I Z A Institute
of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 17564

Gender Identity and Economic Decision Making

Anne Ardila Brenøe
Zeynep Eyibak
Lea Heursen
Eva Ranehill
Roberto A. Weber

DECEMBER 2024

I Z A Institute
of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 17564

Gender Identity and Economic Decision Making

Anne Ardila Brenøe
University of Zurich and IZA

Zeynep Eyibak
University of Zurich

Lea Heursen
Humboldt-Universität zu Berlin

Eva Ranehill
Lund University and University of Gothenburg

Roberto A. Weber
University of Zurich

DECEMBER 2024

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Gender Identity and Economic Decision Making*

Economic research on gender gaps has focused on variation based on the binary classification of "men" and "women". We explore whether a self-reported continuous measure of gender identity (CGI) explains variation in economic decisions and outcomes beyond the relationship with binary gender. We analyze data from four diverse populations (N=8,018), including measures of economic preferences and educational and labor market outcomes. We find that CGI is significantly associated with economic outcomes, with stronger relationships for men than women. Our results indicate that incorporating measures of self-reported gender identity could enhance our understanding of gender gaps in economic behavior and outcomes.

JEL Classification: C91, J16, J2

Keywords: gender identity, non-binary gender, economic preferences, economic outcomes

Corresponding author:
Anne Ardila Brenøe
University of Zurich
Schönberggasse 1
8001 Zurich
Switzerland
E-mail: anne.brenoe@econ.uzh.ch

* We are grateful to participants at several conferences and seminars for helpful comments and suggestions. We thank Johanna Croton, Lukas Diethelm, Asma Hussein, Andia Kozmai, Andreas Lercher, Maximilian Mähr, Viviane Meyer, Chi Nguyen, and Liaoliang Zhang for outstanding research assistance. We thank yousty.ch for generously supporting our data collection among Swiss teenagers and sharing their administrative data with us. This work was supported by the University Research Priority Program "URPP Equality of Opportunity" of the University of Zurich. Lea Heursen thanks the Deutsche Forschungsgemeinschaft through CRC TRR 190 (project number 280092119) for financial support. Eva Ranehill thanks the Tore Browaldh and the Torsten Söderberg foundation for generous support.

1. Introduction

A large and growing body of literature documents a variety of gender gaps in economic behaviors and outcomes. Women tend to select different college majors, career tracks and industries, earn less than men, and are less likely to attain leadership positions in industry or government (Bertrand and Hallock, 2001; Olivetti and Petrongolo, 2008; Blau and Kahn, 2017; Bertrand, 2018; Patnaik, Wiswall and Zafar, 2021). There are also gaps in measures of economic preferences: men often exhibit less risk-aversion, more willingness to compete, and have a stronger tendency to prioritize efficiency over equality (Andreoni and Vesterlund, 2001; Barber and Odean, 2001; Dohmen et al., 2011; Eckel and Grossman, 2008; Filippin and Crosetto, 2016; Fisman, Jakiela and Kariv, 2017; Niederle and Vesterlund, 2011).[1] An important feature of this research is that it studies gender gaps from the perspective of a *binary* classification of individuals as men or women. The focus on this binary categorization of gender is natural given the availability of datasets containing it and the salience of this distinction in social, educational, and economic contexts.

However, a binary classification of individuals as male or female is not the only way in which we may conceptualize the broader notion of *gender*, nor the only way it has been studied in academic research. In the area of gender studies—primarily rooted in sociology and social psychology—gender is often conceptualized as a more complex pattern of behaviors, perceived characteristics, and aspects of individuals' identities, including perceptions about what traits define the concepts of *masculinity* and *femininity* (Hawkesworth, 1997; Pryzgoda and Chrisler, 2000; Westbrook and Saperstein, 2015). As described in a highly influential paper by West and Zimmerman (1987), gender is the collection of behaviors—what an individual "does"—and the degree to which an individual's behavior reflects identification with expectations about what it means to be male or female.

In this study, we draw upon research from these fields to explore whether such broader notions of gender are valuable for understanding gender gaps of interest to economists. Specifically, we provide evidence on whether non-binary, self-reported measures of *gender identity* are helpful for understanding gender gaps beyond any relationship with an indicator for binary gender. We initially present a conceptual framework that provides a basis for why an individual's identity as "masculine" or "feminine" may correlate with gender-stereotypical behaviors, preferences and traits, even after accounting for binary gender classifications. We then empirically investigate these relationships across several measures of preferences, behaviors, and outcomes for which earlier research in economics has found gaps between men and women.

[1] These preference gaps have been associated with gaps in a variety of important economic outcomes, including career choices and salary attainment (Barber and Odean, 2001; Buser, Niederle and Oosterbeek, 2014; Dohmen and Falk, 2011; Dohmen et al., 2011; Fisman, Jakiela and Kariv, 2017; Gärtner, Mollerstrom and Seim, 2017; Zhang, 2013).

There are at least two reasons why measures of self-perceived gender identity may help understand questions of interest to economists. First, despite the observation of differences between men and women in behaviors, such as risk-taking and competitiveness, and in the attainment of a variety of educational and labor-market outcomes, there is also substantial within-gender variation. Better understanding whether men and women who identify as more feminine are less likely to engage in risk taking, enter STEM fields, or obtain high income provides potential value for understanding the sources of economic gender disparities and for improved targeting of policy. Second, non-traditional forms of gender identification are growing in many Western societies. For example, according to the Pew Center, 1.6 percent of U.S. adults identify as transgender or non-binary and this percentage is 5.1 percent for adults under 30 (Brown, 2022; see also, Flores et al., 2016). More nuanced and inclusive measures of gender identity recognize this diversity, including individuals who identify as women or men. Indeed, many datasets widely used by economists are starting to incorporate self-reported gender identity, often with non-binary measures.[2] Because gender identity may motivate individuals' choices and behaviors (Akerlof and Kranton, 2000), it is important to be able to reliably measure such diverse gender identities, in addition to economic outcomes.

As a first step in our analysis, we test the relationship between several measures of non-binary gender identity used in gender studies research and a novel unidimensional question that simply asks individuals where they place themselves on a scale from "very masculine" to "very feminine." This single-item measure of continuous gender identity (CGI) captures a large part of the variation in several richer scales that measure gender identity along separate dimensions for masculinity and femininity, sometimes using scales constructed from dozens of individual items (Bem, 1974; Kachel, Steffens and Niedlich, 2016; Magliozzi, Saperstein and Westbrook, 2016). This part of our analysis replicates and extends our preliminary investigation of the same question in a sample of Swiss students reported in Brenøe et al. (2022), to larger and broader samples. We also document that our CGI measure is fairly stable over two weeks, by eliciting CGI a second time in one of our samples. Given the concordant results across samples and over time, we employ this validated one-question measure of CGI as the primary measure of gender identity in our subsequent research. However, we also elicit and explore other measures of continuous gender identity employed in the broader literature of gender studies.

We include our measure of CGI in four separate data collections using diverse pop-

[2]This includes the 2021 UK Census (https://www.bbc.com/news/uk-wales-55721123, accessed May 9, 2024) and the U.S Census Bureau's Household Pulse Survey (https://www.census.gov/library/stories/2021/08/household-pulse-survey-updates-sex-question-now-asks-sexual-orientation-and-gender-identity.html, accessed May 9, 2024).

ulations.[3] The four samples comprise (i) a sample of Swiss students from the university subject pool in Zurich (*Swiss Uni*), (ii) a broader sample of adult U.S. residents recruited through Prolific Academic (*U.S. Adults*), (iii) a sample of Swiss adolescents recruited from an online platform for apprenticeships (*Swiss Teens*), and (iv) a sample of young Swedish adults recruited through contacts with Swedish secondary schools (*Swedish Youths*). Recruiting from such diverse populations allows us to observe a wide range of economic choices and behaviors for which previous research has found gender gaps, ranging from economic preferences to educational choices and labor market outcomes. The considerable variation in age, from 13 years to 60 years, allows us to observe potential generational differences in reported gender identity or in its association with reported outcomes.

In each sample, we elicit our measure of CGI along with various preference and outcome variables for which previous research has documented gender gaps. In the Swiss Uni sample, we elicit several measures of incentivized and unincentivized preferences. We collected the U.S. Adults sample in two waves. In the first wave, we elicit an incentivized measure of risk-taking, several unincentivized preference measures, and measures of labor market and educational outcomes. In the second wave, we elicit a subset of the economic outcome measures collected in the first wave, together with various alternative measures of gender identity that we use for the re-validation of our novel CGI measure. In the Swiss Teens sample, we use a combination of administrative and survey data to observe occupational search and categorize the occupations by their gender composition and math and language skill requirements as proxies for the associated gender stereotypes. For a subset of these adolescents, we also elicit unincentivized preference measures. In the sample of Swedish Youths, we measure incentivized preferences for equality over efficiency, as well as unincentivized preference measures. We also elicit a measure of the respondents' intended fields of study at university and categorize fields by gender composition to measure the associated gender stereotypes.

In all four samples, we observe heterogeneous responses to the CGI question among both men and women. Not surprisingly, the distributions of responses differ substantially between men and women, with women generally reporting a more feminine and men a

[3]We also had the opportunity to add our CGI measure to a pharmacological experiment implemented by researchers at the Institute of Experimental Psychology, Heinrich Heine University Düsseldorf, which also measured sex hormones. This provides suggestive evidence on the relationship between gender identity and hormones in a sample of students (N=130, 65 women, mean age 24 years). Regressing reported CGI on testosterone, estradiol and progesterone, separately for those identifying as male and female, two relationships among the six tested approached significance at conventional levels. Our point estimates indicate a positive relationship between testosterone levels and masculinity among those identifying as female (p=0.053) and a positive relationship between progesterone and masculinity among those who reported being male (p=0.050). Due to the underpowered nature of the tests, the relationship should be interpreted cautiously. Hormones were measured in saliva pooled across three samples taken five minutes apart at the onset of the experiment. CGI was elicited after the saliva samples were taken, before the onset of the pharmacological experiment. Due to two participants identifying as non-binary, as well as one missing value for testosterone, two for progesterone, and one for gender, our tests comprise between 62 and 64 individuals. Details on the analysis or data are available on request.

more masculine identity. However, the distributions of responses for men and women exhibit substantial variation and overlap.

We then analyze these datasets to identify whether our novel CGI measure correlates with economic preferences as well as educational and labor market outcomes, *beyond* any relationships accounted for by a standard binary gender measure. For each outcome variable and sample, our analysis follows the same basic structure.[4] We first test for a difference in preference and outcome measures between men and women, and only continue the analysis for those outcomes for which we find a statistically significant gap. For these outcomes, we proceed to test the degree to which replacing the binary gender measure with CGI yields a statistically significant relationship of the same sign as the gender gap. A key part of our analysis is to then investigate the degree to which CGI relates to the outcome measure when also controlling for binary gender. Finally, we examine the extent to which reported gender identity correlates with economic choices and outcomes, separately, for men and women.

Expanding the scope of our approach, we also investigate correlations between other non-binary measures of gender identity and a subset of the outcomes we elicit. This allows us to test whether the broad conclusions that arise when using CGI are similar across various continuous measures of gender identity, or if they are specific to our novel and simple measure.

We provide four broad sets of results. First, we consider the statistical significance of CGI in explaining gender gaps in our outcome variables across all 33 tests we conducted. We strongly reject the null hypothesis that CGI has no relationship with the outcomes beyond what can be explained by binary gender alone. These findings are demonstrated in Figure 2, which shows the distributions of p-values that result from the key regressions for each outcome measure, clearly indicating a tendency toward lower p-values than one would expect by chance. Further, while we strongly reject the null hypothesis that CGI has no relationship with the preference and outcome measures we study, the added explanatory power provided is generally modest in absolute terms; that is, introducing CGI into a model that already contains binary gender generally yields modest increases in R^2.

Second, we identify those gender gaps for which we observe the strongest statistical relationships between the relevant outcome and CGI when accounting for binary gender. Given the exploratory nature of our work, we note only those relationships for which CGI is statistically significant at $p < 0.005$ while controlling for binary gender, which is true for 10 of the 33 tests. These include measures of self-reported willingness to take risk, engage in competition, redistributive preferences, the choice of a stereotypically female educational track, income, being a full-time homemaker and weekly average work hours.

[4]We pre-registered the approach for our analysis of the Swiss Uni dataset (https://osf.io/phyt6/). The analysis of the remaining datasets follows the same approach.

Interestingly, we find no significant relationships for incentivized preference measures, including risk and competitiveness, once controlling for binary gender. Thus, our findings suggest that relationships with CGI tend to be strongest for choices and outcomes related to the labor market and for unincentivized preference measures, and weakest for the type of incentivized choices that are widely used to measure economic preferences. The associations between CGI and human capital investment decisions are particularly noteworthy, because those investments happen relatively early in life and have far-reaching implications for subsequent life outcomes.

Third, we study the relationships between CGI and our outcome measures separately for men and women, finding that CGI shows substantially stronger correlations with economic behaviors and outcomes for men than for women. This provides suggestive evidence that the relationship between norms governing gender identity and behavior may be stronger for men. One potential, though speculative, interpretation for the stronger relationship for men is that norms of appropriate economic behavior for women have evolved and relaxed substantially in the past several decades, relative to those for men, at least in the samples we study.

Fourth, we also investigate the relationships between other measures of gender identity used in gender studies research and a subset of our outcome measures related to the labor market, incentivized risk-taking and unincentivized preference measures. In line with the results presented above, we find that scales measuring masculinity show stronger correlations with the type of economic choices and outcomes that we study, after controlling for binary gender, than measures of femininity. Measures of feminine identity add little explanatory value.

Taken together, our findings suggest that measures of continuous gender identity are statistically significantly related to preference and outcome measures often studied in (gender) economics. While the strength of the relationships varies across outcome measures and they rarely provide much additional explanatory power, there are some domains in which the relationships are substantial in magnitude and of economic significance. So, while our results do not indicate that including measures like CGI is critical for having a rich accounting of all the relationships between "gender" and economic preferences and outcomes, there may be populations and decision contexts in which such relationships are particularly important and valuable for improved understanding, prediction, and for policy targeting. For example, we observe particularly strong relationships for men, suggesting that gender identity may be more important for economic behaviors and outcomes in this population.

It is important to note that our work is exploratory and correlational. It is not based on a theory of which behaviors or outcomes should be correlated with gender identity, or on induced variation in such identity to show a causal relationship. Instead, we investigate whether there are relationships between self-reported gender identity and

a wide variety of behaviors and outcomes as an initial step in determining the potential value in studying such relationships.[5] In this regard, our work largely follows much work in gender economics, which documents correlations between binary gender and economic preferences, behavior and outcomes, as a starting point for learning more about these relationships. Because of the exploratory nature of our work and the large number of tests implemented, we employ a pre-specified approach for identifying variables of interest and limit our conclusions about specific relationships involving gender identity to those that are statistically significant at more conservative levels ($p < 0.005$) than those typically employed in economic research (Benjamin et al., 2018).

By focusing on a continuous measure of gender identity, our approach contrasts with most prior economic research on gender, which primarily focuses on differences between men and women.[6] A few notable exceptions acknowledge the potential roles of variation in gender identity for economic behavior. For example, in their influential work on identity, Akerlof and Kranton (2000) recognize that individuals may experience disutility from behaviors that depart from what is expected from members of a social category, such as "woman" and "man." Other empirical research studies whether variation in norms governing gender conduct—e.g., because of variation across cultures or households—influences women's tendencies to demonstrate stereotypically female behavior (Gneezy, Leonard and List, 2009; Bertrand, Kamenica and Pan, 2015; Brenøe, 2022; Bursztyn, González and Yanagizawa-Drott, 2020).

A handful of papers in economics investigate similar questions as the one underlying our research. Burn and Martell (2022) use respondents' answers to various survey items that correlate with binary gender to construct a proxy measure of "gender typicality" (see Fleming, Harris and Halpern, 2017) and then explore its relationship to labor market outcomes, with a focus on explaining differential outcomes for gays and lesbians. They find that gender typicality is related to labor market outcomes for men but not for women, but do not find that this measure explains gaps in outcomes between homosexual and heterosexual individuals. Banan, Santavirta and Sarzosa (2023) take a similar approach, using survey responses on preferences and interests to construct a measure of gender typicality in youth and correlating it with occupational choices, family planning, and health outcomes later in life.

Two recent papers investigate relationships between economic preferences and self-identified gender categories. Fornwagner et al. (2022) focus on differences among cismen, ciswomen, transmen, and transwomen—as well as priming of gender identity—to study

[5]For examples of recent research following a similar empirical approach by systematically investigating correlations between various measures of economic preferences and behaviors, see Chapman et al. (2023) and Stango and Zinman (2023).

[6]An exception to the literature emphasizing gender gaps in means is also provided by Nelson (2015), who emphasizes within-gender variation and the often substantial overlap in the distributions of measures of risk preferences of women and men.

gender and biological sex differences in economic preferences.[7] Overall, they find little evidence of either correlational or causal relationships between both sex and gender and behavior. Coffman, Coffman and Ericson (2024) focus on the economic preferences and beliefs of individuals who identify as non-binary and neither as a woman nor a man on a discrete measure of gender identity. While this research indicates some differences between non-binary individuals and men and women, the results reject simple explanations such as the preferences and beliefs of non-binary individuals lying between those of men and women.

To our knowledge, no other work in economics or related fields measures individuals' self-reported continuous gender identity in samples of men and women and across a variety of populations and investigates the relationships between such identification and a broad set of economic behaviors and outcomes.[8] Moreover, we also provide evidence on the relative value of different gender identity measures, including various measures used in gender studies research.

The rest of this paper is structured as follows. Section 2 provides a brief discussion of the measurement of gender identity outside of economics and a simple framework for how such identity potentially relates to economic behavior and outcomes. This section also reports the results of our validation exercise for our CGI measure. In Section 3, we briefly outline our empirical approach and hypotheses. Thereafter, in Section 4, we describe the different data sets and associated measures, before we present the results in Section 5. Section 6 concludes.

2. Gender Identity: Measurement and Implications for Behavior

Psychologists and other social scientists have long recognized that awareness of gender roles plays an important role in human cognitive and social development (Kohlberg and Kramer, 1969; Gilligan, 1977). While earlier research conceptualized gender as behaviors and identity reflecting concordance with *either* male or female norms, Sandra Bem (1974) advocated for viewing masculinity and femininity as distinct constructs, with the possibility that an individual could exhibit high (or low) concordance with both masculinity and femininity.

Bem (1974) also introduced a scale for measuring the distinct dimensions of masculinity and femininity, which subsequently became widely adopted and referred to as the Bem Sex Role Inventory (BSRI). The inventory contains 60 items eliciting the re-

[7] A small literature in economics studies whether priming experimental participants with gender impacts gender gaps in economic preferences, finding mixed results (see, e.g., Benjamin, Choi and Strickland, 2010, Boschini, Muren and Persson, 2012, Boschini et al., 2018).

[8] There is, however, a growing body of research exploring relationships between non-binary gender and outcomes in other fields in the social sciences, see, for example, Alexander, Bolzendahl and Wängnerud (2021) for a discussion of the use of non-binary measures of gender in political science.

spondents' compliance with traditional gender roles along both masculine (e.g., having a strong personality) and feminine (e.g., loving children) dimensions. Based on the self-reported traits and behaviors, respondents are classified as either masculine, feminine, androgynous, or undifferentiated. Much subsequent research measuring gender identity or conformity with gender norms identifies masculine and feminine dimensions separately, by eliciting respondents' self-reported tendencies to exhibit gender stereotypical behaviors or traits (Mahalik et al., 2003; Mahalik et al., 2005).

More recently, researchers have started questioning the need to rely on lengthy inventories that evaluate masculinity or femininity by measuring adherence to, or self-identification with, characteristics or behaviors whose relationship to gender norms may evolve over time. Thus, some recent approaches measure gender identity directly, by asking individuals the degree to which they perceive themselves as masculine and feminine (Magliozzi, Saperstein and Westbrook, 2016; Solevid et al., 2021). Furthermore, some researchers propose constructions of gender identity with masculinity and femininity as opposite ends of a unidimensional spectrum (Kachel, Steffens and Niedlich, 2016). As an alternative to self-identification, researchers have also proposed data-driven approaches to measure variation in conformity with gendered behavior and norms. For example, Fleming, Harris and Halpern (2017) use the degree to which an individual exhibits traits or behaviors most frequently associated with male or female respondents in a dataset as a measure of gender typicality.

Following the introduction of Bem's measure, as well as other measures of gender identity, several studies have investigated the relationship between gender identity and the tendency to exhibit other behaviors or characteristics typically associated with men or women. For example, self-reported gender identity has been found to correlate with social behaviors like aggressiveness and conformity (Eagly, 1978; Bernard, Bernard and Bernard, 1985; Weisbuch, Beal and O'Neal, 1999), approaches to decision-making (Nezu and Nezu, 1987; Brems and Johnson, 1989), eating disorders (Meyer, Blissett and Oldfield, 2001; Griffiths, Murray and Touyz, 2015) and psychological well-being (Taylor and Hall, 1982; Feather, 1985; Whitley, 1985). We take the existence of these relationships as a starting point for our investigation of whether a measure of continuous gender identity can be helpful for understanding behaviors, traits, and outcomes typically of interest to economists.

2.1. (Re-)Validating a Single-Item Measure of Continuous Gender Identity

Brenøe et al. (2022), introduced a novel measure of continuous gender identity, based on an individual's response to a single question, *"In general, how do you see yourself? Where would you put yourself on this scale from 'Very masculine' to 'Very feminine'?"* where higher scores correspond to a more feminine identity. Brenøe et al. found that responses

to this question correlated positively (negatively) with the feminine (masculine) scores of several two-dimensional measures of gender identity: the Bem Sex Role Inventory (BSRI), the Open Sex Role Inventory (an updated version of the BSRI) and a two-dimensional scale introduced by Magliozzi, Saperstein and Westbrook (2016). They also correlated positively with femininity scores from a unidimensional index proposed by Kachel, Steffens and Niedlich (2016). While this provides evidence that the single-item CGI measure has value for capturing an important part of more complex measures of gender identity, the data in Brenøe et al. were based on a relatively small and homogeneous student sample ($N = 99$).

To provide additional evidence on the relationship between our CGI measure and other scales employed in prior research, we conducted a study of 2,659 adult respondents (aged 20–60) in the U.S. recruited from the platform Prolific Academic. In addition to the measures of gender identity studied in our earlier paper, we collected additional measures used in prior work. Specifically, we included the following eight measures of gender identity:

1. Our novel *Single-item CGI* (1 item, 11-point scale, **CGI**), based on the single question described above, with higher scores indicating a more *feminine* identity.
2. The *Traditional Masculinity-Femininity* scale (Kachel, Steffens and Niedlich, 2016; 6 items, 6-point scales, **TMF**), on which individuals provide several responses ranging from very masculine to very feminine on a unidimensional scale, measuring how respondents personally identify and how their behaviors, attitudes, interests, and appearance would be socially interpreted. The six items are averaged to obtain a score reflecting *femininity*.
3. A *Two-Dimensional Masculinity-Femininity* scale (Magliozzi, Saperstein and Westbrook, 2016; 2 items, 6-point scales, **Magliozzi**), in which participants respond to "how do you see yourself?" separately for feminine and masculine dimensions, with responses ranging from "not at all" to "very." This yields separate scores for *masculinity* and *femininity*.
4. A *Trait-Based Self-Categorization* scale (Solevid et al., 2021; 2 items, 10-point scales, **SOM**), which asks participants the extent to which they believe that they possess masculine and feminine traits, yielding separate scores for *masculinity* and *femininity*.
5. An adapted version of the *Bem Sex-Role Inventory* (Bem, 1979; 10 feminine, 10 masculine, 10 neutral items, 5-point scales, **BSRI**), in which participants self-report the degree to which several characteristics apply personally. We follow the standard scoring approach to obtain separate measures of *masculinity* and *femininity*.
6. The *Open Sex-Role Inventory* (https://openpsychometrics.org/tests/OSRI/; 11 fem-

inine, 11 masculine items, 5-point scales, **OSRI**), which asks similar questions to the BSRI, though with an updated interpretation of masculine and feminine behaviors. Similarly to BSRI, this yields separate scores for *masculinity* and *femininity*.

7. The *Conformity to Masculine Norms Inventory* (Mahalik et al., 2003; 30 items, 5-point scales, **CMNI**), which asks respondents to rate the degree to which they exhibit stereotypically male traits or behaviors (e.g., "I put myself in risky situations"). The items are combined into a single scale reflecting *masculinity*.
8. The *Conformity to Feminine Norms Inventory* (Mahalik et al., 2005; 45 items, 5-point scales, **CFNI**), which asks respondents to rate the degree to which they exhibit stereotypically female traits or behaviors (e.g., "I regularly wear make-up"). The items are combined into a single scale reflecting *femininity*.

All respondents completed the first four measures as well as one from the remaining four inventory-based measures, administered in random order. The choice to let participants complete only one of the inventory-based measures was due to their lengthy nature and substantial overlap in the type of questions asked.

Table 1 shows the correlations between the above measures. As in Brenøe et al. (2022), all the correlations are highly statistically significant and in the anticipated directions. The single-item CGI scale correlates most strongly with the other measures that directly elicit self-reported masculinity and femininity (TMF, Magliozzi, SOM) and less strongly with the measures that elicit gender identity indirectly through evaluations of the applicability of gendered characteristics (BSRI, OSRI, CMNI, CFNI). The final row reports correlations with the first factor from a principal components analysis of all the measures except for CGI, revealing that this aggregated measure correlates highly with CGI. Appendix Figure A1 illustrates the relationships between CGI and each of the other measures in binned scatter plots. This analysis also shows that when elicited separately, the masculinity and femininity dimensions are highly (negatively) correlated, indicating that a large share of the variation between these measures can be captured by a unidimensional scale. Taken together, this analysis corroborates the preliminary evidence from Brenøe et al. (2022), that the single-item CGI measure captures a substantial part of individuals' gender identity, though with a much larger and broader sample of respondents.

To further evaluate the reliability of our CGI measure, we use data from our sample of Swiss university students (N=584), who completed two separate elicitations of the CGI measure two weeks apart. In the first elicitation, completed as part of our main study using this sample, respondents reported CGI on an 11-point scale from "very masculine" to "very feminine." In the second elicitation, they responded on an inverted 12-point scale, from "very feminine" to "very masculine." Despite these changes in the response format, the responses exhibit a high degree of stability (see Appendix Figure A2), with

Table 1: Correlation Matrix of all Gender Scales

Variables	**CGI**	TMF	Magliozzi (f)	Magliozzi (m)	SOM (f)	SOM (m)
TMF	0.923***					
	(0.007)					
Magliozzi (f)	0.901***	0.889***				
	(0.008)	(0.009)				
Magliozzi (m)	-0.899***	-0.880***	-0.846***			
	(0.008)	(0.009)	(0.010)			
SOM (f)	0.879***	0.882***	0.890***	-0.822***		
	(0.009)	(0.009)	(0.009)	(0.011)		
SOM (m)	-0.853***	-0.854***	-0.806***	0.880***	-0.798***	
	(0.010)	(0.010)	(0.011)	(0.009)	(0.012)	
OSRI (f)	0.570***	0.600***	0.627***	-0.500***	0.632***	-0.478***
	(0.032)	(0.031)	(0.030)	(0.034)	(0.030)	(0.034)
OSRI (m)	-0.462***	-0.450***	-0.428***	0.478***	-0.409***	0.508***
	(0.035)	(0.035)	(0.035)	(0.034)	(0.035)	(0.033)
CFNI	0.335***	0.362***	0.348***	-0.338***	0.342***	-0.368***
	(0.037)	(0.036)	(0.037)	(0.037)	(0.037)	(0.036)
CMNI	-0.431***	-0.433***	-0.362***	0.445***	-0.375***	0.446***
	(0.035)	(0.035)	(0.036)	(0.035)	(0.036)	(0.035)
BSRI (f)	0.211***	0.220***	0.252***	-0.193***	0.278***	-0.201***
	(0.038)	(0.038)	(0.037)	(0.038)	(0.037)	(0.038)
BSRI (m)	-0.264***	-0.263***	-0.212***	0.297***	-0.227***	0.297***
	(0.037)	(0.037)	(0.038)	(0.037)	(0.037)	(0.037)
First comp	0.942***	0.953***	0.936***	-0.935***	0.929***	-0.918***
	(0.007)	(0.006)	(0.007)	(0.007)	(0.007)	(0.008)

Notes: The table presents pairwise correlations between gender identity measures with standard errors reported in parantheses. U.S. Adults wave 2 (age 20–60) sample is used. CGI and TMF are unidimensional scales, with higher values indicating higher femininity. "f" refers to the femininity and "m" refers to the masculinity score of the two-dimensional scales BSRI, OSRI, Magliozzi and SOM. CFNI and CMNI are the conformity to feminine and masculine norm inventories. *First comp* combines the first factor from a principal component analysis of all alternate gender identity scales in a given sub-sample excluding CGI. N=2,659 for CGI, TMF, Magliozzi and SOM measures; N=662 for OSRI measures; N=659 for CFNI; N=662 for CMNI; N= 676 for BSRI measures. The significance levels are: *** p<0.005, ** p<0.01, * p<0.05.

a correlation of 0.874 (once the second set of responses is inverted). Thus, aside from strong correlations with alternative measures of gender identity used in prior research, our CGI measure exhibits stability over time.

2.2. Gender Identity and Behavior

Before proceeding to our empirical analysis of the relationships between CGI and economic behaviors and outcomes, we provide a framework for understanding why such relationships may exist, beyond those with an individual's binary categorization as a man or woman. We employ a simple utility framework based on the richer model in Akerlof and Kranton (2000), building on their premise that identity is based on social categories, c, to which an individual belongs. For example, an individual responding to

a survey question eliciting their binary gender classification, may belong either to the category "male" or "female."

Assume that an individual derives utility from the material consequences produced by actions taken, a, as well as from the individual's sense of identity, I,

$$U = U(a, I).$$

The actions correspond, for example, to selecting a risky option or choosing a profession. These actions can impact the individual's utility through standard consequential considerations. Actions can also impact utility through their effect on the individual's identity. Following Akerlof and Kranton (2000), we assume that an individual's identity is influenced by several factors:

$$I = I(a, c, P, \gamma_c).$$

These factors include the individual's actions (a), the social categories to which the individual belongs (c), and social prescriptions (P) regarding appropriate behavior for someone belonging to these social categories. While an individual may belong to several social categories, we are interested in the categorization of individuals as a "woman" or a "man". With respect to how prescriptions influence the individual's gender identity, assume that P represents actions that men or women "should" take.

The final parameter, $\gamma_c \in [0, 1]$, captures the strength of an individual's identification with their binary gender category. That is, a woman can vary in the degree to which she identifies as "feminine" (γ_f) and a man in the degree to which he perceives himself as "masculine" (γ_m). We assume that an individual can strongly identify with a category ($\gamma_c = 1$), strongly reject that category ($\gamma_c = 0$), or exhibit varying degrees of intermediate identification ($\gamma_c \in (0, 1)$). This self-perception, γ_c, determines the degree to which the individual derives utility or disutility from following the behavioral prescriptions for someone with the individual's gender category. For the case of binary gender categories, and in our primary empirical approach, we assume that identification with one's own gender category is inversely related to identification with the other category, such that $\gamma_f = (1 - \gamma_m)$.[9] Gender identity is then a continuous measure with two poles representing either strong identification with one's own or with the other binary category.

To concretely illustrate the above concepts, consider someone selecting between two potential professional tracks—one stereotypically male ($a = 1$), like construction, and one stereotypically female ($a = 0$), like nursing. The prescription (P) is that someone belonging to the category "man" should choose $a = 1$, while someone belonging to the category "woman" should choose $a = 0$. In this case, the individual's sense of gender

[9]The correlation patterns in Table 1 provide support for this assumption. However, it need not necessarily be the case that an individual who identifies as more masculine must also identify as less feminine, as reflected in notions of "androgynous" gender identity (Bem, 1974). In Subsection 5.7, we empirically investigate the value of separate measures of γ_f and γ_m.

identity might be described as $I = 1 - |\gamma_c - 1_{a=P}|$, with γ_c corresponding to the degree to which a man perceives himself as masculine or a woman perceives herself as feminine and the indicator function $1_{a=P}$ taking the value of 1 if the individual takes the action prescribed for his or her category and 0 otherwise.[10] In this example, a man who identifies as very masculine ($\gamma_m = 1$) will obtain a more positive sense of identity when following the behavioral prescriptions for a man ($a = 1$). Conversely, another man who rejects masculine identity ($\gamma_m = 0$) will obtain a stronger sense of identity by rejecting gender stereotypes and adopting actions traditionally prescribed for a woman ($a = 0$). As γ_c varies, the emphasis that an individual places on acting in the manner prescribed for his or her gender category also changes.

Based on this simple framework, we propose two separate channels through which we may observe a relationship between an individual's behavior (a) and the degree to which that individual identifies (γ_c) with their binary gender category. First, if we assume that the individual's actions are (largely) exogenous but identity is malleable, then identity can shift to maximize the utility from following or violating behavioral prescriptions. Thus, a man forced (e.g., by societal expectations) into a stereotypically male career track may, *ceteris paribus*, more strongly identify as masculine to increase his identity-based utility. Second, if the individual's identity is (largely) fixed but the individual has agency over actions, then the actions may be influenced by the strength of gender identity. That is, a man who views himself as very masculine will be more likely to select stereotypically male career paths holding constant the economic benefit from doing so. Regardless of whether the individual adjusts actions to concord with gender identity or vice versa—or, as is more likely the case, that both forces are at play in a setting where actions and identity are determined endogenously—this framework illustrates why we may observe a correlation between behavior and strength of identification with the social categories woman and man, even after accounting for the role played by assignment to one of these categories.

We compare this prediction with one based on a version of this framework in which an individual's strength of identification with gender categories are irrelevant—e.g., when γ_c is absent or when its effect on I is independent of the individual's actions. Under this interpretation, which is closer in spirit to that of Akerlof and Kranton (2000), the individual's utility from identity is determined mainly by social prescriptions for how someone who belongs to their assigned gender category should behave and the individual's actions. Our empirical tests, therefore, aim to identify the degree to which individuals' self-reported strength of identification with their binary gender categories ($\gamma_f = 1 - \gamma_m$) are related to their behavior and outcomes (a), against the null hypothesis of no such

[10] In this simple illustrative example, the individual's overall utility might be represented as $U = v(a) + I$, reflecting additively separable utility from the direct consequences of actions and an additional identity-based component of utility. We provide this example only for illustration, recognizing that there are varied and richer ways of capturing the relationships of interest.

relationship.

3. Empirical Strategy

To explore whether measures of CGI correlate with variation in economic decisions and outcomes, *beyond* what can be accounted for by a binary indicator of gender, we apply the same analytical approach across all outcome variables from the diverse samples. We pre-registered this approach for the analysis of our first sample, Swiss Uni.[11]

We begin by identifying gender gaps of the kind typically studied by economists, by estimating the following regression:

$$Y = \alpha + \beta \; FemBin + \epsilon. \tag{1}$$

We test the null hypothesis that $\beta = 0$, which captures whether a binary classification of gender ($FemBin$) correlates with the relevant behavior, preference, or outcome, Y. $FemBin$ takes the value one for women and zero for men (corresponding to c in the framework in Section 2.2). Our one-sided alternative hypothesis is that β takes the same sign as indicated by prior research on gender preference gaps. For the next steps of our analysis, we retain only those variables for which we reject the null hypothesis at $p < 0.05$.

Conditional on a statistically significant relationship for binary gender, we next investigate the relationship between the outcome measure and CGI. First, we test whether our CGI measure correlates with the elicited outcome variable when it replaces the binary gender measure. To do so, we run regressions of the form,

$$Y = \alpha' + \gamma \; FemCGI + \epsilon. \tag{2}$$

We test the null hypothesis that $\gamma = 0$ against the alternative hypothesis that γ takes the same sign as β, for each outcome Y. We only proceed to the next step when we can reject the null at $p < 0.05$.

Finally, for our principal test, we investigate whether CGI is statistically significantly correlated with the outcome measure when we additionally account for the relationship with binary gender. To test this, we run regressions of the form,

$$Y = \alpha'' + \beta' \; FemBin + \gamma' \; FemCGI + \epsilon, \tag{3}$$

using the same dependent variable as above. Our null hypothesis is that CGI only operates through binary gender, i.e., $\gamma' = 0$. The alternative hypothesis—and our principal test—is that γ' takes the same sign as β from regression (1) for a given outcome Y. This provides evidence on whether CGI correlates with measures of economic preferences and

[11] Our pre-analysis plan is available at https://osf.io/phyt6/.

outcomes, even when accounting for binary gender.

4. Design and Sample Descriptions

This section briefly describes the four different datasets we collected, including the main preference and outcome measures. For a detailed description of each separate part of the data collection, see Appendix B.[12]

Table 2 provides an overview of the main outcome measures elicited in each data collection. A more detailed overview of these variables is available in Appendix Table A1.

Table 2: Overview of Main Outcome Variables

	Dataset			
Outcome measures	**Swiss Uni**	**U.S. Adults**	**Swiss Teens**	**Swedish Youths**
Incentivized preferences				
Risk	Y	Y		
Competitiveness	Y			
Efficiency	Y			Y
Overconfidence	Y			
Non-Incentivized preferences				
Staircase risk	Y			
Risk	Y	Y	Y	Y
Financial risk		Y		
Competitiveness	Y	Y	Y	Y
Redistribution	Y	Y		
Altruism	Y	Y	Y	Y
Educational and labor market outcomes				
Female educational track share		Y	Y	Y
Math/Language skill requirement			Y	
Income		Y		
Full-time homemaker		Y		
Weekly average work hours		Y		
Female industry share		Y		
Managerial responsibilities		Y		
Performance pay		Y		
Wage negotiation		Y		
Work flexibility		Y		

Notes: Appendix Table A1 provides more details on each variable.

We elicited our main variable of interest, our *single-item measure of CGI*, in all samples.[13] In Swiss Uni and U.S. Adults, we additionally measured how a person de-

[12]Full instructions for three of the four data collections are available in Appendix C. The data collection using Swedish Youths was implemented as part of another ongoing study by one of the authors of this paper (Ranehill). The full instructions for this data collection will be made available when the main study is completed.

[13]Across the studies, the timing of when CGI and binary gender were elicited relative to the outcome variables varied (see Appendix B for details).

scribes being seen by others on our masculinity-femininity scale. In U.S Adults, we also elicited two secondary CGI measures asking participants to place themselves on our masculinity-femininity scale relative to men and relative to women, separately. In the Swiss Uni, U.S. Adults, and Swiss Teens samples, we also elicited participants' *binary gender* by asking them to report one of two categories, "male" or "female" which we code as a binary variable taking the value 1 for women and 0 for men.[14] For Swiss Uni, U.S. adults and Swedish Youth we additionally elicited non-binary but discrete gender identity, where participants could select the category "non-binary" or "other."[15]. Due to the small number of individuals providing non-binary gender classifications, we omit participants selecting a category other than "male" or "female" from our primary sample.[16] Further, it reflects our primary research interest in studying the extent to which differences in gender identity may explain heterogeneous economic behavior for individuals who identify as a woman or a man. Across all samples, we also exclude individuals for whom we have multiple binary measures if there is any inconsistency. For the U.S. adults sample, for example, we also observe participants' gender previously reported to Prolific Academic. For either one of these two reasons (non-binary or inconsistency), we exclude a total of 95 participants (less than 2% of all observations). Appendix Table A2 provides a detailed overview of the exclusions made for each sample.

Our first data collection, Swiss Uni, focused primarily on the correlation between CGI and incentivized measures of economic preferences.[17] This dataset was collected in September and October 2021 through an online experiment and comprises 584 student participants from the University of Zurich and the Swiss Federal Institute of Technology (ETH). This dataset primarily elicited preference measures through widely-used incentivized tasks for which earlier studies have reliably documented gender gaps. *Risk preferences* were elicited through a one-shot investment task used by Gneezy and Potters (1997). Our measure of *Competitiveness* follows the design introduced by Niederle and Vesterlund (2007), with slight modifications due to the online setting. Preferences

[14]For U.S. adults, we took the wording from the U.S. Census questionnaire and asked "What is your sex?". In Swiss Uni and Swiss Teens, we asked "Are you male or female?"

[15]For Swiss Uni we asked "What is your current gender?" with the answer options Woman/Man/Transgender/Non-binary/Other/Prefer not to answer. In U.S. Adults (Wave 2) we asked "What gender are you currently?" with the answer options Man (including Trans Male/Trans Man)/Woman (including Trans Female/Trans Woman)/Non-binary/Would rather not say. For Swedish Youth we asked "Do you identify as a woman or a man?" to which participants could answer Woman, Man or Other

[16]In the Appendix, we show that our main results for the U.S. Adults and Swiss Uni samples are very similar when including the responses of individuals who identify as non-binary for our analysis (see Appendix Tables A28 - A31). The U.S. Adults and Swiss Uni samples are the ones in which we elicited both a binary measure of gender and a non-binary but discrete measure of gender identity.

[17]We implemented changes relative to the pre-registration (https://osf.io/phyt6/) mainly because we adapted the experiment, which was originally intended as a laboratory experiment, to an online format due to COVID-19 restrictions. A detailed description of the departure from the initial pre-registration is available in the document "Updates to Pre-registration Final.pdf" available at https://osf.io/phyt6/ and summarized in Appendix B.

for *Equality versus efficiency* were elicited by implementing 15 graphical budget sets involving inter-personal allocations, similarly to Fisman, Kariv and Markovits (2007). Finally, our measure of *Overconfidence* is based on three measures of relative overplacement (Moore and Healy, 2008); for example, participants had to rank their performance on matrix reasoning items (similar to Raven's matrices) relative to a reference group of other participants.

As part of the Swiss Uni data collection, we administered a follow-up survey two weeks after the main experiment that elicited non-incentivized preference measures, such as self-reported measures of risk seeking, competitiveness, attitudes towards redistribution, and altruism. The main purpose of the follow-up survey was to duplicate the elicitation of the CGI measure using a different (reversed) scale to provide the opportunity to account for possible measurement error in our statistical analysis, following the ORIV procedure of Gillen, Snowberg and Yariv (2019).

Our second data collection, U.S. Adults, focused primarily on the correlation between CGI and economic choices and outcomes in a broader sample of U.S. residents of working age. For this analysis, we recruited 3,902 respondents aged 30–60 years in two waves (in March 2022 and March 2023) through Prolific Academic. We restrict our sample to individuals at least 30 years old in order to target a population likely to have at least a few years of work experience.[18] In this sample, we collected a broad set of variables related to demographics, family, education, labor market participation and job attributes, such as study field of college major, employment status, income, and past wage negotiations. We further elicited the incentivized risk measure and the same unincentivized preference measures as in our Swiss Uni sample. As a proxy for real-world (financial) risk taking, we asked whether a respondent actively trades in securities.

Our third data collection, Swiss Teens, focused on educational choices and unincentivized economic preferences among adolescents. We conducted two surveys comprising 1,740 Swiss teenagers. The respondents were recruited through two newsletters sent from the largest Swiss online platform for apprenticeship search. Vocational education and training is the most common type of education after compulsory schooling (9^{th} grade) in Switzerland. Most respondents (91 percent) were 8^{th} and 9^{th} grade students (with an average age of 14.8 years) who were planning to do an apprenticeship after 9^{th} grade. At the time of our surveys (December 2021 and March 2022), the 8^{th} graders were considering potential future apprenticeships. This means exploring different professions at firms through trial apprenticeships, which typically last 1-5 days. The 9^{th} graders were further along—52 percent of them had already signed a contract for their apprenticeships with a specific company. In the first survey, we elicited unincentivized preference measures

[18] In Wave 2, which also serves as the dataset for the validation of our CGI measure reported in Section 2.1, we additionally recruited 751 participants aged 20–29 only for use in this validation analysis. For these respondents, we did not collect the full set of outcome measures.

for risk-seeking, competitiveness, and altruism. Moreover, for respondents in the first survey, we were able to merge their survey responses to profile data from the platform, covering real life decisions in terms of apprenticeship applications. In the second survey, we elicited detailed information on all trial apprenticeships the respondents had done. Based on these data, we construct three main variables that characterize the respondents' occupational preferences in terms of gender composition (female share) and skill requirements (math and language) of their chosen apprenticeship profession.[19]

Our fourth data collection, Swedish Youths, includes 1,041 Swedish students (age 18-19 years) in the final year of the natural science track in Sweden and was part of a larger experiment. The experiment took place in the beginning of 2022, in the weeks before the students applied for their preferred educational field for subsequent university studies. This dataset comprises a measure of incentivized preferences for equality versus efficiency (based on 20 allocation decisions), as well as the unincentivized preference measures for risk, competitiveness, and altruism. Finally, the dataset comprises information on participants' intended fields of undergraduate studies, allowing us to construct a measure of the gender composition of the students' intended field of study using administrative data.

5. Results

In this section, we first describe the distributions of self-reported CGI across our four samples. Thereafter, we study the specific relationships between CGI and the elicited preference measures and educational and labor market outcomes. To provide a broad test of our main hypothesis regarding the general value of CGI, we then present an overview of the statistical significance across all the tests presented in the paper involving CGI and our outcome measures. Finally, we compare the relative strength of the relationship between CGI and the outcome measures to those for the other gender identity scales described in Section 2.

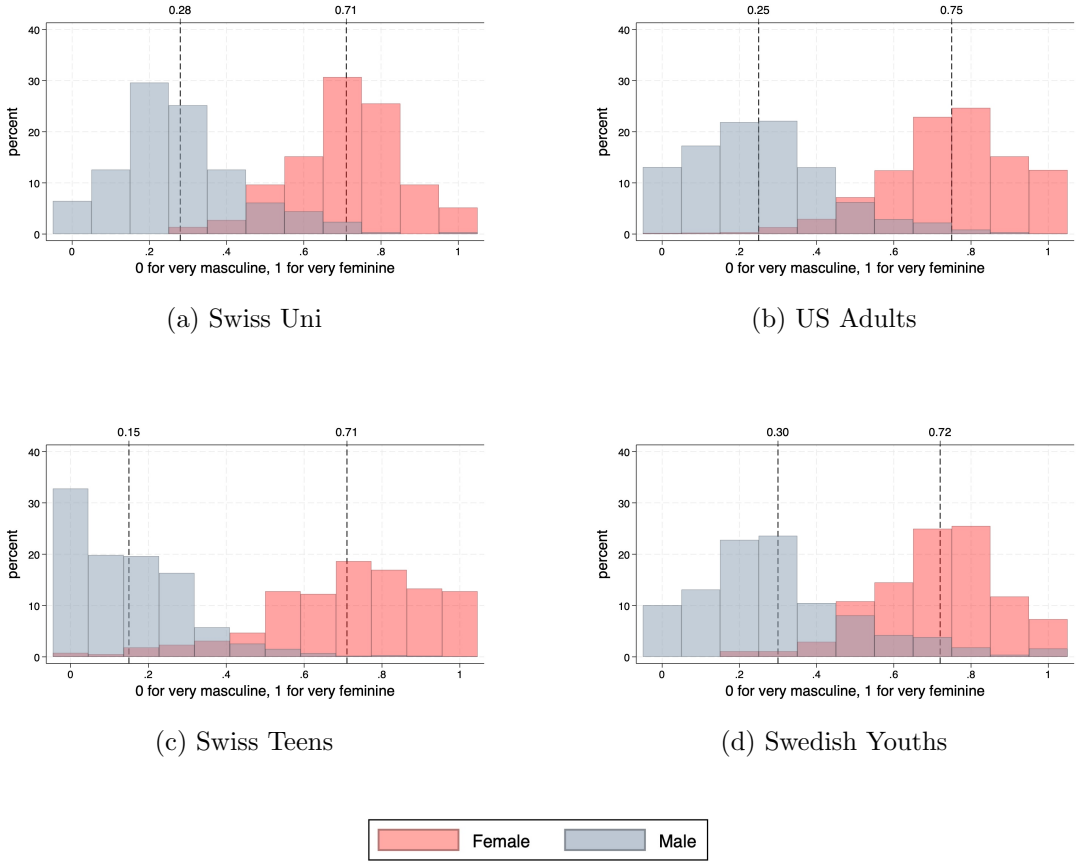
5.1. Continuous Gender Identity in Four Samples

Figure 1 presents the distributions of self-reported gender identity by binary gender across our four samples.[20] Two notable patterns are visible. First, in each sample, as expected, there is a substantial gap in the average gender identity of women and men, with women tending to report stronger feminine identity. Second, in all samples, the within-gender variation is substantial. In three samples, the distributions of CGI for each gender are

[19]We obtain the gender-composition measure for each occupational field using administrative data and the skill requirements from an independent set of expert ratings. For details, see Appendix B.

[20]The distribution of CGI for all participants who report a non-binary gender category in any of our studies (N=78) is shown in Appendix Figure A4. The modal answer is at the mid-point of our scale, with nearly 45% of non-binary respondents choosing this response.

Figure 1: Continuous Gender Identity by Binary Gender



(a) Swiss Uni

(b) US Adults

(c) Swiss Teens

(d) Swedish Youths

Notes: Continuous Gender Identity (CGI) represents our single question, ranging from "very masculine" to "very feminine". CGI is rescaled from 0 to 1 where 1 is very feminine. The dashed lines mark the sample means for women and men. The sample of Swiss Uni (N=584) is collected through online incentivized experiments. The sample of U.S. adults (N=4,653) aged 20–60 is collected through the company Prolific Academic. The sample of Swiss Teens (N=1,740) is collected through the online "job board" company Yousty. The sample of Swedish Youths (N=1,041) is collected through contacts with Swedish high schools.

generally symmetrically distributed and there is substantial overlap in the distributions between women and men. The sample of Swiss Teens (Panel C) differs in this regard, with fat tails in the extremes for both genders but particularly for men.

5.2. Continuous Gender Identity and Economic Preferences

5.2.1 *Incentivized Behavioral Measures*

Table 3 presents the relationships between CGI and our incentivized behavioral measures. In this table and henceforth, we standardize CGI within each sample to have a mean of zero and a standard deviation of one, with a higher value indicating a more feminine identity. As much of our analysis will be shown using similar tables, we describe the structure of Table 3 carefully here. Each column of Panel A tests the relationship between binary gender (female) and the corresponding outcome variable, as specified in Equation (1). We document statistically significant gender differences with the expected sign for our measures of risk, competitiveness, and a tendency to prioritize efficiency over

Table 3: CGI and Incentivized Behavioral Measures

Measure	Risk		Competitiveness	Efficiency		Overconfidence
Sample	Swiss Uni	US Adults	Swiss Uni	Swiss Uni	Swedish Youths	Swiss Uni
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. Binary gender						
Female	-0.378***	-0.215***	-0.128***	-0.145***	-0.161***	-0.002
	(0.081)	(0.032)	(0.041)	(0.023)	(0.017)	(0.083)
R^2	0.036	0.012	0.017	0.062	0.078	0.000
Panel B. Gender identity						
CGI (feminine)	-0.169***	-0.092***	-0.056***	-0.066***	-0.068***	
	(0.040)	(0.016)	(0.020)	(0.012)	(0.009)	
R^2	0.028	0.008	0.013	0.052	0.055	
Panel C. Binary gender and gender identity						
Female	-0.305*	-0.195***	-0.110	-0.110**	-0.135***	
	(0.153)	(0.057)	(0.072)	(0.043)	(0.026)	
CGI (feminine)	-0.045	-0.012	-0.011	-0.022	-0.017	
	(0.074)	(0.029)	(0.036)	(0.021)	(0.013)	
R^2	0.036	0.012	0.017	0.064	0.079	
Panel D. Sample split by binary gender						
Female: CGI (feminine)	-0.039	-0.001	-0.029	-0.042**	-0.014	
	(0.060)	(0.023)	(0.030)	(0.017)	(0.011)	
R^2	0.002	0.000	0.004	0.023	0.003	
Male: CGI (feminine)	-0.017	-0.012	0.013	0.013	-0.010	
	(0.063)	(0.023)	(0.030)	(0.017)	(0.014)	
R^2	0.000	0.000	0.001	0.002	0.001	
Test: male sample split CGI = female sample split CGI						
p-value	0.794	0.717	0.295	0.016	0.696	
Observations	584	3,902	584	584	1,041	584
Mean of Dependent Variable	0	0	0.413	0.538	0.550	0

Notes: The table presents results from regressing the incentivized measures of risk, competitiveness, equality versus efficiency and overconfidence on our standardized (mean=0, SD=1) measure of CGI. CGI is standardized (mean=0, SD=1) separately for the male and female samples in Panel D. Swiss Uni, U.S. Adults (age 30–60), and Swedish Youths samples are used. The estimates in each column and panel come from a separate regression. Risk preference is a standardized measure of participants' investment decisions (mean=0, SD=1). Competitiveness is a dummy that takes the value 1 for those who chose to compete in the competitive task. Our measure of preferences for equality versus efficiency is measured in deciles, with increasing numbers indicating higher priority for efficiency. Overconfidence is measured as relative overplacement and is standardized (mean=0, SD=1). Robust standard errors are reported in parentheses. Significance levels in Panel A are determined by a two-tailed test. Significance levels in Panels B, C and D are determined by a one-tailed test in the direction of the coefficient estimate of female in Panel A. Equality of sample split coefficients in Panel D are tested with a Wald test.*** p<0.005, ** p<0.01, * p<0.05.

equality. For overconfidence, we fail to find a significant gender difference and therefore omit this measure from the remaining analysis, as pre-registered.[21]

In Panel B, we regress the remaining outcome measures on CGI (see Equation (2)). Consistent with the gender differences in Panel A and the relationships between CGI and gender in Figure 1, we see that people who identify as more feminine are also less risk-

[21]While not pre-registered, it is still interesting to know whether reported gender identity correlates with being overconfident. Regressing overconfidence on CGI (standardized), the estimated correlation coefficient is 0.039 with a standard error of 0.044. Also within-gender, we find no evidence that being overconfident correlates with CGI. While some earlier studies find that women tend to be less overconfident, the absence of such a relationship is consistent with the results of a recent meta-study (Bandiera et al., 2022).

loving, less likely to enter competition, and have a lower preference for efficiency relative to equality. These associations are all statistically significant at the $p < 0.005$ level.

Next, in Panel C, we implement our principal test and explore whether CGI has any statistically significant association with our preference measures once we control for binary gender. As indicated in Panel C, the CGI coefficients all have the expected sign, but are small in magnitude and none of them are statistically significantly different from zero, even at the $p < 0.05$ level.

Finally, In Panel D, we investigate the relationships between CGI and the preference measures separately for men and women. Out of the ten estimates in Panel D, only one is statistically significant at $p < 0.01$ (efficiency preferences for women in the Swiss Uni sample). The Wald tests reported at the bottom of Panel D test whether the relationships between CGI and the relevant outcome measures differ significantly for men and women, generally finding no significant differences.[22]

Based on the above analysis, we conclude that there is at best a weak relationship between our measure of CGI and variation in incentivized behavioral measures, once controlling for the portion of the relationship accounted for by binary gender. This is further reinforced by a comparison of the R^2 values in Panels A and C of Table 3, which show little change, if any, with the introduction of our CGI measure.

5.2.2 *Unincentivized Preference Measures*

Table 4 shows the relationships between CGI and unincentivized risk preference and competitiveness measures. Panel A replicates (at $p < 0.05$) gender gaps commonly found in the literature for all measures and samples. Panel B further shows that CGI strongly correlates with all unincentivized risk and competitiveness outcomes in the expected direction.

In contrast to the results for the incentivized behavioral measures, Panel C shows that CGI can account for some of the variation in several of the unincentivized behavioral outcomes beyond binary gender. Across the ten regressions in Table 4, the coefficient for CGI has the expected sign and is statistically significant at $p < 0.05$ in eight cases, four of which have a significance level of $p < 0.005$. In particular, we observe strong relationships between CGI and self-reported risk-seeking and competitiveness for both the U.S. Adults and Swiss Teens samples. We interpret these findings as providing evidence that gender identity can exhibit substantial correlation with unincentivized measures of risk and competitiveness, even when accounting for the relationships with binary gender. This is further supported by the increases in R^2 when comparing Panels A and C in Table 4, which are larger than those observed in Table 3, though still small in absolute

[22] As a complement to the Wald tests in each of our main results tables (Table 3 through Table 7), Appendix Table A32 through Table A36 report complementary regressions of the interactions between CGI and binary gender. The results are generally similar across the two approaches.

Table 4: CGI and Unincentivized Risk and Competitiveness Measures

Measure	Staircase Risk	Risk				Financial Risk	Competitiveness			
Sample	Swiss Uni	Swiss Uni	U.S Adults	Swiss Teens	Swedish Youths	US Adults	Swiss Uni	U.S Adults	Swiss Teens	Swedish Youths
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A. Binary gender										
Female	-0.211*	-0.444***	-0.387***	-0.285***	-0.243***	-0.268***	-0.388***	-0.405***	-0.311***	-0.153*
	(0.083)	(0.081)	(0.031)	(0.071)	(0.062)	(0.015)	(0.081)	(0.031)	(0.071)	(0.062)
R^2	0.011	0.049	0.038	0.020	0.015	0.075	0.038	0.041	0.024	0.006
Panel B. Gender identity										
CGI (feminine)	-0.137***	-0.234***	-0.206***	-0.171***	-0.109***	-0.119***	-0.194***	-0.218***	-0.181***	-0.109***
	(0.044)	(0.040)	(0.016)	(0.036)	(0.031)	(0.008)	(0.040)	(0.016)	(0.036)	(0.029)
R^2	0.019	0.055	0.042	0.029	0.012	0.059	0.038	0.048	0.033	0.012
Panel C. Binary gender and gender identity										
Female	0.031	-0.188	-0.152***	0.124	-0.182*	-0.224***	-0.212	-0.145**	0.087	0.022
	(0.148)	(0.148)	(0.056)	(0.166)	(0.103)	(0.027)	(0.138)	(0.057)	(0.168)	(0.101)
CGI (feminine)	-0.149*	-0.157*	-0.144***	-0.227***	-0.041	-0.027*	-0.108	-0.159***	-0.220***	-0.117**
	(0.078)	(0.073)	(0.029)	(0.085)	(0.051)	(0.013)	(0.068)	(0.029)	(0.085)	(0.048)
R^2	0.019	0.058	0.044	0.030	0.015	0.076	0.042	0.049	0.033	0.012
Panel D. Sample split by binary gender										
Female: CGI (feminine)	-0.050	-0.081	-0.018	-0.051	0.058	-0.001	-0.015	0.068	0.021	0.015
	(0.064)	(0.057)	(0.025)	(0.060)	(0.046)	(0.010)	(0.054)	(0.024)	(0.055)	(0.044)
R^2	0.003	0.007	0.000	0.003	0.003	0.000	0.000	0.005	0.000	0.000
Male: CGI (feminine)	-0.123*	-0.105	-0.143***	-0.154***	-0.097*	-0.029***	-0.110*	-0.253***	-0.238***	-0.163***
	(0.068)	(0.064)	(0.024)	(0.045)	(0.052)	(0.011)	(0.062)	(0.024)	(0.047)	(0.050)
R^2	0.015	0.011	0.020	0.024	0.009	0.003	0.012	0.064	0.056	0.027
Test: male sample split CGI = female sample split CGI										
p-value	0.548	0.867	0.000	0.062	0.027	0.073	0.297	0.000	0.000	0.018
Observations	584	584	3,902	786	1,041	3,902	584	3,902	792	1,041
Mean of Dependent Variable	0	0	0	0	0	0.401	0	0	0	0

Notes: The table presents results from regressing unincentivized measures of risk attitudes and competitiveness on our standardized (mean=0, SD=1) measure of CGI. CGI is standardized (mean=0, SD=1) separately for the male and female samples in Panel D. Swiss Uni, U.S. Adults (age 30-60), Swiss Teens, and Swedish Youths samples are used. The estimates in each column and panel come from a separate regression. Staircase risk is a categorical certainty equivalence measure of risk-taking based on a series of hypothetical allocation decisions. Risk is a self-reported measure of risk-taking. Financial risk is a dummy that takes the value 1 for those who report to actively trade in securities. Competitiveness is a self-reported measure of competitiveness. A higher value means higher risk taking or competitiveness. All measures except financial risk are standardized (mean=0, SD=1). Robust standard errors are reported in parentheses. Significance levels in Panel A are determined by a two-tailed test. Significance levels in Panels B, C and D are determined by a one-tailed test in the direction of the coefficient estimate of female in Panel A. Equality of sample split coefficients in Panel D are tested with a Wald test. *** p<0.005, ** p<0.01, * p<0.05.

terms.

Splitting the sample by binary gender (Panel D) reveals that the relationships between CGI and the unincentivized preference measures are driven by men. All the coefficients for men have the expected sign, nine of ten are statistically significant at $p < 0.05$ and only one is not statistically significant at conventional levels ($p = 0.052$). Six of ten coefficients are highly statistically significant at $p < 0.005$. The Wald tests reported at the bottom of Panel D in Table 4 show that the relationships between CGI and the preference measures are often highly statistically significantly different for men and for women (see, also, Appendix Table A33).

We next consider self-reported attitudes toward redistribution in society and self-reported altruism (Table 5). For all but one regression (altruism in the U.S. sample,

Table 5: CGI and Unincentivized Distributional Preferences

Measure	Redistribution		Altruism			
Sample	Swiss Uni (1)	U.S. Adults (2)	Swiss Uni (3)	U.S. Adults (4)	Swiss Teens (5)	Swedish Youths (6)
Panel A. Binary gender						
Female	0.201*	0.063*	0.249***	0.001	0.210***	0.551***
	(0.083)	(0.032)	(0.082)	(0.032)	(0.071)	(0.060)
R^2	0.010	0.001	0.016	0.000	0.011	0.076
Panel B. Gender identity						
CGI (feminine)	0.106***	0.053***	0.095*		0.074*	0.224***
	(0.041)	(0.017)	(0.041)		(0.035)	(0.032)
R^2	0.011	0.003	0.009		0.005	0.050
Panel C. Binary gender and gender identity						
Female	0.082	-0.072	0.279*		0.412**	0.489***
	(0.155)	(0.055)	(0.166)		(0.165)	(0.099)
CGI (feminine)	0.073	0.083***	-0.018		-0.112	0.042
	(0.076)	(0.029)	(0.084)		(0.082)	(0.052)
R^2	0.012	0.003	0.016		0.013	0.077
Panel D. Sample split by binary gender						
Female: CGI (feminine)	-0.136	-0.115	-0.047		-0.051	0.052
	(0.063)	(0.023)	(0.071)		(0.054)	(0.046)
R^2	0.019	0.013	0.002		0.013	0.003
Male: CGI (feminine)	0.193***	0.196***	0.020		-0.046	0.013
	(0.062)	(0.024)	(0.068)		(0.046)	(0.053)
R^2	0.037	0.038	0.000		0.002	0.000
Test: male sample split CGI = female sample split CGI						
p-value	0.000	0.000	0.477		0.916	0.534
Observations	584	3,902	584	3,902	798	1,041
Mean of Dependent Variable	0	0	0	0	0	0

Notes: The table presents results from regressing unincentivized distributional measures on our standardized (mean=0, SD=1) measure of CGI. CGI is standardized (mean=0, SD=1) separately for the male and female samples in Panel D. Swiss Uni, U.S. Adults (age 30-60), Swiss Teens and Swedish Youths samples are used. The estimates in each column and panel come from a separate regression. Redistribution is a measure of how much economic redistribution one wants in society. Altruism is a measure of how much one would donate out of a windfall gain. A higher value means a greater willingness to redistribute or donate. All outcome measures are standardized (mean=0, SD=1). Robust standard errors are reported in parentheses. Significance levels in Panel A are determined by a two-tailed test. Significance levels in Panels B, C and D are determined by a one-tailed test in the direction of the coefficient estimate of female in Panel A. Equality of sample split coefficients in Panel D are tested with a Wald test.*** p<0.005, ** p<0.01, * p<0.05.

Column 4 in Panel A), we replicate statistically significant gender gaps commonly found in the literature. Further, Panel B indicates that CGI correlates with attitudes towards redistribution and altruism in the expected direction for the five outcome measures that move to this stage of the analysis. However, once controlling for binary gender in Panel C, the relationships between CGI and stated distributional preferences are statistically significant at $p < 0.005$ for only one measure, attitudes toward redistribution among U.S. Adults, and are not statistically significant for other measures. Moreover, the sign for the coefficient for CGI in Panel C has the wrong sign in two cases. As in Table 3, we again observe very small increases in R^2 when comparing Panels A and C.

Splitting the sample by gender in Panel D reveals stronger relationships for men. Two of five regressions for men and none for women yield statistically significant ($p < 0.005$)

coefficients for CGI in the expected direction (redistribution in the Swiss Uni and U.S. Adults samples). These gender differences are statistically significant (see the reported p-values of Wald tests in Table 5 and Appendix Table A34). This is consistent with our observation in Table 4 of stronger relationships between CGI and self-reported preference measures for men than for women.

5.3. Continuous Gender Identity and Educational Choices

Table 6 shows the relationships between CGI and educational outcome measures. Each outcome measure corresponds to the choice of a particular educational track, coded either for the share of women in a study field or apprenticeship (in each case, obtained separately from administrative data), or the math and language requirements of the apprenticeship (evaluated by an independent panel of experts).

Not surprisingly, given the construction of these outcome measures, we find substantial gender gaps in the gender composition (share of women) of the chosen or intended educational track in all three samples (see Panel A).[23] We also find that Swiss female adolescents tend to select apprenticeship tracks with lower math and higher language requirements compared to males. In Panel B, CGI is strongly correlated with all educational outcomes in all samples.

When we control for binary gender in Panel C, all the coefficients for CGI have the expected signs. The relationships with choosing a predominantly female educational track are highly statistically significant at $p < 0.005$ for two samples (U.S. Adults and Swiss Teens). The relationship with lower math and higher language requirements are statistically significant at $p < 0.05$ and $p < 0.01$, respectively. Looking at the changes in R^2 between Panels A and C, we again observe generally small increases, even when the coefficients are highly statistically significant. We thus conclude that CGI appears to be correlated with educational track choices and can account for additional variation in gender-typical educational choices beyond the one already accounted for by binary gender, though the strength and size of these relationships vary across our samples.

Turning to the within-gender relationships in Panel D, the association between CGI and educational choices is statistically significant at $p < 0.05$ in five of ten cases. Three out of these five coefficients of CGI are highly statistically significant at $p < 0.005$. We see no clear pattern that the relationships between CGI and educational choices in Panel C are primarily driven by one of the two genders. Consistently, all Wald tests fail to reject the null hypothesis that the estimated correlation coefficients are the same for boys and girls (see also Appendix Table A35).

[23] Although it is worth noting that the gender share measures are constructed independently from administrative data, and not from the responses in our data. For the intended college major chosen by Swedish Youths, the gap is relatively small, which is likely due to the sample only including students in the science-math high school track.

Table 6: CGI and Educational Choice

Measure	Female Educational Track Share			Skill Requirements	
				Math	Language
Sample	U.S Adults	Swiss Teens	Swedish Youths	Swiss Teens	Swiss Teens
	(1)	(2)	(3)	(4)	(5)
Panel A. Binary gender					
Female	0.125***	0.383***	0.093***	-0.917***	0.913***
	(0.007)	(0.012)	(0.009)	(0.047)	(0.047)
R^2	0.108	0.404	0.090	0.204	0.203
Panel B. Gender identity					
CGI (feminine)	0.061***	0.167***	0.039***	-0.403***	0.404***
	(0.004)	(0.006)	(0.005)	(0.024)	(0.023)
R^2	0.104	0.315	0.063	0.163	0.163
Panel C. Binary gender and gender identity					
Female	0.075***	0.329***	0.079***	-0.762***	0.745***
	(0.014)	(0.023)	(0.014)	(0.084)	(0.083)
CGI (feminine)	0.031***	0.033***	0.009	-0.092*	0.100**
	(0.007)	(0.011)	(0.007)	(0.042)	(0.040)
R^2	0.117	0.408	0.092	0.207	0.206
Panel D. Sample split by binary gender					
Female: CGI (feminine)	0.009*	0.028***	0.007	-0.028	0.135***
	(0.005)	(0.009)	(0.006)	(0.040)	(0.038)
R^2	0.003	0.014	0.002	0.001	0.018
Male: CGI (feminine)	0.024***	0.009	0.005	-0.087**	0.005
	(0.006)	(0.009)	(0.007)	(0.035)	(0.035)
R^2	0.016	0.002	0.001	0.008	0.000
Test: male sample split CGI = female sample split CGI					
p-value	0.098	0.331	0.690	0.105	0.062
Observations	2,279	1,448	1,041	1,433	1,433
Mean of Dependent Variable	0.541	0.442	0.521	0	0

Notes: The table presents results from regressing the educational track measures on our standardized (mean=0, SD=1) measure of CGI. CGI is standardized (mean=0, SD=1) separately for the male and female samples in Panel D. U.S. Adults (age 30–60), Swiss Teens and Swedish Youths samples are used. The estimates in each column and panel come from a separate regression. Female educational track share is the share of women graduating with a bachelor's degree in the chosen field of study in 2020 for U.S. Adults; the share of women from past cohorts of graduates from the apprenticeship for Swiss Teens; the share of women accepted for undergraduate studies in that field the year before our sample made their educational choices for Swedish Youths. Skills requirements are a standardized (mean=0, SD=1) measure based on expert evaluation of the job content in occupations chosen by Swiss Teens to start apprenticeships. Robust standard errors are reported in parentheses. Significance levels in Panel A are determined by a two-tailed test. Significance levels in Panels B, C and D are determined by a one-tailed test in the direction of the coefficient estimate of female in Panel A. Equality of sample split coefficients in Panel D are tested with a Wald test. *** p<0.005, ** p<0.01, * p<0.05.

5.4. Continuous Gender Identity and Labor Market Outcomes

To investigate the relationships between CGI and labor market outcomes, we asked participants in the U.S. Adults sample to provide information on several measures related to labor market outcomes and behaviors (Table A3). To discipline our selection of outcome measures for further study, we implement a selection procedure to minimize false positive relationships between outcome measures and gender for the first step of our analysis. Specifically, for every survey item related to employment and labor market outcomes, we

test whether there is a significant gender gap, correcting for multiple-hypothesis testing (MHT) using the Romano and Wolf (2005) approach. If there is a significant gender gap at the $p < 0.05$ level after this correction, we retain that outcome variable for the second and third stages of our analysis involving CGI and other gender identity measures (in Subsection 5.7). Of the 11 independent labor market outcome measures we collected for U.S. Adults, we obtain eight statistically significant gender gaps using this approach.[24] We then perform the same analysis as for our earlier outcome measures on these eight measures of labor market outcomes. Table 7 presents the results.

Given the procedure for selecting these eight variables, we unsurprisingly observe highly statistically significant gender gaps for all outcomes in Panel A. Compared to men, women earn less, work fewer hours, and are less likely to have managerial responsibilities, be rewarded with variable incentives and to have engaged in wage negotiation. Women are also more likely to be full-time homemakers, work in industries with high female shares, and have jobs that involve flexible work hours. In Panel B, we observe that all the relationships between CGI and these outcomes are also highly statistically significant and have the expected sign.

Panel C shows that, conditional on binary gender, CGI correlates with three outcomes at $p < 0.005$. Accounting for binary gender, individuals who report a more feminine identity earn less, are more likely to be full-time homemakers, and work fewer hours. Two additional outcomes are statistically significant at respectively $p < 0.01$ and $p < 0.05$. This indicates that at least some categories of labor market outcomes have highly statistically significant associations with gender identity, beyond the relationship with binary gender. However, once again we observe that even the relationships that are statistically significant yield modest increases in R^2 in absolute terms—typically less than half a percentage point—when adding CGI as an explanatory variable.

The split-sample analysis in Panel D indicates that the correlation of CGI with the likelihood of being a full-time homemaker is driven primarily by women, while the remaining significant correlations are driven by men.[25] In fact, for men, six out of eight relationships with CGI are significant at $p < 0.05$, of which five are significant at $p < 0.005$. Thus, consistent with earlier observations, we again observe relationships that are stronger for

[24] Appendix Table A3 shows the results from this exploration of potential gender gaps in labor market outcomes for Wave 1 of the U.S. Adults sample, reporting p-values corrected for MHT. We adopted the above selection procedure in Wave 1 of data collection in our U.S. sample; for Wave 2, we only collected labor market-related outcomes that had survived the MHT correction for Wave 1. For the category of variables, *Employment Status*, we find two response categories with substantial gender differences after the MHT correction: whether a respondent is *employed* (by someone else) and whether the respondent is a *full-time homemaker*. Given the relationship between these two response categories (participants could only select one), we retain only the latter one for our analysis (as this more likely reflects an individual's choice to forgo employment). We also performed this analysis on the educational outcomes for U.S. Adults (see Section 5.3), finding the female share in the reported study field to differ substantially between male and female respondents after controlling for MHT.

[25] See the reported p-values of Wald tests in Table 7 and results from a model with female binary gender interacted with CGI in Appendix Table A36.

Table 7: CGI and Labor Market Outcomes

Measure	Income	Full-time Homemaker	Weekly Ave. Work Hours	Female Industry Share	Managerial Responsibilities	Performance Pay	Wage Negotiation	Work Flexibility
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A. Binary gender								
Female	-23.265***	0.099***	-6.644***	0.062***	-0.125***	-0.082***	-0.132***	0.141***
	(1.865)	(0.008)	(0.451)	(0.007)	(0.016)	(0.021)	(0.016)	(0.045)
R^2	0.038	0.043	0.053	0.039	0.016	0.008	0.017	0.005
Panel B. Gender identity								
CGI (feminine)	-10.841***	0.049***	-3.141***	0.025***	-0.061***	-0.043***	-0.066***	0.060***
	(0.958)	(0.004)	(0.222)	(0.003)	(0.008)	(0.010)	(0.008)	(0.022)
R^2	0.033	0.042	0.047	0.026	0.015	0.009	0.017	0.004
Panel C. Binary gender and gender identity								
Female	-16.723***	0.057***	-4.553***	0.060***	-0.078***	-0.038	-0.074**	0.126
	(2.990)	(0.012)	(0.806)	(0.011)	(0.027)	(0.035)	(0.028)	(0.077)
CGI (feminine)	-3.998***	0.026***	-1.278***	0.001	-0.029*	-0.028	-0.035**	0.009
	(1.542)	(0.006)	(0.397)	(0.006)	(0.014)	(0.017)	(0.014)	(0.039)
R^2	0.040	0.047	0.055	0.039	0.018	0.009	0.019	0.005
Panel D. Sample split by binary gender								
Female: CGI (feminine)	2.416	0.025***	-0.155	-0.004	0.027	-0.012	0.002	-0.079
	(1.124)	(0.007)	(0.336)	(0.005)	(0.011)	(0.014)	(0.011)	(0.033)
R^2	0.002	0.006	0.000	0.001	0.003	0.001	0.000	0.006
Male: CGI (feminine)	-6.735***	0.005*	-1.281***	0.005	-0.058***	-0.020	-0.041***	0.085***
	(1.356)	(0.002)	(0.312)	(0.005)	(0.011)	(0.015)	(0.011)	(0.032)
R^2	0.011	0.002	0.009	0.001	0.014	0.002	0.007	0.007
Test: male sample split CGI = female sample split CGI								
p-value	0.000	0.005	0.018	0.145	0.000	0.767	0.009	0.000
Observations	3,902	3,902	3,902	1,994	3,902	1,994	3,902	1,994
Mean of Dependent Variable	61.421	0.062	25.281	0.497	0.400	0.303	0.491	0

Notes: The table regresses the labor market outcomes on our standardized (mean=0, SD=1) measure of CGI. CGI is standardized (mean=0, SD=1) separately for the male and female samples in Panel D. The sample of the U.S. Adults (age 30–60) is used. The estimates in each column and panel come from a separate regression. Income is a self-reported categorical measure in thousands of U.S. dollars. Full-time homemaker is a dummy for working full-time at home. Weekly average work hours is a proxy measure constructed from weeks worked and hours worked, using the product of the two categorical measures and dividing by 52. Female industry share is the share of female employees in a given industry. Managerial responsibilities is a dummy for having managerial responsibilities at work. Performance pay is a dummy that is equal to 1 if the respondent's current or most recent job has performance-related pay. Wage negotiation is a dummy that is equal to 1 if the respondent ever negotiated wage. Work flexibility is a categorical measure of flexibility in working hours where 0 is no flexibility, 0.5 is can adapt work hours and 1 is full flexibility, which is standardized (mean=0, SD=1). Robust standard errors are reported in parentheses. Significance levels in Panel A are determined by a two-tailed test. Significance levels in Panels B, C and D are determined by a one-tailed test in the direction of the coefficient estimate of female in Panel A. Equality of sample split coefficients in Panel D are tested with a Wald test. *** p<0.005, ** p<0.01, * p<0.05.

men.

5.5. Robustness Checks

We test the robustness of the main results in Tables 3 to 7 in four different ways. We report the results from these analyses in detail in the Appendix.

First, we test whether our results are robust to the inclusion of controls. Appendix Tables A4 to A8 show that adding varying controls for the different samples—including age, educational attainment and, for the U.S. Adults sample, ethnicity and geographical fixed effects—to the regressions in Appendix Tables 3 through 7 yields no substantive changes to the results. For example, when we add controls to Table 3 (incentivized behavioral measures) we observe that the coefficients all retain the predicted signs but,

as with Table 3, none are statistically significant at $p < 0.05$ (Appendix Table A4). In Appendix Tables A5 through A8, which correspond to Tables 4 through 7, the statistical significance of some coefficients increases or decreases slightly when adding controls, but there are no substantive changes.

Second, we test for possible non-linearities in the relationships between CGI and the outcome measures reported in Tables 3 through 7. We do so by replacing the linear CGI measure in the regressions with binary indicators for each possible response on the CGI scale. We present the outcome of this analysis in Appendix Tables A9 to A13. For each outcome variable, we focus on an F-test of the joint significance of all the binary CGI indicators when controlling for binary gender in Panel C. These tests sometimes yield greater statistical significance than our primary analysis. For example, the non-linear estimation yields a relationship for the unincentivized risk for Swedish Youths and redistribution for Swiss Uni that are statistically significant at $p < 0.005$ (Appendix Tables A10 and A11). Otherwise, the relationships in which CGI has at least a marginally statistically significant relationship when accounting for binary gender are very similar across both types of analyses. Thus, the linear inclusion of CGI does not seem to mask important non-linearities.

Third, we investigate the potential role of measurement error in CGI, which might bias coefficient estimates for this variable downward. For the first dataset we collected—Swiss Uni—we use the second elicitation of CGI (obtained about two weeks after the main study) to implement the Obviously Related Instrumental Variable Approach (ORIV) by Gillen, Snowberg and Yariv (2019) to account for possible measurement error in the CGI variable.[26] As shown in Appendix Table A14, applying the ORIV method does not substantively change our results compared to the OLS regressions, though there are a few instances where the statistical significance of the estimated coefficients changes.[27]

Fourth, for the sample of U.S. Adults, we perform the analysis for all outcome measures that we elicited in both waves independently for each wave of data collection. Appendix Tables A15 and A16 reveal typically very similar results for all the outcome variables considered in both waves, but collected one year apart and with reversed order in which CGI and outcomes were measured. The signs of the coefficients in Panel C always have the same sign when estimated independently for the two waves and for six of nine cases where a coefficient is statistically significant in one wave it is also statistically significant in the other wave, though the statistical significance sometimes varies across

[26]This approach instruments the original CGI measure with its second slightly modified measurement from the follow-up survey two weeks later.

[27]In particular, the coefficient for staircase risk is no longer statistically significant when using ORIV (though it was significant at $p < 0.05$ in Table 4), while the unincentivized competitiveness measure becomes highly statistically significant ($p < 0.005$) and the unincentivized altruism measure becomes significant at $p < 0.05$ when using ORIV (though neither is statistically significant in our primary analysis).

waves.[28]

5.6. Analysis of Combined Results

Our main analysis provides 33 tests of the relationships between various economic outcomes and binary gender and CGI. The results in Tables 3 through 7 reveal that the relationships between CGI and the outcome measures, when accounting for binary gender, are sometimes highly statistically significant and other times not. In this section, we attempt to discern patterns of significance across these multiple comparisons, and to assess the overall value that a simple measure of gender identity has for understanding economic outcomes.

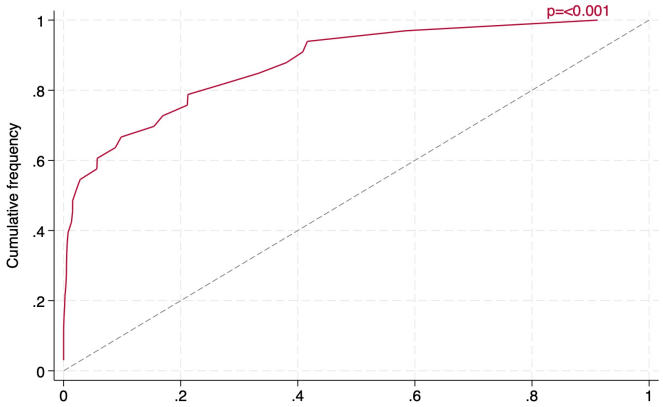
We first look at the full distribution of p-values obtained from our primary tests. Panel A of Figure 2 shows the cumulative distribution of p-values from all the primary tests reported in Subsections 5.2 through 5.4—that is, from all the coefficients for CGI in Panel C in Tables 3 through 7. Under the null hypothesis that CGI provides no additional information once controlling for binary gender, we would expect a uniform distribution of p-values. Our alternative hypothesis predicts a disproportionate prevalence of low p-values. The distribution clearly reveals stronger statistical relationships than one would expect by chance: a Kolmogorov-Smirnov test of the cumulative distribution strongly rejects the null hypothesis of uniformity ($p < 0.001$, one-tailed). Thus, viewing our analysis jointly, we reject that introducing a measure of continuous gender identity yields no significant relationships with the outcomes and behaviors we study, after controlling for binary gender.

Panel B of Figure 2 shows the cumulative distribution of p-values when splitting the sample by binary gender (the separate results from the regressions in Panel D of Tables 3 through 7). This graph clearly illustrates that most of the relationships between the outcome variables and CGI are driven by men. The distribution of p-values for men is skewed toward the left, and is highly statistically significantly different from the uniform distribution (KS test $p < 0.001$). For women, the cumulative distribution is more evenly distributed across all the possible values from 0 to 1, and we can reject the null hypothesis of uniformity only at the $p < 0.05$ level (KS test $p = 0.026$).

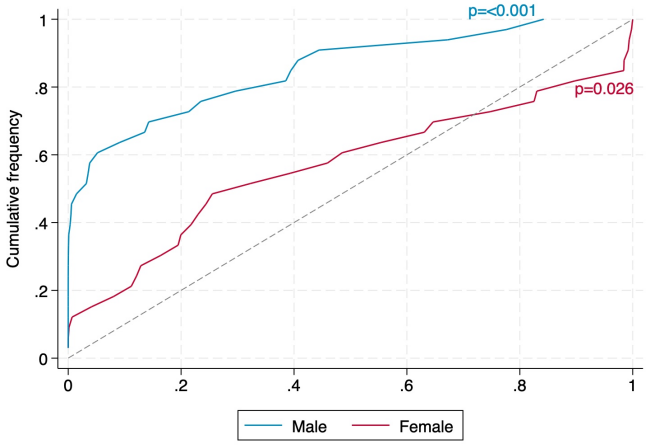
Next, we explore the change in the proportion of variance explained (R^2) in our outcome measures when we add CGI as an explanatory variable (in Panel C of the regressions in Tables 3 to 7), relative to when we only include the binary gender indicator as an explanatory variable (in Panel A of the same tables). Appendix Figure A3 plots this increase in R^2 for all 33 outcome variables. Consistent with our earlier observations,

[28]For example, Wave 2 primarily drives the statistical significance of CGI for having managerial responsibilities in Panel C of Table 7 ($p < 0.05$). Conversely, the statistical significance for the CGI coefficients on financial risk in Panel C of Table 4 ($p < 0.05$) and attitudes to redistribution in Panel C of Table 5 ($p < 0.005$) are primarily driven by Wave 1.

Figure 2: Significance of CGI Across Regressions (Tables 3 through 7)



(a) Distribution of p-values of CGI when accounting for binary gender



(b) Distribution of p-values of CGI by gender

Notes: Panel A presents the cumulative distribution of p-values from all the coefficients for CGI in Panel C in Tables 3–7. Panel B shows the cumulative distribution of p values when splitting the sample by gender (p-values from all the coefficients for CGI in Panel D in Tables 3–7.). These p-values are computed with a one-tailed t-test in the direction of the coefficient estimate for female in panel A of Tables 3–7. The p-values on the plot region are obtained using a Kolmogorov-Smirnov test with the alternative hypothesis that the empirical cumulative distribution of p-values from the regression estimates are greater (smaller p-values) than the uniform cumulative distribution.

the increase is typically small in absolute terms. For example, the increases in R^2 are never larger than one percentage point. However, it is important to note that the original R^2 values in Panel A of Tables 3 to 7 also tend to be small (only four are above 0.1), meaning that the changes in relative terms are often not trivial.

5.7. Comparing Different Measures of Gender Identity

Finally, we investigate whether the broader conclusions that we draw about the relationships between economic outcomes, binary gender and continuous gender identity change depending on the exact measure of gender identity we use. For this, we systematically compare the predictive power of CGI to other gender identity measures for 10 outcomes

studied in Wave 2 of the U.S. Adult sample, where we also collected these alternative measures of gender identity.[29]

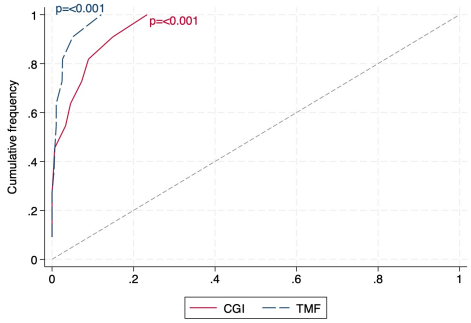
To focus on broader conclusions, rather than comparisons with respect to individual outcome measures, we consider the cumulative distributions of p-values obtained when using each measure of gender identity as an explanatory variable in regressions that also include binary gender—i.e., analogous to the regressions in Panel C of our main analysis. Figure 3 presents these cumulative distributions, while the underlying regressions are reported in greater detail in Appendix Tables A17 through A27. Each graph in Figure 3 shows the CDF of p-values for our measure of CGI (in red) together with an alternate gender identity scale (in blue). For the four alternate measures that elicit masculinity and femininity as two separate dimensions of gender identity, the distributions of p-values of the corresponding coefficients are shown separately. The cumulative distributions of our primary CGI measure differ somewhat across panels because we restrict the analysis in each graph to those sub-samples of respondents for which we also collected the alternate measure of gender identity.

Overall, we see little evidence that these alternate scales have substantially stronger correlations with our outcome measures than our single-item CGI scale. Using two-sided Kolmogorov-Smirnov tests to compare the equality of the distributions in each graph, we reject equality at $p < 0.05$ between the distributions of CGI and the other measures only in four cases —in two of which CGI tends to have lower p-values than the alternate scale (CGI vs Magliozzi femininity, Panel B, p=0.021; CGI vs OSRI femininity, Panel D, $p = 0.004$). Conversely, two scales measuring masculinity tend to have lower p-values than CGI (CGI vs. BSRI masculinity, Panel G, $p = 0.021$; CGI vs. CMNI, Panel F, $p = 0.021$). One reason why these two masculinity scales do well appears to be due to their inclusion of self-evaluations of the respondent's risk-taking propensity. For instance, the BSRI elicits an evaluation of whether the respondent is "willing to take risks" as one of the 10 items indicating masculinity, while the CMNI includes three items measuring risk-taking ("I enjoy taking risks," "I put myself in risky situations," "I take risks"). Responses to these items correlate strongly with our incentivized and unincentivized risk measures and with stated competitiveness (all $p < 0.001$). Thus, we conclude that these alternate measures, despite their substantially lengthier nature, do not generally correlate more strongly than our CGI measure with our outcome measures, beyond a few exceptions in which the scales appear to directly measure a relevant behavioral trait.

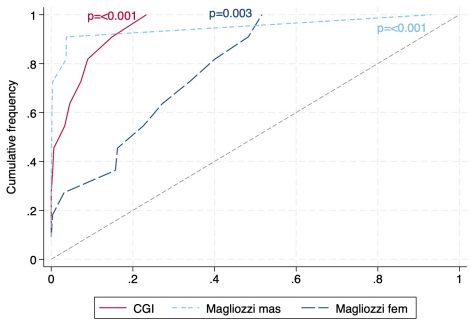
An interesting regularity in Figure 3 is that subscales measuring masculinity (rather than femininity) tend to have stronger relationships with the outcome measures, with distributions of p-values with greater mass on lower values. For the three femininity scales

[29]The 10 outcomes comprise three risk measures (one incentivized and two unincentivized), stated competitiveness, female educational track share, income, full-time homemaker status, weekly average working hours, managerial responsibilities and wage negotiation.

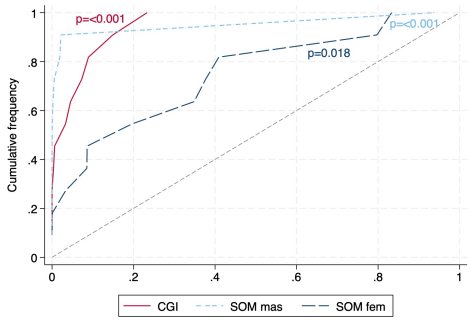
Figure 3: Significance of Different Gender Identity Measures Across Regressions



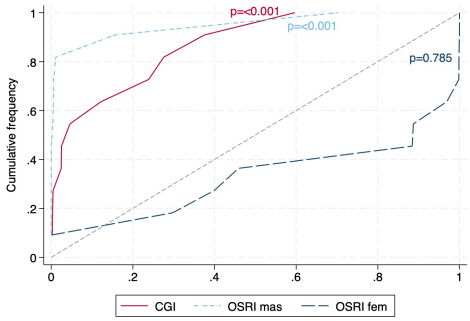
(a) TMF



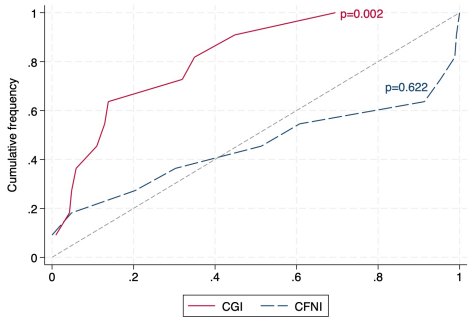
(b) Magliozzi



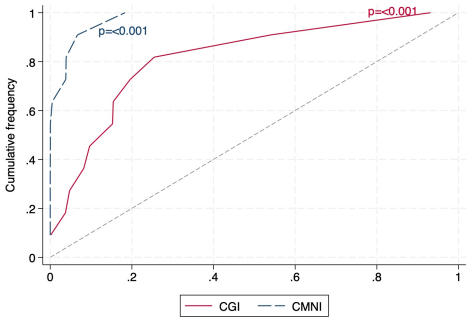
(c) SOM



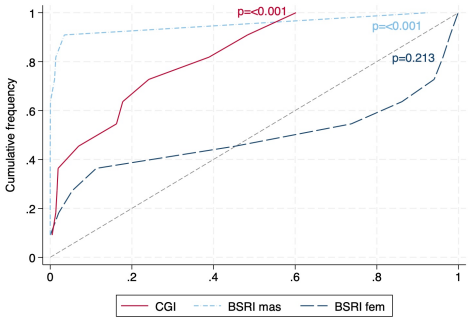
(d) OSRI



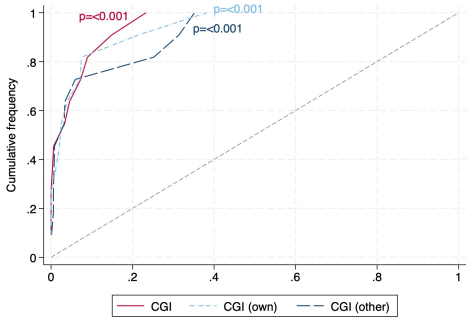
(e) CFNI



(f) CMNI



(g) BSRI



(h) Other CGI Measures

Notes: Each graph shows the cumulative distribution of p-values of gender measures (Panel C of Appendix Tables A17-A27). The p-values on the plot region are obtained using a Kolmogorov-Smirnov test with the alternative hypothesis that the empirical cumulative distribution of p-values from regression estimates are greater (smaller p-values) than the uniform cumulative distribution. The distribution of p-values for CGI varies slightly across panels since each panel uses the same sub-sample of respondents who also reported the alternate gender identity measure.

constructed from self-reported agreement with statements about female gender norms or gender-typical behavior—OSRI (Panel D), CFNI (Panel E), BSRI (Panel G)—we fail to reject that the CDF of p-values is uniform against the alternative hypothesis that it is less than uniform (KS tests $p > 0.213$). Moreover, when a scale contains both femininity and masculinity subscales, the latter tends to yield lower p-values as explanatory variables, and this difference is statistically significant in every case (Magliozzi, Panel B, $p = 0.021$; SOM, Panel C, $p = 0.004$; OSRI, Panel D, $p < 0.001$; BSRI, Panel G, $p = 0.004$). This pattern is generally consistent with our earlier observations that gender identity has stronger relationships with our outcome measures for men than for women.

As a final comparison, we consider the predictive power of our standard CGI measure to two alternate versions of our scale, in which we ask individuals to indicate how others see them on our scale from very masculine to very feminine—CGI (other)—and how they see themselves compared to others with the same binary gender, e.g., compared to other men when the respondent is a man—CGI (own). The CDFs of p-values look very similar (see Panel H of Figure 3) and do not differ significantly ($p > 0.833$ for all three comparisons), indicating that our broad conclusions do not change when eliciting CGI in slightly different ways.

6. Conclusion

We investigate the relationships between gender identity and preferences, behavior and outcomes across several samples of respondents. We focus on outcome measures that are particularly interesting for economists and for which previous research documents evidence of gaps between men and women. We present results from four distinct datasets in which we investigate the association between a validated measure of continuous gender identity (CGI) and economic preferences and outcomes, beyond the explanatory power of binary gender. The datasets comprise respondents across three countries and different age groups and elicit as outcome variables both incentivized and unincentivized measures of economic preferences and several dimensions of educational, labor market and workplace choices and outcomes.

For both genders and across all the samples we study, we find substantial variation in reported gender identity. In each sample, both men and women report varying degrees of masculinity and femininity, and the distributions of self-reported gender identity for men and women always overlap. This suggests that such broad and overlapping distributions of gender identity may provide an opportunity for identity to account for some of the variation in outcomes and behavioral tendencies among men and women.

Overall, we observe a pattern of moderate correlations between continuous gender identity and our outcome measures, once accounting for binary gender. Viewed jointly, the pattern of results across all our tests strongly rejects the null hypothesis of no cor-

relation between gender identity and economic outcome measures beyond binary gender. We observe the strongest statistical relationships for unincentivized measures of risk-seeking and competitiveness, the choice of gender-stereotypical educational tracks, and labor market outcomes like income, seeking work outside of the home, and number of work hours.

Interestingly, we find systematic differences when we consider men and women separately, thereby documenting patterns that have not previously been documented in gender economics. We find that our measure of gender identity is more strongly correlated with economic behaviors and outcomes for men than for women. One possible reason for this pattern could be that norms of appropriate economic behavior for women have relaxed over the last decades as women's possibilities to pursue a career outside of the home, and thereby their economic agency, have increased. Norms about appropriate economic behavior for men may still rely primarily on their role in the labor market. We believe that these findings raise several interesting questions for future research. During the last decades, for instance, women have increasingly been entering STEM fields, while the entrance of men in traditionally female-dominated fields has remained low (Delfino, 2024). To what extent do societal constraints on men's compared to women's gender identity explain such patterns?

Broadly, we interpret our evidence as mixed in terms of the value of measures of self-reported gender identity for economic research. Our results clearly indicate that such measures correlate with economic behaviors and outcomes, even when controlling for binary gender, to a substantially greater degree than what is expected by chance. However, they tend to provide limited improvements in accounting for variation in the outcome measure conditional on binary gender. Our results identify those domains of interest to economists studying gender gaps which have the strongest and the weakest relationships between measures of gender identity and economic behaviors. A true advantage of nuanced measures of self-reported gender identity is that one can study the relationships between economic behaviors and gender identity separately for women and men. Given the novel and perhaps surprising patterns we find in this first systematic investigation of these relationships, we believe that it is a promising avenue for future research. We think that these relationships can point to domains of economic behavior in which societal views on gender-typical behavior may be particularly rigid.

We also provide evidence that our single-item measure of continuous gender identity generally correlates well with other measures used in gender studies. Thus, we find little reason to justify using richer measures, including ones that measure gender identity indirectly, use multiple items, or that elicit separate dimensions of masculinity and femininity. This is particularly true when considering the additional potential costs of including more items in surveys. Of course, we admit that it is entirely possible that such measures may have substantially stronger relationships with other economic behaviors and outcomes

that we do not study, which is a question for future research. We believe that all our findings together suggest that there is some value in utilizing our CGI measure in future research on gender gaps in economics, particularly when the cost is low.

Finally, we believe that our research highlights the value of a growing body of studies that investigate notions of gender and gender identity that depart from traditional classifications as "male" and "female." The substantial variation in identity that we observe across samples suggests that such identification is a potentially important individual characteristic and that future research should further investigate domains in which such identity is both an important influence on outcomes and behavior and where these, in turn, affect individuals' identities.

References

Akerlof, George A., and Rachel E. Kranton. 2000. "Economics and Identity." *Quarterly Journal of Economics*, 115(3): 715–753.

Alexander, Amy C., Catherine Bolzendahl, and Lena Wängnerud. 2021. "Special Issue: Beyond the binary: new approaches to measuring gender in political science research." *European Journal of Politics and Gender*, 4(1).

Andreoni, J., and L. Vesterlund. 2001. "Which is the Fair Sex? Gender Differences in Altruism." *The Quarterly Journal of Economics*, 116(1): 293–312.

Banan, Abigail R, Torsten Santavirta, and Miguel Sarzosa. 2023. "Childhood Gender Nonconformity and Gender Gaps in Life Outcomes." *Unpublished Manuscript*.

Bandiera, Oriana, Nidhi Parekh, Barbara Petrongolo, and Michelle Rao. 2022. "Men are from Mars, and Women Too: A Bayesian Meta-analysis of Overconfidence Experiments." *Economica*, 89(S1).

Barber, Brad M., and Terrance Odean. 2001. "Boys will be Boys: Gender, Overconfidence, and Common Stock Investment." *The Quarterly Journal of Economics*, 116(1): 261–292.

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

Bem, Sandra Lipsitz. 1979. "Theory and measurement of androgyny: A reply to the Pedhazur-Tetenbaum and Locksley-Colten critiques." *Journal of Personality and Social Psychology*, 37(6): 1047–1054.

Benjamin, Daniel J, James J Choi, and A. Joshua Strickland. 2010. "Social Identity and Preferences." *American Economic Review*, 100(4): 1913–1928.

Benjamin, Daniel J., James O. Berger, Magnus Johannesson, Brian A. Nosek, E.-J. Wagenmakers, Richard Berk, ..., and Valen E. Johnson. 2018. "Redefine statistical significance." *Nature Human Behaviour*, 2(1): 6–10. Publisher: Nature Publishing Group.

Bernard, J. L., S. L. Bernard, and M. L. Bernard. 1985. "Courtship Violence and Sex-Typing." *Family Relations*, 34(4): 573.

Bertrand, Marianne. 2018. "Coase Lecture – The Glass Ceiling." *Economica*, 85(338): 205–231.

Bertrand, Marianne, and Kevin F. Hallock. 2001. "The Gender Gap in Top Corporate Jobs." *ILR Review*, 55(1): 3–21.

Bertrand, Marianne, Emir Kamenica, and Jessica Pan. 2015. "Gender Identity and Relative Income within Households." *The Quarterly Journal of Economics*, 130(2): 571–614.

Blau, Francine D., and Lawrence M. Kahn. 2017. "The Gender Wage Gap: Extent, Trends, and Explanations." *Journal of Economic Literature*, 55(3): 789–865.

Bock, Olaf, Ingmar Baetge, and Andreas Nicklisch. 2014. "hroot: Hamburg Registration and Organization Online Tool." *European Economic Review*, 71: 117–120.

Boschini, Anne, Anna Dreber, Emma Von Essen, Astri Muren, and Eva Ranehill. 2018. "Gender and altruism in a random sample." *Journal of Behavioral and Experimental Economics*, 77: 72–77.

Boschini, Anne, Astri Muren, and Mats Persson. 2012. "Constructing gender differences in the economics lab." *Journal of Economic Behavior & Organization*, 84(3): 741–752.

Brems, Christiane, and Mark E. Johnson. 1989. "Problem-Solving Appraisal and Coping Style: The Influence of Sex-Role Orientation and Gender." *The Journal of Psychology*, 123(2): 187–194.

Brenøe, Anne Ardila. 2022. "Brothers increase women's gender conformity." *Journal of Population Economics*, 35(4): 1859–1896.

Brenøe, Anne Ardila, Lea Heursen, Eva Ranehill, and Roberto A. Weber. 2022. "Continuous Gender Identity and Economics." *AEA Papers and Proceedings*, 112: 573–577.

Brown, Anna. 2022. "About 5% of young adults in the U.S. say their gender is different from their sex assigned at birth." https://pewrsr.ch/3Qi2Ejd (Accessed on 2024-05-10).

Burn, Ian, and Michael E. Martell. 2022. "Gender typicality and sexual minority labour market differentials." *British Journal of Industrial Relations*, 60(4): 784–814.

Bursztyn, Leonardo, Alessandra L. González, and David Yanagizawa-Drott. 2020. "Misperceived Social Norms: Women Working Outside the Home in Saudi Arabia." *American Economic Review*, 110(10): 2997–3029.

Buser, Thomas, Muriel Niederle, and Hessel Oosterbeek. 2014. "Gender, Competitiveness, and Career Choices." *The Quarterly Journal of Economics*, 129(3): 1409–1447.

Buser, Thomas, Muriel Niederle, and Hessel Oosterbeek. 2021. "Can Competitiveness predict Education and Labor Market Outcomes? Evidence from Incentivized Choice and Survey Measures." Working Paper. Working Paper Series. National Bureau of Economic Research. https://doi.org/10.3386/w28916.

Chapman, Jonathan, Mark Dean, Pietro Ortoleva, Erik Snowberg, and Colin Camerer. 2023. "Econographics." *Journal of Political Economy Microeconomics*, 1(1): 115–161. Publisher: The University of Chicago Press.

Chen, Daniel L., Martin Schonger, and Chris Wickens. 2016. "oTree—An open-source platform for laboratory, online, and field experiments." *Journal of Behavioral and Ex-*

perimental Finance, 9: 88–97.

Coffman, Katherine B., Lucas C. Coffman, and Keith Marzilli Ericson. 2024. "Non-Binary Gender Economics." Working Paper. Working Paper Series. National Bureau of Economic Research. https://doi.org/10.3386/w32222.

Condon, David M., and William Revelle. 2014. "The international cognitive ability resource: Development and initial validation of a public-domain measure." *Intelligence*, 43: 52–64.

Delfino, Alexia. 2024. "Breaking Gender Barriers: Experimental Evidence on Men in Pink-Collar Jobs." *American Economic Review*, 114(6): 1816–1853.

Dohmen, Thomas, and Armin Falk. 2011. "Performance Pay and Multidimensional Sorting: Productivity, Preferences, and Gender." *American Economic Review*, 101(2): 556–590.

Dohmen, Thomas, Armin Falk, David Huffman, Uwe Sunde, Jürgen Schupp, and Gert G. Wagner. 2011. "Individual risk attitudes: Measurement, determinants, and behavioral consequences." *Journal of the European Economic Association*, 9(3): 522–550.

Eagly, Alice H. 1978. "Sex differences in influenceability." *Psychological Bulletin*, 85(1): 86–116.

Eckel, Catherine C., and Philip J. Grossman. 2008. "Forecasting risk attitudes: An experimental study using actual and forecast gamble choices." *Journal of Economic Behavior & Organization*, 68(1): 1–17.

Exley, Christine L, and Judd B Kessler. 2022. "The Gender Gap in Self-Promotion*." *The Quarterly Journal of Economics*, 137(3): 1345–1381.

Falk, Armin, Anke Becker, Thomas Dohmen, David Huffman, and Uwe Sunde. 2023. "The Preference Survey Module: A Validated Instrument for Measuring Risk, Time, and Social Preferences." *Management Science*, 69(4): 1935–1950.

Feather, N. T. 1985. "Masculinity, femininity, self-esteem, and subclinical depression." *Sex Roles*, 12(5-6): 491–500.

Filippin, Antonio, and Paolo Crosetto. 2016. "A Reconsideration of Gender Differences in Risk Attitudes." *Management Science*, 62(11): 3138–3160.

Fisman, Raymond, Pamela Jakiela, and Shachar Kariv. 2017. "Distributional preferences and political behavior." *Journal of Public Economics*, 155: 1–10.

Fisman, Raymond, Shachar Kariv, and Daniel Markovits. 2007. "Individual Preferences for Giving." *American Economic Review*, 97(5): 1858–1876.

Fleming, Paul J., Kathleen Mullan Harris, and Carolyn Tucker Halpern. 2017. "De-

scription and Evaluation of a Measurement Technique for Assessment of Performing Gender." *Sex Roles*, 76(11-12): 731–746.

Flores, Andrew R, Jody L Herman, Gary J Gates, and Taylor N T Brown. 2016. "How many adults identify as transgender in the United States?"

Fornwagner, Helena, Brit Grosskopf, Alexander Lauf, Vanessa Schöller, and Silvio Städter. 2022. "On the robustness of gender differences in economic behavior." *Scientific Reports*, 12(1): 21549.

Gillen, Ben, Erik Snowberg, and Leeat Yariv. 2019. "Experimenting with Measurement Error: Techniques with Applications to the Caltech Cohort Study." *Journal of Political Economy*, 127(4): 1826–1863.

Gilligan, Carol. 1977. "In a Different Voice: Women's Conceptions of Self and of Morality." *Harvard Educational Review*, 47(4): 481–517.

Gneezy, U., and J. Potters. 1997. "An Experiment on Risk Taking and Evaluation Periods." *The Quarterly Journal of Economics*, 112(2): 631–645.

Gneezy, Uri, Kenneth L. Leonard, and John A. List. 2009. "Gender Differences in Competition: Evidence From a Matrilineal and a Patriarchal Society." *Econometrica*, 77(5): 1637–1664.

Griffiths, Scott, Stuart B. Murray, and Stephen Touyz. 2015. "Extending the masculinity hypothesis: An investigation of gender role conformity, body dissatisfaction, and disordered eating in young heterosexual men." *Psychology of Men & Masculinity*, 16(1): 108–114.

Gärtner, Manja, Johanna Mollerstrom, and David Seim. 2017. "Individual risk preferences and the demand for redistribution." *Journal of Public Economics*, 153: 49–55.

Hawkesworth, Mary. 1997. "Confounding Gender." *Signs: Journal of Women in Culture and Society*, 22(3): 649–685.

Kachel, Sven, Melanie C. Steffens, and Claudia Niedlich. 2016. "Traditional Masculinity and Femininity: Validation of a New Scale Assessing Gender Roles." *Frontiers in Psychology*, 7.

Kohlberg, L., and R. Kramer. 1969. "Continuities and Discontinuities in Childhood and Adult Moral Development." *Human Development*, 12(2): 93–120.

Magliozzi, Devon, Aliya Saperstein, and Laurel Westbrook. 2016. "Scaling Up: Representing Gender Diversity in Survey Research." *Socius: Sociological Research for a Dynamic World*, 2: 237802311666435.

Mahalik, James R., Benjamin D. Locke, Larry H. Ludlow, Matthew A. Diemer, Ryan P. J.

Scott, Michael Gottfried, and Gary Freitas. 2003. "Development of the Conformity to Masculine Norms Inventory." *Psychology of Men & Masculinity*, 4(1): 3–25.

Mahalik, James R., Elisabeth B. Morray, Aimée Coonerty-Femiano, Larry H. Ludlow, Suzanne M. Slattery, and Andrew Smiler. 2005. "Development of the Conformity to Feminine Norms Inventory." *Sex Roles*, 52(7-8): 417–435.

Meyer, Caroline, Jackie Blissett, and Claire Oldfield. 2001. "Sexual orientation and eating psychopathology: The role of masculinity and femininity." *International Journal of Eating Disorders*, 29(3): 314–318.

Moore, Don A., and Paul J. Healy. 2008. "The trouble with overconfidence." *Psychological Review*, 115(2): 502–517.

Nelson, Julie A. 2015. "Are women really more risk-averse than men? A re-analysis of the literature using expanded methods." *Journal of Economic Surveys*, 29(3): 566–585.

Nezu, Arthur M., and Christine M. Nezu. 1987. "Psychological distress, problem solving, and coping reactions: Sex role differences." *Sex Roles*, 16(3-4): 205–214.

Niederle, M., and L. Vesterlund. 2007. "Do Women Shy Away From Competition? Do Men Compete Too Much?" *The Quarterly Journal of Economics*, 122(3): 1067–1101.

Niederle, Muriel, and Lise Vesterlund. 2011. "Gender and Competition." *Annual Review of Economics*, 3(1): 601–630.

Olivetti, Claudia, and Barbara Petrongolo. 2008. "Unequal Pay or Unequal Employment? A Cross-Country Analysis of Gender Gaps." *Journal of Labor Economics*, 26(4): 621–654.

Patnaik, Arpita, Matthew Wiswall, and Basit Zafar. 2021. "College Majors 1." In *The Routledge Handbook of the Economics of Education*. Routledge. Num Pages: 43.

Pryzgoda, Jayde, and Joan C. Chrisler. 2000. "Definitions of Gender and Sex: The Subtleties of Meaning." *Sex Roles*, 43(7): 553–569.

Romano, Joseph P., and Michael Wolf. 2005. "Stepwise Multiple Testing as Formalized Data Snooping." *Econometrica*, 73(4): 1237–1282.

Solevid, Maria, Lena Wängnerud, Monika Djerf-Pierre, and Elias Markstedt. 2021. "Gender gaps in political attitudes revisited: the conditional influence of non-binary gender on left–right ideology and GAL-TAN opinions." *European Journal of Politics and Gender*, 4(1): 93–112.

Stango, Victor, and Jonathan Zinman. 2023. "We Are All Behavioural, More, or Less: A Taxonomy of Consumer Decision-Making." *The Review of Economic Studies*, 90(3): 1470–1498.

Taylor, Marylee C., and Judith A. Hall. 1982. "Psychological androgyny: Theories, methods, and conclusions." *Psychological Bulletin*, 92(2): 347–366.

Weisbuch, Max, Daniel Beal, and Edgar C. O'Neal. 1999. "How Masculine Ought I Be? Men's Masculinity and Aggression." *Sex Roles*, 40(7): 583–592.

West, Candace, and Don H. Zimmerman. 1987. "Doing Gender." *Gender & Society*, 1(2): 125–151.

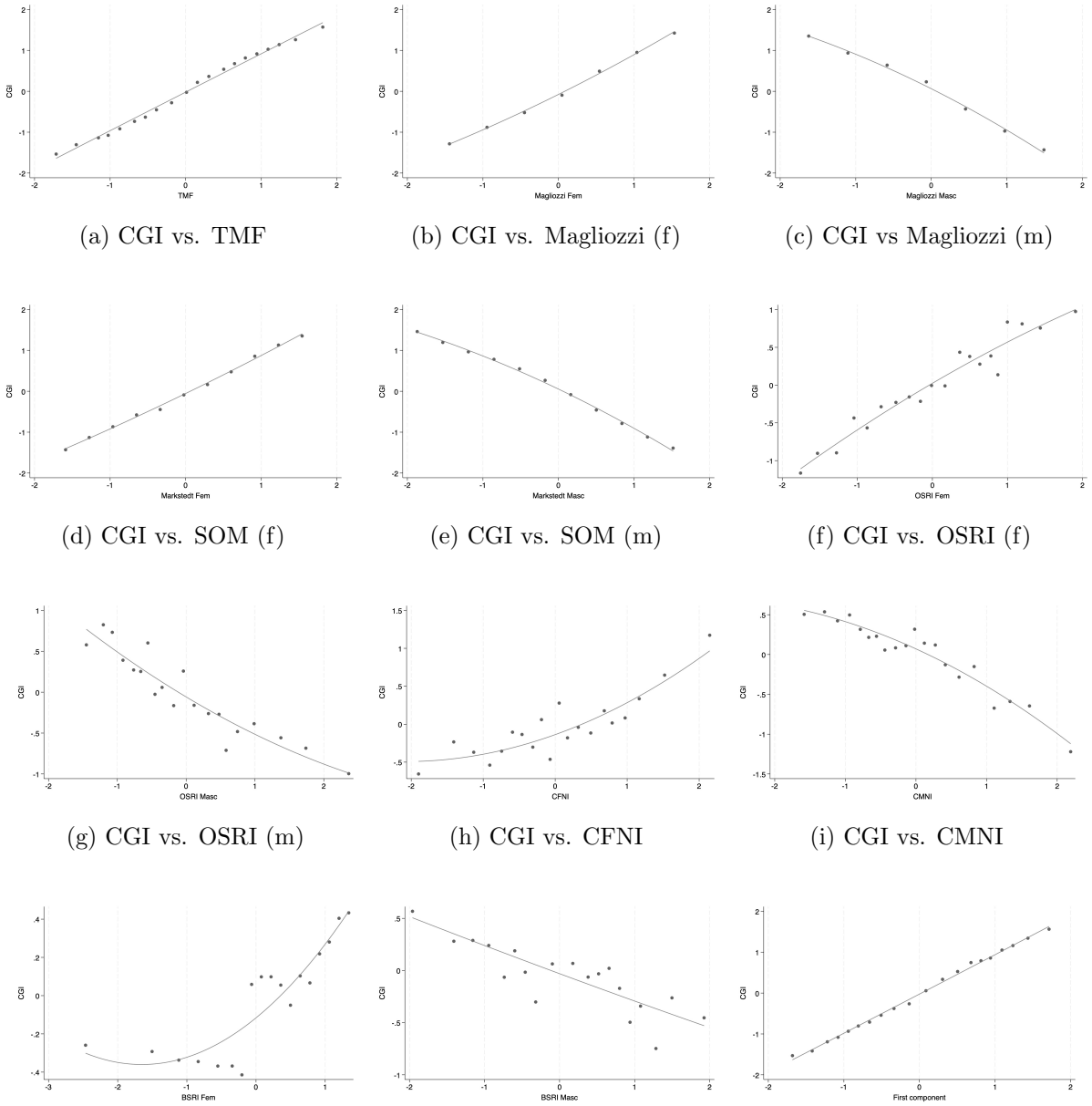
Westbrook, Laurel, and Aliya Saperstein. 2015. "New Categories Are Not Enough: Rethinking the Measurement of Sex and Gender in Social Surveys." *Gender & Society*, 29(4): 534–560.

Whitley, Bernard E. 1985. "Sex-role orientation and psychological well-being: Two meta-analyses." *Sex Roles*, 12(1-2): 207–225.

Zhang, Y. Jane. 2013. "Can Experimental Economics Explain Competitive Behavior Outside the Lab?" *SSRN Electronic Journal*.

A. Appendix A: Tables and Figures

Figure A1: Binned Scatterplots of CGI vs. Alternative Gender Identity Measures



(a) CGI vs. TMF (b) CGI vs. Magliozzi (f) (c) CGI vs Magliozzi (m)

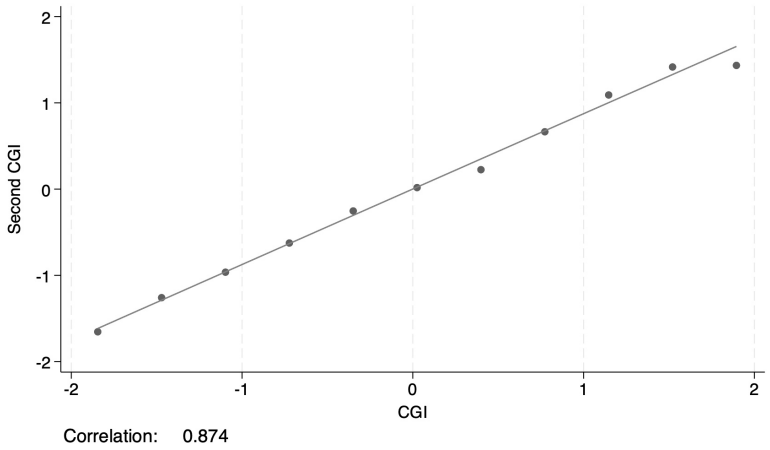
(d) CGI vs. SOM (f) (e) CGI vs. SOM (m) (f) CGI vs. OSRI (f)

(g) CGI vs. OSRI (m) (h) CGI vs. CFNI (i) CGI vs. CMNI

(j) CGI vs. BSRI (f) (k) CGI vs. BSRI (m) (l) CGI vs. First component

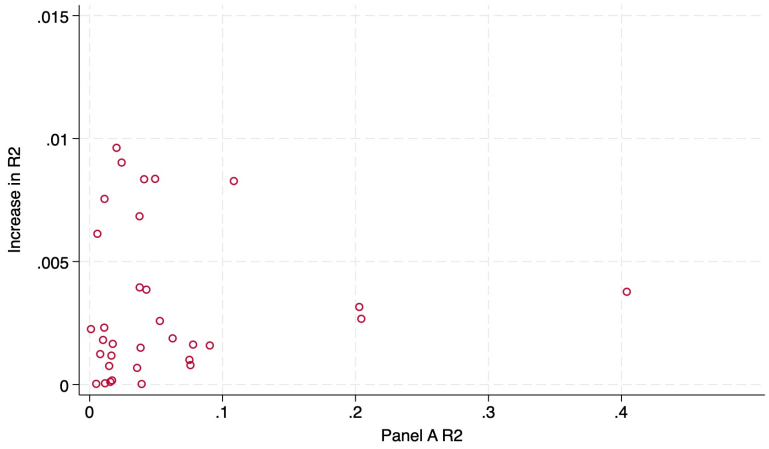
Notes: The figure presents the binscatter plots of CGI against alternative gender identity measures. U.S. Adults wave 2 (age 20–60) sample is used. CGI and TMF are unidimensional scales, with higher values indicating higher femininity. "f" refers to the femininity and "m" refers to the masculinity score of the two-dimensional scales BSRI, OSRI, Magliozzi and SOM. CFNI and CMNI are the conformity to feminine and masculine norm inventories. *First comp* combines the first factor from a principal component analysis of all gender identity scales (excluding CGI). N=2,659 for CGI, TMF, Magliozzi and SOM measures; N=662 for OSRI measures; N=659 for CFNI; N=662 for CMNI; N= 676 for BSRI measures. Quadratic line fit is added to the binscatter plot, where each scatter point represents the average of both gender identity measures in each bin.

Figure A2: Scatterplot of Separate Measurements of CGI



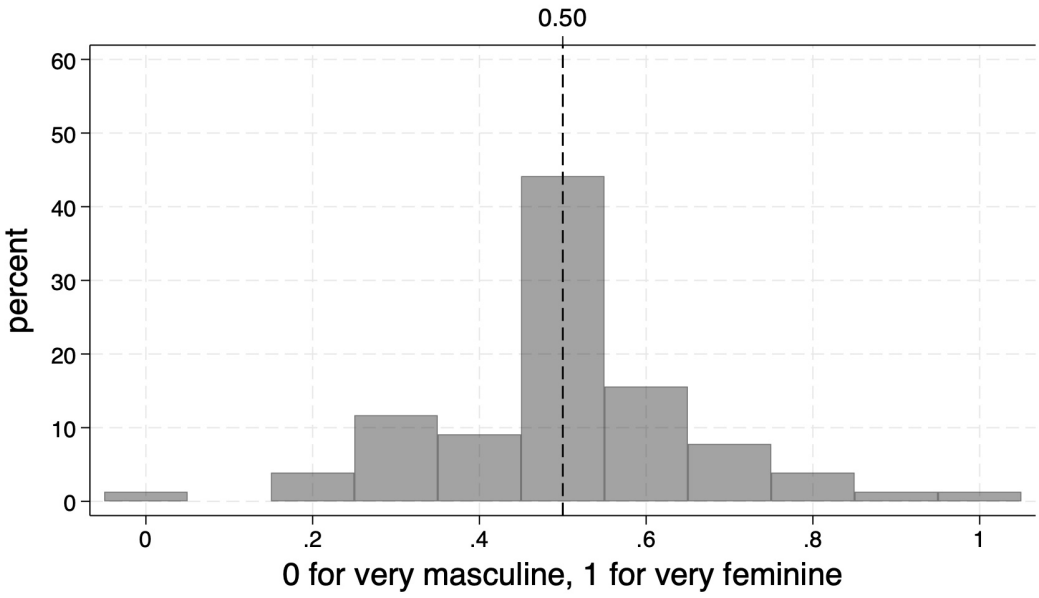
Notes: Swiss Uni sample (N=584) is used. CGI represents our first measurement, ranging from "very masculine" to "very feminine". Second CGI represents the same question asked two weeks later, with a slightly modified scale and direction ("very feminine" to "very masculine"). Responses for second CGI are inverted and both measures are standardized (Mean=0, SD=1). Quadratic line fit is added to the binscatter plot, where each scatter point represents the average of both CGI measures in each bin.

Figure A3: Improvement in R^2



Notes: The figure plots R^2 obtained in a regression of an outcome variable on a binary gender indicator (Panel A) against the increase in R^2 when CGI is added as an explanatory variable (Panel C). Each dot corresponds to one of our 33 outcome measures presented in Tables 3 through 7.

Figure A4: CGI of Non-Binary Participants



Notes: Histogram of CGI of participants identifying as non-binary. Swiss Uni (N=8), US Adults wave 2 (N=58) and Swedish Youths (N=12) samples are used. CGI represents our single question, ranging from "very masculine" to "very feminine". CGI is rescaled from 0 to 1 where 1 is very feminine. The dashed line marks the sample mean.

Table A1: Detailed Overview of Main Outcome Variables

Variable	**Description**
Independent variables	
Sex	Binary classification of the sexes taking the value 1 for women.
Continuous gender identity (CGI)	The participant's self-reported gender identity on a scale from 0 (very masculine) to 10 (very feminine).
Incentivized preferences	
Risk	Measured through the Gneezy and Potters (1997) investment task. Our measure of risk taking is the amount a participant chooses to invest in the risky investment. Standardized based on an investment ranging from 0 to 100.
Competitiveness	Measured using the Niederle and Vesterlund (2007) approach. Binary variable taking the value 1 if the participant chose to compete.
Efficiency	Measured using the Fisman, Kariv and Markovits (2007) approach. We use the value of ρ split by decile as our measure, with lower values indicating stronger equality versus efficiency focus.
Overconfidence	Measured as relative overplacement following Gillen, Snowberg and Yariv (2019). The variable we use is a standardized summary index of three overplacement measures (participants rank guess with respect to relative performance on a series of matrix reasoning items, relative performance on the competitive task (in session and in group).
Non-Incentivized preferences	
Staircase risk	Risk preferences measured as the certainty equivalence arising from the hypothetical risk module proposed by Falk et al. (2023). Standardized.
Risk	Self-reported general willingness to take risks based on the question validated in Dohmen et al. (2011). Standardized based on an answer scale from 1-10, where higher numbers indicate higher risk tolerance.
Financial risk	Binary variable taking the value of 1 if the respondent actively trades in securities.
Competitiveness	Self-reported general willingness to compete based on the question validated in Buser, Niederle and Oosterbeek (2021). Standardized based on an answer scale from 1-10, where higher numbers indicate higher competitiveness.
Redistribution	Self-reported preference for redistribution in society. Standardized based on an answer scale from 1-10, where higher numbers indicate preference for more redistribution.
Altruism	2 measures depending on dataset. 1. Self-reported amount donated from windfall gain. Standardized measure based on a donation ranging from 0-2400 (US Adults, Swiss Uni, Swiss Teens). 2. Self-reported willingness to donate to charity on a scale from 0-10. Standardized (Swedish Youth).
Educational/Ocupational outcomes	
Female educational track share	For U.S. Adults: share of women graduating with a bachelor's degree in a chosen college major field of study in 2020. For Swiss Teens: share of female apprentice graduates within the specific profession from 2019–2021. For Swedish Youths: share of women admitted in previous year (2022) to chosen educational field.
Math/language skill requirement	Math and language skills requirements with respect to occupations corresponding to the apprenticeships chosen by Swiss Teens, based on expert evaluation of job content. Standardized.
Income	Categorical variable based on self-reported income during the last 12 months (Categories 0-5000, 5001-10.000, 10.001-25.000, 25.001-50.000, 50.001-100.000, 100.001-250.000, >250.001)
Full-time homemaker	Binary variable taking the value of 1 if the participant reports being a full-time homemaker.
Weekly average work hours	Proxy measure constructed from self-reported number of weeks worked and average weekly hours.
Female industry share	Share of women working in the respondent's reported industry.
Managerial responsibilities	Binary variable taking the value of 1 if the survey respondent has managerial responsibilities at current/most recent employment.
Performance pay	Binary variable taking the value of 1 if the survey respondent has performance-related pay at current/most recent employment.
Wage negotiation	Binary variable taking the value of 1 if the respondent ever negotiated the wage.
Work flexibility	Categorical measure of flexibility in working hours where 0=No flexibility, 0.5=Can adapt work hours or Choose fix work hours, 1=Full flexibility. Standardized.

Notes: This table gives a brief description of each variable used for the analysis presented in tables 3 through 7. Please see Appendix B for more details. All standardized variables have a mean=0 and SD=1.

Table A2: Sample Selection

	Swiss Uni	U.S wave 1 (30-60)	U.S wave 2 (20-29)	U.S wave 2 (30-60)	Swiss Teens	Swedish Youth
Initial sample	597	2,002	800	2000	1,755	1,053
Non-binary gender	8	-	30	28	-	12
Inconsistent binary gender	5	8	12	25	15	-
Attention check fail	-	0	7	39	-	-
Final sample	584	1,994	751	1,908	1,740	1,041

Notes: This table provides an overview of how many observations were not included in the final sample for which reasons. *Inconsistent binary gender* means that there was any inconsistency across several measures of binary gender or binary sex we had from a respondent, for example when we compared self-reports in our survey to records obtained from Prolific Academic. *Non-binary gender* means that a person reported a non-binary gender category, for example "other". *Attention check fail* means that a respondent failed at least one of two attention checks.

Table A3: Descriptive Statistics: Education and Labor Market

	Mean		Difference
	Male	Female	
	(1)	(2)	(3)
Education			
No Schooling Completed	0.00	0.01	0.01
Regular High School Diploma	0.09	0.08	-0.01
GED or Alternative Credential	0.02	0.02	0.00
Some College Credit (less than 1 year)	0.06	0.07	0.01
1 or more Years of College Credit, No Degree	0.11	0.10	-0.00
Associate's Degree	0.08	0.11	0.03
Bachelor's Degree	0.42	0.38	-0.04
Master's Degree	0.16	0.18	0.02
Doctorate Degree	0.03	0.02	-0.01
Professional Degree Beyond a Bachelor's Degree	0.02	0.02	-0.00
Female Educational Track Share	0.49	0.61	0.11***
Labor Market			
Employment Status (respondents selected one response category)			
Employed	0.73	0.59	-0.14**
Self-employed	0.14	0.15	0.01
Relative Assisting on a Farm or Business	0.00	0.00	-0.00
In Full Time Education (at school, university, etc.)	0.01	0.01	-0.00
Full Time Homemaker	0.01	0.11	0.10**
Currently on Child-care Leave or other Leave	0.00	0.00	0.00
Unable to Work due to Long Term Illness or Disability	0.03	0.04	0.01
Retired	0.01	0.02	0.01
Unemployed	0.06	0.07	0.01
Other	0.00	0.01	0.00
Sector and Industry			
Private Sector Employee	0.83	0.82	-0.01
Female Industry	0.47	0.53	0.06**
Work Conditions			
Income (in thousand US dollars)	73.83	52.25	-21.58**
Weekly Ave. Work Hours	29.16	21.55	-7.61**
Work flexibility (0 No flexibility, 1 Full flexibility)	0.45	0.50	0.05*
Changes to Working Arrangements (0 No, 1 Yes)	0.32	0.29	-0.03
Take off an Hour for Personal Matters (0 Very difficult, 4 Very easy)	3.14	3.07	-0.07
Managerial Responsibilities (0 No, 1 Yes)	0.47	0.34	-0.14**
Perfomance Pay (0 No, 1 Yes)	0.34	0.26	-0.08**
Wage Negotiation (0 No, 1 Yes)	0.48	0.36	-0.11**
Observations	999	995	

Notes: Multiple hypothesis test correction was done separately for the family of variables Labor Market and Education. US Sample - Wave 1 is used. Significance levels: *** $p < 0.005$, ** $p < 0.01$, * $p < 0.05$

Table A4: CGI and Incentivized Behavioral Measures (with controls)

Measure	Risk		Competitiveness	Efficiency		Overconfidence
Sample	Swiss Uni	US Adults	Swiss Uni	Swiss Uni	Swedish Youths	Swiss Uni
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. Binary gender						
Female	-0.383***	-0.211***	-0.135***	-0.147***	-0.149***	0.006
	(0.082)	(0.032)	(0.041)	(0.023)	(0.017)	(0.083)
R^2	0.038	0.022	0.028	0.066	0.159	0.010
Panel B. Gender identity						
CGI (feminine)	-0.169***	-0.090***	-0.057***	-0.067***	-0.058***	
	(0.040)	(0.016)	(0.021)	(0.012)	(0.009)	
R^2	0.030	0.019	0.023	0.055	0.136	
Panel C. Binary gender and gender identity						
Female	-0.321*	-0.197***	-0.127*	-0.113***	-0.136***	
	(0.153)	(0.057)	(0.071)	(0.043)	(0.026)	
CGI (feminine)	-0.038	-0.009	-0.005	-0.021	-0.008	
	(0.075)	(0.029)	(0.036)	(0.021)	(0.013)	
R^2	0.039	0.022	0.028	0.068	0.159	
Panel D. Sample split by binary gender						
Female: CGI (feminine)	-0.020	0.000	-0.022	-0.042**	-0.006	
	(0.060)	(0.024)	(0.029)	(0.017)	(0.012)	
R^2	0.022	0.013	0.043	0.033	0.133	
Male: CGI (feminine)	-0.016	-0.020	0.015	0.013	0.004	
	(0.063)	(0.024)	(0.030)	(0.017)	(0.000)	
R^2	0.009	0.014	0.005	0.004	0.145	
Test: male sample split CGI = female sample split CGI						
p-value	0.964	0.518	0.358	0.014	0.551	
Observations	584	3,902	584	584	1,041	584
Mean of Dependent Variable	0	0	0.413	0.538	0.550	0

Notes: The table presents results from regressing the incentivized measures of risk, competitiveness, equality versus efficiency and overconfidence on our standardized (mean=0, SD=1) measure of CGI. CGI is standardized (mean=0, SD=1) separately for the male and female samples in Panel D. Swiss Uni, U.S. Adults (age 30-60), and Swedish Youths samples are used. The controls are as follows: quadratic age and education level for Swiss Uni; quadratic age, race, census divisions and wave for U.S Adults; school for Swedish Youths. The estimates in each column and panel come from a separate regression. Risk preference is a standardized measure of participants' investment decisions (mean=0, SD=1). Competitiveness is a dummy that takes the value 1 for those who chose to compete in the competitive task. Our measure of preferences for equality versus efficiency is measured in deciles, with increasing numbers indicating higher priority for efficiency. Overconfidence is measured as relative overplacement and standardized (mean=0, SD=1). Robust standard errors are reported in parentheses. Significance levels in Panel A are determined by a two-tailed test. Significance levels in Panels B, C and D are determined by a one-tailed test in the direction of the coefficient estimate of female in Panel A. *** p<0.005, ** p<0.01, * p<0.05.

Table A5: CGI and Unincentivized Risk and Competitiveness Measures (with controls)

Measure	Staircase Risk	Risk				Financial Risk	Competitiveness			
Sample	Swiss Uni	Swiss Uni	U.S Adults	Swiss Teens	Swedish Youths	US Adults	Swiss Uni	U.S Adults	Swiss Teens	Swedish Youths
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A. Binary gender										
Female	-0.206*	-0.436***	-0.389***	-0.287***	-0.236***	-0.267***	-0.396***	-0.407***	-0.311***	-0.143*
	(0.083)	(0.081)	(0.031)	(0.071)	(0.065)	(0.015)	(0.082)	(0.031)	(0.071)	(0.063)
R^2	0.015	0.053	0.053	0.026	0.062	0.087	0.044	0.053	0.024	0.072
Panel B. Gender identity										
CGI (feminine)	-0.137***	-0.232***	-0.203***	-0.168***	-0.104***	-0.118***	-0.197***	-0.216***	-0.181***	-0.096***
	(0.044)	(0.040)	(0.016)	(0.036)	(0.033)	(0.008)	(0.041)	(0.016)	(0.036)	(0.030)
R^2	0.023	0.059	0.056	0.034	0.059	0.070	0.044	0.058	0.033	0.076
Panel C. Binary gender and gender identity										
Female	0.051	-0.169	-0.171***	0.090	-0.177*	-0.224***	-0.220	-0.161***	0.094	0.000
	(0.149)	(0.148)	(0.057)	(0.168)	(0.106)	(0.027)	(0.139)	(0.057)	(0.170)	(0.106)
CGI (feminine)	-0.158*	-0.163*	-0.133***	-0.208**	-0.040	-0.026*	-0.108	-0.150***	-0.224***	-0.096*
	(0.078)	(0.073)	(0.029)	(0.086)	(0.053)	(0.013)	(0.069)	(0.029)	(0.086)	(0.050)
R^2	0.023	0.062	0.059	0.034	0.062	0.088	0.048	0.060	0.033	0.076
Panel D. Sample split by binary gender										
Female: CGI (feminine)	-0.064	-0.088	-0.026	-0.048	0.036	0.001	-0.011	0.062	0.016	0.006
	(0.065)	(0.059)	(0.025)	(0.060)	(0.051)	(0.010)	(0.055)	(0.024)	(0.056)	(0.049)
R^2	0.029	0.011	0.015	0.003	0.057	0.019	0.012	0.016	0.001	0.098
Male: CGI (feminine)	-0.125*	-0.108*	-0.130***	-0.142***	-0.083	-0.034***	-0.111*	-0.246***	-0.237***	-0.155
	(0.069)	(0.065)	(0.025)	(0.045)	(0.000)	(0.012)	(0.062)	(0.025)	(0.048)	(0.000)
R^2	0.017	0.016	0.036	0.036	0.103	0.015	0.015	0.074	0.057	0.129
Test: male sample split CGI = female sample split CGI										
p-value	0.656	0.916	0.003	0.090	0.102	0.027	0.278	0.000	0.000	0.040
Observations	584	584	3,902	786	1,041	3,902	584	3,902	792	1,041
Mean of Dependent Variable	0	0	0	0	0	0.401	0	0	0	0

Notes: The table presents results from regressing unincentivized measures of risk attitudes and competitiveness on our standardized (mean=0, SD=1) measure of CGI. CGI is standardized (mean=0, SD=1) separately for the male and female samples in Panel D. Swiss Uni, U.S. Adults (age 30-60), Swiss Teens, and Swedish Youths samples are used. The controls are as follows: quadratic age, race, census divisions and wave for U.S Adults; quadratic age and education level for Swiss Uni; grade level and survey round for Swiss Teens; school for Swedish Youth. The estimates in each column and panel come from a separate regression. Staircase risk is a categorical certainty equivalence measure of risk-taking based on a series of hypothetical allocation decisions. Risk is a self-reported measure of risk-taking. Financial risk is a dummy that takes the value 1 for those who report to actively trade in securities. Competitiveness is a self-reported measure of competitiveness. A higher value means higher risk taking or competitiveness. All measures except financial market risk are standardized (mean 0, SD 1). Robust standard errors are reported in parentheses. Significance levels in Panel A are determined by a two-tailed test. Significance levels in Panels B, C and D are determined by a one-tailed test in the direction of the coefficient estimate of female in Panel A. *** p<0.005, ** p<0.01, * p<0.05.

Table A6: CGI and Unincentivized Distributional Preferences (with controls)

Measure	Redistribution		Altruism			
Sample	Swiss Uni (1)	U.S. Adults (2)	Swiss Uni (3)	U.S. Adults (4)	Swiss Teens (5)	Swedish Youths (6)
Panel A. Binary gender						
Female	0.209*	0.067*	0.250***	-0.003	0.205***	0.540***
	(0.082)	(0.032)	(0.084)	(0.032)	(0.070)	(0.063)
R^2	0.027	0.026	0.020	0.014	0.033	0.127
Panel B. Gender identity						
CGI (feminine)	0.111***	0.055***	0.097**		0.067*	0.222***
	(0.040)	(0.017)	(0.041)		(0.035)	(0.034)
R^2	0.028	0.027	0.014		0.027	0.106
Panel C. Binary gender and gender identity						
Female	0.085	-0.072	0.274		0.459***	0.462***
	(0.156)	(0.055)	(0.172)		(0.166)	(0.102)
CGI (feminine)	0.076	0.085***	-0.015		-0.140	0.052
	(0.076)	(0.029)	(0.086)		(0.083)	(0.053)
R^2	0.029	0.028	0.020		0.036	0.129
Panel D. Sample split by binary gender						
Female: CGI (feminine)	-0.149	-0.105	-0.058		-0.073	0.056
	(0.062)	(0.023)	(0.073)		(0.054)	(0.049)
R^2	0.044	0.050	0.029		0.036	0.074
Male: CGI (feminine)	0.198***	0.190***	0.023		-0.052	0.001
	(0.062)	(0.024)	(0.070)		(0.046)	(0.000)
R^2	0.061	0.053	0.014		0.012	0.090
Test: male sample split CGI = female sample split CGI						
p-value	0.000	0.000	0.398		0.928	0.402
Observations	584	3,902	584	3,902	798	1,041
Mean of Dependent Variable	0	0	0	0	0	0

Notes: The table presents results from regressing unincentivized distributional measures on our standardized (mean=0, SD=1) measure of CGI. CGI is standardized (mean=0, SD=1) separately for the male and female samples in Panel D. Swiss Uni, U.S. Adults (age 30-60), Swiss Teens and Swedish Youth samples are used. The controls are as follows: quadratic age, race, census divisions and wave for U.S adults; quadratic age and education level for Swiss Uni; grade level and survey round for Swiss Teens; school for Swedish Youths. The estimates in each column and panel come from a separate regression. The table presents results from regressing self-reported preferences on redistribution and altruism on our standardized measure of CGI. Redistribution is a measure of how much economic redistribution one wants in society. Altruism is a measure of how much one would donate out of a windfall gain. A higher value means the greater willingness to redistribute or donate. All outcome measures are standardized (mean=0, SD=1). Robust standard errors are reported in parentheses. Significance levels in Panel A are determined by a two-tailed test. Significance levels in Panels B, C and D are determined by a one-tailed test in the direction of the coefficient estimate of female in Panel A. *** p<0.005, ** p<0.01, * p<0.05.

Table A7: CGI and Educational Choice (with controls)

Measure	Female Educational Track Share			Skill Requirements	
				Math	Language
Sample	U.S Adults	Swiss Teens	Swedish Youths	Swiss Teens	Swiss Teens
	(1)	(2)	(3)	(4)	(5)
Panel A. Binary gender					
Female	0.126***	0.384***	0.090***	-0.917***	0.915***
	(0.007)	(0.012)	(0.010)	(0.047)	(0.046)
R^2	0.120	0.411	0.154	0.220	0.212
Panel B. Gender identity					
CGI (feminine)	0.061***	0.168***	0.037***	-0.408***	0.403***
	(0.004)	(0.006)	(0.005)	(0.023)	(0.023)
R^2	0.114	0.323	0.129	0.182	0.171
Panel C. Binary gender and gender identity					
Female	0.078***	0.328***	0.077***	-0.737***	0.760***
	(0.014)	(0.023)	(0.014)	(0.083)	(0.083)
CGI (feminine)	0.029***	0.034***	0.009	-0.107***	0.092*
	(0.007)	(0.011)	(0.007)	(0.041)	(0.040)
R^2	0.127	0.415	0.155	0.223	0.215
Panel D. Sample split by binary gender					
Female: CGI (feminine)	0.010*	0.026***	0.008	-0.045	0.115***
	(0.005)	(0.010)	(0.007)	(0.041)	(0.041)
R^2	0.016	0.017	0.076	0.009	0.023
Male: CGI (feminine)	0.022***	0.014	0.007	-0.098***	0.019
	(0.006)	(0.009)	(0.000)	(0.035)	(0.036)
R^2	0.038	0.027	0.126	0.042	0.014
Test: male sample split CGI = female sample split CGI					
p-value	0.172	0.721	0.715	0.110	0.230
Observations	2,279	1,448	1,041	1,433	1,433
Mean of Dependent Variable	0.541	0.442	0.521	0	0

Notes: The table presents results from regressing the educational track measures on our standardized (mean=0, SD=1) measure of CGI. CGI is standardized (mean=0, SD=1) separately for the male and female samples in Panel D. U.S. Adults (age 30–60), Swiss Teens and Swedish Youths samples are used. The controls are as follows: quadratic age, race, census divisions and wave for U.S adults; grade level and survey round for Swiss Teens; school for Swedish Youths. The estimates in each column and panel come from a separate regression. Female educational track share is the share of women graduating with a bachelor's degree in the chosen field of study in 2020 for U.S. Adults; the share of women from past cohorts of graduates from the apprenticeship for Swiss Teens; the share of women accepted for undergraduate studies in that field the year before our sample made their educational choices for Swedish Youths.. Skills requirements are a standardized (mean=0, SD=1) measure based on expert evaluation of the job content in occupations chosen by Swiss Teens to start apprenticeships. Robust standard errors are reported in parentheses. Significance levels in Panel A are determined by a two-tailed test. Significance levels in Panels B, C and D are determined by a one-tailed test in the direction of the coefficient estimate of female in Panel A. *** p<0.005, ** p<0.01, * p<0.05.

Table A8: CGI and Labor Market Outcomes (with controls)

Measure	Income	Full-time Homemaker	Weekly Ave. Work Hours	Female Industry Share	Managerial Responsibilities	Performance Pay	Wage Negotiation	Work Flexibility
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A. Binary gender								
Female	-23.153***	0.100***	-6.607***	0.062***	-0.126***	-0.084***	-0.133***	0.142***
	(1.863)	(0.008)	(0.452)	(0.007)	(0.016)	(0.021)	(0.016)	(0.045)
R^2	0.058	0.050	0.061	0.050	0.030	0.014	0.048	0.014
Panel B. Gender identity								
CGI (feminine)	-10.754***	0.050***	-3.127***	0.026***	-0.061***	-0.043***	-0.066***	0.062***
	(0.964)	(0.004)	(0.223)	(0.004)	(0.008)	(0.010)	(0.008)	(0.023)
R^2	0.052	0.049	0.055	0.037	0.029	0.015	0.028	0.013
Panel C. Binary gender and gender identity								
Female	-16.584***	0.058***	-4.500***	0.059***	-0.081***	-0.041	-0.086***	0.119
	(2.974)	(0.012)	(0.811)	(0.012)	(0.027)	(0.035)	(0.027)	(0.077)
CGI (feminine)	-4.017***	0.026***	-1.289***	0.002	-0.028*	-0.026	-0.028*	0.014
	(1.547)	(0.006)	(0.399)	(0.006)	(0.014)	(0.018)	(0.014)	(0.039)
R^2	0.059	0.054	0.063	0.050	0.031	0.015	0.049	0.014
Panel D. Sample split by binary gender								
Female: CGI (feminine)	2.591	0.028***	-0.117	-0.004	0.023	-0.012	-0.006	-0.084
	(1.146)	(0.007)	(0.339)	(0.005)	(0.011)	(0.014)	(0.011)	(0.033)
R^2	0.023	0.022	0.011	0.013	0.015	0.016	0.029	0.025
Male: CGI (feminine)	-6.751***	0.006*	-1.410***	0.005	-0.051***	-0.015	-0.029**	0.087***
	(1.382)	(0.003)	(0.316)	(0.005)	(0.011)	(0.015)	(0.011)	(0.032)
R^2	0.037	0.010	0.027	0.020	0.034	0.016	0.043	0.021
Test: male sample split CGI = female sample split CGI								
p-value	0.000	0.004	0.006	0.209	0.000	0.942	0.157	0.000
Observations	3,902	3,902	3,902	1,994	3,902	1,994	3,902	1,994
Mean of Dependent Variable	61.421	0.062	25.281	0.497	0.400	0.303	0.491	0

Notes: The table presents results from regressing the labor market outcomes on our standardized (mean=0, SD=1) measure of CGI. CGI is standardized (mean=0, SD=1) separately for the male and female samples in Panel D. The sample of the U.S. Adults (age 30-60) is used. The controls are as follows: quadratic age, race, census divisions and wave. The estimates in each column and panel come from a separate regression. Income is a self-reported categorical measure in thousand U.S. dollars. Full-time homemaker is a dummy for working full-time at home. Weekly average work hours is a proxy measure constructed from weeks worked and hours worked, using the product of the two categorical measures and dividing by 52. Female industry share is the share of female employees in a given industry. Managerial responsibilities is a dummy for having managerial responsibilities at work. Performance pay is a dummy that is equal to 1 if the respondent's current or most recent job has performance-related pay. Wage negotiation is a dummy that is equal to 1 if the respondent ever negotiated wage. Work flexibility is a categorical measure of flexibility in working hours where 0 is no flexibility, 0.5 is can adapt work hours and 1 is full flexibility, which is standardized (mean=0, SD=1). Robust standard errors are reported in parentheses. Significance levels in Panel A are determined by a two-tailed test. Significance levels in Panels B, C and D are determined by a one-tailed test in the direction of the coefficient estimate of female in Panel A. *** $p<0.005$, ** $p<0.01$, * $p<0.05$.

Table A9: CGI and Incentivized Behavioral Measures (with CGI dummies)

Measure	Risk		Competitiveness	Efficiency		Overconfidence
Sample	Swiss Uni	US Adults	Swiss Uni	Swiss Uni	Swedish Youth	Swiss Uni
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. Binary gender						
F-stat (overall)	21.513	45.689	9.963	38.626	86.394	0.001
p-value	0.000	0.000	0.002	0.000	0.000	0.978
Panel B. Gender identity						
F-stat (overall)	2.551	5.937	2.650	4.285	8.186	
p-value	0.005	0.000	0.004	0.000	0.000	
Panel C. Binary gender and gender identity						
F-stat (overall)	2.890	5.940	2.549	4.919	9.557	
p-value	0.001	0.000	0.004	0.000	0.000	
Test: Binary sex=0						
F-test	4.226	4.122	1.553	7.910	19.186	
p-value	0.040	0.042	0.213	0.005	0.000	
Test: All CGI dummies=0						
F-test	1.113	1.716	1.620	1.198	1.428	
p-value	0.350	0.071	0.097	0.289	0.163	
Observations	584	3,902	584	584	1,041	584

Notes: The table presents results from regressing the incentivized measures of risk, competitiveness, equality versus efficiency and overconfidence on binary indicators for each possible response on our measure of CGI. Swiss Uni, U.S. Adults (age 30-60), and Swedish Youths samples are used. The estimates in each column and panel come from a separate regression. Risk preference is a standardized measure of participants' investment decisions standardized (mean=0, SD=1). Competitiveness is a dummy that takes the value 1 for those who chose to compete in the competitive task. Our measure of preferences for equality versus efficiency is measured in deciles, with increasing numbers indicating higher priority for efficiency. Overconfidence is measured as relative overplacement and standardized (mean=0, SD=1).

Table A10: CGI and Unincentivized Risk and Competitiveness Measures (with CGI dummies)

Measure	Staircase Risk	Risk				Financial Risk	Competitiveness			
Sample	Swiss Uni	Swiss Uni	US Adults	Swiss Teens	Swedish Youth	US Adults	Swiss Uni	US Adults	Swiss Teens	Swedish Youth
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A. Binary gender										
F-stat (overall)	6.549	30.068	152.059	16.247	15.415	316.622	22.681	166.561	19.221	6.109
p-value	0.011	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.014
Panel B. Gender identity										
F-stat (overall)	2.093	4.106	19.268	3.247	6.596	27.665	3.740	35.439	4.928	6.490
p-value	0.023	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Panel C. Binary gender and gender identity										
F-stat (overall)	1.909	3.827	18.843	3.267	6.312	30.430	3.544	32.958	4.933	5.907
p-value	0.036	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Test: Binary sex=0										
F-test	0.183	0.794	10.288	2.692	1.528	48.612	1.307	6.219	1.836	0.020
p-value	0.669	0.373	0.001	0.101	0.217	0.000	0.253	0.013	0.176	0.887
Test: All CGI dummies=0										
F-test	1.399	1.170	5.717	2.149	5.233	1.347	1.546	17.385	3.913	5.468
p-value	0.177	0.308	0.000	0.019	0.000	0.199	0.119	0.000	0.000	0.000
Observations	584	584	3,902	786	1,041	3,902	584	3,902	792	1,041

Notes: The table presents results from regressing unincentivized measures of risk attitudes and competitiveness on binary indicators for each possible response on our measure of CGI. Swiss Uni, U.S. Adults (age 30-60), Swiss Teens, and Swedish Youths samples are used. The estimates in each column and panel come from a separate regression. Staircase risk is a categorical certainty equivalence measure of risk-taking based on a series of hypothetical allocation decisions. Risk is a self-reported measure of risk-taking. Financial risk is a dummy that takes the value 1 for those who report to actively trade in securities. Competitiveness is a self-reported measure of competitiveness. A higher value means higher risk-taking or competitiveness. All measures except financial market risk are standardized (mean 0, SD 1).

Table A11: CGI and Unincentivized Distributional Preferences (with CGI dummies)

Measure	Redistribution		Altruism			
Sample	Swiss Uni (1)	U.S. Adults (2)	Swiss Uni (3)	U.S. Adults (4)	Swiss Teens (5)	Swedish Youth (6)
Panel A. Binary gender						
F-stat (overall)	5.910	3.888	9.189	0.002	8.786	83.852
p-value	0.015	0.049	0.003	0.966	0.003	0.000
Panel B. Gender identity						
F-stat (overall)	3.165	14.081	1.809		1.904	6.768 ’
p-value	0.001	0.000	0.056		0.036	0.000
Panel C. Binary gender and gender identity						
F-stat (overall)	2.990	12.840	2.428		1.840	8.994
p-value	0.001	0.000	0.006		0.039	0.000
Test: Binary sex=0						
F-test	0.454	0.710	2.225		1.598	18.738
p-value	0.501	0.400	0.136		0.207	0.000
Test: All CGI dummies=0						
F-test	2.416	13.809	1.374		1.316	0.706
p-value	0.008	0.000	0.189		0.217	0.720
Observations	584	3,902	584	3,902	798	1,041

Notes. The table presents results from regressing unincentivized distributional measures on binary indicators for each possible response on our measure of CGI. Swiss Uni, U.S. Adults (age 30–60), Swiss Teens and Swedish Youth samples are used. The estimates in each column and panel come from a separate regression. The table presents results from regressing self-reported preferences on redistribution and altruism on our standardized measure of CGI. Redistribution is a measure of how much economic redistribution one wants in society. Altruism is a measure of how much one would donate out of a windfall gain. A higher value means the greater willingness to redistribute or donate. All outcome measures are standardized (mean=0, SD=1).

Table A12: CGI and educational choices (with CGI dummies)

Measure	Female Educational Track Share			Skill Requirements	
				Math	Language
Sample	US Adults	Swiss Teens	Swedish Youth	Swiss Teens	Swiss Teens
	(1)	(2)	(3)	(4)	(5)
Panel A. Binary gender					
F-stat (overall)	280.099	970.728	102.972	381.072	382.547
p-value	0.000	0.000	0.000	0.000	0.000
Panel B. Gender identity					
F-stat (overall)	28.648	78.492	8.327	32.268	34.834 ’
p-value	0.000	0.000	0.000	0.000	0.000
Panel C. Binary gender and gender identity					
F-stat (overall)	29.282	90.441	10.000	34.407	37.707
p-value	0.000	0.000	0.000	0.000	0.000
Test: Binary sex=0					
F-test	21.977	134.059	24.566	46.668	63.649
p-value	0.000	0.000	0.000	0.000	0.000
Test: All CGI dummies=0					
F-test	3.064	2.806	0.780	2.208	2.744
p-value	0.001	0.002	0.648	0.015	0.002
Observations	2,279	1,448	1,041	1,433	1,433

Notes: The table presents results from regressing the educational track measures on binary indicators for each possible response on our measure of CGI. U.S. Adults (age 30–60), Swiss Teens and Swedish Youths samples are used. The estimates in each column and panel come from a separate regression. Female educational track share is the share of women graduating with a bachelor's degree in the chosen field of study in 2020 for U.S. Adults; the share of women from past cohorts of graduates from the apprenticeship for Swiss Teens; the share of women accepted for undergraduate studies in that field the year before our sample made their educational choices for Swedish Youths. Skills requirements are a standardized (mean=0, SD=1) measure based on expert evaluation of the job content in occupations chosen by Swiss Teens to start apprenticeships.

Table A13: CGI and Labor Market Outcomes (with CGI dummies)

Measure	Income	Full-time Homemaker	Weekly Ave. Work Hours	Female Industry Share	Managerial Responsibilities	Performance Pay	Wage Negotiation	Work Flexibility
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A. Binary gender								
F-stat (overall)	155.642	172.645	217.216	81.057	65.005	15.972	68.948	9.929
p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.002
Panel B. Gender identity								
F-stat (overall)	16.270	19.780	23.098	6.684	10.460	2.299	9.929	2.095
p-value	0.000	0.000	0.000	0.000	0.000	0.011	0.000	0.022
Panel C. Binary gender and gender identity								
F-stat (overall)	17.923	18.351	23.383	8.230	9.944	2.112	9.309	2.074
p-value	0.000	0.000	0.000	0.000	0.000	0.017	0.000	0.019
Test: Binary sex=0								
F-test	33.204	29.015	24.915	22.699	4.074	0.244	2.898	1.780
p-value	0.000	0.000	0.000	0.000	0.044	0.621	0.089	0.182 ’
Test: All CGI dummies=0								
F-test	4.571	2.853	3.299	1.085	4.202	0.757	3.172	1.274
p-value	0.000	0.002	0.000	0.370	0.000	0.671	0.000	0.240
Observations	3,902	3,902	3,902	1,994	3,902	1,994	3,902	1,994

Notes: The table presents results from regressing the labor market outcomes on binary indicators for each possible response on our measure of CGI. The sample of the U.S. Adults (age 30-60) is used. The estimates in each column and panel come from a separate regression. Income is a self-reported categorical measure in thousand U.S. dollars. Full-time homemaker is a dummy for working full-time at home. Weekly average work hours is a proxy measure constructed from weeks worked and hours worked, using the product of the two categorical measures and dividing by 52. Female industry share is the share of female employees in a given industry. Managerial responsibilities is a dummy for having managerial responsibilities at work. Performance pay is a dummy that is equal to 1 if the respondent's current or most recent job has performance-related pay. Wage negotiation is a dummy that is equal to 1 if the respondent ever negotiated wage. Work flexibility is a categorical measure of flexibility in working hours where 0 is no flexibility, 0.5 is can adapt work hours and 1 is full flexibility, which is standardized (mean=0, SD=1).

Table A14: ORIV Analysis - Swiss Uni Sample

	Incentivized Measures			Unincentivized Measures				
Measure	Risk	Competitiveness	Efficiency	Staircase Risk	Risk	Competitiveness	Redistribution	Alturism
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A. Binary gender								
Female	-0.378***	-0.128***	-0.145***	-0.211*	-0.444***	-0.388***	0.201*	0.249***
	(0.081)	(0.041)	(0.023)	(0.083)	(0.081)	(0.081)	(0.083)	(0.082)
R^2	0.036	0.017	0.062	0.011	0.049	0.038	0.010	0.016
Panel B. Gender identity								
ORIV CGI (feminine)	-0.139***	-0.069***	-0.070***	-0.111**	-0.254***	-0.252***	0.129***	0.168***
	(0.046)	(0.023)	(0.013)	(0.047)	(0.046)	(0.046)	(0.046)	(0.048)
Panel C. Binary gender and gender identity								
Female	-0.553**	-0.055	-0.116*	-0.111	-0.110	0.077	-0.034	-0.086
	(0.199)	(0.097)	(0.057)	(0.200)	(0.189)	(0.193)	(0.208)	(0.198)
ORIV CGI (feminine)	0.108	-0.045	-0.018	-0.062	-0.205*	-0.286***	0.144	0.206*
	(0.110)	(0.054)	(0.031)	(0.111)	(0.106)	(0.107)	(0.113)	(0.112)
Panel D. Sample split by binary gender								
Female: ORIV CGI (feminine)	0.086	-0.075	-0.035	-0.020	-0.102	-0.126	-0.121	0.108
	(0.095)	(0.046)	(0.026)	(0.095)	(0.094)	(0.091)	(0.094)	(0.096)
Male: ORIV CGI (feminine)	0.045	0.024	0.015	-0.055	-0.142*	-0.217**	0.287***	0.133
	(0.090)	(0.041)	(0.023)	(0.088)	(0.085)	(0.090)	(0.088)	(0.089)
Observations	584	584	584	584	584	584	584	584
Mean of Dependent Variable	0	0.413	0.538	0	0	0	0	0

Notes: The table presents results from Obviously Related Instrumental Variable Approach (ORIV) approach by Gillen, Snowberg and Yariv (2019). For the estimates presented in Panels B, C and D our standardized single-item CGI measure (10-point scale) was instrumented with a similar question from the follow-up survey (11-point scale). Swiss Uni sample is used. The estimates in each column and panel come from a separate regression. Risk preference is a standardized measure of participants' investment decisions (mean=0, SD=1). Competitiveness is a dummy that takes the value 1 for those who chose to compete in the competitive task. Our measure of preferences for equality versus efficiency is measured in deciles, with increasing numbers indicating higher priority for efficiency. Staircase risk is a categorical certainty equivalence measure of risk-taking based on a series of hypothetical allocation decisions. Risk is a self-reported measure of risk-taking. Competitiveness is a self-reported measure of competitiveness. A higher value means higher risk-taking or competitiveness. All measures except incentivized competitiveness and efficiency are standardized (mean=0, SD=1). Robust standard errors are reported in parentheses. Significance levels in Panel A are determined by a two-tailed test. Significance levels in Panels B, C and D are determined by a one-tailed test in the direction of the coefficient estimate of female in Panel A. *** p<0.005, ** p<0.01, * p<0.05.

Table A15: CGI and Preference Measures by U.S Waves

	Incentivized Risk		Unincentivized Risk		Financial Risk		Competitiveness		Redistribution		Alturism	
	Wave 1 (1)	Wave 2 (2)	Wave 1 (3)	Wave 2 (4)	Wave 1 (5)	Wave 2 (6)	Wave 1 (7)	Wave 2 (8)	Wave 1 (9)	Wave 2 (10)	Wave 1 (11)	Wave 2 (12)
Panel A. Binary gender												
Female	-0.210***	-0.220***	-0.345***	-0.434***	-0.260***	-0.277***	-0.351***	-0.461***	0.038	0.091*	-0.004	0.007
	(0.045)	(0.046)	(0.044)	(0.045)	(0.021)	(0.021)	(0.044)	(0.045)	(0.045)	(0.046)	(0.045)	(0.046)
R^2	0.011	0.012	0.030	0.047	0.069	0.082	0.031	0.053	0.000	0.002	0.000	0.000
Panel B. Gender identity												
CGI (feminine)	-0.088***	-0.095***	-0.165***	-0.245***	-0.118***	-0.121***	-0.189***	-0.245***	0.055**	0.051*		
	(0.022)	(0.023)	(0.023)	(0.023)	(0.011)	(0.011)	(0.023)	(0.022)	(0.023)	(0.024)		
R^2	0.008	0.009	0.027	0.060	0.057	0.062	0.036	0.060	0.003	0.003		
Panel C. Binary gender and gender identity												
Female	-0.195*	-0.200*	-0.225***	-0.091	-0.201***	-0.244***	-0.132	-0.177*	-0.150	0.022		
	(0.079)	(0.081)	(0.079)	(0.082)	(0.037)	(0.038)	(0.078)	(0.083)	(0.077)	(0.079)		
CGI (feminine)	-0.010	-0.012	-0.074*	-0.207***	-0.036*	-0.020	-0.136***	-0.171***	0.116***	0.041		
	(0.040)	(0.041)	(0.041)	(0.042)	(0.019)	(0.019)	(0.040)	(0.042)	(0.041)	(0.042)		
R^2	0.011	0.012	0.032	0.060	0.071	0.082	0.037	0.062	0.005	0.003		
Panel D. Sample split by binary gender												
Female: CGI (feminine)	0.004	-0.006	-0.019	-0.019	-0.015	0.013	0.072	0.065	-0.115	-0.115		
	(0.033)	(0.033)	(0.037)	(0.034)	(0.015)	(0.014)	(0.033)	(0.035)	(0.034)	(0.031)		
R^2	0.000	0.000	0.000	0.000	0.001	0.001	0.005	0.004	0.013	0.013		
Male: CGI (feminine)	-0.013	-0.007	-0.066*	-0.215***	-0.027*	-0.035*	-0.222***	-0.273***	0.226***	0.160***		
	(0.033)	(0.032)	(0.034)	(0.035)	(0.016)	(0.016)	(0.034)	(0.034)	(0.034)	(0.034)		
R^2	0.000	0.000	0.004	0.046	0.003	0.005	0.049	0.075	0.051	0.025		
Observations	1,994	1,908	1,994	1,908	1,994	1,908	1,994	1,908	1,994	1,908	1,994	1,908
Mean of Dependent Variable	0	0	0	0	0.420	0.381	0	0	0	0	0	0

Notes: The table presents results from regressing preference measures separately by waves for the U.S Adults (age 30–60) sample. The estimates in each column and panel come from a separate regression. Incentivized risk preference is a measure of participants' investment decisions (mean=0, SD=1). Unincentivized risk is a self-reported measure of risk-taking. Financial risk is a dummy that takes the value 1 for those who report to actively trade in securities. Competitiveness is a self-reported measure of competitiveness. A higher value means higher risk-taking or competitiveness. Redistribution is a measure of how much economic redistribution one wants in society. Altruism is a measure of how much one would donate out of a windfall gain. A higher value means greater willingness to redistribute or donate. All measures except financial risk are standardized (mean=0, SD=1). Robust standard errors are reported in parentheses. Significance levels in Panel A are determined by a two-tailed test. Significance levels in Panels B, C and D are determined by a one-tailed test in the direction of the coefficient estimate of female in Panel A. *** p<0.005, ** p<0.01, * p<0.05.

Table A16: CGI and Educational and Labor Market Outcomes by U.S Waves

	Female Educational Track Share		Income		Full-time Homemaker		Weekly Average Work Hours		Managerial Responsibilities		Wage Negotiation	
	Wave 1 (1)	Wave 2 (2)	Wave 1 (3)	Wave 2 (4)	Wave 1 (5)	Wave 2 (6)	Wave 1 (7)	Wave 2 (8)	Wave 1 (9)	Wave 2 (10)	Wave 1 (11)	Wave 2 (12)
Panel A. Binary gender												
Female	0.113***	0.139***	-21.580***	-25.045***	0.103***	0.096***	-7.614***	-5.632***	-0.137***	-0.114***	-0.114***	-0.150***
	(0.010)	(0.011)	(2.707)	(2.557)	(0.011)	(0.011)	(0.640)	(0.634)	(0.022)	(0.022)	(0.022)	(0.022)
R^2	0.092	0.129	0.031	0.048	0.046	0.040	0.066	0.040	0.019	0.014	0.013	0.023
Panel B. Gender identity												
CGI (feminine)	0.053***	0.071***	-10.223***	-11.579***	0.051***	0.048***	-3.670***	-2.639***	-0.060***	-0.063***	-0.055***	-0.071***
	(0.005)	(0.006)	(1.349)	(1.360)	(0.006)	(0.006)	(0.315)	(0.313)	(0.011)	(0.011)	(0.011)	(0.011)
R^2	0.080	0.133	0.028	0.041	0.045	0.040	0.062	0.035	0.015	0.016	0.012	0.021
Panel C. Binary gender and gender identity												
Female	0.081***	0.071***	-14.596***	-18.710***	0.060***	0.053**	-4.854***	-4.023***	-0.116***	-0.031	-0.071	-0.101*
	(0.019)	(0.022)	(4.291)	(4.134)	(0.015)	(0.019)	(1.121)	(1.164)	(0.036)	(0.040)	(0.037)	(0.041)
CGI (feminine)	0.019*	0.042***	-4.318*	-3.822*	0.026***	0.026***	-1.706***	-0.970*	-0.013	-0.050**	-0.026	-0.029
	(0.009)	(0.011)	(2.137)	(2.224)	(0.008)	(0.010)	(0.552)	(0.573)	(0.018)	(0.020)	(0.019)	(0.020)
R^2	0.095	0.143	0.033	0.049	0.050	0.043	0.071	0.041	0.020	0.017	0.014	0.024
Panel D. Sample split by binary gender												
Female: CGI (feminine)	0.003	0.018*	1.792	3.101	0.032***	0.019*	-0.524	0.191	0.041	0.014	0.006	-0.005
	(0.007)	(0.009)	(1.583)	(1.604)	(0.010)	(0.010)	(0.477)	(0.476)	(0.015)	(0.016)	(0.015)	(0.016)
R^2	0.000	0.010	0.001	0.004	0.010	0.004	0.001	0.000	0.007	0.001	0.000	0.000
Male: CGI (feminine)	0.019*	0.028***	-6.443***	-7.359***	0.001	0.010*	-1.438***	-1.275***	-0.051***	-0.070***	-0.035*	-0.028*
	(0.008)	(0.009)	(1.927)	(1.896)	(0.002)	(0.005)	(0.446)	(0.434)	(0.015)	(0.016)	(0.016)	(0.016)
R^2	0.009	0.022	0.010	0.014	0.000	0.007	0.012	0.009	0.010	0.019	0.005	0.003
Observations	1,225	1,054	1,994	1,908	1,994	1,908	1,994	1,908	1,994	1,908	1,994	1,908
Mean of Dependent Variable	0.548	0.532	63.065	59.703	0.061	0.062	25.363	25.195	0.405	0.395	0.421	0.564

Notes: The table presents results from regressing educational and labor market outcomes separately by waves for the U.S Adults (age 30 –60) sample. The estimates in each column and panel come from a separate regression. Female educational track share is the share of women graduating with a bachelor's degree in the chosen field of study in 2020. Income is a self-reported categorical measure in thousand U.S. dollars. Full-time homemaker is a dummy for working full-time at home. Weekly average work hours is a proxy measure constructed from weeks worked and hours worked, using the product of the two categorical measures and dividing by 52. Female industry share is the share of female employees in a given industry. Managerial responsibilities is a dummy for having managerial responsibilities at work. Performance pay is a dummy that is equal to 1 if the respondent's current or most recent job has performance-related pay. Wage negotiation is a dummy that is equal to 1 if the respondent ever negotiated wage. Robust standard errors are reported in parentheses. Significance levels in Panel A are determined by a two-tailed test. Significance levels in Panels B, C and D are determined by a one-tailed test in the direction of the coefficient estimate of female in Panel A. *** p<0.005, ** p<0.01, * p<0.05.

Table A17: Other Gender Identity Measures and Incentivized Risk

Gender Measure	CGI main	TMF	Maglioz fem	Maglioz mas	SOM fem	SOM mas	OSRI fem	OSRI mas	CFNI	CMNI	BSRI fem	BSRI mas	CGI other	CGI own
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
Panel A. Binary gender														
Female	-0.235***	-0.235***	-0.235***	-0.235***	-0.235***	-0.235***	-0.205**	-0.205**	-0.317***	-0.141	-0.270***	-0.270***	-0.235***	-0.235***
	(0.039)	(0.039)	(0.039)	(0.039)	(0.039)	(0.039)	(0.077)	(0.077)	(0.077)	(0.078)	(0.077)	(0.077)	(0.039)	(0.039)
R^2	0.014	0.014	0.014	0.014	0.014	0.014	0.011	0.011	0.025	0.005	0.018	0.018	0.014	0.014
Panel B. Gender identity														
GI	-0.104***	-0.108***	-0.101***	0.116***	-0.079***	0.115***	-0.064	0.139***	-0.001	0.089**	0.003	0.185***	-0.103***	-0.079***
	(0.019)	(0.019)	(0.019)	(0.020)	(0.020)	(0.020)	(0.039)	(0.041)	(0.040)	(0.037)	(0.038)	(0.038)	(0.019)	(0.019)
R^2	0.011	0.012	0.010	0.013	0.006	0.013	0.004	0.019	0.000	0.008	0.000	0.034	0.011	0.006
Panel C. Binary gender and gender identity														
Female	-0.194***	-0.172**	-0.203***	-0.137*	-0.280***	-0.140*	-0.200*	-0.103	-0.354***	-0.080	-0.278***	-0.201**	-0.199***	-0.225***
	(0.067)	(0.066)	(0.065)	(0.066)	(0.060)	(0.059)	(0.091)	(0.087)	(0.082)	(0.085)	(0.078)	(0.078)	(0.065)	(0.050)
GI	-0.025	-0.039	-0.020	0.060*	0.029	0.062*	-0.005	0.117**	0.057	0.072*	0.025	0.164***	-0.022	-0.008
	(0.034)	(0.033)	(0.033)	(0.034)	(0.030)	(0.030)	(0.046)	(0.046)	(0.041)	(0.041)	(0.039)	(0.039)	(0.033)	(0.025)
R^2	0.014	0.014	0.014	0.015	0.014	0.015	0.011	0.022	0.028	0.009	0.019	0.044	0.014	0.014
Panel D. Sample split by binary gender														
Female: GI	0.013	-0.002	0.013	0.038	0.066	0.025	-0.016	0.039	0.073	0.026	0.029	0.226***	0.015	0.012
	(0.029)	(0.029)	(0.029)	(0.029)	(0.029)	(0.029)	(0.057)	(0.058)	(0.057)	(0.051)	(0.058)	(0.052)	(0.028)	(0.028)
R^2	0.000	0.000	0.000	0.001	0.004	0.001	0.000	0.002	0.005	0.001	0.001	0.051	0.000	0.000
Male: GI	-0.038	-0.041	-0.033	0.033	-0.019	0.060*	0.004	0.147**	0.035	0.092*	0.020	0.092	-0.039	-0.022
	(0.027)	(0.027)	(0.027)	(0.028)	(0.026)	(0.028)	(0.050)	(0.057)	(0.056)	(0.052)	(0.053)	(0.059)	(0.027)	(0.027)
R^2	0.001	0.002	0.001	0.001	0.000	0.004	0.000	0.022	0.001	0.009	0.000	0.009	0.001	0.001
Observations	2,659	2,659	2,659	2,659	2,659	2,659	662	662	659	662	676	676	2,659	2,659
Mean of Dependent Variable	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Notes: This table presents results from regressing the incentivized measure of risk on all 14 measures of gender identity collected in this sample. U.S. Adults Wave 2 sample (age 20–60) is used. The estimates in each column and panel come from a separate regression. Risk preference is a standardized measure of participants' investment decisions (mean=0, SD=1). A higher value means higher risk-taking. *N* differs across columns since each participant responded to only one of the four lengthy gender identity inventories (OSRI, CFNI, CMNI, BSRI). See Subsection 2.1 for a detailed description of the alternate gender identity measures. CGI other and CGI own indicate how respondents think that others see them and how they see themselves compared to others with the same binary gender on a scale from 0-very masculine to 10-very feminine. For CGI, TMF, Maglioz fem, SOM fem, CFNI, BSRI fem, CGI other, CGI own higher values indicate greater femininity. For these measures significance levels in Panels B, C and D are determined by a one-tailed test in the direction of the coefficient estimate of female in Panel A. For Maglioz mas, SOM mas, OSRI mas, CMNI, BSRI mas higher values indicate greater masculinity. For these measures, significance levels in Panels B, C and D are determined by a one-tailed test in the opposite direction of the coefficient estimate of female in Panel A. Significance levels in Panel A are determined by a two-tailed test for all measures. All gender identity measures are standardized (mean=0, SD=1) within the relevant sample.** p<0.005, ** p<0.01, * p<0.05.

Table A18: Other Gender Identity Measures and Unincentivized Risk

Gender Measure	CGI main	TMF	Maglioz fem	Maglioz mas	SOM fem	SOM mas	OSRI fem	OSRI mas	CFNI	CMNI	BSRI fem	BSRI mas	CGI other	CGI own
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
Panel A. Binary gender														
Female	-0.418***	-0.418***	-0.418***	-0.418***	-0.418***	-0.418***	-0.425***	-0.425***	-0.484***	-0.335***	-0.422***	-0.422***	-0.418***	-0.418***
	(0.038)	(0.038)	(0.038)	(0.038)	(0.038)	(0.038)	(0.076)	(0.076)	(0.076)	(0.077)	(0.075)	(0.075)	(0.038)	(0.038)
R^2	0.044	0.044	0.044	0.044	0.044	0.044	0.045	0.045	0.058	0.028	0.045	0.045	0.044	0.044
Panel B. Gender identity														
GI	-0.218***	-0.216***	-0.168***	0.251***	-0.167***	0.247***	0.038	0.335***	-0.111***	0.377***	0.034	0.465***	-0.199***	-0.191***
	(0.020)	(0.020)	(0.020)	(0.019)	(0.020)	(0.019)	(0.039)	(0.035)	(0.041)	(0.033)	(0.040)	(0.033)	(0.020)	(0.020)
R^2	0.048	0.047	0.028	0.063	0.028	0.061	0.001	0.112	0.012	0.142	0.001	0.216	0.040	0.037
Panel C. Binary gender and gender identity														
Female	-0.185**	-0.196***	-0.416***	-0.030	-0.399***	-0.106	-0.724***	-0.164	-0.461***	-0.022	-0.444***	-0.236***	-0.279***	-0.293***
	(0.068)	(0.065)	(0.064)	(0.065)	(0.060)	(0.060)	(0.092)	(0.084)	(0.079)	(0.084)	(0.076)	(0.070)	(0.065)	(0.049)
GI	-0.142***	-0.137***	-0.001	0.239***	-0.012	0.207***	0.252	0.299***	-0.035	0.372***	0.069	0.440***	-0.086**	-0.098***
	(0.035)	(0.034)	(0.033)	(0.034)	(0.031)	(0.031)	(0.047)	(0.040)	(0.042)	(0.039)	(0.039)	(0.034)	(0.034)	(0.026)
R^2	0.050	0.050	0.044	0.063	0.044	0.062	0.086	0.118	0.060	0.142	0.049	0.230	0.046	0.049
Panel D. Sample split by binary gender														
Female: GI	0.024	0.011	0.057	0.076***	0.084	0.091***	0.216	0.270***	-0.131*	0.351***	0.036	0.451***	0.028	0.060
	(0.029)	(0.028)	(0.028)	(0.029)	(0.028)	(0.029)	(0.059)	(0.056)	(0.062)	(0.050)	(0.053)	(0.046)	(0.028)	(0.029)
R^2	0.001	0.000	0.003	0.006	0.007	0.008	0.047	0.073	0.017	0.123	0.001	0.203	0.001	0.004
Male: GI	-0.184***	-0.165***	-0.052*	0.218***	-0.085***	0.197***	0.203	0.284***	0.064	0.342***	0.105	0.429***	-0.127***	-0.207***
	(0.030)	(0.029)	(0.030)	(0.028)	(0.029)	(0.030)	(0.052)	(0.050)	(0.055)	(0.051)	(0.059)	(0.051)	(0.029)	(0.028)
R^2	0.034	0.027	0.003	0.047	0.007	0.039	0.041	0.081	0.004	0.117	0.011	0.184	0.016	0.043
Observations	2,659	2,659	2,659	2,659	2,659	2,659	662	662	659	662	676	676	2,659	2,659
Mean of Dependent Variable	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Notes: This table presents results from regressing the unincentivized measure of risk on all 14 measures of gender identity collected in this sample. U.S. Adults Wave 2 sample (age 20–60) is used. The estimates in each column and panel come from a separate regression. Risk preference is a standardized measure of self-reported willingness to take risks (mean=0, SD=1). A higher value means higher risk-taking. *N* differs across columns since each participant responded to only one of the four lengthy gender identity inventories (OSRI, CFNI, CMNI, BSRI). See Subsection 2.1 for a detailed description of the alternate gender identity measures. CGI other and CGI own indicate how respondents think that others see them and how they see themselves compared to others with the same binary gender on a scale from 0-very masculine to 10-very feminine. For CGI, TMF, Maglioz fem, SOM fem, CFNI, BSRI fem, CGI other, CGI own higher values indicate greater femininity. For these measures significance levels in Panels B, C and D are determined by a one-tailed test in the direction of the coefficient estimate of female in Panel A. For Maglioz mas, SOM mas, OSRI mas, CMNI, BSRI mas higher values indicate greater masculinity. For these measures, significance levels in Panels B, C and D are determined by a one-tailed test in the opposite direction of the coefficient estimate of female in Panel A. Significance levels in Panel A are determined by a two-tailed test for all measures. All gender identity measures are standardized (mean=0, SD=1) within the relevant sample. .** p<0.005, ** p<0.01, * p<0.05.

Table A19: Other Gender Identity Measures and Financial Risk

Gender Measure	CGI main	TMF	Maglioz fem	Maglioz mas	SOM fem	SOM mas	OSRI fem	OSRI mas	CFNI	CMNI	BSRI fem	BSRI mas	CGI other	CGI own
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
Panel A. Binary gender														
Female	-0.277***	-0.277***	-0.277***	-0.277***	-0.277***	-0.277***	-0.286***	-0.286***	-0.321***	-0.209***	-0.292***	-0.292***	-0.277***	-0.277***
	(0.021)	(0.021)	(0.021)	(0.021)	(0.021)	(0.021)	(0.043)	(0.043)	(0.042)	(0.043)	(0.042)	(0.042)	(0.021)	(0.021)
R^2	0.082	0.082	0.082	0.082	0.082	0.082	0.086	0.086	0.107	0.047	0.091	0.091	0.082	0.082
Panel B. Gender identity														
GI	-0.121***	-0.124***	-0.114***	0.134***	-0.118***	0.131***	-0.061***	0.143***	-0.067***	0.169***	-0.060***	0.074***	-0.117***	-0.105***
	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.023)	(0.022)	(0.022)	(0.019)	(0.022)	(0.022)	(0.011)	(0.011)
R^2	0.062	0.065	0.055	0.077	0.060	0.073	0.015	0.086	0.018	0.124	0.015	0.024	0.058	0.047
Panel C. Binary gender and gender identity														
Female	-0.244***	-0.227***	-0.270***	-0.175***	-0.240***	-0.186***	-0.323***	-0.197***	-0.313***	-0.085	-0.280***	-0.273***	-0.262***	-0.249***
	(0.038)	(0.038)	(0.036)	(0.039)	(0.035)	(0.035)	(0.052)	(0.050)	(0.046)	(0.046)	(0.043)	(0.043)	(0.038)	(0.029)
GI	-0.020	-0.031	-0.005	0.062***	-0.024	0.059***	0.032	0.098***	-0.012	0.152***	-0.036	0.047*	-0.009	-0.021
	(0.019)	(0.019)	(0.018)	(0.019)	(0.017)	(0.017)	(0.027)	(0.025)	(0.024)	(0.021)	(0.022)	(0.021)	(0.019)	(0.014)
R^2	0.082	0.083	0.082	0.087	0.083	0.087	0.089	0.118	0.108	0.130	0.097	0.101	0.082	0.083
Panel D. Sample split by binary gender														
Female: GI	0.013	-0.004	0.016	0.019	0.008	0.027*	0.028	0.122***	-0.032	0.104***	-0.045	-0.001	0.007	0.005
	(0.014)	(0.014)	(0.014)	(0.015)	(0.014)	(0.015)	(0.029)	(0.029)	(0.030)	(0.030)	(0.029)	(0.028)	(0.014)	(0.014)
R^2	0.001	0.000	0.001	0.002	0.000	0.004	0.004	0.081	0.006	0.056	0.012	0.000	0.000	0.000
Male: GI	-0.035*	-0.031*	-0.021	0.052***	-0.035*	0.049***	0.025	0.064*	0.011	0.167***	-0.025	0.097***	-0.018	-0.038**
	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)	(0.032)	(0.033)	(0.033)	(0.026)	(0.033)	(0.031)	(0.016)	(0.016)
R^2	0.005	0.004	0.002	0.011	0.005	0.010	0.003	0.016	0.001	0.112	0.003	0.038	0.001	0.006
Observations	1,908	1,908	1,908	1,908	1,908	1,908	472	472	475	475	486	486	1,908	1,908
Mean of Dependent Variable	0.381	0.381	0.381	0.381	0.381	0.381	0.388	0.388	0.402	0.362	0.370	0.370	0.381	0.381

Notes: This table presents results from regressing an unincentivized measure of financial market risk on all 14 measures of gender identity collected in this sample. U.S. Adults Wave 2 sample (age 30–60) is used. The estimates in each column and panel come from a separate regression. Financial risk is a dummy that takes the value 1 for those who report to actively trade in securities. N differs across columns since each participant responded to only one of the four lengthy gender identity inventories (OSRI, CFNI, CMNI, BSRI). See Subsection 2.1 for a detailed description of the alternate gender identity measures. CGI other and CGI own indicate how respondents think that others see them and how they see themselves compared to others with the same binary gender on a scale from 0-very masculine to 10-very feminine. For CGI, TMF, Maglioz fem, SOM fem, CFNI, BSRI fem, CGI other, CGI own higher values indicate greater femininity. For these measures significance levels in Panels B, C and D are determined by a one-tailed test in the direction of the coefficient estimate of female in Panel A. For Maglioz mas, SOM mas, OSRI mas, CMNI, BSRI mas higher values indicate greater masculinity. For these measures, significance levels in Panels B, C and D are determined by a one-tailed test in the opposite direction of the coefficient estimate of female in Panel A. Significance levels in Panel A are determined by a two-tailed test for all measures. All gender identity measures are standardized (mean=0, SD=1) within the relevant sample. .** p<0.005, ** p<0.01, * p<0.05.

Table A20: Other Gender Identity Measures and Competitiveness

Gender Measure	CGI main	TMF	Maglioz fem	Maglioz mas	SOM fem	SOM mas	OSRI fem	OSRI mas	CFNI	CMNI	BSRI fem	BSRI mas	CGI other	CGI own
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
Panel A. Binary gender														
Female	-0.455***	-0.455***	-0.455***	-0.455***	-0.455***	-0.455***	-0.378***	-0.378***	-0.513***	-0.469***	-0.464***	-0.464***	-0.455***	-0.455***
	(0.038)	(0.038)	(0.038)	(0.038)	(0.038)	(0.038)	(0.077)	(0.077)	(0.076)	(0.076)	(0.075)	(0.075)	(0.038)	(0.038)
R^2	0.052	0.052	0.052	0.052	0.052	0.052	0.036	0.036	0.066	0.055	0.054	0.054	0.052	0.052
Panel B. Gender identity														
GI	-0.230***	-0.239***	-0.194***	0.268***	-0.180***	0.268***	-0.017	0.249***	-0.022	0.405***	0.040	0.500***	-0.220***	-0.223***
	(0.019)	(0.019)	(0.019)	(0.019)	(0.019)	(0.019)	(0.039)	(0.037)	(0.040)	(0.035)	(0.039)	(0.033)	(0.019)	(0.019)
R^2	0.053	0.057	0.038	0.072	0.032	0.072	0.000	0.062	0.000	0.164	0.002	0.250	0.048	0.050
Panel C. Binary gender and gender identity														
Female	-0.237***	-0.198***	-0.402***	-0.059	-0.440***	-0.117*	-0.550***	-0.199*	-0.559***	-0.157	-0.489***	-0.265***	-0.287***	-0.289***
	(0.069)	(0.066)	(0.067)	(0.069)	(0.062)	(0.060)	(0.095)	(0.084)	(0.082)	(0.081)	(0.075)	(0.068)	(0.067)	(0.050)
GI	-0.134***	-0.159***	-0.033	0.244***	-0.010	0.224***	0.146	0.206***	0.070	0.372***	0.078	0.472***	-0.104***	-0.130***
	(0.035)	(0.033)	(0.034)	(0.034)	(0.032)	(0.030)	(0.048)	(0.041)	(0.042)	(0.040)	(0.038)	(0.034)	(0.034)	(0.026)
R^2	0.058	0.061	0.052	0.072	0.052	0.073	0.049	0.070	0.070	0.169	0.060	0.267	0.056	0.062
Panel D. Sample split by binary gender														
Female: GI	0.062	0.034	0.100	0.047	0.132	0.049*	0.178	0.259***	0.053	0.326***	0.026	0.511***	0.051	0.056
	(0.030)	(0.029)	(0.030)	(0.029)	(0.029)	(0.029)	(0.058)	(0.049)	(0.061)	(0.049)	(0.055)	(0.044)	(0.029)	(0.029)
R^2	0.004	0.001	0.010	0.002	0.017	0.002	0.032	0.067	0.003	0.106	0.001	0.261	0.003	0.003
Male: GI	-0.226***	-0.225***	-0.141***	0.270***	-0.136***	0.284***	0.073	0.146***	0.087	0.376***	0.145	0.429***	-0.185***	-0.273***
	(0.030)	(0.029)	(0.029)	(0.029)	(0.030)	(0.028)	(0.055)	(0.055)	(0.054)	(0.056)	(0.056)	(0.054)	(0.029)	(0.028)
R^2	0.051	0.051	0.020	0.073	0.019	0.081	0.005	0.021	0.008	0.141	0.021	0.184	0.034	0.075
Observations	2,659	2,659	2,659	2,659	2,659	2,659	662	662	659	662	676	676	2,659	2,659
Mean of Dependent Variable	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Notes: This table presents results from regressing the unincentivized measure of competitiveness on all 14 measures of gender identity collected in this sample. U.S. Adults Wave 2 sample (age 20–60) is used. The estimates in each column and panel come from a separate regression. Competitiveness is a self-reported measure of competitiveness. A higher value mean higher competitiveness. N differs across columns since each participant responded to only one of the four lengthy gender identity inventories (OSRI, CFNI, CMNI, BSRI). See Subsection 2.1 for a detailed description of the alternate gender identity measures. CGI other and CGI own indicate how respondents think that others see them and how they see themselves compared to others with the same binary gender on a scale from 0-very masculine to 10-very feminine. For CGI, TMF, Maglioz fem, SOM fem, CFNI, BSRI fem, CGI other, CGI own higher values indicate greater femininity. For these measures significance levels in Panels B, C and D are determined by a one-tailed test in the direction of the coefficient estimate of female in Panel A. For Maglioz mas, SOM mas, OSRI mas, CMNI, BSRI mas higher values indicate greater masculinity. For these measures, significance levels in Panels B, C and D are determined by a one-tailed test in the opposite direction of the coefficient estimate of female in Panel A. Significance levels in Panel A are determined by a two-tailed test for all measures. All gender identity measures are standardized (mean=0, SD=1) within the relevant sample. .** p<0.005, ** p<0.01, * p<0.05.

Table A21: Other Gender Identity Measures and Redistribution

Gender Measure	CGI main	TMF	Maglioz fem	Maglioz mas	SOM fem	SOM mas	OSRI fem	OSRI mas	CFNI	CMNI	BSRI fem	BSRI mas	CGI other	CGI own
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
Panel A. Binary gender														
Female	0.088*	0.088*	0.088*	0.088*	0.088*	0.088*	0.126	0.126	-0.029	0.110	0.147	0.147	0.088*	0.088*
	(0.039)	(0.039)	(0.039)	(0.039)	(0.039)	(0.039)	(0.078)	(0.078)	(0.078)	(0.078)	(0.077)	(0.077)	(0.039)	(0.039)
R^2	0.002	0.002	0.002	0.002	0.002	0.002	0.004	0.004	0.000	0.003	0.005	0.005	0.002	0.002
Panel B. Gender identity														
GI	0.051**	0.063***	0.067***	-0.018	0.074***	-0.012	0.143***	-0.121***	-0.170***	-0.145***	0.139***	-0.086*	0.040*	0.040*
	(0.021)	(0.021)	(0.020)	(0.020)	(0.021)	(0.021)	(0.040)	(0.041)	(0.041)	(0.039)	(0.038)	(0.040)	(0.021)	(0.021)
R^2	0.003	0.004	0.005	0.000	0.005	0.000	0.020	0.015	0.029	0.021	0.019	0.007	0.002	0.002
Panel C. Binary gender and gender identity														
Female	0.011	-0.041	-0.057	0.171**	-0.066	0.160***	-0.067	0.025	0.092	-0.014	0.106	0.116	0.066	0.061
	(0.065)	(0.064)	(0.063)	(0.065)	(0.059)	(0.057)	(0.094)	(0.087)	(0.081)	(0.084)	(0.077)	(0.078)	(0.065)	(0.049)
GI	0.047	0.080*	0.090***	0.051	0.099***	0.048	0.162***	-0.116**	-0.185***	-0.148***	0.130***	-0.073*	0.013	0.021
	(0.035)	(0.034)	(0.033)	(0.034)	(0.032)	(0.031)	(0.049)	(0.046)	(0.043)	(0.043)	(0.039)	(0.040)	(0.035)	(0.027)
R^2	0.003	0.004	0.005	0.003	0.006	0.003	0.021	0.015	0.031	0.021	0.022	0.011	0.002	0.002
Panel D. Sample split by binary gender														
Female: GI	-0.111	-0.133	-0.101	0.155	-0.116	0.169	0.008	-0.046	-0.245***	-0.034	0.054	-0.021	-0.103	-0.120
	(0.028)	(0.029)	(0.028)	(0.027)	(0.029)	(0.029)	(0.059)	(0.061)	(0.059)	(0.058)	(0.052)	(0.055)	(0.029)	(0.029)
R^2	0.012	0.018	0.010	0.024	0.013	0.029	0.000	0.002	0.060	0.001	0.003	0.000	0.011	0.014
Male: GI	0.158***	0.209***	0.196***	-0.113***	0.216***	-0.130***	0.220***	-0.141**	-0.094*	-0.213***	0.207***	-0.127*	0.117***	0.149***
	(0.029)	(0.028)	(0.027)	(0.029)	(0.026)	(0.028)	(0.052)	(0.056)	(0.057)	(0.053)	(0.055)	(0.058)	(0.028)	(0.028)
R^2	0.025	0.044	0.038	0.013	0.047	0.017	0.048	0.020	0.009	0.045	0.043	0.016	0.014	0.022
Observations	2,659	2,659	2,659	2,659	2,659	2,659	662	662	659	662	676	676	2,659	2,659
Mean of Dependent Variable	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Notes: This table presents results from regressing the unincentivized measure of redistribution on all 14 measures of gender identity collected in this sample. U.S. Adults Wave 2 sample (age 20–60) is used. The estimates in each column and panel come from a separate regression. Redistribution is a measure of how much economic redistribution one wants in society. A higher value means a greater willingness to redistribute. *N* differs across columns since each participant responded to only one of the four lengthy gender identity inventories (OSRI, CFNI, CMNI, BSRI). See Subsection 2.1 for a detailed description of the alternate gender identity measures. CGI other and CGI own indicate how respondents think that others see them and how they see themselves compared to others with the same binary gender on a scale from 0-very masculine to 10-very feminine. For CGI, TMF, Maglioz fem, SOM fem, CFNI, BSRI fem, CGI other, CGI own higher values indicate greater femininity. For these measures significance levels in Panels B, C and D are determined by a one-tailed test in the direction of the coefficient estimate of female in Panel A. For Maglioz mas, SOM mas, OSRI mas, CMNI, BSRI mas higher values indicate greater masculinity. For these measures, significance levels in Panels B, C and D are determined by a one-tailed test in the opposite direction of the coefficient estimate of female in Panel A. Significance levels in Panel A are determined by a two-tailed test for all measures. All gender identity measures are standardized (mean=0, SD=1) within the relevant sample. .** p<0.005, ** p<0.01, * p<0.05.

Table A22: Other Gender Identity Measures and Female Educational Track Share

Gender Measure	CGI main	TMF	Maglioz fem	Maglioz mas	SOM fem	SOM mas	OSRI fem	OSRI mas	CFNI	CMNI	BSRI fem	BSRI mas	CGI other	CGI own
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
Panel A. Binary gender														
Female	0.131***	0.131***	0.131***	0.131***	0.131***	0.131***	0.128***	0.128***	0.129***	0.143***	0.122***	0.122***	0.131***	0.131***
	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.018)	(0.018)	(0.019)	(0.018)	(0.019)	(0.019)	(0.009)	(0.009)
R^2	0.117	0.117	0.117	0.117	0.117	0.117	0.117	0.117	0.115	0.141	0.097	0.097	0.117	0.117
Panel B. Gender identity														
GI	0.067***	0.065***	0.065***	-0.066***	0.063***	-0.065***	0.039***	-0.048***	0.029***	-0.067***	0.028***	0.000	0.060***	0.057***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.009)	(0.009)	(0.009)	(0.009)	(0.010)	(0.010)	(0.005)	(0.005)
R^2	0.124	0.117	0.115	0.118	0.108	0.115	0.044	0.067	0.024	0.122	0.020	0.000	0.099	0.089
Panel C. Binary gender and gender identity														
Female	0.062***	0.072***	0.075***	0.070***	0.083***	0.076***	0.121***	0.106***	0.122***	0.105***	0.117***	0.128***	0.097***	0.097***
	(0.017)	(0.017)	(0.016)	(0.016)	(0.014)	(0.015)	(0.022)	(0.020)	(0.019)	(0.020)	(0.020)	(0.019)	(0.016)	(0.012)
GI	0.042***	0.036***	0.035***	-0.037***	0.031***	-0.036***	0.006	-0.024*	0.014*	-0.044***	0.021*	0.014	0.021**	0.027***
	(0.009)	(0.008)	(0.008)	(0.008)	(0.007)	(0.007)	(0.011)	(0.010)	(0.009)	(0.010)	(0.010)	(0.010)	(0.008)	(0.006)
R^2	0.133	0.130	0.128	0.129	0.128	0.132	0.118	0.130	0.121	0.183	0.108	0.101	0.121	0.129
Panel D. Sample split by binary gender														
Female: GI	0.016**	0.014*	0.016**	-0.016**	0.010*	-0.020***	0.015	-0.009	0.024*	-0.030**	0.036**	0.024	0.003	0.012*
	(0.007)	(0.007)	(0.006)	(0.006)	(0.006)	(0.006)	(0.011)	(0.013)	(0.010)	(0.013)	(0.015)	(0.013)	(0.006)	(0.006)
R^2	0.010	0.007	0.009	0.009	0.004	0.014	0.009	0.003	0.022	0.036	0.039	0.017	0.000	0.005
Male: GI	0.031***	0.028***	0.025***	-0.027***	0.029***	-0.028***	-0.003	-0.029*	0.003	-0.048***	0.005	0.003	0.021***	0.031***
	(0.008)	(0.008)	(0.007)	(0.007)	(0.007)	(0.007)	(0.014)	(0.013)	(0.013)	(0.013)	(0.014)	(0.014)	(0.007)	(0.007)
R^2	0.027	0.022	0.018	0.020	0.023	0.021	0.000	0.024	0.000	0.061	0.001	0.000	0.012	0.026
Observations	1,531	1,531	1,531	1,531	1,531	1,531	395	395	369	387	380	380	1,531	1,531
Mean of Dependent Variable	0.532	0.532	0.532	0.532	0.532	0.532	0.535	0.535	0.523	0.539	0.531	0.531	0.532	0.532

Notes: This table presents results from regressing female educational track share on all 14 measures of gender identity collected in this sample. U.S. Adults Wave 2 sample (age 20–60) is used. The estimates in each column and panel come from a separate regression. Female educational track share is the share of women graduating with a bachelor's degree in the chosen field of study in 2020. *N* differs across columns since each participant responded to only one of the four lengthy gender identity inventories (OSRI, CFNI, CMNI, BSRI). See Subsection 2.1 for a detailed description of the alternate gender identity measures. CGI other and CGI own indicate how respondents think that others see them and how they see themselves compared to others with the same binary gender on a scale from 0-very masculine to 10-very feminine. For CGI, TMF, Maglioz fem, SOM fem, CFNI, BSRI fem, CGI other, CGI own higher values indicate greater femininity. For these measures significance levels in Panels B, C and D are determined by a one-tailed test in the direction of the coefficient estimate of female in Panel A. For Maglioz mas, SOM mas, OSRI mas, CMNI, BSRI mas higher values indicate greater masculinity. For these measures, significance levels in Panels B, C and D are determined by a one-tailed test in the opposite direction of the coefficient estimate of female in Panel A. Significance levels in Panel A are determined by a two-tailed test for all measures. All gender identity measures are standardized (mean=0, SD=1) within the relevant sample. .** p<0.005, ** p<0.01, * p<0.05.

Table A23: Other Gender Identity Measures and Income

Gender Measure	CGI main	TMF	Maglioz fem	Maglioz mas	SOM fem	SOM mas	OSRI fem	OSRI mas	CFNI	CMNI	BSRI fem	BSRI mas	CGI other	CGI own
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
Panel A. Binary gender														
Female	-20.376***	-20.376***	-20.376***	-20.376***	-20.376***	-20.376***	-13.953***	-13.953***	-26.463***	-17.840***	-23.488***	-23.488***	-20.376***	-20.376***
	(2.099)	(2.099)	(2.099)	(2.099)	(2.099)	(2.099)	(4.010)	(4.010)	(4.127)	(4.165)	(4.503)	(4.503)	(2.099)	(2.099)
R^2	0.034	0.034	0.034	0.034	0.034	0.034	0.018	0.018	0.058	0.027	0.040	0.040	0.034	0.034
Panel B. Gender identity														
GI	-9.401***	-9.439***	-8.770***	9.819***	-8.424***	9.081***	-2.135	9.227***	4.753	6.432***	1.414	10.259***	-9.214***	-8.142***
	(1.112)	(1.110)	(1.088)	(1.082)	(1.095)	(1.084)	(1.936)	(2.019)	(1.958)	(1.958)	(2.157)	(2.315)	(1.116)	(1.117)
R^2	0.029	0.029	0.025	0.032	0.023	0.027	0.002	0.031	0.007	0.014	0.001	0.030	0.028	0.022
Panel C. Binary gender and gender identity														
Female	-15.093***	-14.756***	-17.640***	-13.016***	-18.290***	-15.495***	-17.576***	-7.305	-33.137***	-15.105***	-24.538***	-20.053***	-15.860***	-16.849***
	(3.281)	(3.366)	(3.270)	(3.262)	(3.005)	(3.088)	(5.280)	(4.284)	(4.393)	(4.612)	(4.562)	(4.350)	(3.333)	(2.529)
GI	-3.232*	-3.480*	-1.709	4.528***	-1.349	3.236*	3.059	7.641***	10.185	3.257	3.342	8.144***	-2.787	-2.762*
	(1.759)	(1.788)	(1.699)	(1.694)	(1.579)	(1.595)	(2.556)	(2.166)	(2.035)	(2.166)	(2.142)	(2.267)	(1.787)	(1.357)
R^2	0.035	0.035	0.034	0.036	0.034	0.036	0.020	0.035	0.088	0.030	0.043	0.058	0.035	0.036
Panel D. Sample split by binary gender														
Female: GI	3.734	2.099	3.180	-1.878	4.211	-1.857	-2.424	0.327	3.635	2.599	1.393	2.961	2.895	3.359
	(1.300)	(1.398)	(1.258)	(1.185)	(1.194)	(1.396)	(3.080)	(2.400)	(2.496)	(2.385)	(2.703)	(3.091)	(1.335)	(1.379)
R^2	0.006	0.002	0.005	0.002	0.008	0.002	0.003	0.000	0.007	0.003	0.001	0.003	0.004	0.005
Male: GI	-7.265***	-5.829***	-4.921***	7.892***	-5.248***	6.920***	6.214	11.263***	16.269	3.279	5.380	13.698***	-6.148***	-7.559***
	(1.541)	(1.583)	(1.580)	(1.617)	(1.530)	(1.594)	(2.757)	(2.863)	(2.944)	(2.998)	(3.323)	(3.212)	(1.622)	(1.546)
R^2	0.014	0.009	0.007	0.017	0.008	0.013	0.013	0.042	0.069	0.003	0.007	0.045	0.010	0.016
Observations	2,659	2,659	2,659	2,659	2,659	2,659	662	662	659	662	676	676	2,659	2,659
Mean of Dependent Variable	54.914	54.914	54.914	54.914	54.914	54.914	53.074	53.074	56.153	54.728	55.692	55.692	54.914	54.914

Notes: This table presents results from regressing income on all 14 measures of gender identity collected in this sample. U.S. Adults Wave 2 sample (age 20–60) is used. The estimates in each column and panel come from a separate regression. Income is a self-reported categorical measure in thousand U.S. dollars. N differs across columns since each participant responded to only one of the four lengthy gender identity inventories (OSRI, CFNI, CMNI, BSRI). See Subsection 2.1 for a detailed description of the alternate gender identity measures. CGI other and CGI own indicate how respondents think that others see them and how they see themselves compared to others with the same binary gender on a scale from 0-very masculine to 10-very feminine. For CGI, TMF, Maglioz fem, SOM fem, CFNI, BSRI fem, CGI other, CGI own higher values indicate greater femininity. For these measures significance levels in Panels B, C and D are determined by a one-tailed test in the direction of the coefficient estimate of female in Panel A. For Maglioz mas, SOM mas, OSRI mas, CMNI, BSRI mas higher values indicate greater masculinity. For these measures, significance levels in Panels B, C and D are determined by a one-tailed test in the opposite direction of the coefficient estimate of female in Panel A. Significance levels in Panel A are determined by a two-tailed test for all measures. All gender identity measures are standardized (mean=0, SD=1) within the relevant sample. .** p<0.005, ** p<0.01, * p<0.05.

Table A24: Other Gender Identity Measures and Full-time Homemaker

Gender Measure	CGI main	TMF	Maglioz fem	Maglioz mas	SOM fem	SOM mas	OSRI fem	OSRI mas	CFNI	CMNI	BSRI fem	BSRI mas	CGI other	CGI own
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
Panel A. Binary gender														
Female	0.096***	0.096***	0.096***	0.096***	0.096***	0.096***	0.103***	0.103***	0.123***	0.080***	0.080***	0.080***	0.096***	0.096***
	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.022)	(0.022)	(0.024)	(0.020)	(0.021)	(0.021)	(0.011)	(0.011)
R^2	0.040	0.040	0.040	0.040	0.040	0.040	0.044	0.044	0.057	0.032	0.028	0.028	0.040	0.040
Panel B. Gender identity														
GI	0.048***	0.047***	0.045***	-0.046***	0.043***	-0.048***	0.031***	-0.019*	0.021*	-0.029***	0.017*	-0.033***	0.047***	0.041***
	(0.006)	(0.006)	(0.006)	(0.006)	(0.005)	(0.006)	(0.009)	(0.009)	(0.013)	(0.009)	(0.009)	(0.011)	(0.006)	(0.006)
R^2	0.040	0.039	0.035	0.036	0.031	0.039	0.016	0.006	0.007	0.017	0.005	0.018	0.038	0.028
Panel C. Binary gender and gender identity														
Female	0.053**	0.056***	0.068***	0.064***	0.077***	0.056***	0.100***	0.108***	0.123***	0.068***	0.077***	0.070***	0.059***	0.076***
	(0.019)	(0.019)	(0.018)	(0.021)	(0.018)	(0.017)	(0.025)	(0.025)	(0.024)	(0.021)	(0.021)	(0.021)	(0.018)	(0.015)
GI	0.026***	0.024**	0.017*	-0.019*	0.012	-0.026***	0.003	0.005	-0.000	-0.015*	0.011	-0.026*	0.023**	0.015*
	(0.010)	(0.010)	(0.009)	(0.011)	(0.009)	(0.009)	(0.010)	(0.010)	(0.013)	(0.009)	(0.009)	(0.011)	(0.009)	(0.008)
R^2	0.043	0.043	0.041	0.042	0.041	0.044	0.044	0.045	0.057	0.036	0.030	0.039	0.042	0.042
Panel D. Sample split by binary gender														
Female: GI	0.019*	0.020*	0.014	-0.012	0.008	-0.022*	0.003	0.014	0.012	-0.022	0.023	-0.041*	0.022*	0.012
	(0.010)	(0.011)	(0.010)	(0.010)	(0.011)	(0.010)	(0.018)	(0.021)	(0.022)	(0.016)	(0.016)	(0.019)	(0.010)	(0.010)
R^2	0.004	0.004	0.002	0.002	0.001	0.005	0.000	0.002	0.001	0.006	0.006	0.018	0.005	0.002
Male: GI	0.010*	0.008*	0.006*	-0.009*	0.008*	-0.009*	0.001	-0.002	-0.014	-0.008	-0.003	-0.009	0.003	0.010*
	(0.005)	(0.005)	(0.003)	(0.005)	(0.004)	(0.005)	(0.002)	(0.004)	(0.008)	(0.007)	(0.005)	(0.010)	(0.004)	(0.005)
R^2	0.007	0.005	0.003	0.006	0.004	0.006	0.000	0.000	0.016	0.005	0.000	0.004	0.001	0.007
Observations	1,908	1,908	1,908	1,908	1,908	1,908	472	472	475	475	486	486	1,908	1,908
Mean of Dependent Variable	0.062	0.062	0.062	0.062	0.062	0.062	0.064	0.064	0.072	0.053	0.062	0.062	0.062	0.062

Notes: This table presents results from regressing full-time homemaker on all 14 measures of gender identity collected in this sample. U.S. Adults Wave 2 sample (age 30–60) is used. The estimates in each column and panel come from a separate regression. Full-time homemaker is a dummy for working full-time at home. N differs across columns since each participant responded to only one of the four lengthy gender identity inventories (OSRI, CFNI, CMNI, BSRI). See Subsection 2.1 for a detailed description of the alternate gender identity measures. CGI other and CGI own indicate how respondents think that others see them and how they see themselves compared to others with the same binary gender on a scale from 0-very masculine to 10-very feminine. For CGI, TMF, Maglioz fem, SOM fem, CFNI, BSRI fem, CGI other, CGI own higher values indicate greater femininity. For these measures significance levels in Panels B, C and D are determined by a one-tailed test in the direction of the coefficient estimate of female in Panel A. For Maglioz mas, SOM mas, OSRI mas, CMNI, BSRI mas higher values indicate greater masculinity. For these measures, significance levels in Panels B, C and D are determined by a one-tailed test in the opposite direction of the coefficient estimate of female in Panel A. Significance levels in Panel A are determined by a two-tailed test for all measures. All gender identity measures are standardized (mean=0, SD=1) within the relevant sample. .** p<0.005, ** p<0.01, * p<0.05.

Table A25: Other Gender Identity Measures and Weekly Average Work Hours

Gender Measure	CGI main	TMF	Maglioz fem	Maglioz mas	SOM fem	SOM mas	OSRI fem	OSRI mas	CFNI	CMNI	BSRI fem	BSRI mas	CGI other	CGI own
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
Panel A. Binary gender														
Female	-5.632***	-5.632***	-5.632***	-5.632***	-5.632***	-5.632***	-4.437***	-4.437***	-7.655***	-4.815***	-5.573***	-5.573***	-5.632***	-5.632***
	(0.634)	(0.634)	(0.634)	(0.634)	(0.634)	(0.634)	(1.242)	(1.242)	(1.303)	(1.268)	(1.266)	(1.266)	(0.634)	(0.634)
R^2	0.040	0.040	0.040	0.040	0.040	0.040	0.027	0.027	0.069	0.030	0.039	0.039	0.040	0.040
Panel B. Gender identity														
GI	-2.639***	-2.765***	-2.424***	2.860***	-2.573***	2.703***	-0.367	1.563***	0.048	1.507**	-0.553	2.609***	-2.757***	-2.432***
	(0.313)	(0.311)	(0.317)	(0.312)	(0.313)	(0.315)	(0.629)	(0.597)	(0.668)	(0.633)	(0.635)	(0.643)	(0.314)	(0.314)
R^2	0.035	0.038	0.029	0.041	0.033	0.037	0.001	0.013	0.000	0.012	0.002	0.034	0.038	0.030
Panel C. Binary gender and gender identity														
Female	-4.023***	-3.354***	-4.963***	-2.841*	-4.168***	-3.614***	-6.004***	-3.804**	-8.748***	-4.305***	-5.548***	-4.724***	-3.395***	-4.309***
	(1.164)	(1.102)	(1.099)	(1.184)	(1.028)	(1.022)	(1.456)	(1.429)	(1.361)	(1.383)	(1.284)	(1.279)	(1.150)	(0.884)
GI	-0.970*	-1.391**	-0.413	1.677***	-0.928*	1.298**	1.362	0.697	1.570	0.624	-0.074	2.139***	-1.356**	-0.986*
	(0.573)	(0.541)	(0.548)	(0.583)	(0.506)	(0.508)	(0.727)	(0.688)	(0.691)	(0.690)	(0.642)	(0.655)	(0.569)	(0.437)
R^2	0.041	0.043	0.040	0.044	0.041	0.043	0.033	0.029	0.079	0.031	0.039	0.060	0.043	0.042
Panel D. Sample split by binary gender														
Female: GI	0.191	-0.255	0.426	0.080	0.303	0.427	1.529	-0.171	0.136	0.834	-1.633*	2.831***	-0.344	0.021
	(0.476)	(0.468)	(0.495)	(0.468)	(0.470)	(0.460)	(0.896)	(0.986)	(0.909)	(0.894)	(0.872)	(0.887)	(0.466)	(0.479)
R^2	0.000	0.000	0.001	0.000	0.000	0.001	0.012	0.000	0.000	0.003	0.012	0.037	0.001	0.000
Male: GI	-1.275***	-1.320***	-0.886*	1.955***	-1.356***	1.293***	0.787	1.212	2.932	0.380	1.521	1.307	-1.207***	-1.529***
	(0.434)	(0.414)	(0.418)	(0.438)	(0.414)	(0.445)	(0.815)	(0.812)	(0.914)	(0.893)	(0.899)	(0.929)	(0.446)	(0.433)
R^2	0.009	0.010	0.004	0.022	0.010	0.010	0.004	0.009	0.047	0.001	0.013	0.010	0.008	0.013
Observations	1,908	1,908	1,908	1,908	1,908	1,908	472	472	475	475	486	486	1,908	1,908
Mean of Dependent Variable	25.195	25.195	25.195	25.195	25.195	25.195	26.160	26.160	25.585	24.177	24.873	24.873	25.195	25.195

Notes: This table presents results from regressing weekly average work hours on all 14 measures of gender identity collected in this sample. U.S. Adults Wave 2 sample (age 30–60) is used. The estimates in each column and panel come from a separate regression. Weekly average work hours is a proxy measure constructed from weeks worked and hours worked, using the product of the two categorical measures and dividing by 52. *N* differs across columns since each participant responded to only one of the four lengthy gender identity inventories (OSRI, CFNI, CMNI, BSRI). See Subsection 2.1 for a detailed description of the alternate gender identity measures. CGI other and CGI own indicate how respondents think that others see them and how they see themselves compared to others with the same binary gender on a scale from 0-very masculine to 10-very feminine. For CGI, TMF, Maglioz fem, SOM fem, CFNI, BSRI fem, CGI other, CGI own higher values indicate greater femininity. For these measures significance levels in Panels B, C and D are determined by a one-tailed test in the direction of the coefficient estimate of female in Panel A. For Maglioz mas, SOM mas, OSRI mas, CMNI, BSRI mas higher values indicate greater masculinity. For these measures, significance levels in Panels B, C and D are determined by a one-tailed test in the opposite direction of the coefficient estimate of female in Panel A. Significance levels in Panel A are determined by a two-tailed test for all measures. All gender identity measures are standardized (mean=0, SD=1) within the relevant sample. .** p<0.005, ** p<0.01, * p<0.05.

Table A26: Other Gender Identity Measures and Managerial Responsibilities

Gender Measure	CGI main	TMF	Maglioz fem	Maglioz mas	SOM fem	SOM mas	OSRI fem	OSRI mas	CFNI	CMNI	BSRI fem	BSRI mas	CGI other	CGI own
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
Panel A. Binary gender														
Female	-0.114***	-0.114***	-0.114***	-0.114***	-0.114***	-0.114***	-0.049	-0.049	-0.178***	-0.078	-0.149***	-0.149***	-0.114***	-0.114***
	(0.022)	(0.022)	(0.022)	(0.022)	(0.022)	(0.022)	(0.044)	(0.044)	(0.044)	(0.045)	(0.044)	(0.044)	(0.022)	(0.022)
R^2	0.014	0.014	0.014	0.014	0.014	0.014	0.003	0.003	0.033	0.006	0.023	0.023	0.014	0.014
Panel B. Gender identity														
GI	-0.063***	-0.059***	-0.046***	0.066***	-0.046***	0.073***	0.051	0.065***	0.020	0.086***	0.050	0.144***	-0.059***	-0.050***
	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.022)	(0.023)	(0.022)	(0.022)	(0.020)	(0.020)	(0.011)	(0.011)
R^2	0.016	0.015	0.009	0.018	0.009	0.022	0.011	0.018	0.002	0.030	0.010	0.086	0.014	0.011
Panel C. Binary gender and gender identity														
Female	-0.031	-0.050	-0.115***	-0.015	-0.107***	-0.001	-0.160***	0.013	-0.219***	-0.010	-0.172***	-0.096*	-0.053	-0.084***
	(0.040)	(0.039)	(0.038)	(0.040)	(0.036)	(0.035)	(0.054)	(0.050)	(0.047)	(0.049)	(0.045)	(0.044)	(0.039)	(0.030)
GI	-0.050**	-0.039*	0.001	0.059***	-0.004	0.072***	0.097	0.068***	0.058	0.084***	0.065	0.134***	-0.037*	-0.022
	(0.020)	(0.020)	(0.019)	(0.020)	(0.018)	(0.018)	(0.027)	(0.026)	(0.024)	(0.024)	(0.021)	(0.021)	(0.020)	(0.015)
R^2	0.017	0.016	0.014	0.018	0.014	0.022	0.030	0.019	0.046	0.030	0.040	0.095	0.015	0.015
Panel D. Sample split by binary gender														
Female: GI	0.014	0.006	0.018	0.004	0.027	0.024	0.026	0.074*	0.010	0.006	0.019	0.117***	0.014	0.031
	(0.016)	(0.015)	(0.016)	(0.016)	(0.015)	(0.016)	(0.032)	(0.033)	(0.032)	(0.034)	(0.029)	(0.028)	(0.016)	(0.016)
R^2	0.001	0.000	0.001	0.000	0.003	0.003	0.003	0.025	0.000	0.000	0.002	0.061	0.001	0.004
Male: GI	-0.070***	-0.049***	-0.017	0.067***	-0.030*	0.072***	0.124	0.052	0.103	0.130***	0.111	0.148***	-0.057***	-0.067***
	(0.016)	(0.017)	(0.017)	(0.016)	(0.017)	(0.016)	(0.030)	(0.032)	(0.032)	(0.031)	(0.028)	(0.030)	(0.016)	(0.016)
R^2	0.019	0.010	0.001	0.018	0.004	0.021	0.064	0.011	0.042	0.068	0.049	0.088	0.013	0.018
Observations	1,908	1,908	1,908	1,908	1,908	1,908	472	472	475	475	486	486	1,908	1,908
Mean of Dependent Variable	0.395	0.395	0.395	0.395	0.395	0.395	0.362	0.362	0.394	0.415	0.409	0.409	0.395	0.395

Notes: This table presents results from regressing managerial responsibilities on all 14 measures of gender identity collected in this sample. U.S. Adults Wave 2 sample (age 30–60) is used. The estimates in each column and panel come from a separate regression. Managerial responsibilities is a dummy for having managerial responsibilities at work. *N* differs across columns since each participant responded to only one of the four lengthy gender identity inventories (OSRI, CFNI, CMNI, BSRI). See Subsection 2.1 for a detailed description of the alternate gender identity measures. CGI other and CGI own indicate how respondents think that others see them and how they see themselves compared to others with the same binary gender on a scale from 0-very masculine to 10-very feminine. For CGI, TMF, Maglioz fem, SOM fem, CFNI, BSRI fem, CGI other, CGI own higher values indicate greater femininity. For these measures significance levels in Panels B, C and D are determined by a one-tailed test in the direction of the coefficient estimate of female in Panel A. For Maglioz mas, SOM mas, OSRI mas, CMNI, BSRI mas higher values indicate greater masculinity. For these measures, significance levels in Panels B, C and D are determined by a one-tailed test in the opposite direction of the coefficient estimate of female in Panel A. Significance levels in Panel A are determined by a two-tailed test for all measures. All gender identity measures are standardized (mean=0, SD=1) within the relevant sample. .** p<0.005, ** p<0.01, * p<0.05.

Table A27: Other Gender Identity Measures and Wage Negotiation

Gender Measure	CGI main	TMF	Maglioz fem	Maglioz mas	SOM fem	SOM mas	OSRI fem	OSRI mas	CFNI	CMNI	BSRI fem	BSRI mas	CGI other	CGI own
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
Panel A. Binary gender														
Female	-0.150***	-0.150***	-0.150***	-0.150***	-0.150***	-0.150***	-0.118**	-0.118**	-0.189***	-0.069	-0.219***	-0.219***	-0.150***	-0.150***
	(0.022)	(0.022)	(0.022)	(0.022)	(0.022)	(0.022)	(0.046)	(0.046)	(0.045)	(0.045)	(0.044)	(0.044)	(0.022)	(0.022)
R^2	0.023	0.023	0.023	0.023	0.023	0.023	0.014	0.014	0.037	0.005	0.048	0.048	0.023	0.023
Panel B. Gender identity														
GI	-0.071***	-0.077***	-0.064***	0.089***	-0.054***	0.086***	0.020	0.098***	-0.027	0.065***	0.005	0.117***	-0.074***	-0.065***
	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.023)	(0.022)	(0.023)	(0.022)	(0.023)	(0.022)	(0.011)	(0.011)
R^2	0.021	0.024	0.016	0.032	0.012	0.030	0.002	0.039	0.003	0.017	0.000	0.055	0.022	0.017
Panel C. Binary gender and gender identity														
Female	-0.101*	-0.075	-0.137***	-0.007	-0.174***	-0.041	-0.211***	-0.036	-0.194***	-0.019	-0.228***	-0.180***	-0.090*	-0.114***
	(0.041)	(0.040)	(0.039)	(0.041)	(0.036)	(0.036)	(0.054)	(0.051)	(0.047)	(0.049)	(0.045)	(0.045)	(0.040)	(0.031)
GI	-0.029	-0.046*	-0.008	0.086***	0.015	0.070***	0.081	0.090***	0.007	0.061**	0.025	0.099***	-0.036*	-0.027*
	(0.020)	(0.020)	(0.019)	(0.020)	(0.018)	(0.018)	(0.027)	(0.025)	(0.024)	(0.024)	(0.023)	(0.023)	(0.020)	(0.015)
R^2	0.024	0.026	0.023	0.032	0.023	0.031	0.032	0.040	0.037	0.018	0.051	0.086	0.025	0.024
Panel D. Sample split by binary gender														
Female: GI	-0.005	-0.019	0.000	0.050***	0.013	0.043***	0.038	0.080**	-0.025	0.103***	0.009	0.133***	-0.013	-0.010
	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)	(0.033)	(0.032)	(0.034)	(0.029)	(0.032)	(0.031)	(0.016)	(0.016)
R^2	0.000	0.001	0.000	0.010	0.001	0.007	0.006	0.026	0.002	0.043	0.000	0.072	0.001	0.000
Male: GI	-0.028*	-0.034*	-0.009	0.045***	0.006	0.047***	0.089	0.081**	0.040	0.021	0.040	0.059*	-0.029*	-0.030*
	(0.016)	(0.016)	(0.016)	(0.016)	(0.015)	(0.016)	(0.031)	(0.032)	(0.031)	(0.031)	(0.032)	(0.032)	(0.016)	(0.016)
R^2	0.003	0.005	0.000	0.009	0.000	0.009	0.033	0.027	0.007	0.002	0.007	0.015	0.004	0.004
Observations	1,908	1,908	1,908	1,908	1,908	1,908	472	472	475	475	486	486	1,908	1,908
Mean of Dependent Variable	0.564	0.564	0.564	0.564	0.564	0.564	0.542	0.542	0.592	0.592	0.533	0.533	0.564	0.564

Notes: This table presents results from regressing wage negotiation on all 14 measures of gender identity collected in this sample. U.S. Adults Wave 2 sample (age 30–60) is used. The estimates in each column and panel come from a separate regression. Wage negotiation is a dummy that is equal to 1 if the respondent has ever engaged in wage negotiations. N differs across columns since each participant responded to only one of the four lengthy gender identity inventories (OSRI, CFNI, CMNI, BSRI). See Subsection 2.1 for a detailed description of the alternate gender identity measures. CGI other and CGI own indicate how respondents think that others see them and how they see themselves compared to others with the same binary gender on a scale from 0-very masculine to 10-very feminine. For CGI, TMF, Maglioz fem, SOM fem, CFNI, BSRI fem, CGI other, CGI own higher values indicate greater femininity. For these measures significance levels in Panels B, C and D are determined by a one-tailed test in the direction of the coefficient estimate of female in Panel A. For Maglioz mas, SOM mas, OSRI mas, CMNI, BSRI mas higher values indicate greater masculinity. For these measures, significance levels in Panels B, C and D are determined by a one-tailed test in the opposite direction of the coefficient estimate of female in Panel A. Significance levels in Panel A are determined by a two-tailed test for all measures. All gender identity measures are standardized (mean=0, SD=1) within the relevant sample. .** p<0.005, ** p<0.01, * p<0.05.

Table A28: CGI and U.S. Adults Preference Measures with and without Non-binary Participants

	Incentivized Risk		Unincentivized Risk		Financial Risk		Competitiveness		Redistribution		Alturism	
	exclude nonbinary (1)	include nonbinary (2)	exclude nonbinary (3)	include nonbinary (4)	exclude nonbinary (5)	include nonbinary (6)	exclude nonbinary (7)	include nonbinary (8)	exclude nonbinary (9)	include nonbinary (10)	exclude nonbinary (11)	include nonbinary (12)
Panel A. Binary gender												
Female	-0.220***	-0.215***	-0.434***	-0.431***	-0.277***	-0.280***	-0.461***	-0.464***	0.091*	0.106*	0.007	0.002
	(0.046)	(0.045)	(0.045)	(0.044)	(0.021)	(0.021)	(0.045)	(0.044)	(0.046)	(0.045)	(0.046)	(0.045)
R^2	0.012	0.012	0.047	0.046	0.082	0.083	0.053	0.054	0.002	0.003	0.000	0.000
Panel B. Gender identity												
CGI (feminine)	-0.095***	-0.097***	-0.245***	-0.243***	-0.121***	-0.120***	-0.245***	-0.243***	0.051*	0.051*		
	(0.023)	(0.023)	(0.023)	(0.023)	(0.011)	(0.011)	(0.022)	(0.022)	(0.024)	(0.024)		
R^2	0.009	0.009	0.060	0.059	0.062	0.061	0.060	0.059	0.003	0.003		
Panel C. Binary gender and gender identity												
Female	-0.200*	-0.173*	-0.091	-0.096	-0.244***	-0.252***	-0.177*	-0.203*	0.022	0.067		
	(0.081)	(0.078)	(0.082)	(0.079)	(0.038)	(0.037)	(0.083)	(0.081)	(0.079)	(0.077)		
CGI (feminine)	-0.012	-0.026	-0.207***	-0.204***	-0.020	-0.017	-0.171***	-0.159***	0.041	0.024		
	(0.041)	(0.040)	(0.042)	(0.041)	(0.019)	(0.019)	(0.042)	(0.040)	(0.042)	(0.041)		
R^2	0.012	0.012	0.060	0.060	0.082	0.084	0.062	0.062	0.003	0.003		
Panel D. Sample split by binary gender												
Female: CGI (feminine)	-0.006	-0.010	-0.019	-0.030	0.013	0.015	0.065	0.077	-0.115	-0.144		
	(0.033)	(0.033)	(0.034)	(0.034)	(0.014)	(0.014)	(0.035)	(0.034)	(0.031)	(0.031)		
R^2	0.000	0.000	0.000	0.001	0.001	0.001	0.004	0.006	0.013	0.021		
Male: CGI (feminine)	-0.007	-0.020	-0.215***	-0.207***	-0.035*	-0.034*	-0.273***	-0.278***	0.160***	0.170***		
	(0.032)	(0.032)	(0.035)	(0.035)	(0.016)	(0.016)	(0.034)	(0.033)	(0.034)	(0.034)		
R^2	0.000	0.000	0.046	0.043	0.005	0.005	0.075	0.077	0.025	0.029		
Observations	1,908	1,936	1,908	1,936	1,908	1,936	1,908	1,936	1,908	1,936	1,908	1,936
Mean of Dependent Variable	0	0	0	0	0.381	0.379	0	0	0	0	0	0

Notes: The table presents the results from regressing incentivized and unincentivized preference measures on our standardized CGI measure (mean=0, SD=1). CGI is standardized (mean=0, SD=1) separately for the male and female samples in Panel D. U.S. Adults (age 30-60) sample is used. In the first column for each outcome measure, we exclude individuals identifying as non-binary and in the second column, we include individuals identifying as non-binary. Binary gender is measured with the question "What is your sex?" with answer options "female" and "male" which every participant had to answer. Non-binary gender is measured with the question"What gender are you currently?" with answer options "Man (including Tans Male/Trans Man)", "Woman (including Trans Female/Trans Woman)", "Non-binary" and "Would rather not say". The estimates in each column and panel come from a separate regression. Incentivized risk preference is a measure of participants' investment decisions. Unincentivized risk is a self-reported measure of risk-taking. Financial risk is a dummy that takes the value 1 for those who report to actively trade in securities. Competitiveness is a self-reported measure of competitiveness. A higher value means higher risk taking or competitiveness. All measures except financial risk are standardized (mean=0, SD=1). Robust standard errors are reported in parentheses. Significance levels in Panel A are determined by a two-tailed test. Significance levels in Panels B, C and D are determined by a one-tailed test in the direction of the coefficient estimate of female in Panel A. *** p<0.005, ** p<0.01, * p<0.05.

Table A29: CGI and U.S. Adults Educational and Labor Market Outcomes with and without Non-binary Participants

	Female Educational Track Share		Income		Full-time Homemaker		Weekly Average Work Hours		Managerial Responsibilities		Wage Negotiation	
	exclude nonbinary (1)	include nonbinary (2)	exclude nonbinary (3)	include nonbinary (4)	exclude nonbinary (5)	include nonbinary (6)	exclude nonbinary (7)	include nonbinary (8)	exclude nonbinary (9)	include nonbinary (10)	exclude nonbinary (11)	include nonbinary (12)
Panel A. Binary gender												
Female	0.139***	0.139***	-25.045***	-24.971***	0.096***	0.094***	-5.632***	-5.468***	-0.114***	-0.114***	-0.150***	-0.150***
	(0.011)	(0.011)	(2.557)	(2.535)	(0.011)	(0.011)	(0.634)	(0.631)	(0.022)	(0.022)	(0.022)	(0.022)
R^2	0.129	0.128	0.048	0.048	0.040	0.039	0.040	0.037	0.014	0.014	0.023	0.023
Panel B. Gender identity												
CGI (feminine)	0.071***	0.071***	-11.579***	-11.552***	0.048***	0.048***	-2.639***	-2.646***	-0.063***	-0.062***	-0.071***	-0.072***
	(0.006)	(0.006)	(1.360)	(1.349)	(0.006)	(0.006)	(0.313)	(0.310)	(0.011)	(0.011)	(0.011)	(0.011)
R^2	0.133	0.131	0.041	0.041	0.040	0.039	0.035	0.035	0.016	0.016	0.021	0.021
Panel C. Binary gender and gender identity												
Female	0.071***	0.072***	-18.710***	-18.398***	0.053**	0.050**	-4.023***	-3.446***	-0.031	-0.040	-0.101*	-0.101*
	(0.022)	(0.022)	(4.134)	(3.970)	(0.019)	(0.018)	(1.164)	(1.136)	(0.040)	(0.038)	(0.041)	(0.039)
CGI (feminine)	0.042***	0.041***	-3.822*	-4.011*	0.026***	0.027***	-0.970*	-1.234*	-0.050**	-0.045**	-0.029	-0.030
	(0.011)	(0.011)	(2.224)	(2.141)	(0.010)	(0.009)	(0.573)	(0.559)	(0.020)	(0.019)	(0.020)	(0.020)
R^2	0.143	0.142	0.049	0.049	0.043	0.043	0.041	0.040	0.017	0.017	0.024	0.024
Panel D. Sample split by binary gender												
Female: CGI (feminine)	0.018*	0.017*	3.101	3.452	0.019*	0.022*	0.191	0.167	0.014	0.013	-0.005	-0.003
	(0.009)	(0.009)	(1.604)	(1.581)	(0.010)	(0.010)	(0.476)	(0.470)	(0.016)	(0.016)	(0.016)	(0.016)
R^2	0.010	0.009	0.004	0.005	0.004	0.005	0.000	0.000	0.001	0.001	0.000	0.000
Male: CGI (feminine)	0.028***	0.029***	-7.359***	-8.058***	0.010*	0.010*	-1.275***	-1.583***	-0.070***	-0.066***	-0.028*	-0.031*
	(0.009)	(0.009)	(1.896)	(1.860)	(0.005)	(0.004)	(0.434)	(0.437)	(0.016)	(0.016)	(0.016)	(0.016)
R^2	0.022	0.023	0.014	0.016	0.007	0.006	0.009	0.014	0.019	0.017	0.003	0.004
Observations	1,054	1,064	1,908	1,936	1,908	1,936	1,908	1,936	1,908	1,936	1,908	1,936
Mean of Dependent Variable	0.532	0.533	59.703	59.278	0.062	0.061	25.195	25.101	0.395	0.394	0.564	0.562

Notes: The table presents the results from regressing educational and labor market outcomes on our standardized CGI measure (mean=0, SD=1). CGI is standardized (mean=0, SD=1) separately for the male and female samples in Panel D. Adults (age 30-60) sample is used. In the first column for each outcome measure, we exclude individuals identifying as non-binary and in the second column, we include individuals identifying as non-binary. U.S. Binary gender is measured with the question "What is your sex?" with answer options "female" and "male" which every participant had to answer. Non-binary gender is measured with the question "What gender are you currently?" with answer options "Man (including Tans Male/Trans Man)", "Woman (including Trans Female/Trans Woman)", "Non-binary" and "Would rather not say". The estimates in each column and panel come from a separate regression. Female educational track share is the share of women graduating with a bachelor's degree in the chosen field of study in 2020. Income is a self-reported categorical measure in thousands of U.S. dollars. Full-time homemaker is a dummy for working full-time at home. Weekly average work hours is a proxy measure constructed from weeks worked and hours worked, using the product of the two categorical measures and dividing by 52. Managerial responsibilities is a dummy for having managerial responsibilities at work. Wage negotiation is a dummy that is equal to 1 if the respondent ever negotiated wage. Robust standard errors are reported in parentheses. Significance levels in Panel A are determined by a two-tailed test. Significance levels in Panels B, C and D are determined by a one-tailed test in the direction of the coefficient estimate of female in Panel A. *** p<0.005, ** p<0.01, * p<0.05.

Table A30: CGI and Swiss Uni Incentivized Behavioral Measures with and without Non-binary Participants

	Risk		Competitiveness		Efficiency		Overconfidence	
	exclude nonbinary (1)	include nonbinary (2)	exclude nonbinary (3)	include nonbinary (4)	exclude nonbinary (5)	include nonbinary (6)	exclude nonbinary (7)	include nonbinary (8)
Panel A. Binary gender								
Female	-0.378***	-0.380***	-0.128***	-0.130***	-0.145***	-0.144***	-0.002	0.005
	(0.081)	(0.081)	(0.041)	(0.040)	(0.023)	(0.023)	(0.083)	(0.083)
R^2	0.036	0.036	0.017	0.017	0.062	0.061	0.000	0.000
Panel B. Gender identity								
CGI (feminine)	-0.169***	-0.168***	-0.056***	-0.056***	-0.066***	-0.067***		
	(0.040)	(0.039)	(0.020)	(0.020)	(0.012)	(0.011)		
R^2	0.028	0.028	0.013	0.013	0.052	0.053		
Panel C. Binary gender and gender identity								
Female	-0.305*	-0.317*	-0.110	-0.115	-0.110*	-0.103*		
	(0.153)	(0.151)	(0.072)	(0.071)	(0.043)	(0.043)		
CGI (feminine)	-0.045	-0.039	-0.011	-0.009	-0.022	-0.025		
	(0.074)	(0.073)	(0.036)	(0.035)	(0.021)	(0.021)		
R^2	0.036	0.037	0.017	0.018	0.064	0.063		
Panel D. Sample split by binary gender								
Female: CGI (feminine)	-0.039	-0.043	-0.029	-0.031	-0.042**	-0.044***		
	(0.060)	(0.060)	(0.030)	(0.029)	(0.017)	(0.017)		
R^2	0.002	0.002	0.004	0.004	0.023	0.025		
Male: CGI (feminine)	-0.017	-0.008	0.013	0.017	0.013	0.011		
	(0.063)	(0.062)	(0.030)	(0.029)	(0.017)	(0.017)		
R^2	0.000	0.000	0.001	0.001	0.002	0.002		
Observations	584	590	584	590	584	590	584	590
Mean of Dependent Variable	0	0	0.413	0.414	0.538	0.537	0	0

Notes: The table presents results from regressing incentivized behavioral measures on our standardized CGI measure (mean=0, SD=1). CGI is standardized (mean=0, SD=1) separately for the male and female samples in Panel D. Swiss Uni sample is used. In the first column for each outcome measure, we exclude individuals identifying as non-binary and in the second column, we include individuals identifying as non-binary. Binary gender is measured with the question "Are you male or female?" with answer options "female" and "male" which every participant had to answer. Non-binary gender is measured with the question "What is your current gender?" with answer options "Woman", "Man", "Transgender","Non-binary", "Prefer not to answer" and "Other". The estimates in each column and panel come from a separate regression. Risk preference is a standardized measure of participants' investment decisions (mean=0, SD=1). Competitiveness is a dummy that takes the value 1 for those who chose to compete in the competitive task. Our measure of preferences for equality versus efficiency is measured in deciles, with increasing numbers indicating higher priority for efficiency. Overconfidence is measured as relative overplacement and is standardized (mean=0, SD=1). Robust standard errors are reported in parentheses. Significance levels in Panel A are determined by a two-tailed test. Significance levels in Panels B, C and D are determined by a one-tailed test in the direction of the coefficient estimate of female in Panel A. .*** p<0.005, ** p<0.01, * p<0.05.

Table A31: CGI and Swiss Uni Unincentivized Behavioral Measures with and without Non-binary Participants

	Staircase Risk		Risk		Competitiveness		Redistribution		Alturism	
	exclude nonbinary (1)	include nonbinary (2)	exclude nonbinary (3)	include nonbinary (4)	exclude nonbinary (5)	include nonbinary (6)	exclude nonbinary (7)	include nonbinary (8)	exclude nonbinary (9)	include nonbinary (10)
Panel A. Binary gender										
Female	-0.211*	-0.217**	-0.444***	-0.441***	-0.388***	-0.390***	0.201*	0.196*	0.249***	0.257***
	(0.083)	(0.082)	(0.081)	(0.081)	(0.081)	(0.081)	(0.083)	(0.082)	(0.082)	(0.082)
R^2	0.011	0.012	0.049	0.049	0.038	0.038	0.010	0.010	0.016	0.017
Panel B. Gender identity										
CGI (feminine)	-0.137***	-0.140***	-0.234***	-0.233***	-0.194***	-0.195***	0.106***	0.101**	0.095*	0.099**
	(0.044)	(0.043)	(0.040)	(0.039)	(0.040)	(0.040)	(0.041)	(0.040)	(0.041)	(0.041)
R^2	0.019	0.020	0.055	0.054	0.038	0.038	0.011	0.010	0.009	0.010
Panel C. Binary gender and gender identity										
Female	0.031	0.028	-0.188	-0.187	-0.212	-0.214	0.082	0.092	0.279	0.283
	(0.148)	(0.146)	(0.148)	(0.147)	(0.138)	(0.137)	(0.155)	(0.154)	(0.166)	(0.164)
CGI (feminine)	-0.149*	-0.151*	-0.157*	-0.156*	-0.108	-0.108	0.073	0.064	-0.018	-0.016
	(0.078)	(0.077)	(0.073)	(0.072)	(0.068)	(0.068)	(0.076)	(0.076)	(0.084)	(0.083)
R^2	0.019	0.020	0.058	0.057	0.042	0.042	0.012	0.011	0.016	0.017
Panel D. Sample split by binary gender										
Female: CGI (feminine)	-0.050	-0.056	-0.081	-0.092	-0.015	-0.007	-0.136	-0.130	-0.047	-0.039
	(0.064)	(0.064)	(0.057)	(0.057)	(0.054)	(0.054)	(0.063)	(0.062)	(0.071)	(0.070)
R^2	0.003	0.003	0.007	0.009	0.000	0.000	0.019	0.017	0.002	0.002
Male: CGI (feminine)	-0.123*	-0.121*	-0.105	-0.095	-0.110*	-0.118*	0.193***	0.177***	0.020	0.015
	(0.068)	(0.067)	(0.064)	(0.064)	(0.062)	(0.062)	(0.062)	(0.062)	(0.068)	(0.068)
R^2	0.015	0.015	0.011	0.009	0.012	0.014	0.037	0.031	0.000	0.000
Observations	584	590	584	590	584	590	584	590	584	590
Mean of Dependent Variable	0	0	0	0	0	0	0	0	0	0

Notes: The table presents results from regressing unincentivized behavioral measures on our standardized CGI measure (mean=0, SD=1). CGI is standardized (mean=0, SD=1) separately for the male and female samples in Panel D. Swiss Uni sample is used. In the first column for each outcome measure, we exclude individuals identifying as non-binary and in the second column, we include individuals identifying as non-binary. Binary gender is measured with the question "Are you male or female?" with answer options "female" and "male" which every participant had to answer. Non-binary gender is measured with the question "What is your current gender?" with answer options "Woman", "Man", "Transgender","Non-binary", "Prefer not to answer" and "Other". The estimates in each column and panel come from a separate regression. Staircase risk is a categorical certainty equivalence measure of risk-taking based on a series of hypothetical allocation decisions. Risk is a self-reported measure of risk-taking. Competitiveness is a self-reported measure of competitiveness. A higher value means higher risk taking or competitiveness. Redistribution is a measure of how much economic redistribution one wants in society. Altruism is a measure of how much one would donate out of a windfall gain. A higher value means a greater willingness to redistribute or donate. All measures are standardized (mean=0, SD=1). Robust standard errors are reported in parentheses. Significance levels in Panel A are determined by a two-tailed test. Significance levels in Panels B, C and D are determined by a one-tailed test in the direction of the coefficient estimate of female in Panel A. *** p<0.005, ** p<0.01, * p<0.05.

Table A32: CGI and Incentivized Behavioral Measures (with Binary Gender and CGI Interaction)

Measure	Risk		Competitiveness	Efficiency		Overconfidence
Sample	Swiss Uni	US Adults	Swiss Uni	Swiss Uni	Swedish Youths	Swiss Uni
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. Binary gender						
Female	-0.378***	-0.215***	-0.128***	-0.145***	-0.161***	-0.002
	(0.081)	(0.032)	(0.041)	(0.023)	(0.017)	(0.083)
R^2	0.036	0.012	0.017	0.062	0.078	0.000
Panel B. Gender identity						
CGI (feminine)	-0.169***	-0.092***	-0.056***	-0.066***	-0.068***	
	(0.040)	(0.016)	(0.020)	(0.012)	(0.009)	
R^2	0.028	0.008	0.013	0.052	0.055	
Panel C. Binary gender and gender identity						
Female	-0.305*	-0.195***	-0.110	-0.110**	-0.135***	
	(0.153)	(0.057)	(0.072)	(0.043)	(0.026)	
CGI (feminine)	-0.045	-0.012	-0.011	-0.022	-0.017	
	(0.074)	(0.029)	(0.036)	(0.021)	(0.013)	
R^2	0.036	0.012	0.017	0.064	0.079	
Panel D. Sample split by binary gender						
Female: CGI (feminine)	-0.039	-0.001	-0.029	-0.042**	-0.014	
	(0.060)	(0.023)	(0.030)	(0.017)	(0.011)	
R^2	0.002	0.000	0.004	0.023	0.003	
Male: CGI (feminine)	-0.017	-0.012	0.013	0.013	-0.010	
	(0.063)	(0.023)	(0.030)	(0.017)	(0.014)	
R^2	0.000	0.000	0.001	0.002	0.001	
Panel E. Binary gender and CGI interaction						
Female	-0.301*	-0.196***	-0.101	-0.098*	-0.134***	
	(0.152)	(0.057)	(0.072)	(0.043)	(0.026)	
CGI (feminine)	-0.029	-0.022	0.021	0.021	-0.013	
	(0.106)	(0.041)	(0.048)	(0.027)	(0.018)	
Female*CGI	-0.037	0.021	-0.075	-0.098*	-0.010	
	(0.146)	(0.057)	(0.073)	(0.042)	(0.027)	
R^2	0.036	0.012	0.017	0.064	0.079	
Observations	584	3,902	584	584	1,041	584
Mean of Dependent Variable	0	0	0.413	0.538	0.550	0

Notes: The table presents results from regressing the incentivized measures of risk, competitiveness, equality versus efficiency and overconfidence on our standardized (mean=0, SD=1) measure of CGI. Swiss Uni, U.S. Adults (age 30-60), and Swedish Youths samples are used. The estimates in each column and panel come from a separate regression. CGI is standardized (mean=0, SD=1) separately for the male and female samples in Panel D. Risk preference is a standardized measure of participants' investment decisions (mean=0, SD=1). Competitiveness is a dummy that takes the value 1 for those who chose to compete in the competitive task. Our measure of preferences for equality versus efficiency is measured in deciles, with increasing numbers indicating higher priority for efficiency. Overconfidence is measured as relative overplacement and standardized (mean=0, SD=1). Robust standard errors are reported in parentheses. Significance levels in Panel A are determined by a two-tailed test. Significance levels in Panels B, C and D are determined by a one-tailed test in the direction of the coefficient estimate of female in Panel A. For Panel E, significance levels are determined using a one-tailed test in the direction of the coefficient estimate of female in Panel A, except for the interaction term, which is determined using a two-tailed test. *** p<0.005, ** p<0.01, * p<0.05.

Table A33: CGI and Unincentivized Risk and Competitiveness Measures (with Binary Gender and CGI Interaction)

Measure	Staircase Risk	Risk				Financial Risk	Competitiveness			
Sample	Swiss Uni	Swiss Uni	U.S Adults	Swiss Teens	Swedish Youths	US Adults	Swiss Uni	U.S Adults	Swiss Teens	Swedish Youths
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A. Binary gender										
Female	-0.211*	-0.444***	-0.387***	-0.285***	-0.243***	-0.268***	-0.388***	-0.405***	-0.311***	-0.153*
	(0.083)	(0.081)	(0.031)	(0.071)	(0.062)	(0.015)	(0.081)	(0.031)	(0.071)	(0.062)
R^2	0.011	0.049	0.038	0.020	0.015	0.075	0.038	0.041	0.024	0.006
Panel B. Gender identity										
CGI (feminine)	-0.137***	-0.234***	-0.206***	-0.171***	-0.109***	-0.119***	-0.194***	-0.218***	-0.181***	-0.109***
	(0.044)	(0.040)	(0.016)	(0.036)	(0.031)	(0.008)	(0.040)	(0.016)	(0.036)	(0.029)
R^2	0.019	0.055	0.042	0.029	0.012	0.059	0.038	0.048	0.033	0.012
Panel C. Binary gender and gender identity										
Female	0.031	-0.188	-0.152***	0.124	-0.182*	-0.224***	-0.212	-0.145**	0.087	0.022
	(0.148)	(0.148)	(0.056)	(0.166)	(0.103)	(0.027)	(0.138)	(0.057)	(0.168)	(0.101)
CGI (feminine)	-0.149*	-0.157*	-0.144***	-0.227***	-0.041	-0.027*	-0.108	-0.159***	-0.220***	-0.117**
	(0.078)	(0.073)	(0.029)	(0.085)	(0.051)	(0.013)	(0.068)	(0.029)	(0.085)	(0.048)
R^2	0.019	0.058	0.044	0.030	0.015	0.076	0.042	0.049	0.033	0.012
Panel D. Sample split by binary gender										
Female: CGI (feminine)	-0.050	-0.081	-0.018	-0.051	0.058	-0.001	-0.015	0.068	0.021	0.015
	(0.064)	(0.057)	(0.025)	(0.060)	(0.046)	(0.010)	(0.054)	(0.024)	(0.055)	(0.044)
R^2	0.003	0.007	0.000	0.003	0.003	0.000	0.000	0.005	0.000	0.000
Male: CGI (feminine)	-0.123*	-0.105	-0.143***	-0.154***	-0.097*	-0.029***	-0.110*	-0.253***	-0.238***	-0.163***
	(0.068)	(0.064)	(0.024)	(0.045)	(0.052)	(0.011)	(0.062)	(0.024)	(0.047)	(0.050)
R^2	0.015	0.011	0.020	0.024	0.009	0.003	0.012	0.064	0.056	0.027
Panel E. Binary gender and CGI interaction										
Female	0.020	-0.191	-0.163***	0.156	-0.209*	-0.227***	-0.229	-0.171***	0.150	-0.007
	(0.152)	(0.147)	(0.057)	(0.165)	(0.105)	(0.027)	(0.140)	(0.057)	(0.165)	(0.104)
CGI (feminine)	-0.190*	-0.167	-0.245***	-0.411***	-0.131*	-0.049***	-0.169*	-0.410***	-0.605***	-0.210***
	(0.105)	(0.103)	(0.041)	(0.120)	(0.071)	(0.019)	(0.095)	(0.039)	(0.120)	(0.064)
Female*CGI	0.095	0.024	0.213***	0.311	0.228*	0.048	0.142	0.533***	0.648***	0.236*
	(0.161)	(0.145)	(0.059)	(0.168)	(0.104)	(0.027)	(0.138)	(0.058)	(0.167)	(0.100)
R^2	0.019	0.058	0.044	0.030	0.015	0.076	0.042	0.049	0.033	0.012
Observations	584	584	3,902	786	1,041	3,902	584	3,902	792	1,041
Mean of Dependent Variable	0	0	0	0	0	0.401	0	0	0	0

Notes: The table presents results from regressing unincentivized measures of risk attitudes and competitiveness on our standardized (mean=0, SD=1) measure of CGI. CGI is standardized (mean=0, SD=1) separately for the male and female samples in Panel D. Swiss Uni, U.S. Adults (age 30-60), Swiss Teens, and Swedish Youths samples are used. The estimates in each column and panel come from a separate regression. Staircase risk is a categorical certainty equivalence measure of risk-taking based on a series of hypothetical allocation decisions. Risk is a self-reported measure of risk-taking. Financial risk is a dummy that takes the value 1 for those who report to actively trade in securities. Competitiveness is a self-reported measure of competitiveness. A higher value means higher risk taking or competitiveness. All measures except financial market risk are standardized (mean 0, SD 1). Robust standard errors are reported in parentheses. Significance levels in Panel A are determined by a two-tailed test. Significance levels in Panels B, C and D are determined by a one-tailed test in the direction of the coefficient estimate of female in Panel A. For Panel E, significance levels are determined using a one-tailed test in the direction of the coefficient estimate of female in Panel A, except for the interaction term, which is determined using a two-tailed test. *** $p<0.005$, ** $p<0.01$, * $p<0.05$.

Table A34: CGI and Unincentivized Distributional Preferences (with Binary Gender and CGI Interaction)

Measure	Redistribution		Altruism			
Sample	Swiss Uni (1)	U.S. Adults (2)	Swiss Uni (3)	U.S. Adults (4)	Swiss Teens (5)	Swedish Youths (6)
Panel A. Binary gender						
Female	0.201*	0.063*	0.249***	0.001	0.210***	0.551***
	(0.083)	(0.032)	(0.082)	(0.032)	(0.071)	(0.060)
R^2	0.010	0.001	0.016	0.000	0.011	0.076
Panel B. Gender identity						
CGI (feminine)	0.106***	0.053***	0.095*		0.074*	0.224***
	(0.041)	(0.017)	(0.041)		(0.035)	(0.032)
R^2	0.011	0.003	0.009		0.005	0.050
Panel C. Binary gender and gender identity						
Female	0.082	-0.072	0.279*		0.412**	0.489***
	(0.155)	(0.055)	(0.166)		(0.165)	(0.099)
CGI (feminine)	0.073	0.083***	-0.018		-0.112	0.042
	(0.076)	(0.029)	(0.084)		(0.082)	(0.052)
R^2	0.012	0.003	0.016		0.013	0.077
Panel D. Sample split by binary gender						
Female: CGI (feminine)	-0.136	-0.115	-0.047		-0.051	0.052
	(0.063)	(0.023)	(0.071)		(0.054)	(0.046)
R^2	0.019	0.013	0.002		0.013	0.003
Male: CGI (feminine)	0.193***	0.196***	0.020		-0.046	0.013
	(0.062)	(0.024)	(0.068)		(0.046)	(0.053)
R^2	0.037	0.038	0.000		0.002	0.000
Panel E. Binary gender and CGI interaction						
Female	0.148	-0.045	0.292*		0.414**	0.481***
	(0.155)	(0.055)	(0.167)		(0.164)	(0.097)
CGI (feminine)	0.315***	0.337***	0.033		-0.122	0.017
	(0.101)	(0.041)	(0.112)		(0.121)	(0.073)
Female*CGI	-0.562***	-0.538***	-0.117		0.017	0.063
	(0.152)	(0.058)	(0.169)		(0.165)	(0.102)
R^2	0.012	0.003	0.016		0.013	0.077
Observations	584	3,902	584	3,902	798	1,041
Mean of Dependent Variable	0	0	0	0	0	0

Notes: The table presents results from regressing unincentivized distributional measures on our standardized (mean=0, SD=1) measure of CGI. CGI is standardized (mean=0, SD=1) separately for the male and female samples in Panel D. Swiss Uni, U.S. Adults (age 30-60), Swiss Teens and Swedish Youth samples are used. The estimates in each column and panel come from a separate regression. The table presents results from regressing self-reported preferences on redistribution and altruism on our standardized measure of CGI. Redistribution is a measure of how much economic redistribution one wants in society. Altruism is a measure of how much one would donate out of a windfall gain. A higher value means the greater willingness to redistribute or donate. All outcome measures are standardized (mean=0, SD=1). Robust standard errors are reported in parentheses. Significance levels in Panel A are determined by a two-tailed test. Significance levels in Panels B, C and D are determined by a one-tailed test in the direction of the coefficient estimate of female in Panel A. For Panel E, significance levels are determined using a one-tailed test in the direction of the coefficient estimate of female in Panel A, except for the interaction term, which is determined using a two-tailed test. *** $p<0.005$, ** $p<0.01$, * $p<0.05$.

Table A35: CGI and Educational Choice (with Binary Gender and CGI Interaction)

Measure	Female Educational Track Share			Skill Requirements	
				Math	Language
Sample	U.S Adults	Swiss Teens	Swedish Youths	Swiss Teens	Swiss Teens
	(1)	(2)	(3)	(4)	(5)
Panel A. Binary gender					
Female	0.125***	0.383***	0.093***	-0.917***	0.913***
	(0.007)	(0.012)	(0.009)	(0.047)	(0.047)
R^2	0.108	0.404	0.090	0.204	0.203
Panel B. Gender identity					
CGI (feminine)	0.061***	0.167***	0.039***	-0.403***	0.404***
	(0.004)	(0.006)	(0.005)	(0.024)	(0.023)
R^2	0.104	0.315	0.063	0.163	0.163
Panel C. Binary gender and gender identity					
Female	0.075***	0.329***	0.079***	-0.762***	0.745***
	(0.014)	(0.023)	(0.014)	(0.084)	(0.083)
CGI (feminine)	0.031***	0.033***	0.009	-0.092*	0.100**
	(0.007)	(0.011)	(0.007)	(0.042)	(0.040)
R^2	0.117	0.408	0.092	0.207	0.206
Panel D. Sample split by binary gender					
Female: CGI (feminine)	0.009*	0.028***	0.007	-0.028	0.135***
	(0.005)	(0.009)	(0.006)	(0.040)	(0.038)
R^2	0.003	0.014	0.002	0.001	0.018
Male: CGI (feminine)	0.024***	0.009	0.005	-0.087**	0.005
	(0.006)	(0.009)	(0.007)	(0.035)	(0.035)
R^2	0.016	0.002	0.001	0.008	0.000
Panel E. Binary gender and CGI interaction					
Female	0.078***	0.328***	0.079***	-0.763***	0.744***
	(0.014)	(0.023)	(0.014)	(0.082)	(0.084)
CGI (feminine)	0.041***	0.020	0.007	-0.171**	0.010
	(0.010)	(0.018)	(0.009)	(0.067)	(0.070)
Female*CGI	-0.024	0.022	0.006	0.136	0.156
	(0.014)	(0.023)	(0.014)	(0.084)	(0.084)
R^2	0.117	0.408	0.092	0.207	0.206
Observations	2,279	1,448	1,041	1,433	1,433
Mean of Dependent Variable	0.541	0.442	0.521	0	0

Notes: The table presents results from regressing the educational track measures on our standardized (mean=0, SD=1) measure of CGI. CGI is standardized (mean=0, SD=1) separately for the male and female samples in Panel D. U.S. Adults (age 30–60), Swiss Teens and Swedish Youths samples are used. The estimates in each column and panel come from a separate regression. Female educational track share is the share of women graduating with a bachelor's degree in the chosen field of study in 2020 for U.S. Adults; the share of women from past cohorts of graduates from the apprenticeship for Swiss Teens; the share of women accepted for undergraduate studies in that field the year before our sample made their educational choices for Swedish Youths.. Skills requirements are a standardized (mean=0, SD=1) measure based on expert evaluation of the job content in occupations chosen by Swiss Teens to start apprenticeships. Robust standard errors are reported in parentheses. Significance levels in Panel A are determined by a two-tailed test. Significance levels in Panels B, C and D are determined by a one-tailed test in the direction of the coefficient estimate of female in Panel A. For Panel E, significance levels are determined using a one-tailed test in the direction of the coefficient estimate of female in Panel A, except for the interaction term, which is determined using a two-tailed test. *** p<0.005, ** p<0.01, * p<0.05.

Table A36: CGI and Labor Market Outcomes (with Binary Gender and CGI Interaction)

Measure	Income	Full-time Homemaker	Weekly Ave. Work Hours	Female Industry Share	Managerial Responsibilities	Performance Pay	Wage Negotiation	Work Flexibility
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A. Binary gender								
Female	-23.265***	0.099***	-6.644***	0.062***	-0.125***	-0.082***	-0.132***	0.141***
	(1.865)	(0.008)	(0.451)	(0.007)	(0.016)	(0.021)	(0.016)	(0.045)
R^2	0.038	0.043	0.053	0.039	0.016	0.008	0.017	0.005
Panel B. Gender identity								
CGI (feminine)	-10.841***	0.049***	-3.141***	0.025***	-0.061***	-0.043***	-0.066***	0.060***
	(0.958)	(0.004)	(0.222)	(0.003)	(0.008)	(0.010)	(0.008)	(0.022)
R^2	0.033	0.042	0.047	0.026	0.015	0.009	0.017	0.004
Panel C. Binary gender and gender identity								
Female	-16.723***	0.057***	-4.553***	0.060***	-0.078***	-0.038	-0.074**	0.126
	(2.990)	(0.012)	(0.806)	(0.011)	(0.027)	(0.035)	(0.028)	(0.077)
CGI (feminine)	-3.998***	0.026***	-1.278***	0.001	-0.029*	-0.028	-0.035**	0.009
	(1.542)	(0.006)	(0.397)	(0.006)	(0.014)	(0.017)	(0.014)	(0.039)
R^2	0.040	0.047	0.055	0.039	0.018	0.009	0.019	0.005
Panel D. Sample split by binary gender								
Female: CGI (feminine)	2.416	0.025***	-0.155	-0.004	0.027	-0.012	0.002	-0.079
	(1.124)	(0.007)	(0.336)	(0.005)	(0.011)	(0.014)	(0.011)	(0.033)
R^2	0.002	0.006	0.000	0.001	0.003	0.001	0.000	0.006
Male: CGI (feminine)	-6.735***	0.005*	-1.281***	0.005	-0.058***	-0.020	-0.041***	0.085***
	(1.356)	(0.002)	(0.312)	(0.005)	(0.011)	(0.015)	(0.011)	(0.032)
R^2	0.011	0.002	0.009	0.001	0.014	0.002	0.007	0.007
Panel E. Binary gender and CGI Interaction								
Female	-17.508***	0.055***	-4.647***	0.061***	-0.085***	-0.038	-0.078***	0.150*
	(2.960)	(0.012)	(0.809)	(0.011)	(0.027)	(0.035)	(0.028)	(0.077)
CGI (feminine)	-11.425***	0.009*	-2.173***	0.009	-0.098***	-0.032	-0.069***	0.134***
	(2.300)	(0.004)	(0.529)	(0.008)	(0.019)	(0.024)	(0.019)	(0.049)
Female*CGI	15.740***	0.036**	1.896*	-0.017	0.146***	0.010	0.072**	-0.278***
	(3.053)	(0.013)	(0.800)	(0.012)	(0.027)	(0.035)	(0.028)	(0.078)
R^2	0.040	0.047	0.055	0.039	0.018	0.009	0.019	0.005
Observations	3,902	3,902	3,902	1,994	3,902	1,994	3,902	1,994
Mean of Dependent Variable	61.421	0.062	25.281	0.497	0.400	0.303	0.491	0

Notes: The table presents results from regressing the labor market outcomes on our standardized (mean=0, SD=1) measure of CGI. CGI is standardized (mean=0, SD=1) separately for the male and female samples in Panel D. The sample of the U.S. Adults (age 30-60) is used. The estimates in each column and panel come from a separate regression. Income is a self-reported categorical measure in thousand U.S. dollars. Full-time homemaker is a dummy for working full-time at home. Weekly average work hours is a proxy measure constructed from weeks worked and hours worked, using the product of the two categorical measures and dividing by 52. Female industry share is the share of female employees in a given industry. Managerial responsibilities is a dummy for having managerial responsibilities at work. Performance pay is a dummy that is equal to 1 if the respondent's current or most recent job has performance-related pay. Wage negotiation is a dummy that is equal to 1 if the respondent ever negotiated wage. Work flexibility is a categorical measure of flexibility in working hours where 0 is no flexibility, 0.5 is can adapt work hours and 1 is full flexibility, which is standardized (mean=0, SD=1). Robust standard errors are reported in parentheses. Significance levels in Panel A are determined by a two-tailed test. Significance levels in Panels B, C and D are determined by a one-tailed test in the direction of the coefficient estimate of female in Panel A. For Panel E, significance levels are determined using a one-tailed test in the direction of the coefficient estimate of female in Panel A, except for the interaction term, which is determined using a two-tailed test. *** $p<0.005$, ** $p<0.01$, * $p<0.05$.

B. Appendix B

B.1. Online experiment with Swiss students—*Swiss Uni*

B.1.1 *Participants and Procedures*

Our first data collection focused primarily on the explanatory power of CGI for gender gaps in economic preferences. This experiment was pre-registered (https://osf.io/phyt6/). Changes were implemented compared to the pre-registration mainly because we adapted and intended laboratory experiment to an online format due to Covid-19 restrictions. Most importantly, we increased the sample size based on an expectation of additional noise in the online setting, a few secondary measures were excluded for a shorter experiment more suitable to the online setting, and the competitive task was exchanged to prevent cheating, or any beliefs thereof. A detailed description of the departure from the initial pre-registration is available in the document "Updates to Pre-registration Final.pdf" available at https://osf.io/phyt6/.

The online experiment was conducted in September and October 2021. The experiment was implemented in English using o-Tree (Chen, Schonger and Wickens, 2016). The 597 participants were students recruited from a subject pool consisting mainly of students at the University of Zurich (UZH) and the Swiss Federal Institute of Technology (ETH), using the software h-root (Bock, Baetge and Nicklisch, 2014). Participants received a variable amount based on the outcome of their decisions in the incentivized tasks. During the experiment, payoffs were measured in Experimental Currency Units (ECU) with an exchange rate of 20 ECU to 1 CHF. A follow-up survey programmed with Qualtrics was conducted two weeks after the main experiment. Participation in the follow-up survey was strongly incentivized—participants received a fixed payment of 50 CHF only for successfully completing both the laboratory session and the online survey. Thus, the dropout rate was extremely small (less than 1%). The main purpose of the follow-up survey was to duplicate the measure of CGI using a slightly modified scale to account for possible measurement error, following Gillen, Snowberg and Yariv (2019). We considered this approach because self-reported gender may be susceptible to measurement error. In the experiment and follow-up survey, we collected several incentivized and un-incentivized preference measures. We list our main measures below.[30] All the

[30]Our main measure of CGI and the incentivized preference measures are relevant for our primary pre-registered analysis. For secondary analysis, we additionally elicited a two-dimensional measure (measuring masculinity and femininity on separate dimensions) in first and third person—how others see a respondent—following Magliozzi, Saperstein and Westbrook (2016). The questionnaire also included an unincentivized measure of overconfidence following Gillen, Snowberg and Yariv (2019), an unincentivized measure of willingness to engage in self-promotion following the design of Exley and Kessler (2022), questions about field of study, family and sibling structure, parental division of household work, parental education, occupation, a short version of the Big-5 personality inventory, and perceived gender identity (male, female, transgender, other), sex at birth, sexual orientation, and relationship status.

study material (experimental decision screens and the follow-up survey) are reproduced in Appendix C.

B.1.2 *Elicited Measures*

CGI. In the main experiment, CGI was elicited in a final survey after completion of the incentivized parts of the study. The question came right after demographic questions (age, binary gender, nationality and mother tongue) on a new screen. The precise question was "In general, how do you see yourself? Where would you put yourself on this scale from "0-Very masculine" to "10-Very feminine"?". On the same screen, we also asked how most people see a respondent on this scale ("In general, how do most people see you? Where would most people put you on this scale from "0-Very masculine" to "11-Very feminine"?"). In the follow-up survey, CGI was elicited as the first question right after initial instructions. It was asked in the following way "In general, how do you see yourself? Where would you put yourself on this scale from "0-Very feminine" to "11-Very masculine"?".

Incentivized preference measures. Our main outcome variables were elicited at the onset of the experiment and comprise incentivized preference measures of attitudes to risk, competitiveness, preference for equality versus efficiency, and overconfidence that are frequently used in the literature, and for which earlier studies have documented a gender gap.

Risk preferences were elicited through a one-shot investment task (Gneezy and Potters, 1997), in which participants allocate between 0 and 100 ECU to a risky investment. The investment has a 50% probability of success, in which case it returns 2.5 times the invested amount. If the investment fails, the invested ECU are lost. Our outcome variable for risk seeking is the amount invested. Based on earlier research, we expect women to invest less than men.

We elicited *Competitiveness* following the design introduced by Niederle and Vesterlund (2007), although with a different competitive task. The experimental task chosen for the online implementation is the matrix task used by Buser, Niederle and Oosterbeek (2021), in which participants are asked to identify the two numbers in a 3x3 matrix that sum to a target number. We used this task since participants could easily solve the arithmetic task from the original design with the aid of a calculator. Participants were incentivized to solve as many tasks as possible across three rounds of three minutes each. In Round 1, participants received a piece-rate payment of 10 ECU per correct exercise. In Round 2, participants were compensated under a tournament scheme—they were randomly assigned to groups of four and the participant who solved the most exercises within a group earned 40 ECU per correct calculation, with the other three group members earning nothing. Ties were randomly broken. Participants did not find out

how they performed relative to other group members until the end of the experiment. In Round 3, participants chose between the piece-rate or tournament payment schemes. If they choose the tournament scheme, their performance was compared to their group members' previous performance in Round 2. One round was randomly selected for payout. Our outcome variable for competitiveness is the binary choice in Round 3, with the choice of the tournament indicating a preference for competition. Based on prior research, we expect women to compete less than men.

We elicited preferences for *Equality versus efficiency* by implementing 15 graphical budget sets, similarly to Fisman, Kariv and Markovits (2007). In each choice, a participant distributes ECU between him- or herself and another randomly assigned participant with the relative price of giving varying across choices. Our main estimate of interest is the parameter ρ, which measures the equality-efficiency tradeoff from a CES utility function. To determine ρ, we use the maximum likelihood (ML) estimation method provided by Fisman et al adapted to our setting with 15 predetermined budget sets instead of 50 randomly chosen ones. Positive values of ρ, specifically $0 < \rho \leq 1$, indicate distributional preferences that are weighted towards efficiency (increasing total income), while negative values of ρ indicate weighting toward equality (reducing differences in income). Our outcome variable for this preference is the decile to which a participant's estimated ρ belongs, following the approach proposed in Fisman, Jakiela and Kariv (2017) to deal with outliers. Lower deciles indicate a relatively lower weight placed on efficiency relative to equity. Prior research documents that women often prioritize equality to a greater extent than men (Fisman, Jakiela and Kariv, 2017). Hence, we expect women to have lower scores than men.

We elicit *Overconfidence* using three measures of relative overplacement (Moore and Healy, 2008). First, following Gillen, Snowberg and Yariv (2019), participants solve a selection of eight matrix reasoning items (similar to Raven's matrices) provided by Condon and Revelle (2014). Participants have 45 seconds to solve each puzzle. One of the eight puzzles is randomly chosen for payout, which is 50 ECU if it is solved correctly and 0 if it is not. Participants then guess their performance rank within a randomly chosen reference group of 26 study participants, including themselves, receiving an additional 20 ECU if they guess within two ranks of their actual rank.[31] Second, we elicited two additional measures of overplacement during the measurement of competitive preferences, during which we asked participants to guess their relative rank, first within the group of four contestants and then within a randomly chosen reference group of 26 study participants. For a correct rank-out-of-four-guess, and a rank-out-of-26-guess within two ranks of their actual rank, participants receive 20 ECU, respectively. We construct our measure of overconfidence as a standardized summary index of these three

[31] They also guess how many of the eight tasks they solved correctly (overestimation) and, if this guess is correct, they again receive 20 ECU. As our focus is on overplacement, we do not use this measure.

overplacement measures and reverse the sign such that a higher score indicates greater overconfidence. Based on prior research, we expect women to be less overconfident than men.

Unincentivized preference measures. In the follow-up survey we also elicited non-incentivized preference measures such as self-reported measures of risk seeking, competitiveness, overprecision, preference for redistribution, and altruism. As hypothetical measures of risk we use i) a question from Dohmen et al. (2011) asking participants to report their general willingness to take risks on a scale from 0 to 10, and ii) the staircase measure of hypothetical choices between a lottery and a safe value, varying the amount of money in the safe option across choices, as proposed in Falk et al. (2023). The measure of competitiveness follows a similar logic as the Dohmen et al. (2011) general risk measure (Buser, Niederle and Oosterbeek, 2021). As a hypothetical measure of altruism participants stated how much of an unexpected CHF 2400 windfall they would donate to a good cause (Dohmen et al., 2011). Finally, we asked participants to state how much redistribution they want in society on a scale ranging from 1 (no redistribution) to 10 (full redistribution).

B.2. Prolific Survey with U.S. Adults—*U.S. Adults*

B.2.1 *Participants and Procedures*

To study whether CGI explains variation in economic choices and outcomes in a broader sample and across a more varied set of outcomes, we also administered an online survey to U.S. residents of working age in two waves. The respondents were recruited and paid through Prolific Academic. We limited recruitment to participants reporting English as their primary language and with an approval rate for prior studies above 98 percent. Only participants who passed two attention checks were allowed to complete the study in wave 1, and those who failed attention checks were later excluded from the sample in wave 2. Recruitment was stratified on gender and age. In wave 1, conducted in March 2022, we recruited 800 respondents in the age brackets 30-39 and 40-49 years, and 400 respondents in the age bracket 50-60 years, giving slightly lower weight to older participants who are closer to exiting the labor market. Participants received $1.25 for survey completion and could earn up to $2.5 in an incentivized risky investment task. In March 2023, we recruited a new sample of 2,802 participant from the Prolific subject pool again stratified on gender and age, excluding previous participants. In wave 2, we also recruited adults in the age bracket 20-29. We collected 800 participants from each of the age groups, 20-29, 30-39 and 40-49, and 400 participants aged 50-60, due to the lower number of available respondents in the latter age group.

B.2.2 *Elicited Measures*

CGI. In wave 1, CGI was elicited at the beginning of the survey right after demographic questions (age, U.S state of residence, ZIP code, sex) and on a separate screen. In addition to self-view, we also asked where most people would put a respondent on that scale from 0-"very masculine" to 10-"very feminine" (*CGI (other)*). Both measures were elicited prior to any of the outcome measures and on the same screen Additionally, at the very end of the survey, we asked respondents to place themselves on a scale from 0-"very masculine" to 10-"very feminine" relative to 1) women and 2) men. The placement relative to others with the same gender is the *CGI (own)* measure reported in the main text. In wave 2, CGI was elicited after all of the outcome measures. It was part of the second block of the survey in which different gender identity measures were administered in random order. All four versions of CGI (standard, as seen by most people, relative to men, relative to women) were asked together on the same screen.

Other measures. All other survey items fall into four domains: Demographics, Education, Family, Employment and Work, and Preferences. We briefly list all the variables collected here. The exact wording of all questions and answer options is available in Appendix C. Moreover, Table A1 describes all the main variables that are used for the analysis presented in the Results section of the paper.

The set of *Demographics* include age, sex, race/ethnicity, current state of residence and zip code. The question "What is your sex?" with answer options "male" and "female" was taken from the U.S. census questionnaire. The set of *Family and Education* variables consists of relationship status, sexual orientation, children, division of housework, level of education (highest degree completed) and major field of study when applicable. *Employment and Work* includes employment status, job search behavior (if applicable), sector of employment, industry, income, weeks worked last year, usual hours worked, flexibility of working hours, changes to working arrangements, the ability to take off an hour for personal matters, managerial responsibilities, performance pay, and experience with wage negotiation. The female share in industry and educational track share were constructed from publicly available datasets for the year 2020 (U.S. Bureau of Labor Statistics, National Center for Education Statistics). The educational track share is the share of women who graduated with a bachelor's degree in a field of study in 2020.

We also collected three types of preference measures. First, we asked the same incentivized risk task as in the Swiss Uni Sample with an investment endowment of 1 USD. Second, as a proxy for risk-taking, we asked whether a respondent actively trades in securities. Third, we elicited the same unincentivized measures of willingness to take risk (the general questions from Dohmen et al. (2011)), competitiveness, attitude towards redistribution and altruism as in our Swiss Uni sample (see section B.1.2 **Unincentivized**

Preference Measures for a detailed description).

B.3. Online survey with Swiss Adolescents—*Swiss Teens*

B.3.1 *Participants and Procedures*

Our third dataset is based on two surveys comprising 1,755 Swiss adolescents. We recruited these respondents through two newsletters sent from the largest Swiss online platform for apprenticeship search. This platform covers around 90 percent of all online postings for apprenticeships lasting two to four years, which form the most important part of the vocational education and training system in the German-speaking part of Switzerland. About two-thirds of a birth cohort do an apprenticeship after compulsory education, which allows them to combine vocational school with on-the-job training. Most respondents (91 percent) were 8^{th} and 9^{th} grade students (with an average age of 14.8 years) who were planning to do an apprenticeship after compulsory schooling (9^{th} grade). At the time of our surveys (December 2021 and March 2022), the 8^{th} graders had just started considering which apprenticeship they would like to do in the future. They would subsequently apply for trial apprenticeships, which typically last 1-5 days and allow the student to experience a specific apprenticeship at a specific firm. The 9^{th} graders were further along—52 percent of them had already signed a contract for their apprenticeships with a specific company.

B.3.2 *Elicited Measures*

CGI. In the first survey, we elicited CGI on the first page after the consent form; the CGI question was asked after the traditional binary gender classification (Are you male or female?). In the second survey, we asked our CGI question at the very end of the survey. The CGI question was in both surveys asked on a scale from 0-"Very feminine" to 11-"Very masculine"; for the analysis we reverse the scale.

Other measures. In the first survey, after the CGI question, we elicited the same unincentivized measures of willingness to take risk (the general questions from Dohmen et al. (2011)), competitiveness, attitude towards redistribution and altruism as in our Swiss Uni and U.S. Adults samples (see section B.1.2 **Unincentivized Preference Measures** for a detailed description). Moreover, for these respondents we were also able to merge their survey responses to their administrative user profile data from the platform. From this administrative data, we observe all applications submitted for trial apprenticeships and apprenticeships by a respondent through the platform. Thus, as a complement to the measures elicited in our survey, we observe relevant real-life behaviors. While the platform covers around 90 percent of all online apprenticeship postings, it only covers around

a quarter of all trial apprenticeships, as these are much more common to organize informally through informal networks. Given this and because trial apprenticeships mainly take place in the spring, our administrative data on occupational search predominantly captures 9th graders. Therefore, we conducted our second survey to elicit all trial apprenticeships and one-day information events students had attended among 8th and 9th graders by March 2022. Attendance at these trial apprenticeships and information events (we refer to these combined as experiences henceforth) count as excused school absences and represent an important part of students' occupational choice process.

For each application (observed in the administrative data set) and experience (reported in the second survey), we know the exact classification for the "apprenticeship profession" and merge this to characteristics about the specific apprenticeship. For the analysis of the relationship between CGI and occupational choices, we focus on three main outcome variables that characterize the student's occupational preferences in terms of gender composition and skill requirements. Since a student may apply to multiple (trial) apprenticeships, we take the average of each characteristic across all considered (trial) apprenticeships. First, we consider the female share of apprentice graduates from 2019–2021 (i.e., the gender composition of past apprentice cohorts, using the administrative LABB (*Längsschnittanalysen im Bildungsbereich*) data from the Swiss Federal Statistical Office). Second, we consider the math and school language skill requirements, based on expert evaluations of the academic requirements associated with each apprenticeship; these data come from https://anforderungsprofile.ch. We standardize the latter two measures to have a mean of zero and a standard deviation of one.

B.4. Experiment with Swedish Secondary School Students—*Swedish Youths*

B.4.1 *Participants and Procedures*

We had the opportunity to insert our measure of CGI into an experiment with Swedish students (age 18-19 years). The experiment was implemented in several schools and all 1,063 respondents were students in the final year of the natural science track in Swedish secondary education. It was implemented either in the respondents' school or online and took place in the beginning of 2022, in the weeks before the students applied for their preferred educational field for subsequent university studies.

B.4.2 *Elicited Measures*

CGI. *CGI* was elicited in a final survey after completion of the incentivized parts of the study. The question came right after demographic questions (age and a discrete measure of gender identity "Do you identify as a woman or man?" with answer options "Woman", "Man" and "Other").

Incentivized preferences. Preferences for *Equality versus efficiency* were elicited at the end of the main study through the implementation of 20 graphical budget sets, similarly to Fisman, Kariv and Markovits (2007). In each choice, a participant distributes ECU between him- or herself and another randomly assigned participant with the relative price of giving varying across choices. Our main estimate of interest is the parameter ρ, which measures the equality-efficiency tradeoff from a CES utility function. To determine ρ, we use the maximum likelihood (ML) estimation method provided by Fisman et al adapted to our setting with 20 predetermined budget sets instead of 50 randomly chosen ones. Positive values of ρ, specifically $0 < \rho \leq 1$, indicate distributional preferences that are weighted towards efficiency (increasing total income), while negative values of ρ indicate weighting toward equality (reducing differences in income). Our outcome variable for this preference is the decile to which a participant's estimated ρ belongs, following the approach proposed in Fisman, Jakiela and Kariv (2017) to deal with outliers. Lower deciles indicate a relatively lower weight placed on efficiency relative to equity.

Unincentivized preferences. The end survey also included unincentivized preferences for risk (general willingness to take risks from Dohmen et al. (2011)) and competitiveness (Buser, Niederle and Oosterbeek, 2021), as well as a measure asking participants to report their willingness to donate to charity on a scale from 0-10. Finally, the data set also comprises information on the participant's intended field of undergraduate studies. We use register data from Statistics Sweden on the actual gender composition of these study fields in 2022 to generate a variable measuring the gender composition of the respondent's intended field of study.

C. Appendix C

This Appendix reproduces the original full instructions and survey questions that we used for our primary data collection.

The data collection using Swedish Youths was implemented as part of another ongoing study by one of the authors of this paper (Ranehill). The full instructions for this data collection will be made available when that main study is completed.

Order in which Study Materials are Presented

1. Swiss Uni: Experiment
2. Swiss Uni: Follow-Up Survey
3. U.S. Adults: Wave 1
4. U.S. Adults: Wave 2
5. Swiss Teens Survey 1: English Translation
6. Swiss Teens Survey 1: German Original
7. Swiss Teens Survey 2: English Translation
8. Swiss Teens Survey 2: German Original