (Not at all interested)
loylead	Country needs most loyalty towards its leaders	1 (Agree strongly) – 5 (Disagree strongly)
hmsfmlsh	Ashamed if close family member gay or lesbian	1 (Agree strongly) – 5 (Disagree strongly)
sgnptit	Signed petition last 12 months	1 = yes, 0 = no
bctprd	Boycotted certain products last 12 months	1 = yes, 0 = no
hmsacld	Gay and lesbian couples right to adopt children	1(Agree strongly)-5(Disagree strongly)
cptppola	Confident in own ability to participate in politics	1 (Not at all confident) – 5 (Completely confident)
clsprty	Feel closer to a particular party than all other parties	1 = yes, 0 = no
imbgeco	Immigration bad or good for country's economy	0 (Bad for the economy) – 10 (Good for the economy)
badge	Worn or displayed campaign badge/sticker last 12 months	1=yes 0=no
actrolga	Able to take active role in political group	1 (Not at all able) – 5 (Completely able)
donprty	Donated to or participated in political party or pressure group last 12 months	1 = yes, 0 = no
wrclmch	How worried about climate change	1(Not at all worried)-5(Extremely worried)