‘What’s the tea’ with gender nonconformity?*

Abigail R. Banan† Torsten Santavirta‡ Miguel Sarzosa§

December 30, 2021

Preliminary - Please do not circulate

Abstract

This paper explores the relationship between gender nonconformity and long-term life outcomes. To do so, we take advantage of a unique longitudinal dataset comprising administrative records and surveys that follows the entire cohort born in 1953 in Stockholm, Sweden. We use a rich barrage of survey questions on behaviors and preferences collected in primary school to construct an index of early gender conformity. We find evidence indicating a strong link between gender nonconformity during early adolescence and life outcomes. Female nonconformers have substantially better school careers, labor market outcomes and are more likely to end up in male dominated occupations as compared to their gender conforming counterparts. The opposite pattern is found for male gender nonconformers; they perform substantially worse at school, earn less as adults and are more likely to be unemployed relative to their gender conforming male peers. We also see a positive correlation between gender nonconforming boys and mental health and addiction issues later in life. Furthermore, we find that gender nonconforming girls tend to delay childbearing in adulthood. Furthermore, we find evidence that gender nonconforming students make classroom social networks more cohesive and less fractured.

JEL Classification: I21, J13, J24
Keywords: gender nonconformity, long-term outcomes

*This paper was supported by the Handelsbanken Research Foundation grant number P2016-0093. The content is solely the responsibility of the authors and does not necessarily represent the official views of the Handelsbanken Research Foundation. We are grateful to Billur Aksoy and the participants in the Purdue Labor Lunch (West Lafayette, 2021), and the CSQIEP Seminar on Economics of LGBTQ+ Individuals (2021).

†Corresponding Author. Department of Economics, Purdue University. 403 W. State St, West Lafayette, IN 47907. Email: abanan@purdue.edu. Phone: (765)-494-5426

‡Helsinki Graduate School of Economics and Swedish School of Social and Municipal Affairs, University of Helsinki. Helsinki, Finland. Email: torsten.santavirta@helsinki.fi

§Department of Economics, Purdue University. 403 W. State St, West Lafayette, IN 47907. Email: msarzosa@purdue.edu.

1 Introduction

Most individuals in a given culture have expectations in some form or another about men and women. These expectations can be anything with regards to how they dress, how they act, body language, and personal preferences. Some individuals do not fit entirely those societal preconceptions and expectations. They are considered gender nonconforming. Gender nonconformity is conceptually different from sexual orientation or gender identity. According to the American Psychological Association, gender nonconforming "describes an individual whose gender identity or gender expression differs from the *gender norms* associated with the sex they were assigned at birth" [APA, 2018, emphasis added]. That is, a person is considered to be gender nonconforming when they identify themselves as a given gender but breaks from the norms of what society expects to see from that gender.

Gender conformity has become subject to debate as the awareness of discrimination issues increases; moreover, some people simply worry that gender nonconforming children will face negative life outcomes. *The New York Times* recently reported that authorities in China fear they are experiencing a "crisis of masculinity" because "boys are getting more effeminate [May, 2021]." The fear is that there could be negative consequences if boys are not "masculine" enough. Chinese officials are claiming that there could be social issues with national security if too many boys are gender nonconforming and are now looking into policies that would aim to "toughen up" boys, while parents have begun to enroll their sons in "boot camps, hoping they will become 'real men' [May, 2021]." Moreover, North Carolina recently introduced Senate Bill 514, which would "compel state employees to immediately notify parents in writing if their child displays ‘gender nonconformity’ [Press, 2021].”[1]

The alleged ‘crisis’ in China and Senate Bill 514 in North Carolina are just a couple examples of why we need to better understand childhood gender nonconformity, and it’s relationship with adult outcomes. Plenty of work has been done in the psychology literature that demonstrates the negative outcomes associated with gender nonconformity.[2] These psychological studies commonly focus on mental health issues such as depression, anxiety, or post-traumatic stress disorder that are a result of negative feedback from society (e.g.,bullying, victimization, or discrimination) [Landolt et al., 2004, Ehrensaft, 2011, Roberts et al., 2012, 2013, Gordon et al., 2017]. The mechanisms discussed in those scholarly works are fairly straightforward: society responds negatively to children’s gender nonconformity, so the children suffer later on as a consequence.

This paper goes beyond the immediate mental health consequences societal and peer rejection may imprint on gender nonconforming children. We explore the extent to which gender nonconformity has long term effects on individuals. To do so, we utilize principal component analysis to construct an index of gender conformity. We construct this index using survey questions that were answered during childhood and include the gender of friends, preferred school subject, and hobbies. Based on this measure, we define nonconformity in relation to the stereotypical individauls by setting a cut off at the twentieth percentile of the gender conformity index distribution and explore

[1]The bill aims to prevent transgender identified youths from starting to transition prior to the age of 21. However, the North Carolina Senate Bill assumes that gender nonconformity is purely a symptom of gender dysphoria when the majority of gender nonconforming youths do not go on to identify as transgender. In fact, many individuals in the LGBTQ+ community, not only transgender individuals, are gender nonconforming as children [Bailey et al., 1993, Rieger et al., 2008]. Senate Bill 514 demonstrates a clear misunderstanding of gender nonconformity in children; moreover, the fundamental confusion around gender nonconformity in children seems to be an international phenomenon.

[2]See, for instance, [Landolt et al., 2004, Skidmore et al., 2006, Lippa, 2008, Rieger et al., 2008, Ehrensaft, 2011, Adelson, 2012, Rieger and Savin-Williams, 2012, Roberts et al., 2012, 2013, Gordon et al., 2017, MacMullin et al., 2021]

the association between childhood gender nonconformity and adult educational, career, marriage market, mental health, and fertility outcomes. In supplemental analyses we treat gender nonconformity as a latent factor and recover the association between nonconformity and life outcomes for the whole distribution of nonconformity/conformity from the least conforming to the most conforming/stereotypical individuals.

We find evidence indicating that gender nonconformity is positively correlated with future grades and career outcomes for girls while the opposite is found in boys. We also see a positive correlation between gender nonconforming boys and mental health and addiction issues later in life. Additionally, we find that gender-nonconforming girls tend to delay childbearing in adulthood. Furthermore, we find that household characteristics such as having a mother with a professional work position or a college-educated father increase gender-nonconformity, and that having older brothers decreases gender-nonconformity among boys while increasing gender-nonconformity in girls. Our results are novel because they are based on measure of gender conformity that is free from personal- or memory- related biases unlike the previous literature; moreover, our new gender conformity index is used in conjunction with a longitudinal dataset that provides a unique opportunity to examine adult outcomes in children identified as gender nonconforming

First, we contribute to the psychology literature that studies gender conformity. Our measure of gender conformity is not self-reported. By utilizing multiple survey questions that are only indirectly related to gender conformity and which were responded to during early adolescence, we avoid confounding due to recall bias or personal bias. Additionally, any preconceived notions the authors may have about gender conformity are not biasing the results because we use principal component analysis, which allows the variation in the data to speak for itself. Second, economists have looked at gender conformity but have faced limitations to their breadth of study. Brenøe [2018] examines how sibling composition impacts gender conformity; however, this paper measures gender conformity through occupational and partner choice, which means the authors cannot comment on gender conformity during childhood. Brenøe [2021] explores the underlying sources of gender nonconformity and provides causal evidence of only one predictor. In this paper however we are able to examine the relationship of many household characteristics and gender conformity. Although not being exogenous, we can look at correlation between gender conformity and a variety of life outcomes while controlling for any potentially confounding household characteristics.

The remainder of this article is structured as follows. Section 2 describes the data and the process for constructing our measure of gender conformity. Section 3 reviews the regression analysis of the paper. Section 4 discusses the results of the correlation analysis. Section 5 explores gender nonconformity as a latent variable. Section 6 investigates potential links between classroom networks and gender nonconformity, and Section 7 offers concluding thoughts.

2 Data: The Stockholm Birth Cohort

We use data from the Stockholm Birth Cohort Study (SBC) which follows the cohort of children born in 1953 who were living in the Stockholm metropolitan area on 1 November 1963. This cohort study links individuals across two longitudinal data sets. The first is the Stockholm Metropolitan Study 1953–85, which consists of all children born in 1953 who were living in the Stockholm metropolitan area on 1 November 1963. The second is The Swedish Work and Mortality Database, an administrative data set which includes information on education, income, work, unemployment and mortality for all individuals living in Sweden in 1980 or 1990.[3] The SBC follows the individuals until 2009.

The SBC study includes an in-class school survey that was conducted in 1966 when the cohort members were 13 years old. Students at practically all schools in the Stockholm county filled out two questionnaires, including questions on the gender of their closest friends, the favorite school subjects and extracurricular activities, which we use to construct our measure of gender conformity.

2.1 The gender conformity index

To construct a measure of gender nonconformity, we used answers from five survey questions included in the school study. The common feature pertaining to these questions is that their answers are heavily skewed in opposite directions depending on the gender of the respondents. We utilize principal component analysis to extract the latent underlying common factor driving the variation in these observed variables. We consider this common factor to be a measure of gender conformity.

2.1.1 Observed measures

Gender of friends. The first question we use is “with whom do you spend most of your time?” This question had five possible choices. Students could state that they

[3] The linked data sets were matched based on 13 variables which are available in both data sets: county, municipality, sex, birth month, marital status, employment, profession, socio-economic index, number of apartments in the building, year of construction of the building, quality of the construction, index of overcrowding and occupation of the property’s manager. To verify the matches, additional data on birth year of one or both parents were used. For 96% of the original cohort, data were matched. See [Stenberg and Vågerö, 2006] for a description of the data set and the matching procedure. Codebooks of the data are available online at: https://www.stockholmbirthcohort.su.se.

spend their time "mostly with boys," "mostly with girls," "mostly with a group of boys and girls," "mostly adults," or "mostly on my own." Figure 1 shows that students tend to play with other same-sex students, corroborating the already existing ample evidence of gender homophily, especially among pre- and primary schoolers (more citations on this) [Stehlé et al., 2013]. A deviation from the gender-homophily pattern could be an indication of gender nonconformity.

Preferred school subject. The next question we considered is "which subject do you like best at school?" Figure 2 shows the histogram of student responses to this question. Again, the figure shows a clear difference in responses between male and female students. Boys more frequently selected subjects like history and math than girls, and girls more frequently selected subjects like home-economics, language arts, music, and religion than boys. This difference in class preferences that we find has been seen in other works. Many studies have noted that young girls and boys tend to prefer different subjects in school [Henderson et al., 1999, Jones et al., 2000, Lupart et al., 2004, Buccheri et al., 2011]. Moreover, several papers have looked at the question of why girls tend to pursue certain career paths and not others.[4] Additionally, many articles point out that young women typically pursue different majors in college than young men.[5]

Hobby interests. The next three variables considered are continuous variables that demonstrate the child's interest level in various hobby categories. The questionnaire

[4]See, for instance, [Linn and Hyde, 1989, Eccles et al., 1990, Dick and Rallis, 1991, Aros et al., 1998, Jones et al., 2000, Lupart et al., 2004, Gadassi and Gati, 2009, Su et al., 2009, Weisgram et al., 2010, Buccheri et al., 2011, Buser et al., 2014, Gati and Perez, 2014, Sanyal, 2020]

[5]See, for instance, [Paglin and Rufolo, 1990, England and Li, 2006, Dickson, 2010, Riegle-Crumb et al., 2012, Buser et al., 2014, Gemici and Wiswall, 2014, Ahn et al., 2019, Porter and Serra, 2020]

gave ten options for each of the given categories and the students were meant to respond with whether they found the topics "very interesting," "interesting," "dull," or "very dull." These responses were given a corresponding numerical value of 5, 4, 2, or 1 points respectively; thus, a score of 10 meant the student had very little interest in that hobby category and a score of 50 meant that the student was very interested in that hobby category.

The first hobby category is "outdoor interests," which include activities such as gymnastics, biking, basketball, running, sailing, or coaching. Figure 3(a) shows that that boys had a higher average of interest in this category. The second hobby category was "domestic interests," which include activities like cooking, baking, and making clothes. On average the girls in the study reported having a higher interest in domestic hobbies than the boys in the study, which is shown in Figure 3(b)—have in mind that the children in the sample were all born in 1953. The final category is "technical interests," which include activities such as building a model railway, repairing a bicycle, building a radio, or reading about space ships. As Figure 3(c) shows, boys had a higher reported interest in technical hobbies than girls on average.

Again, the differences in responses between boys and girls that we see in these survey questions are not new in the literature. We know from earlier studies that boys and girls tend to have different interests in activities and toys.[6] For instance, boys more frequently report sports as one of their favorite activities when compared to girls, and this finding holds across cultures being studied, which aligns with the results of Figure 3(a) [Gibbons et al., 1997]. Other studies have shown that boys tend to be 'thing-oriented' while girls tend to be 'people-oriented', and the findings in Figure 3(b) and

[6]See, for instance, [Ashmore et al., 1986, Gibbons et al., 1997, Aros et al., 1998, Lippa, 2010, Niederle, 2014]

Figure 3(c) concur with those results [Aros et al., 1998, Lippa, 2010].[7]

2.1.2 The construction of the index

To consolidate these variables that clearly show a difference in expected preferences for boys and girls, we utilize principal component analysis. Principal component analysis (PCA) is regularly used for dimensionality reduction, which fits the needs of this project. This analysis allows us to exploit the variation in the previously mentioned questions and create a value to represent a latent variable, gender conformity. Moreover, PCA ensures that any bias the authors may have will not factor into the index that is created. The data is able to speak for itself. We use the first component to build our measure of interest. We normalize the values of the factor produced by that first component to the $[0, 1]$ interval for simplicity. Figure 5 shows the distribution of component scores for boys and girls separately. This figure can be interpreted as a gender-typical scale since girls are clustered near zero and boys are clustered near one.

This sub-section describes the construction of our continuous one-dimensional index of gender conformity. When considering Figure 5, we can think of the boys who are near one and the girls who are near zero as expressing more gender-typical behaviors. Moreover, the boys who are near zero do not appear to express gender-typical preferences, and the girls near one do not seem to express gender-typical preferences; the implication here is that these children would be the gender nonconformers of the sample. Therefore, the final step to creating the value which we refer to as a "gender conformity index" (GCI) is to flip the girls' factor distribution around by subtracting those scores

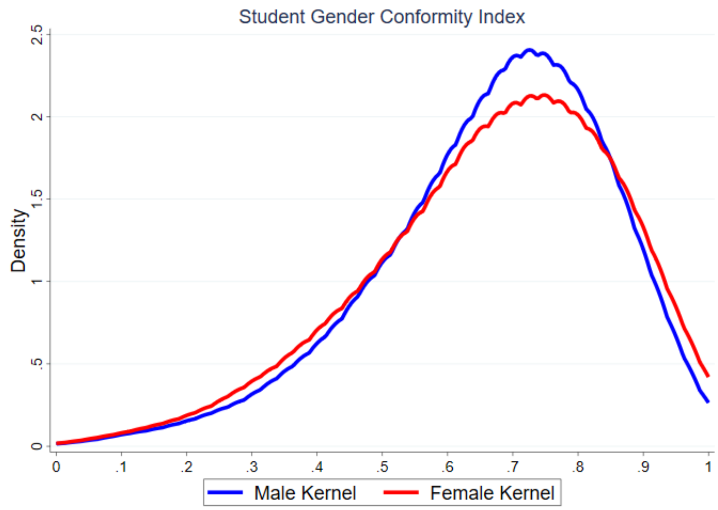
[7] While these differences are evident in youth studies, the reason for these differences are unclear Lippa [2010]. Academic debate over biology versus socialization and environment as the key mechanism for gender differences has yet to be settled, but interactions between nature and nurture likely contribute to the differences in males and females [Lippa, 2010].

from one. By doing so we have all of the girls who express feminine characteristics and the boys who express masculine characteristics near one, and the boys who express feminine characteristics and the girls who express masculine characteristics near zero. Put another way, the children who are near one would be considered "extremely gender conforming," and the children who are near zero are considered "extremely gender non-conforming" regardless of their gender. Figure 1 shows the distribution of the gender conformity index by gender. We thus define the gender-nonconforming individuals in relation to the same-gender model (stereotypical) individuals. To operationalize gender nonconformity (GNC) we set the threshold for conformity at the twentieth percentile of the GCI distribution and construct a binary variable that takes on value one for the bottom 20 percent of the GCI distribution and zero otherwise.[8] We test the robustness of our results to the use of other thresholds, and in Section 5, we also provide results using a continuous GNC factor.

Our method for constructing an index of adolescent gender conformity is advantageous for many reasons. First, being based on survey responses at age 13, it is clean from recall bias, which is particularly important as our education outcomes of interest are student performance and schooling choices in high school. Even though the jury is out on the extent of memory biases in retrospective measures [Gottschalk, 2003, Ross, 1980, Bailey et al., 1993, Rieger et al., 2008, Skidmore et al., 2006], and innate pre-dispositions are by definition not immutable rendering gender identity still plastic at age 13 [Berenbaum, 2018]. In order to understand adolescent outcomes one needs to use predetermined measures of GNC. Second, personal bias is not likely either since

[8] We test the internal reliability of the variables being used to construct this index (i.e. the five survey questions and child's sex), and we have a Cronbach's Alpha equal to 0.75, which is considered respectable according to the psychology and medical literature [Nunnally, 1978, Tavakol and Dennick, 2011]. Moreover, we have a Kaiser-Meyer-Olkin measure of sampling adequacy equal to 0.67, which is considered to be satisfactory enough to merit conducting principal component analysis [Kaiser and Rice, 1974].

Figure 1



Note: Data from Stockholm Birth Cohort. As in Figure 5, the score for girls was subtracted from one in order to flip the girls' distribution, so that the bulk of the girls' data was clustered near one. This allows us to make the assumption that the boys and girls near zero are gender nonconforming since they are different than the norm set by their peers. To create this figure an in-built kernel density command was used on the modified first component variable while also splitting the data by gender.

survey questions do not directly ask about gender conformity, but rather preferences on indirect topics (TS: Is personal bias an issue in earlier studies, if so, citations here). We decide whether an individual is a gender conformer based on how they answer seemingly unrelated questions relative to their peers, so their individual biases do not dictate their assignment. [9] Third, the use of PCA ensures that any preconceived notions the authors may have are not factoring into the assignment and only the common variation across the manifested measure is being picked up by the index. Lastly, the survey took place at school, which gave it a mandatory character and resulted in a non-response rate of only 9% (the percentage of pupils absent on that particular school day).[10] The low non-

[9]It warrants menion that by gauging femininity/masculinity based on interests, gender composition of reference group and stereotypical/anti-stereotypical favorite subject at school we capture one of three central aspects that psychologists refer to as "core" femininity/masculinity, namely gender identity while the other two, gender-role adoption and gender-role preference are left aside [Constantinople, 1973]. For our purposes, gender identity is the critical one, since this aspect is based on social comparisons and behavior while the other two are based on non-relative statements [Kachel et al., 2016].

[10]School was compulsory in Sweden for all children aged between 6-16 in 1966.

response rate in combination with the complete coverage of schools in the Stockholm county is likely to increase the external validity of our study, which is an advantage as compared to previously used measures of gender conformity.

2.2 Description of key measures

This section will provide a brief description of key measures used in the main analysis of the study. For details on how each measure was constructed, please see Appendix B. The SBC data contains information on various household and life outcomes of interest for the individuals in the study. We look at educational, career, occupational, marriage, mental health, and fertility outcomes in addition to household measures related to the individual's siblings and parents.

Educational measures. The SBC data contains information on educational outcomes on individuals in the study. We utilize the data on individuals grades, post secondary study, and information on the areas of study. "GPA Grade 9" is a continuous variable from 100 (lowest) to 500 (highest) indicating the individual's "average marks in spring term of ninth form of elementary school." "Any Post Secondary" is a binary variables that states if an individual in this study attended any post secondary education. "High School Dropout" is a binary variable that states if an individual in this study dropped out of high school. "STEM secondary track" is a binary variable that identifies whether an individual listed a STEM field as their first choice of study. The fields we classify as being male dominated are any of the natural sciences or technical fields given.

Income, career, and occupational measures. The SBC data contains information on career and occupational outcomes for individuals in the study. "log income in 1980" is a continuous measure that indicates earnings for the year 1980. We also have similar income variables for the years 1990 and 2000. Moreover, "average log income 1990-95" and "average log income 1990-2001" provide the average earnings over the listed six and twelve year time periods, respectively. The individuals in this study were between the ages 37 and 48 from 1990-2001, so these measures give us a sense of their income at a time when they would likely be more established in their careers if they did pursue a career. If an individual worked 35 hours per week or more in 1980, then we classified them as "full time in 1980." If they worked less than 35 hours per week in 1980, we classified them as "part time in 1980." We also include a variable "not employed in 1980" for individuals who did not have gainful employment at that time. "Professional" is a binary variable that indicates if the individual in the study had a professional career. "legal or business," "STEM," "blue collar," "clerical support," "teacher-other health," and "service and sales" are binary variables that state whether an individual worked in that specific occupation type between 1970 and 1980. "did not work" is a binary measure indicating whether an individual did not work between 1970 and 1980.

Marriage, mental health, and fertility measures. The SBC data contains information on marriage and health outcomes for individuals in the study. For this study, we consider how gender conformity might impact long run marriage outcomes. "Married by 1980" are a binary variable states whether an individual in the study was married by 1980. We also consider a similar divorce measure. Furthermore, we consider the relationship between gender conformity and mental illness as well as substance abuse. "Mental health" and "addiction" are binary measures that state whether an individual was ever hospitalized for any mental health disorders or alcohol and/or drug related disease between 1981-2008. The data contained birth records for each of the female individuals, but not male individual in the study; therefore, the fertility measures used in our analysis only pertain to females. "Total births" is a continuous measure for the number of children that an individual had. "Age at first birth" provides the age at which an individual in this study first gave birth conditional on her ever having had a child.

Household measures. The SBC data contains information on the home life of individuals when they were children. We examine information on the individual's siblings, father, and mother. In our analysis we consider whether siblings have any influence on gender conformity. The SBC data contains information the number of brothers and sisters an individual in the study has; moreover, we also know if their siblings are older or younger. We also investigate whether a mother or father figure was present in the individual's home. "fatherless" is a binary variable that states whether there was a father present in the household; this variable covers instances in which either a parent has died or the parents are not living together and a step-parent is not present. A corresponding version of this variable for mothers is considered as well. Also, we look into the relationship between gender conformity and the mother working. We go a step further than simply considering whether the mother works outside the home. "Professional Mother" is a binary variable that tells us whether the mother worked outside the home in a skilled professional position, while "Working Mother" is a binary variable that tells us whether the mother worked outside the home in any job. Because the children in this study were born in 1953, we wanted to examine if having a matriarchal household would have any relationship with gender conformity in children; therefore, we use a variable called "Head of House Female." This measure is a binary variable that tells us if the individual had a female head of their house while growing up. Lastly, we include measures on the parents education, and expand our investigation to think about how the difference in education between the individual's parents may influence gender conformity. "More Educated Father" is a binary variable that indicate if the individual's father had completed more school than the mother while "More Educated Mother accounts for the inverse situation. Furthermore, "Father Less than HS" is a binary variable that indicates whether the individual's father did not finish high school, and "Father Any College" is a binary variable that states whether the individual's father attended college at all. Again, corresponding variables for mothers are included in the data.

2.3 Descriptive Statistics

Table 1 through 3 provide some descriptive statistics on the individuals in the study. Table 1 shows that the only significant difference between boys and girls is the household variable for the child having multiple older sisters. Table 2 shows that the only significant difference between gender conforming and nonconforming students is that nonconforming children have more educated fathers. Table 3 gives descriptive statistics on fertility outcomes, which only includes female individuals in the study. In the next section, we will explore these household characteristics further.

Table 1: Descriptive Statistics Split by Student's Sex

Variable	Male	Female	Difference	S.E.
Gender conformity				
Gender nonconforming child	0.202	0.201	0.001	0.008
Gender conformity index	0.669	0.668	0.001	0.004
Household Characteristics				
Motherless	0.011	0.011	0.001	0.002
Fatherless	0.058	0.057	0.001	0.005
Professional mother	0.041	0.037	0.004	0.004
Working mother	0.186	0.185	0.001	0.008
Older brother	0.343	0.357	-0.015	0.010
Older sister	0.326	0.333	-0.007	0.010
Multiple older brothers	0.077	0.077	0.000	0.005
Multiple older sisters	0.063	0.073	-0.010*	0.005
Female head of household	0.072	0.079	-0.007	0.005
More educated father	0.228	0.222	0.006	0.009
More educated mother	0.016	0.014	0.001	0.002
Mother less than HS	0.932	0.937	-0.005	0.005
Mother any college	0.017	0.015	0.002	0.003
Father less than HS	0.736	0.746	-0.009	0.009
Father any college	0.088	0.082	0.006	0.006
Life outcomes				
GPA in grade 9	319.681	323.991	-4.310**	1.596
Upper secondary dropout	0.418	0.428	-0.011	0.011
STEM secondary track	0.727	0.446	0.280***	0.010
Any post secondary	0.350	0.351	-0.001	0.010
Log income in 1980	6.260	5.823	0.437***	0.026
Log income in 1990	1.849	1.419	0.430***	0.013
Log income in 2000	1.999	1.703	0.296***	0.016
Average log income 1990-95	1.846	1.485	0.361***	0.012
Average log income 1990-2001	1.922	1.602	0.320***	0.012
Full time in 1980	0.799	0.524	0.275***	0.010
Part time in 1980	0.084	0.265	-0.181***	0.008
Not employed in 1980	0.117	0.211	-0.094***	0.008
Professional	0.150	0.080	0.069***	0.008
Legal or business	0.174	0.185	-0.011	0.008
STEM	0.150	0.053	0.097***	0.006
Blue collar	0.390	0.099	0.291***	0.008
Clerical support	0.049	0.199	-0.150***	0.007
Teacher-other health	0.071	0.227	-0.156***	0.007
Service and sales	0.092	0.143	-0.051***	0.007
Did not work	0.075	0.095	-0.020***	0.006
Married by 1980	0.251	0.430	-0.179***	0.010
Divorced by 1980	0.019	0.045	-0.026***	0.004
Mental health	0.089	0.070	0.018**	0.006
Addiction	0.053	0.025	0.029***	0.004
Observations	4,639	4,794		

Note: Data from Stockholm Birth Cohort. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 2: Descriptive Statistics Split by Student's Gender Conformity

Variable	Conforming	Nonconforming	Difference	S.E.
Gender conformity				
Gender conformity index	0.739	0.392	0.347***	0.003
Student's Sex				
Female	0.508	0.507	0.001	0.013
Household characteristics				
Motherless	0.010	0.015	-0.004	0.003
Fatherless	0.057	0.059	-0.002	0.006
Professional mother	0.037	0.046	-0.010	0.005
Working mother	0.186	0.186	-0.000	0.010
Older brother	0.347	0.360	-0.013	0.012
Older sister	0.327	0.337	-0.010	0.012
Multiple older brothers	0.076	0.079	-0.003	0.007
Multiple older sisters	0.068	0.070	-0.003	0.006
Female head of household	0.076	0.075	0.001	0.007
More educated father	0.218	0.254	-0.037***	0.011
More educated mother	0.014	0.017	-0.003	0.003
Mother less than HS	0.937	0.925	0.011	0.006
Mother any college	0.015	0.021	-0.005	0.003
Father less than HS	0.749	0.709	0.040***	0.011
Father any college	0.080	0.106	-0.027***	0.007
Life Outcomes				
GPA in grade 9	319.681	330.658	-10.977***	1.992
Upper secondary dropout	0.432	0.386	0.046***	0.013
STEM secondary track	0.593	0.550	0.042***	0.013
Any post secondary	0.338	0.402	-0.064***	0.012
Log income in 1980	6.049	5.995	0.054	0.032
Log income in 1990	1.635	1.615	0.020	0.017
Log income in 2000	1.852	1.835	0.017	0.021
Average log income 1990-95	1.667	1.647	0.020	0.016
Average log income 1990-2001	1.762	1.746	0.016	0.016
Full time in 1980	0.659	0.661	-0.002	0.012
Part time in 1980	0.177	0.169	0.008	0.010
Not employed in 1980	0.163	0.169	-0.006	0.010
Professional	0.115	0.125	-0.011	0.009
Legal or business	0.178	0.186	-0.008	0.010
STEM	0.105	0.083	0.023**	0.008
Blue collar	0.245	0.230	0.015	0.011
Clerical support	0.125	0.125	-0.001	0.009
Teacher-other health	0.145	0.170	-0.025**	0.009
Service and sales	0.118	0.117	0.001	0.008
Did not work	0.084	0.089	-0.005	0.007
Married by 1980	0.346	0.326	0.020	0.012
Divorced by 1980	0.031	0.039	-0.008	0.005
Mental health	0.078	0.086	-0.008	0.007
Addiction	0.036	0.050	-0.014**	0.005
Observations	7,531	1,902		

Note: Data from Stockholm Birth Cohort. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 3: Descriptive Statistics Split by Student's Gender Conformity (Female Only)

Variable	Conforming	Nonconforming	Difference	S.E.
Fertility outcomes				
Childless	0.127	0.114	0.013	0.012
Total births	1.668	1.610	0.058	0.040
Teenage mother	0.027	0.013	0.013*	0.006
Age at first birth	23.800	24.231	-0.431**	0.133
Observations	3,829	965		

Note: Data from Stockholm Birth Cohort. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

3 Regression Analysis

This paper estimates the correlation between being gender nonconforming and life outcomes. To adjust for potential confounders, we estimate:

$$\mathrm{y}_{is} = \alpha_0 + \omega_s + \mathrm{GNC}_{is}\beta + \mathbf{x}_{is}\alpha_1 + \epsilon_{is}. \tag{1}$$

The scalar y_{is} represents the various life outcomes considered in this analysis, which include education, career, marriage, mental health, and fertility outcomes. The subscripts i and s refer to individual i in school s. Moreover, GNC_{is} is a dummy variable that indicates if an individual is classified as being a gender nonconforming child. Regressions include school-level fixed effects ω_s in addition to several controls $\mathbf{x}_{is}$.

We want to be careful about the controls included in our main regressions since gender conformity may be correlated with non-obvious household characteristics. Therefore, we investigate the relationships between an array of home-life characteristics that might correlate with gender conformity. Again, we include school-level fixed-effects. We are simply regressing individuals' household characteristics on the child's gender conformity measures. We do these regressions by individual household variable and also all together. The results of these regressions can be found in Tables 1 and 2.

We then use these household characteristics that correlate with gender conformity as controls in our main regression. The main regressions include household variables that correlate with gender conformity such as parents' level of education, the existence of older siblings, whether a parent was missing from the household, and whether the mother worked or had a professional position. Furthermore, we control for the traditional socioeconomic status indicators such as owning a home, size of the home, quality of the home, whether the home was overcrowded. The main coefficient of interest is β; we can learn something about the relationship between being a gender nonconforming child and later in life outcomes.

We use the binary variable for whether a child is gender nonconforming. We define a child as being gender nonconforming if they fall in the bottom twentieth percentile of students in the gender conformity index distribution. These regressions were run for the overall sample as well as the subsample of boys and girls separately due to the suspicion that the effects for gender nonconforming boys would be different than gender nonconforming girls. We hypothesize that conflicting forces for boys and girls would lead the overall effects to attenuate, so we ran the regressions separately.

4 Main Results

In this section we discuss the results of the analysis. We find differences in the relationship of gender nonconformity and life outcomes between boys and girls.

Table 4: Correlation of own gender nonconformity with education outcomes

	Gender Nonconforming Male	S.E.	Gender Nonconforming Female	S.E.
Educational outcomes				
GPA in grade 9	-6.000**	(2.788)	21.809***	(2.636)
Upper secondary dropout	0.032*	(0.018)	-0.074***	(0.018)
STEM secondary track	-0.137***	(0.016)	0.058***	(0.018)
Any post secondary	0.020	(0.016)	0.061***	(0.017)

Note: Data from Stockholm Birth Cohort. This table provides the results of the regressions for equation 1. These regressions control for school fixed effects and sociodemographic background covariates such as parents' level of education, a dummy for family living in an owner-occupied home, size of the home, quality of the home, whether the home was overcrowded; moreover, these regressions include household variables that in our data (or typically in the literature) are found to correlate with gender conformity such as the existence of older siblings, whether a parent was missing from the household, and whether the mother worked or had a professional position. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

4.1 Life outcome correlation results

Education. Table 4 reveals that gender nonconformity is positively correlated with student performance for girls while negatively so for boys. The magnitude of the coefficients is sizable. We find that GNC girls have roughly 0.29 standard deviations higher GPA in grade 9 of compulsory school than gender conforming girls. We can also see that the upper secondary dropout rate for girls is negatively correlated with being GNC, meaning if a girl is GNC, she is 17.3% more likely to drop out of school after completion of compulsory school in grade 9 than gender conforming girls. Moreover, GNC girls in this study were 17.4% more likely to take any post secondary study than comparable gender conforming girls.

However, educational outcomes are the opposite for GNC boys. We see 0.078 standard deviations lower GPA in grade 9 and a 7.7% higher likelihood of dropping out of school after completion of compulsory school in grade 9 as compared to gender conforming boys. As to choice of upper secondary school programme, GNC girls in this study were

13.0% more likely to apply to a STEM program for their first choice than comparable gender conforming girls. GNC boys again were 18.9% less likely to opt for a STEM program as compared to their gender conforming counterparts.

Table 5: Correlation of own gender nonconformity on career occupational outcomes

	Gender Nonconforming Male	S.E.	Gender Nonconforming Female	S.E.
Log income outcomes				
Log income in 1980	-0.158***	(0.042)	0.027	(0.051)
Log income in 1990	-0.104***	(0.024)	0.048**	(0.022)
Log income in 2000	-0.100***	(0.032)	0.036	(0.027)
Average log income 1990-95	-0.102***	(0.024)	0.041**	(0.020)
Average log income 1990-2001	-0.100***	(0.025)	0.040**	(0.020)
Career outcomes				
Full time in 1980	-0.039**	(0.015)	0.030	(0.019)
Part time in 1980	0.014	(0.011)	-0.023	(0.017)
Not employed in 1980	0.025**	(0.012)	-0.008	(0.015)
Professional	-0.019	(0.014)	0.029**	(0.012)
Occupational outcomes†				
Legal or business	0.019	(0.014)	0.001	(0.014)
STEM	-0.094***	(0.016)	0.025***	(0.007)
Blue collar	-0.010	(0.018)	-0.005	(0.011)
Clerical support	0.027***	(0.007)	-0.016	(0.015)
Teacher-other health	0.010	(0.009)	0.029*	(0.015)
Service and sales	0.029**	(0.010)	-0.020	(0.013)
Did not work	0.019*	(0.009)	-0.014	(0.011)

Note: Data from Stockholm Birth Cohort. This table provides the results of the regressions for equation 1. These regressions control for school fixed effects and sociodemographic background covariates such as parents' level of education, a dummy for family living in an owner-occupied home, size of the home, quality of the home, whether the home was overcrowded; moreover, these regressions include household variables that in our data (or typically in the literature) are found to correlate with gender conformity such as the existence of older siblings, whether a parent was missing from the household, and whether the mother worked or had a professional position. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.
† Results from multinomial logit. Marginal effect reported

Labor market. After having documented a link between gender nonconformity and education, and importantly, opposite patterns for males and females, we move on to looking at labor market outcomes in Table 5. The results show that GNC girls went on to earn 2.7%, 4.9%, and 3.7% more relative to gender conforming girls in 1980, 1990, and 2000. Moreover, we find that GNC girls earned 4.1% and 4.2% more than their gender conforming female peers on average for the 1990 to 1995 and 1990 to 2001 time frames respectively. GNC boys show a different earnings story. We see that GNC boys went on to earn 14.6%, 9.9%, and 9.5% less relative to gender conforming boys in 1980, 1990, and 2000. GNC boys earned 9.7% and 9.5% less than gender conforming boys on average for the years 1990 to 1995 and 1990 to 2001, respectively.

Relative to their respective gender conforming peers, we find that GNC girls worked full time 5.7% more in 1980, and GNC boys worked full time 4.9% less in 1980. Additionally, we see 36.1% more professional careers for GNC girls relative to comparable gender-conforming girls. As for the work they go into, we see that GNC girls are going into STEM fields 58.4% more than gender conforming girls, and GNC boys are going into STEM fiends 23.5% less than their gender conforming male peers. We also find that GNC boys are going into clerical work, service and sales, and not working 14.9%, 11.4%, and 6.8% more than the gender conforming boys in this study.

Marriage, mental health, and fertility. Lastly, we look at demographic and health outcomes in Table 6. The results show that GNC girls were not statistically significantly less likely to get married but were nevertheless 37.7% more likely to get divorced by age 27 than their gender conforming peers. We also find that GNC boys had more mental disorders and addiction later in life than gender conforming boys. GNC boys showed a 26% increase in mental health issues and a 45.1% increase in addiction. Furthermore, GNC girls seem be be more likely to postpone childbirth relative to gender conforming girls. We do not have the corresponding parental data for boys that we do

Table 6: Correlation of own gender nonconformity on marriage, mental health, and fertility outcomes

	Gender Nonconforming Male	S.E.	Gender Nonconforming Female	S.E.
Marriage outcomes				
Married by 1980	-0.023	(0.017)	-0.011	(0.019)
Divorced by 1980	0.003	(0.005)	0.017**	(0.008)
Mental health outcomes				
Mental health	0.023**	(0.011)	-0.003	(0.010)
Addiction	0.024***	(0.009)	0.009	(0.006)
Birth Outcomes				
Childless			-0.013	(0.011)
Total births			-0.024	(0.031)
Age at first birth			0.247*	(0.136)

Note: Data from Stockholm Birth Cohort. This table provides the results of the regressions for equation 1. These regressions control for school fixed effects and sociodemographic background covariates such as parents' level of education, a dummy for family living in an owner-occupied home, size of the home, quality of the home, whether the home was overcrowded; moreover, these regressions include household variables that in our data (or typically in the literature) are found to correlate with gender conformity such as the existence of older siblings, whether a parent was missing from the household, and whether the mother worked or had a professional position.*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

for girls, which is why columns one and two are empty in the table.

5 Gender Nonconformity as a Latent Variable

Defining nonconformity by use of a threshold in terms of distance from the stereotypical modal same-sex individuals on the conformity index distribution is a convenient and space conserving way of analyzing correlations between gender nonconforming behavior and life outcomes. However, in order to stay loyal to the research agenda in which this study is centered, we are cautious of not being too tightly tied around a binary categorization of conformity. In this section we take a continuous approach to nonconformity and recover the correlations between nonconformity and life outcomes for the whole range of the conformity distribution by treating gender conformity as a latent factor. In Section 2.1.2 we inferred our gender nonconforming index from behaviors and preferences that *indirectly* reveal an underlying characteristic. That is, we observe manifestations of gender nonconformity while the driver of those behaviors and preferences remain unobserved. In that sense, akin to the issue of underlying skills and test scores, we can frame our empirical problem as one of a latent factor whose influence is manifested in some observed behaviors or scores [Heckman et al., 2006]. Thus, we can apply the empirical techniques developed for the identification of the unobserved heterogeneity that drives the manifest variables we do observe [Carneiro et al., 2003, Hansen et al., 2004].[11] That is, we consider a linear relation between the unobserved factor θ and the manifest variables $\mathbf{T}$ [Bartholomew et al., 2011]. Namely,

$$\mathbf{T} = \mathbf{X}_T \beta^T + \alpha^T \theta^A + \mathbf{e^T}, \tag{2}$$

where $\mathbf{X}_T$ is a matrix with all observable controls for each measurement and α^T is a vector containing the loadings of unobserved factor A in each mean measurement T.

There are multiple advantages of pursuing such approach. First, the factor is treated as unobserved. We estimate its distribution $f_{\theta^A}(\cdot)$ using a mixture of normals, which provides flexibility to fit the factor's true underlying distribution without strong functional assumptions [Judd, 1998]. Second, we identify the factor while controlling for predetermined characteristics $\mathbf{X}_T$ that have an influence on reporting of the measures we use. Third, the factor is continuous. That way, we observe the effects of different degrees of gender nonconformity on adult outcomes. Fourth, in describing the relation

[11] In Appendix C, we present details on the identification in models with unobserved heterogeneity.

with adult outcomes y, we use a two-factor model of the form

$$y = \mathbf{x}_Y \beta^Y + \alpha^{Y,A}\theta^A + \alpha^{Y,B}\theta^B + e^Y, \tag{3}$$

where the additional factor θ^B is cognitive skills. Therefore, we estimate the relation between the gender(non)-conformity factor and later outcomes controlling for any effect cognitive skills might have on them.[12]

Table 5 in Appendix A presents the estimates of measurement system (2). As indicated in Appendix C, identification of the unobserved heterogeneity requires at least three measures per factor [Carneiro et al., 2003]. We estimate use $f_{\theta^B}(\cdot)$ using three measures that relate to IQ dimensions collected when students are 13 years old. Namely, we use the scores on verbal, numeric and spatial tests to identify a cognitive skill factor. We identify a *masculinity* factor $f_{\theta^A}(\cdot)$ separate for the sample of men and women. As manifest scores, we use the reported preferences for outdoor, domestic or technical activities. Note that in both cases—in men and women—higher scores of this measures imply behaviors or tastes that are more in line with the typical male. As controls for the measurement system $\mathbf{X}_T$, we use having a female head of household, education of the father and ownership of the dwelling in which they live.

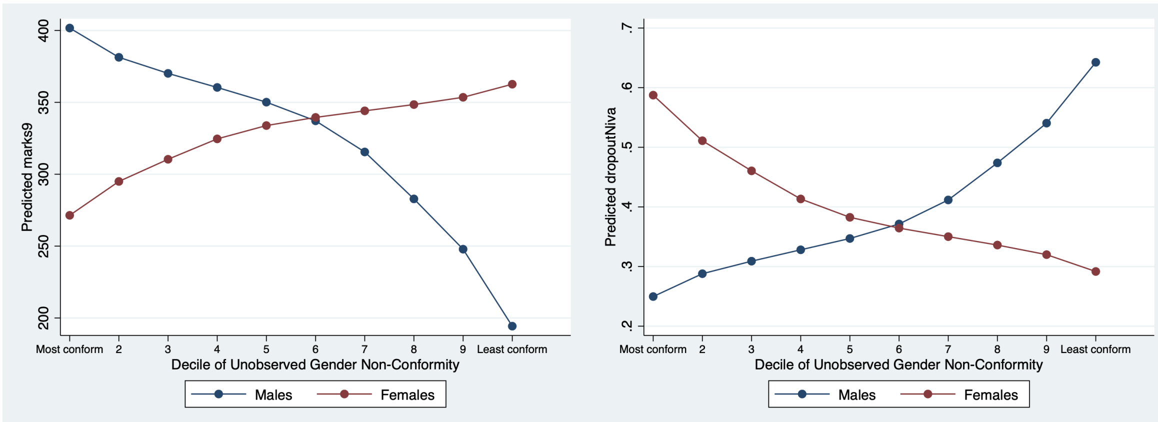
With the cognitive skills and *masculinity* factors' estimated distributions at hand, we set out to estimate their relation with adult outcomes in equation (3). To ease interpretation, and given that the masculinity is unobserved, we rely on simulations of the expected outcome as a function of that dimension of unobserved heterogeneity:

$$E_{\theta^B}[Y|\theta^A] = E[\mathbf{x}\beta] + \alpha^{Y,A}\theta^{Y,A} \tag{4}$$

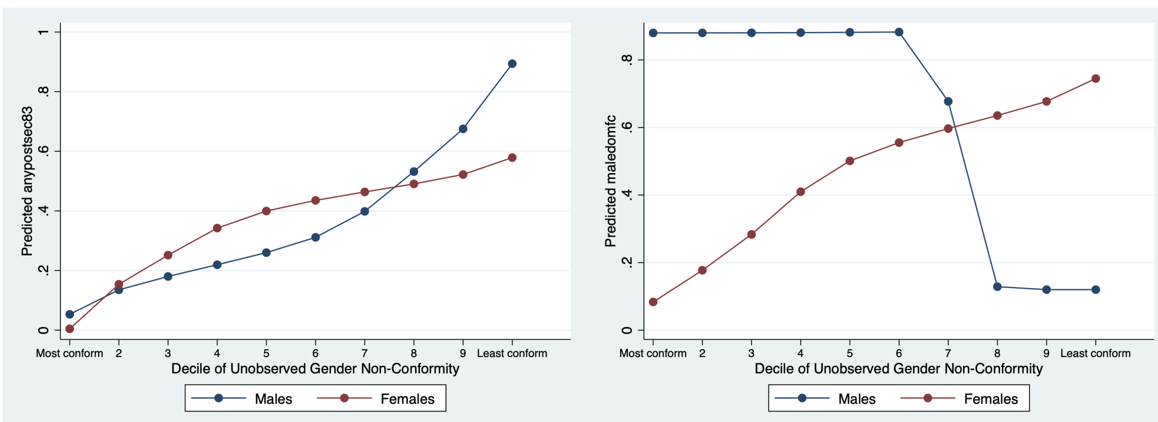
[12] A linear model like (3) can fit poorly data from outcomes in which we observe low incidence rates. In those cases, we also fit a nonlinear specification. In particular, we estimate a Probit version of (3).

Figure 2: Gender Nonconformity and Education Outcomes

(a) GPA Grades 9 (b) High School Drop Out



(c) Any Postsecondary (d) STEM Track in Upper Secondary

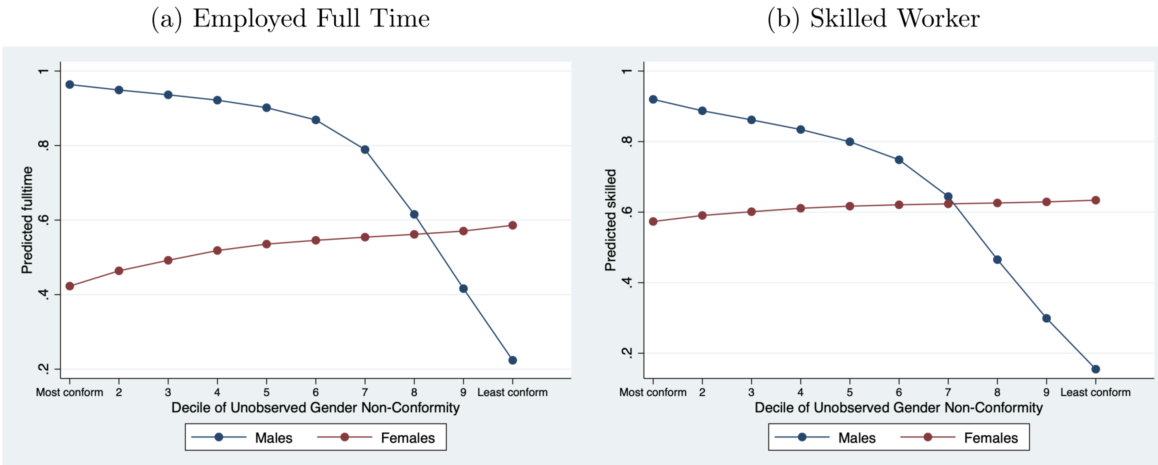


Note: All panels present the $E_{\theta^B}[Y|\theta^A] = E[\mathbf{x}\beta] + \alpha^{Y,A}\theta^{Y,A}$ in the vertical axis product of 20,000 simulations based on the findings of the empirical model (3). The horizontal axis contains the deciles of the gender-nonconforming factor.

Thus, we randomly draw 20,000 θ^A and θ^B from the distributions $f_{\theta^A}(\cdot)$ and $f_{\theta^B}(\cdot)$ estimated in the first-step estimations and construct $E_{\theta^B}[Y|\theta^A]$. This way, from the simulated graphs, we see how the unobserved masculinity relate with to the outcome variable. To translate the masculinity factor to a gender-nonconfomity measure, in the figures, we reverse its values among men.

Our results from the latent factor models go in line with those of the regression analyses using the GNC indicator. Figures 2a show that gender-nonconformity is associated

Figure 3: Gender Nonconformity and Labor Market Outcomes



Note: All panels present the $E_{\theta^B}[Y|\theta^A] = E[\mathbf{x}\beta] + \alpha^{Y,A}\theta^{Y,A}$ in the vertical axis product of 20,000 simulations based on the findings of the empirical model (3). The horizontal axis contains the deciles of the gender-nonconforming factor.

with lower GPA scores by 9^{th} grade for boys, but higher scores for girls. The gaps are sizable. The least gender-conforming boys score 2.7 standard deviations less than the most gender-conforming boys ((194.3-401.7)/76.8=2.7, where 76.8 is the standard deviation of GPA scores). On the contrary, the least gender-conforming girls score 1.2 standard deviations *more* than the most gender-conforming girls. Likewise, in Figure 2b, we see differences in high school completion by gender-conformity. Interestingly, the most gender-conforming girls and the least gender-conforming boys have similar high school dropout rates (58.7% for gender-conforming women and 64.3% for gender-nonconforming boys). While the least gender-conforming girls have similar high school dropout rates to gender-conforming boys (29.2% and 25%, respectively).

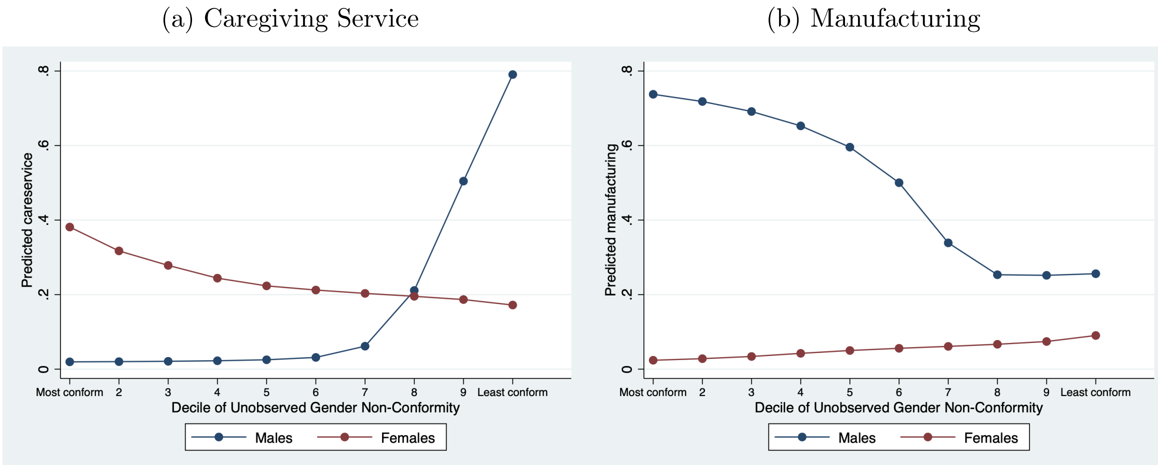
Figure 3 presents the relation between the gender-nonconformity factor and labor market outcomes. We find that gender-nonconforming men and women earn less that gender-conforming ones. However, the difference within men is significantly larger. While the average log-income of the gender-conforming man is 8.78, that of the average gender-nonconforming one is 5. This responds to multiple reasons. For instance, Figures 3a and 3b show that gender-nonconforming men are substantially less likely to hold a full-time job and are very unlikely to select skilled occupations. Furthermore, Figure 2d shows that gender-nonconforming men are 86% less likely to apply for a STEM field in upper secondary or vocational school, which are known for initiating professional paths that tend to result in better paid occupations. Figures 4 attest to that. They show that gender-nonconforming men are more likely to sort into (out of) lower (higher) paying occupations. While virtually none of the men in the first seven deciles of the gender-nonconforming factor work in caregiving occupations, 79% of those in the tenth decile do. In the same way, Figure 4b shows that 74% of the most gender-conforming men choose occupations in manufacturing, while only 25.6% of the least gender-conforming do.

On the other hand, gender-nonconforming women are 7 times more likely to choose a STEM field in upper secondary than their gender-conforming counterparts. These large gaps in education choices for women do not translate into large differences in the incidence of working in a skilled occupation. However, we do find that gender-nonconforming women are 16.3 percentage points (or 38.5%) *more* likely to work in a full time job and 21 percentage points (or 55%) *less* likely to work in a caregiving occupation than the most gender-conforming women.

We also find differences in fertility outcomes for gender-nonconforming women. In particular, Figure 5a shows that gender-nonconforming women postpone their first birth by about one year and a half relative to the most gender-conforming women.

Finally, we find a significant differences in the incidence of hospitalization due to mental heath issues across males' gender-nonconforming spectrum. Figure 5b indicates that

Figure 4: Gender Nonconformity and Occupational Outcomes



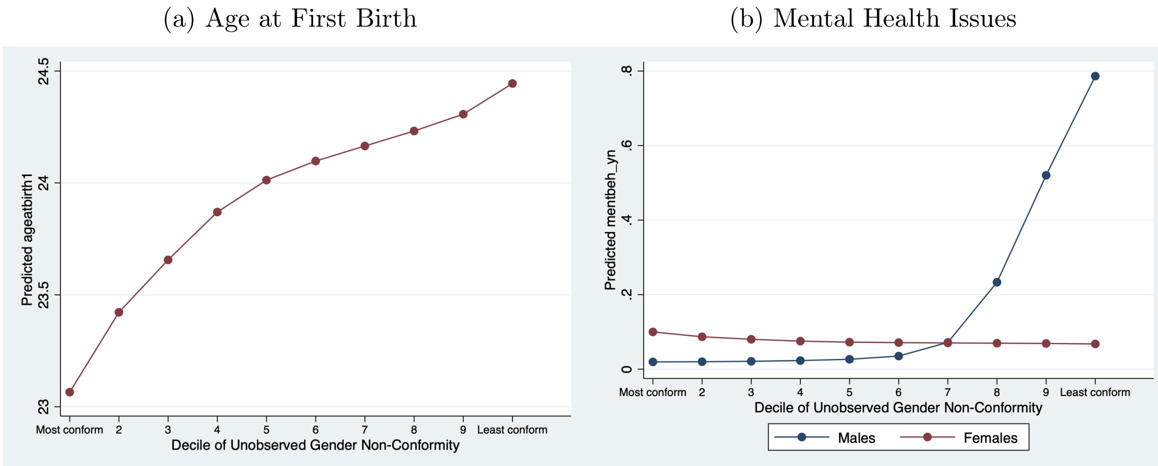
Note: Panel 4a present the $E[\mathbf{x}\hat{\beta} + \hat{\alpha}^{Y,A}\theta^{Y,A} + \hat{\alpha}^{Y,B}\theta^{Y,B} + \varepsilon > 0]$ in the vertical axis product of 20,000 simulations based on the findings of the empirical model (3). The horizontal axis contains the deciles of the gender-nonconforming factor.

almost all of the males' hospitalizations occur among gender-nonconforming men. In fact, while hospitalization are very close to zero among men in the first seven deciles of the gender-nonconforming distribution, 78% of the least gender-conforming (i.e., those in the 10^{th} decile of the gender-nonconforming distribution) we hospitalized due to mental health issues at some point between 1981 and 2008.

6 GNC Students and the Classroom Network

GNC students are special in that—to some extents—they have traits of both genders. Thus, having more GNC students in the classroom might create different social dynamics. In this Section, we set out to explore one of those social dynamics: social network cohesion. To do so, we use the friendship nomination data the SBC collected in 1966. Given that we observe an entire cohort, we are able to see all the classroom friendship networks taking place at that time. We want to inquire whether having more GNC stu-

Figure 5: Gender Nonconformity Fertility and Mental Health Outcomes



Note: All panels present the $E_{\theta^B}[Y|\theta^A] = E[\mathbf{x}\beta] + \alpha^{Y,A}\theta^{Y,A}$ in the vertical axis product of 20,000 simulations based on the findings of the empirical model (3). The horizontal axis contains the deciles of the gender-nonconforming factor.

dents changes the shape of the classrooms' network. Thus, we regress network-cohesion statistics (i.e., clustering and diameter) of classroom c in school s, n_{cs}, with the fraction of GNC students in the classroom $\mathrm{G\bar{N}C}_{cs}$.

$$\mathrm{n}_{cs} = \alpha_0 + \mathrm{G\bar{N}C}_{cs}\beta + \omega_s + \epsilon_{cs}$$

Our identification assumption relies on the random allocation of students to classrooms [Santavirta and Sarzosa, 2019]. So, after controlling for school fixed-effects, the number of GNC peers that happened to be allocated to a given classroom is orthogonal to students' characteristics. Recall that our gender conformity index includes a variable for homophily in the principal component analysis used to construct it; thus, a concern may be that the results we find from these regressions are simply mechanical. The idea would be that gender nonconformers do not follow the homophily norm (i.e. they play with children who are not their gender), which would lead to less clustered classrooms

Table 7: Impact of peer's gender nonconformity on classroom networks

	All Students	Separated by Gender	All Students	Separated by Gender
Clustering				
Class GNC Fraction	-0.102*		-0.084	
	(0.060)		(0.059)	
Class GNC Fraction Male		-0.025		-0.016
		(0.040)		(0.041)
Class GNC Fraction Female		-0.115***		-0.104***
		(0.040)		(0.040)
Relative Diameter				
Class GNC Fraction	-0.488**		-0.489**	
	(0.204)		(0.203)	
Class GNC Fraction Male		-0.147		-0.162
		(0.137)		(0.140)
Class GNC Fraction Female		-0.365***		-0.351**
		(0.138)		(0.137)
PCA Includes Homophily?	Yes	Yes	No	No

Note: Data from Stockholm Birth Cohort. This table provides the results of the regressions for equation 6. These regressions control for school fixed effects. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

and a smaller relative diameter. Therefore, we did our principal component analysis separately without the homophily variable to see if our results still hold. The gender conformity index did not change for most individuals by more than 0.0001 of a point, and removing the homophily variable did not cause individuals to get classified into the gender nonconformer differently than before. We include both sets of results (with and without homophily) in Table 7.

We find that as the fraction of gender nonconforming students in the classroom increases, the classroom networks are significantly less clustered and the relative diameter of the classroom is significantly less. That means that the presence of GNC students makes classrooms' social networks more cohesive and less fractured. When we split gender nonconformity fractions by gender, we find that the results are driven exclusively by the gender nonconforming girls. That is, GNC girls and not boys serve as bridges between social cliques–usually defined by gender in this age group. GNC girls help connect students who otherwise would belong to disjoint social networks, despite being in the same classroom.

Additionally, our results do not change when we exclude homophily from the principle component analysis, which reassures us that the results are not mechanically driven by the homophily input variable.

7 Conclusions and Discussion

Our paper provides new information on the relationship between childhood gender nonconformity and life outcomes by using a novel approach to measuring gender nonconformity in children. Using several survey questions, we are able to create a gender conformity index. Furthermore, since the survey questions used were answered during childhood, we can note that our measure of gender conformity is not subject to personal- or memory- bias for which past attempts at measuring gender conformity have been criticized.

Correlation discussion. Childhood gender nonconformity seems to be associated with worse outcomes for boys than for girls. Various potential mechanisms could contribute to the differences we find in life outcomes for GNC boys and girls. We see GNC girls making better grades while GNC boys make worse grades relative to their gender conforming peers. We find that GNC boys tend to choose STEM fields less while GNC girls tend to do the opposite; moreover, GNC girls also delay having children. This difference in career choices coupled with GNC girls postponing motherhood could help explain differences in income seen in GNC boys and girls.

GNC girls delaying children delays the "child penalty" that working women traditionally suffer more intensely than men; this "child penalty" is the significant reduction in women's pay that typically comes after having the first child [Angelov et al., 2016, Doepke and Kindermann, 2019, Kleven et al., 2019a,b]. Additionally, since GNC boys are opting out of STEM career paths, they tend to earn less since those fields ordinarily pay less, conversely GNC girls tend to choose STEM fields that normally pay more [Pan, 2015, Blau and Kahn, 2017, Cortes and Pan, 2018]. Again, this cohort of children was born in 1953; therefore, when they were old enough to enter the workforce, they were joining the labor market at a time when female dominated and gender neutral career paths were valued less than they are today relative to male dominated career paths like STEM, which leads to another potential mechanism for the differences in outcomes for boys and girls.

The next potential mechanism that could contribute to results we find on life outcomes for gender nonconforming boys and girls in this study is the social movement that took place during their early life. The second wave of feminism in Sweden began in the 1960s and continued through the 1970s and into the 1980s [Gelb, 1989, Forsås-Scott, 2021]. Thus, the bulk of this cohort's formative years occurred during a period of major social transition. In addition to other issues, the second feminist wave pushed for gender equality in the work place (i.e. equal opportunity and equal pay for equal work). Growing up during a time frame when women and young girls were being empowered by feminist groups to pursue careers outside of what was traditionally expected of them and to challenge the status quo could have encouraged the young GNC girls in this study to be more ambitious, to go to college, and to pursue STEM careers. The GNC girls in this study were likely receiving some indirect encouragement for their behavior and preferences from social activism while GNC boys were not likely receiving that same indirect validation.

Another potential reason for the differences that we find in life outcomes between GNC boys and girls could be negative feedback from peers. As most children that have difficulty fitting in because they are different in some regard to the rest of their peers, GNC girls and boys are potentially subject to bullying and discrimination [Sarzosa and Urzúa, 2021]. Furthermore, GNC boys may potentially experience different forms of bullying than GNC girls. We know from previous studies on bullying that boys and girls tend to have different experiences with bullying. Boys are more likely to be victimized through physical bullying than girls [Scheithauer et al., 2006, Carbone-Lopez et al., 2010, Silva et al., 2013]. Also, girls are more likely to be victims of indirect bullying such as teasing or online bullying [Carbone-Lopez et al., 2010, Smith et al., 2019]. These noted differences in bullying might explain why GNC boys have more negative life outcomes than GNC girls. Furthermore, the larger point that should be noted is that this is likely a self-fulfilling prophecy; people think that gender nonconformity or "otherness" is potentially bad for a child, as we have seen from recent reports about the "masculinity" crisis; so they shame, victimize, or reject that child. Gender nonconformers tend to be the target of bullying, discrimination, and victimization, which leads to negative outcomes. Awful negative experiences without any outside encouragement from social groups could lead to a lackluster educational performance that snowballs into an below average labor market outcomes later in life. We can see from our results on mental health in Table 6 that GNC boys, but not girls, had significantly more issues with mental illness and addiction, which gives some evidence that these potential mechanisms could hold.

Concluding remarks. The results of this paper provide evidence that youthful gender nonconformity is linked to individual life outcomes. While we speculate on mechanisms that could explain the results we find, further work needs to be done to substantiate these hypotheses and explore the causal mechanisms that could be leading to these life outcomes in GNC children.

References

M. Adelson, Stewart L. Practice parameter on gay, lesbian, or bisexual sexual orientation, gender nonconformity, and gender discordance in children and adolescents. *Journal of the American Academy of Child and Adolescent Psychiatry*, 51(9):957–974, 2012. ISSN 0890-8567. 2

T. Ahn, P. Arcidiacono, A. Hopson, and J. R. Thomas. Equilibrium grade inflation with implications for female interest in stem majors. 2019. 5

N. Angelov, P. Johansson, and E. Lindahl. Parenthood and the gender gap in pay. *Journal of labor economics*, 34(3):545–579, 2016. ISSN 0734-306X. 7

APA. A glossary: Defining transgender terms, Sep 2018. URL `https://www.apa.org/monitor/2018/09/ce-corner-glossary`. 1

J. R. Aros, G. A. Henly, and N. T. Curtis. Occupational sextype and sex differences in vocational preference-measured interest relationships. *Journal of vocational behavior*, 53(2):227–242, 1998. ISSN 0001-8791. 4, 6, 2.1.1

R. D. Ashmore, F. K. Del Boca, and A. J. Wohlers. Three - gender stereotypes. In *The Social Psychology of Female-Male Relations*, pages 69–119. Elsevier Inc, 1986. ISBN 1483216209. 6

J. Bailey, J. S. Miller, and L. Willerman. Maternally rated childhood gender nonconformity in homosexuals and heterosexuals. *Archives of sexual behavior*, 22(5):461–469, 1993. ISSN 0004-0002. 1, 2.1.2

D. Bartholomew, M. Knott, and I. Moustaki. *Latent Variable Models and Factor Analysis: A Unified Approach*. Wiley Series in Probability and Statistics. West Sussex, United Kingdom, 2011. 5, C

S. A. Berenbaum. Evidence needed to understand gender identity: Commentary on turban and ehrensaft (2018). *J Child Psychol Psychiatry*, 59(12):1244–1247, 2018. doi: 10.1111/jcpp.12997. 2.1.2

F. D. Blau and L. M. Kahn. The gender wage gap: Extent, trends, and explanations. *Journal of economic literature*, 55(3):789–865, 2017. ISSN 0022-0515. 7

A. A. Brenøe. Origins of gender norms: sibling gender composition and women's choice of occupation and partner. *University of Zurich, Department of Economics, Working Paper*, (294), 2018. 1, A

A. A. Brenøe. Brothers increase womenâĂŹs gender conformity. *Journal of Population Economics*, pages 1–38, 2021. 1, A

G. Buccheri, N. A. Gürber, and C. Brühwiler. The impact of gender on interest in science topics and the choice of scientific and technical vocations. *International journal of science education*, 33(1):159–178, 2011. ISSN 0950-0693. 2.1.1, 4

T. Buser, M. Niederle, and H. Oosterbeek. Gender, competitiveness, and career choices. *The Quarterly journal of economics*, 129(3):1409–1447, 2014. ISSN 0033-5533. 4, 5

K. Carbone-Lopez, F.-A. Esbensen, and B. T. Brick. Correlates and consequences of peer victimization: Gender differences in direct and indirect forms of bullying. *Youth violence and juvenile justice*, 8 (4):332–350, 2010. ISSN 1541-2040. 7

P. Carneiro, K. T. Hansen, and J. Heckman. Estimating Distributions of Treatment Effects with an Application to the Returns to Schooling and Measurement of the Effects of Uncertainty on College Choice. *International Economic Review*, 44(2):361–422, Apr. 2003. 5, 5, 14, 15

A. Constantinople. Masculinity-femininity: an exception to a famous dictumm? *Psychological Bulletin*, 80(5):389–407, 1973. doi: 10.1037/h0035334. 9

P. Cortes and J. Pan. Occupation and gender. *The Oxford handbook of women and the economy*, pages 425–452, 2018. 7

T. Dick and S. Rallis. Factors and influences on high school students' career choices. *Journal for research in mathematics education*, 22(4):281–292, 1991. ISSN 0021-8251. 4

L. Dickson. Race and gender differences in college major choice. *The Annals of the American Academy of Political and Social Science*, 627(1):108–124, 2010. ISSN 0002-7162. 5

M. Doepke and F. Kindermann. Bargaining over babies: Theory, evidence, and policy implications. *The American economic review*, 109(9):3264–3306, 2019. ISSN 0002-8282. 7

J. S. Eccles, J. E. Jacobs, and R. D. Harold. Gender role stereotypes, expectancy effects, and parents' socialization of gender differences. *Journal of social issues*, 46(2):183–201, 1990. ISSN 0022-4537. 4

D. Ehrensaft. Boys will be girls, girls will be boys: Children affect parents as parents affect children in gender nonconformity. *Psychoanalytic psychology*, 28(4):528–548, 2011. ISSN 0736-9735. 1, 2

P. England and S. Li. Desegregation stalled: The changing gender composition of college majors, 1971-2002. *Gender & society*, 20(5):657–677, 2006. ISSN 0891-2432. 5

H. Forsås-Scott. Feminism in the nordics: An historical overview to 1990, Mar 2021. URL https://nordics.info/show/artikel/feminism-in-the-nordics-an-historical-overview-to-2010/. 7

R. Gadassi and I. Gati. The effect of gender stereotypes on explicit and implicit career preferences. *The Counseling psychologist*, 37(6):902–922, 2009. ISSN 0011-0000. 4

I. Gati and M. Perez. Gender differences in career preferences from 1990 to 2010: Gaps reduced but not eliminated. *Journal of counseling psychology*, 61(1):63–80, 2014. ISSN 0022-0167. 4

J. Gelb. *Feminism and politics: A comparative perspective*. Univ of California Press, 1989. 7

A. Gemici and M. Wiswall. Evolution of gender differences in post-secondary human capital investments: College majors. *International economic review (Philadelphia)*, 55(1):23–56, 2014. ISSN 0020-6598. 5

J. L. Gibbons, M. Lynn, and D. A. Stiles. Cross-national gender differences in adolescents' preferences for free-time activities. *Cross-cultural research*, 31(1):55–69, 1997. ISSN 1069-3971. 2.1.1, 6

A. R. Gordon, N. Krieger, C. A. Okechukwu, S. Haneuse, M. Samnaliev, B. M. Charlton, and S. B. Austin. Decrements in health-related quality of life associated with gender nonconformity among u.s. adolescents and young adults. *Quality of life research*, 26(8):2129–2138, 2017. ISSN 0962-9343. 1, 2

L. Gottschalk. Same-sex sexuality and childhood gender non-conformity: A spurious connection. *Journal of Gender Studies*, 12(1):35–50, 2003. doi: 10.1080/0958923032000067808. URL https://doi.org/10.1080/0958923032000067808. 2.1.2

K. T. Hansen, J. J. Heckman, and K. J. Mullen. The effect of schooling and ability on achievement test scores. *Journal of Econometrics*, 121(1-2):39–98, July 2004. 5

J. Heckman, J. Stixrud, and S. Urzua. The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior. *Journal of Labor Economics*, 24(3):411–482, July 2006. 5, C

B. B. Henderson, M. H. Marx, and Y. C. Kim. Academic interests and perceived competence in american, japanese, and korean children. *Journal of cross-cultural psychology*, 30(1):32–50, 1999. ISSN 0022-0221. 2.1.1

M. G. Jones, A. Howe, and M. J. Rua. Gender differences in students' experiences, interests, and attitudes toward science and scientists. *Science education (Salem, Mass.)*, 84(2):180–192, 2000. ISSN 0036-8326. 2.1.1, 4

K. L. Judd. *Numerical Methods in Economics*. The MIT Press, Cambridge, Massachusetts, 1998. 5

S. Kachel, M. C. Steffens, and C. Niedlich. Traditional masculinity and femininity: Validation of a new scale assessing gender roles. *Frontiers in Psychology*, 7:956, 2016. ISSN 1664-1078. doi: 10.3389/fpsyg.2016.00956. URL https://www.frontiersin.org/article/10.3389/fpsyg.2016.00956. 9

H. F. Kaiser and J. Rice. Little jiffy, mark iv. *Educational and psychological measurement*, 34(1):111–117, 1974. ISSN 0013-1644. 8

H. Kleven, C. Landais, J. Posch, A. Steinhauer, and J. Zweimüller. Child penalties across countries: Evidence and explanations. *AEA papers and proceedings*, 109:122–126, 2019a. ISSN 2574-0768. 7

H. Kleven, C. Landais, and J. E. Søgaard. Children and gender inequality: Evidence from denmark. *American economic journal. Applied economics*, 11(4):181–209, 2019b. ISSN 1945-7782. 7

I. Kotlarski. On characterizing the gamma and the normal distribution. *Pacific Journal of Mathematics*, 20(1):69–76, 1967. C

M. A. Landolt, K. Bartholomew, C. Saffrey, D. Oram, and D. Perlman. Gender nonconformity, childhood rejection, and adult attachment: A study of gay men. *Archives of sexual behavior*, 33(2):117–128, 2004. ISSN 0004-0002. 1, 2

M. C. Linn and J. S. Hyde. Gender, mathematics, and science. *Educational researcher*, 18(8):17–27, 1989. ISSN 0013-189X. 4

R. A. Lippa. The relation between childhood gender nonconformity and adult masculinity-femininity and anxiety in heterosexual and homosexual men and women. *Sex roles*, 59(9):684–693, 2008. ISSN 0360-0025. 2

R. A. Lippa. Gender differences in personality and interests: When, where, and why?: Gender differences in personality and interests. *Social and personality psychology compass*, 4(11):1098–1110, 2010. ISSN 1751-9004. 6, 2.1.1, 7

J. L. Lupart, E. Cannon, and J. A. Telfer. Gender differences in adolescent academic achievement, interests, values and life-role expectations. *High ability studies*, 15(1):25–42, 2004. ISSN 1359-8139. 2.1.1, 4

L. N. MacMullin, L. M. Bokeloh, A. N. Nabbijohn, A. Santarossa, A. I. R. van der Miesen, D. E. Peragine, and D. P. VanderLaan. Examining the relation between gender nonconformity and psychological well-being in children: The roles of peers and parents. *Archives of sexual behavior*, 50(3):823–841, 2021. ISSN 0004-0002. 2

T. May. A 'masculinity crisis'? china says the boys are not all right, Feb 2021. URL https://www.nytimes.com/2021/02/05/world/asia/china-masculinity-schoolboys.html. 1

M. Niederle. *Gender*. National Bureau of Economic Research, 2014. 6

J. Nunnally. Construction of tests for special purposes. *Psychometric Theory. 2nd ed. New York: McGraw-Hill*, 1978. 8

C. Olivetti, E. Patacchini, and Y. Zenou. Mothers, peers, and gender-role identity. *Journal of the European Economic Association*, 18(1):266–301, 2020.

M. Paglin and A. Rufolo. Heterogeneous human capital, occupational choice, and male-female earnings differences. *Journal of labor economics*, 8(1):123–144, 1990. ISSN 0734-306X. 5

J. Pan. Gender segregation in occupations: The role of tipping and social interactions. *Journal of labor economics*, 33(2):365–408, 2015. ISSN 0734-306X. 7

C. Porter and D. Serra. Gender differences in the choice of major: The importance of female role models. *American economic journal. Applied economics*, 12(3):226–254, 2020. ISSN 1945-7782. 5

T. A. Press. N.c. bill would ban treatment for trans people under 21, Apr 2021. URL https://www.nbcnews.com/feature/nbc-out/n-c-bill-would-ban-treatment-trans-people-under-21-n1263146. 1

G. Rieger and R. C. Savin-Williams. Gender nonconformity, sexual orientation, and psychological well-being. *Archives of sexual behavior*, 41(3):611–621, 2012. ISSN 0004-0002. 2

G. Rieger, J. A. W. Linsenmeier, L. Gygax, and J. M. Bailey. Sexual orientation and childhood gender nonconformity: Evidence from home videos. *Developmental psychology*, 44(1):46–58, 2008. ISSN 0012-1649. 1, 2, 2.1.2

C. Riegle-Crumb, B. King, E. Grodsky, and C. Muller. The more things change, the more they stay the same? prior achievement fails to explain gender inequality in entry into stem college majors over time. *American educational research journal*, 49(6):1048–1073, 2012. ISSN 0002-8312. 5

A. L. Roberts, M. Rosario, H. L. Corliss, K. C. Koenen, and S. B. Austin. Childhood gender nonconformity: A risk indicator for childhood abuse and posttraumatic stress in youth. *Pediatrics (Evanston)*, 129(3):410–417, 2012. ISSN 0031-4005. 1, 2

P. Roberts, Andrea L., P. Rosario, Margaret, S. Slopen, Natalie, P. Calzo, Jerel P., and S. Austin, S. Bryn. Childhood gender nonconformity, bullying victimization, and depressive symptoms across adolescence and early adulthood: An 11-year longitudinal study. *Journal of the American Academy of Child and Adolescent Psychiatry*, 52(2):143–152, 2013. ISSN 0890-8567. 1, 2

M. W. Ross. Retrospective distortion in homosexual research. *Archives of Sexual Behavior*, 9(6): 523–531, 1980. doi: 10.1007/BF01542156. 2.1.2

I. Sagar-Ouriaghli, E. Godfrey, L. Bridge, L. Meade, and J. S. Brown. Improving mental health service utilization among men: a systematic review and synthesis of behavior change techniques within interventions targeting help-seeking. *American journal of men's health*, 13(3):1557988319857009, 2019.

T. Santavirta and M. Sarzosa. Effects of disruptive peers in endogeneous social networks. 2019. 6

A. Sanyal. Impact of teacher-student gender-matching on girls' stem skills and preference. Unpublished Manuscript, 2020. 4

M. Sarzosa and S. Urzua. Implementing Factor Models for Unobserved Heterogeneity in Stata. *The Stata Journal*, 16(1):197–228, Jan. 2016. 13

M. Sarzosa and S. Urzúa. Bullying among adolescents: The role of skills. *Quantitative economics*, 12 (3):945–980, 2021. ISSN 1759-7323. 7

H. Scheithauer, T. Hayer, F. Petermann, and G. Jugert. Physical, verbal, and relational forms of bullying among german students: age trends, gender differences, and correlates. *Aggressive behavior*, 32(3):261–275, 2006. ISSN 0096-140X. 7

M. A. I. Silva, B. Pereira, D. Mendonça, B. Nunes, and W. A. de Oliveira. The involvement of girls and boys with bullying: An analysis of gender differences. *International journal of environmental research and public health*, 10(12):6820–6831, 2013. ISSN 1661-7827. 7

W. C. Skidmore, J. A. W. Linsenmeier, and J. M. Bailey. Gender nonconformity and psychological distress in lesbians and gay men. *Archives of sexual behavior*, 35(6):685–697, 2006. ISSN 0004-0002. 2, 2.1.2

P. K. Smith, L. López-Castro, S. Robinson, and A. Görzig. Consistency of gender differences in bullying in cross-cultural surveys. *Aggression and violent behavior*, 45:33–40, 2019. ISSN 1359-1789. 7

J. Stehlé, F. Charbonnier, T. Picard, C. Cattuto, and A. Barrat. Gender homophily from spatial behavior in a primary school: A sociometric study. *Social Networks*, 35:604–613, Sept. 2013. doi: 10.1016/j.socnet.2013.08.003. 2.1.1

S.-r. Stenberg and D. Vågerö. Cohort profile: The stockholm birth cohort of 1953. *International journal of epidemiology*, 35(3):546–548, 2006. ISSN 0300-5771. 3

R. Su, J. Rounds, and P. I. Armstrong. Men and things, women and people: A meta-analysis of sex differences in interests. *Psychological bulletin*, 135(6):859–884, 2009. ISSN 0033-2909. 4

Y. Takahashi. Gender differences in the cost of corrections in group work. *Unpublished manuscript*, 2021.

M. O. Tang, J. L. Oliffe, P. M. Galdas, A. Phinney, and C. S. Han. College menâĂŹs depression-related help-seeking: A gender analysis. *Journal of Mental Health*, 23(5):219–224, 2014.

M. Tavakol and R. Dennick. Making sense of cronbach's alpha. *International journal of medical education*, 2:53–55, 2011. ISSN 2042-6372. 8

S. Urzua. Racial Labor Market Gaps. *Journal of Human Resources*, 43(4):919, 2008. C

D. L. Vogel, S. R. Heimerdinger-Edwards, J. H. Hammer, and A. Hubbard. âĂIJboys don't cryâĂİ: Examination of the links between endorsement of masculine norms, self-stigma, and help-seeking attitudes for men from diverse backgrounds. *Journal of counseling psychology*, 58(3):368, 2011.

D. Wasserman. *Oxford textbook of suicidology and suicide prevention*. Oxford University Press, 2020.

E. S. Weisgram, R. S. Bigler, and L. S. Liben. Gender, values, and occupational interests among children, adolescents, and adults. *Child development*, 81(3):778–796, 2010. ISSN 0009-3920. 4

Appendix

A Tables and Figures

Figure 1: Distribution of Gender Homophily

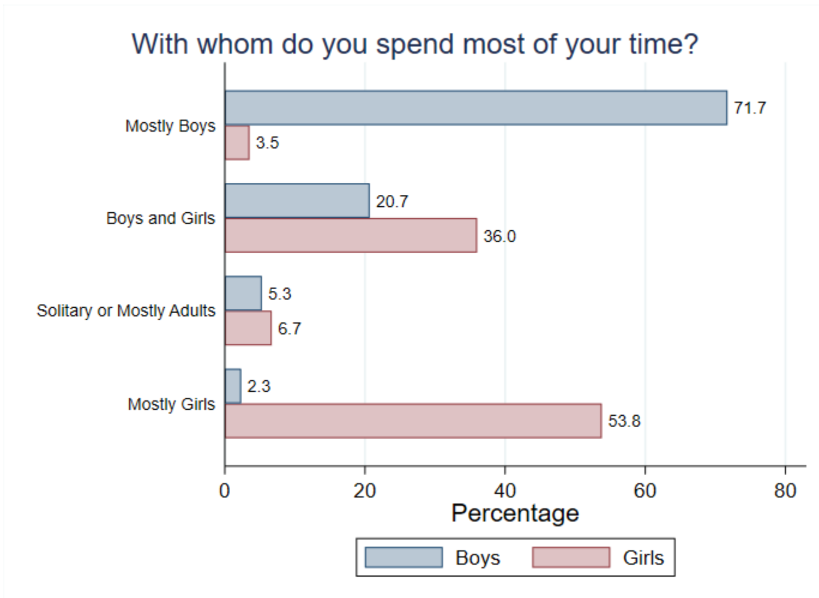


Figure 2: Gender Distribution of the Preferred School Subject

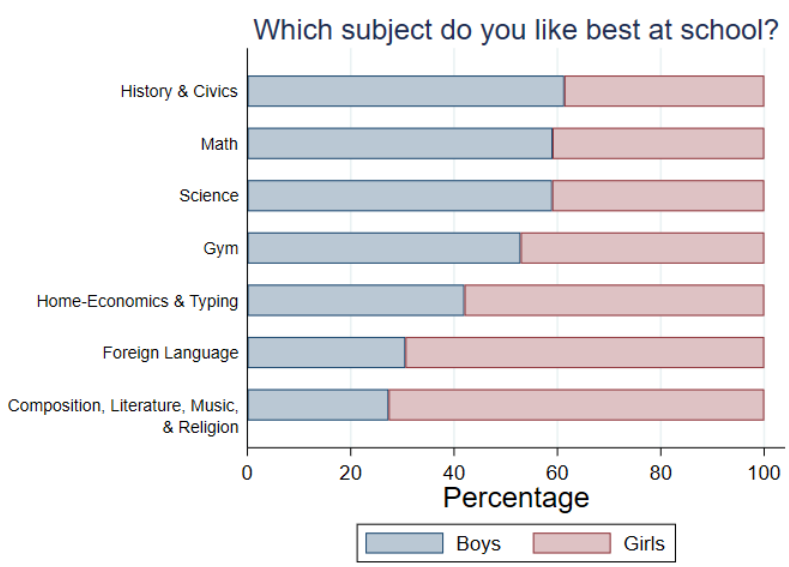
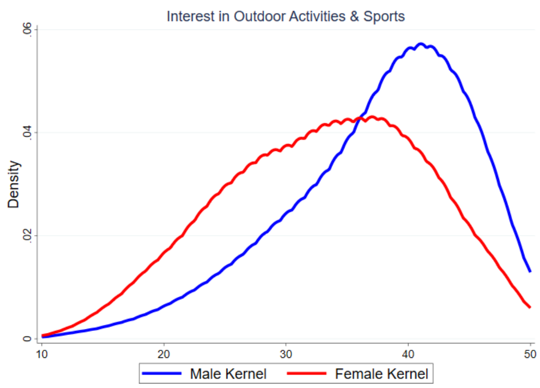
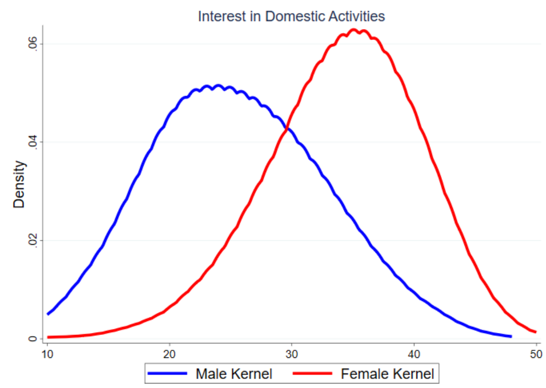


Figure 3: Hobby Preference by Gender

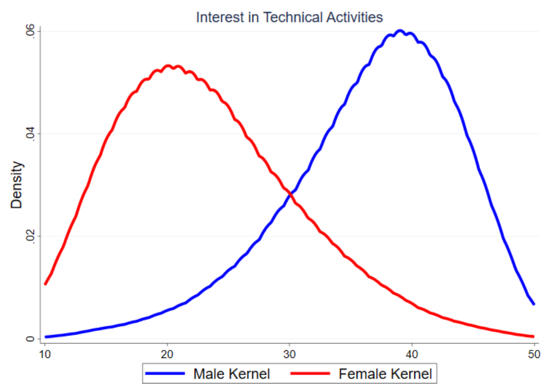
(a)



(b)

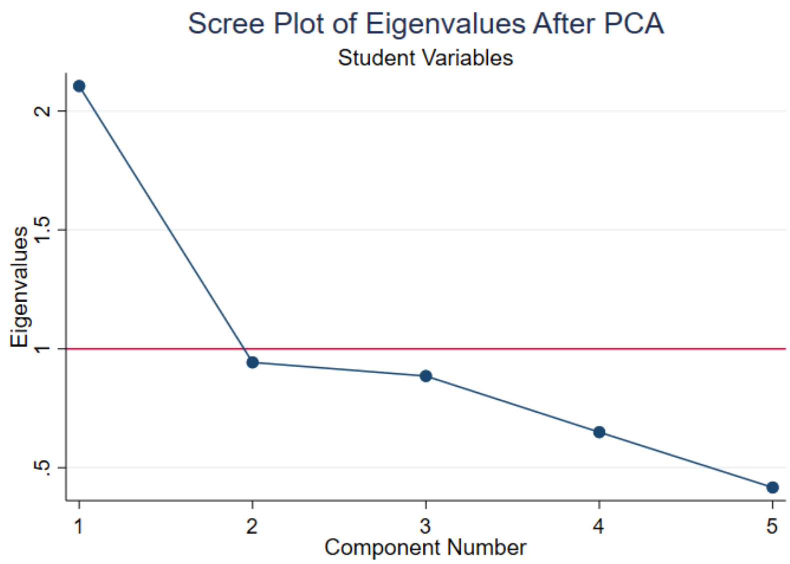


(c)



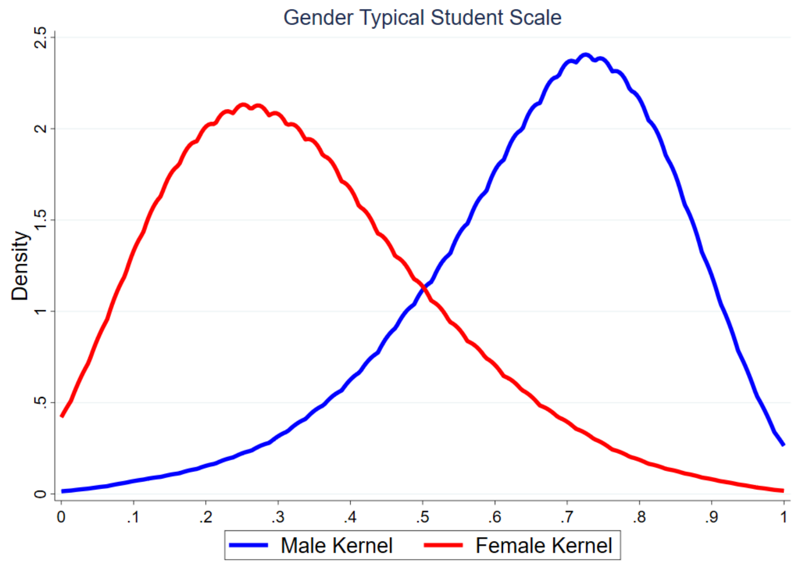
Note: Data from Stockholm Birth Cohort. The given three variables in the data provide a numerical score from 10 to 50, where 10 indicates very little interest in the hobby category and 50 indicates very high interest. Each figure was constructed using an in-built kernel density command while also splitting the data by gender.

Figure 4



Note: Data from Stockholm Birth Cohort. To create this figure, an in-built scree plot command was used after running principal component analysis.

Figure 5: Distributions of Gender Conformity Factor by Gender of the Respondent



Note: Data from Stockholm Birth Cohort. After running principal component analysis, the first component was used to create this figure since it accounted for the majority of the variation in the data. To create this figure an in-built kernel density command was used on the first component variable while also splitting the data by gender.

Household background results. We consider if an association exists between household characteristics and gender conformity, so we can appropriately control for these confounding variables in our main specification. Past papers have found that certain household traits such as family composition influence gender conformity in girls [Brenøe, 2018, 2021]. Our paper considers an array of household characteristics in addition to the previously studied traits to see what other characteristics may be related to childhood gender nonconformity. Equation 5 gives the specification. Again, we include school-level fixed-effects. We are simply regressing individuals' household characteristics on their given gender conformity measures.

$$\text{GNC}_{is} = \alpha_0 + \omega_s + \mathbf{x}_{is}\alpha_1 + \epsilon_{is} \tag{5}$$

Table 1 gives correlation results between household characteristics and childhood gender conformity. The top panel shows how individual household characteristics correlate with gender nonconformity while the bottom panel has the results for regressions with all the household characteristics included.

In Table 1 we see that household characteristics such as not having a mother in the home, mothers having professional careers, and having a college educated father are all correlated with being a GNC child. When we separate the regressions by gender we see that the only household trait that is significantly correlated with gender conformity for boys is having an older brother. Having an older brother is associated with being more GNC for boys, while the opposite is seen for girls. Girls' gender conformity seems to be associated with many variables. Not having a mother or having a professional working mother, having a father who is more educated than their mother, or having a college educated father are all significantly correlated with being a GNC girl. Moreover, having a father who did not finish high school is associated with more gender conformity in girls. When we run multivariate regressions with the household characteristics for boys and girls separately, boys' conformity is still significantly correlated with having an older brother while girls' gender conformity remains significantly related to their mother's characteristics and having an older brother. Table 1 yields similar results to Table 2, which again shows are results hold regardless of the twentieth percentile cutoff decision.

Table 1: Correlation of household characteristics on gender nonconformity

Household Characteristic	Gender Nonconforming Male	S.E.	Gender Nonconforming Female	S.E.
Panel A: Univariate linear regression				
Motherless	0.024	(0.056)	0.102*	(0.056)
Professional mother	0.005	(0.030)	0.061*	(0.031)
Older brother	0.050***	(0.013)	-0.024**	(0.012)
Older sister	0.005	(0.013)	0.009	(0.012)
More educated father	0.018	(0.015)	0.040***	(0.015)
Father less than HS	-0.012	(0.014)	-0.043***	(0.014)
Father any college	0.041*	(0.022)	0.059***	(0.022)
Panel B: Multiple linear regression				
Motherless	0.024	(0.056)	0.112**	(0.056)
Professional mother	0.006	(0.031)	0.055*	(0.032)
Older brother	0.049***	(0.013)	-0.025**	(0.012)
Older sister	0.004	(0.013)	0.011	(0.012)
More educated father	0.029	(0.035)	0.016	(0.036)
Father less than HS	0.022	(0.035)	-0.017	(0.036)
Father any college	0.036	(0.025)	0.035	(0.025)

Note: Data from Stockholm Birth Cohort. Each cell in Panel A reports the coefficient from a single univariate panel regression using demeaned variables (school mean). Each column in Panel B reports coefficients from a single multiple linear regression that uses demeaned variables (school mean). We use xtreg command in Stata 17 for these panel data regressions. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 2: Correlation of household characteristics on gender conformity

Household Characteristic	Gender Conformity Index Male	S.E.	Gender Conformity Index Female	S.E.
Panel A: Univariate linear regression				
Motherless	-0.018	(0.024)	-0.045*	(0.026)
Professional mother	-0.012	(0.013)	-0.037**	(0.014)
Older brother	-0.030***	(0.005)	0.011*	(0.006)
Older sister	0.001	(0.005)	-0.001	(0.006)
More educated father	-0.003	(0.006)	-0.023***	(0.007)
Father less than HS	0.002	(0.006)	0.023***	(0.007)
Father any college	-0.013	(0.010)	-0.031***	(0.010)
Panel B: Multiple linear regression				
Motherless	-0.019	(0.024)	-0.051*	(0.026)
Professional mother	-0.013	(0.013)	-0.035**	(0.015)
Older brother	-0.030***	(0.005)	0.011*	(0.006)
Older sister	0.001	(0.005)	-0.002	(0.006)
More educated father	-0.008	(0.015)	-0.017	(0.017)
Father less than HS	-0.008	(0.015)	0.002	(0.017)
Father any college	-0.012	(0.011)	-0.017	(0.012)

Note: Each cell in Panel A reports the coefficient from a single univariate panel regression using demeaning each variable by subtracting from it the school means. Each column in Panel B reports coefficients from a single multiple linear regression that demeans each variable by subtracting from it the school mean. (We used xtreg command in Stata for these panel data regressions). *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Correlations robustness. Here, to once more confirm that our results in Section 4 are not influenced by our binary definition of gender nonconformity vs. conformity, we use the underlying continuous gender conformity index as the measure of gender conformity. Equation 6 gives the specification using the gender conformity index:

$$\mathrm{y}_{is} = \alpha_0 + \omega_s + \mathrm{GCI}_{is}\beta + \mathbf{x}_{is}\alpha_1 + \epsilon_{is}. \tag{6}$$

The model is otherwise identical to equation 1 except for the main explanatory variable being the continuous gender conformity index instead of the binary gender nonconforming variable. The gender conformity index and the binary gender nonconforming variable are negatively correlated by definition. Hence, we expect the coefficients from equation 1 and equation 6 to move have opposite signs.

Table 3 presents the results using the gender nonconformity index as opposed to the dichotomous GNC variable. Table 3 shows that the results hold regardless of the decision to make the distributional cutoff below the twentieth percentile (recall that the variables GNC and GCI move in opposite directions; thus, the coefficients found in Table 3 have opposite signs to those presented in Tables 4, 5 and 6).

Table 3: Correlation of own gender conformity on life outcomes

	Gender Conformity Index Male	S.E.	Gender Conformity Index Female	S.E.
Educational outcomes				
GPA in grade 9	18.949***	(6.515)	-55.494***	(5.675)
Upper secondary dropout	-0.057	(0.043)	0.192***	(0.038)
STEM secondary track	0.444***	(0.038)	-0.135***	(0.039)
Any post secondary	-0.092**	(0.038)	-0.159***	(0.036)
Income outcomes				
Log income in 1980	0.431***	(0.097)	-0.141	(0.110)
Log income in 1990	0.297***	(0.056)	-0.082*	(0.049)
Log income in 2000	0.318***	(0.075)	-0.144**	(0.058)
Average log income 1990-95	0.300***	(0.055)	-0.116***	(0.043)
Average log income 1990-2001	0.310***	(0.058)	-0.131***	(0.043)
Career outcomes				
Full time in 1980	0.118***	(0.036)	-0.061	(0.041)
Part time in 1980	-0.055**	(0.025)	0.028	(0.036)
Not employed in 1980	-0.064**	(0.029)	0.034	(0.033)
Professional	0.028	(0.034)	-0.037	(0.025)
Occupational outcomes				
Legal or business	-0.087***	(0.033)	-0.060*	(0.031)
STEM	0.239***	(0.031)	-0.091***	(0.018)
Blue collar	0.067	(0.041)	-0.010	(0.024)
Clerical support	-0.074***	(0.019)	0.094***	(0.032)
Teacher-other health	-0.016	(0.022)	-0.038	(0.034)
Service and sales	-0.074***	(0.025)	0.068**	(0.028)
Did not work	-0.055**	(0.023)	0.038	(0.024)
Marriage outcomes				
Married by 1980	0.092**	(0.039)	-0.011	(0.040)
Divorced by 1980	-0.008	(0.012)	-0.037**	(0.017)
Mental health outcomes				
Mental health	-0.094***	(0.025)	0.036*	(0.021)
Addiction	-0.083***	(0.020)	-0.004	(0.013)
Birth outcomes				
Childless			0.038	(0.024)
Total births			0.034	(0.067)
Teenage mother			0.031**	(0.012)
Age at first birth			-0.495*	(0.290)

Note: Data from Stockholm Birth Cohort. This table provides the results of the regressions for equation 6. These regressions control for school fixed effects and sociodemographic background covariates such as parents' level of education, a dummy for family living in an owner-occupied home, size of the home, quality of the home, whether the home was overcrowded; moreover, these regressions include household variables that in our data (or typically in the literature) are found to correlate with gender conformity such as the existence of older siblings, whether a parent was missing from the household, and whether the mother worked or had a professional position.*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 4: Multinomial Logit Results: Average Marginal Effects

	Law-business	STEM	Blue collar	Clerical	Teacher-other health	Service-sales	Did not work
GCI Boys	-0.092**	0.246***	0.062	-0.063***	-0.023	-0.067**	-0.062**
	(0.031)	(0.033)	(0.041)	(0.017)	(0.021)	(0.024)	(0.021)
GCI Girls	-0.053	-0.081***	-0.011	0.095**	-0.045	0.065*	0.032
	(0.030)	(0.017)	(0.023)	(0.032)	(0.032)	(0.028)	(0.023)

	Part time in 1980	Full time in 1980	Not employed in 1980
GCI Boys	-0.041	0.102**	-0.060*
	(0.023)	(0.034)	(0.027)
GCI Girls	0.038	-0.072	0.034
	(0.035)	(0.039)	(0.033)

	Not Professional	Professional
GCI Boys	-0.016	0.016
	(0.033)	(0.033)
GCI Girls	0.041	-0.041
	(0.023)	(0.023)

Standard errors in parentheses

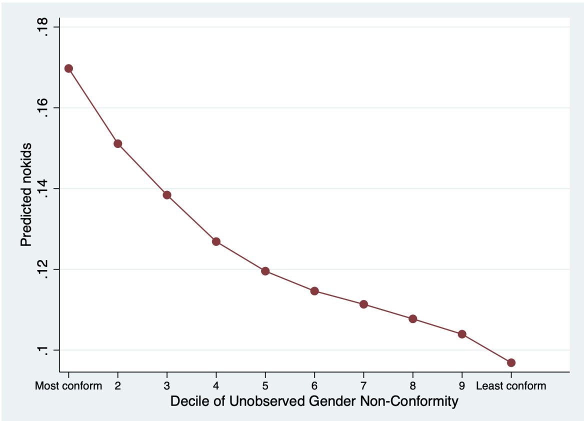
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 5: System of Manifest Variables and the Identification of Latent Cognitive Skills and Masculity Factor

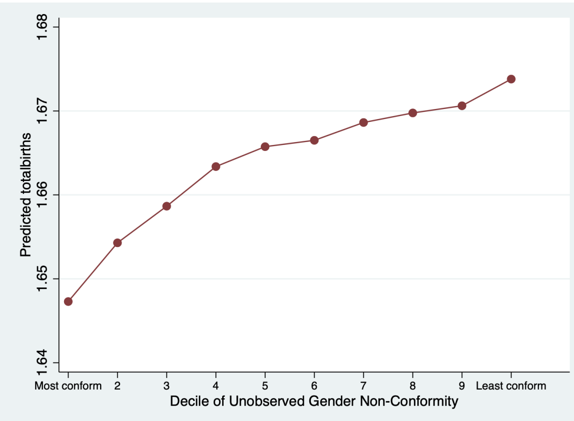
	Cognitive Scores			Male-biased Behaviors					
					(Males)			(Females)	
	Spatial	Verbal	Numeric	Outdoor	Domestic	Technical	Outdoor	Domestic	Technical
Cognitive	0.731***	0.711***	1						
	(0.020)	(0.019)							
Masculinity				5.435***	-0.147*	1	3.538***	-0.163***	1
				(0.463)	(0.084)		(0.175)	(0.040)	
Female	-1.532***	-0.160	-1.444***						
	(0.127)	(0.112)	(0.139)						
Female Head HHld	0.018	0.354	-0.294	0.298	-0.266	-0.361	-0.735*	-0.271	0.027
	(0.248)	(0.219)	(0.273)	(0.315)	(0.386)	(0.375)	(0.399)	(0.311)	(0.381)
FatherEduc: HS	1.638***	2.505***	2.437***	-0.504**	0.095	-0.071	-0.123	-0.686***	0.617**
	(0.175)	(0.155)	(0.193)	(0.226)	(0.264)	(0.257)	(0.272)	(0.224)	(0.273)
FatherEduc: College	2.131***	4.119***	4.384***	-0.649**	0.576*	-0.590*	0.144	-1.368***	1.635***
	(0.233)	(0.207)	(0.257)	(0.313)	(0.349)	(0.340)	(0.381)	(0.305)	(0.372)
Homeowner	0.562***	0.381**	0.273	-0.664***	-0.055	-0.013	-0.239	-0.750***	0.635**
	(0.169)	(0.150)	(0.186)	(0.222)	(0.256)	(0.249)	(0.278)	(0.218)	(0.267)
Constant	25.073***	24.936***	23.010***	37.901***	25.624***	36.968***	33.365***	34.591***	23.040***
	(0.207)	(0.184)	(0.228)	(0.122)	(0.123)	(0.120)	(0.136)	(0.103)	(0.127)
Obs.			10967			5403			5575

Figure 6: Gender Nonconformity and Fertility Outcomes

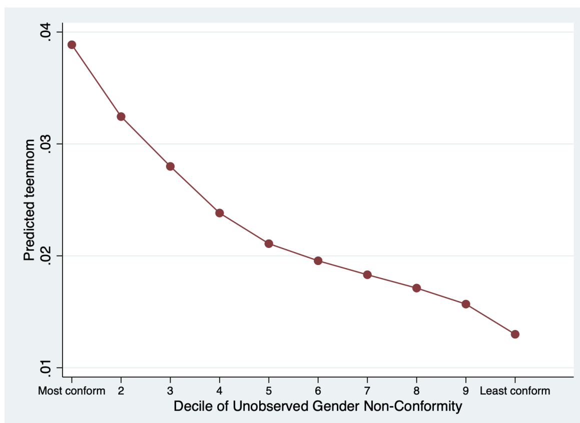
(a) Childless



(b) Total Births



(c) Teenage Mother



A.1 Additional Tables and Figures from Unobserved Heterogeneity Model

B Construction of Variables

B.1 Outcome Variables

Educational Variables

- GPA in grade 9 - This variable is a continuous variable from 100 (lowest) to 500 (highest) indicating the individual's "average marks in spring term of 9th form of elementary school."

- Upper secondary dropout - This binary variable indicates if an individual in this study dropped out of school before completing a minimum two year upper secondary education. The data contains education variables on the amount of schooling that an individual had completed by 2000. If the individual is listed as less than 11 years of schooling, then they are classified as having dropped out of upper secondary school.
- STEM secondary track - This binary variable indicates whether an individual listed a male dominated upper secondary school or vocational program as this first choice of study after having completed compulsory school. An individual who chose a STEM program is assigned as having a "STEM secondary track" choice of study.
- Any post secondary - This binary variable indicates whether an individual attended any form of post secondary study as of 1983. The data provides data on level of education that the individual had started or completed by 1983. If the individual is listed as any outcome besides "no post-upper secondary school course or program registered," then we assigned the individual as having "any post secondary" education.

Income Variables

- Log income in 1980 - This continuous variable states the log income of the individual in the study in the year 1980. The SBC data contained income information on the individuals in the study. To construct this variable, we took the log of the income listed in the year 1980.
- Log income in 1990 - This continuous variable states the log income of the individual in the study in the year 1990. The SBC data contained income information on the individuals in the study. To construct this variable, we took the log of the income listed in the year 1990.
- Log income in 2000 - This continuous variable states the log income of the individual in the study in the year 2000. The SBC data contained income information on the individuals in the study. To construct this variable, we took the log of the income listed in the year 2000.
- Average log income 1990-95 - This continuous variable states the average log income of the individual in the study from the years 1990 to 1995. The SBC data contained income information on the individuals in the study. To construct this variable, we took the log of the average of those yearly incomes listed.

- Average log income 1990-2001 - This continuous variable states the average log income of the individual in the study from the years 1990 to 2001. The SBC data contained income information on the individuals in the study. To construct this variable, we took the log of the average of those yearly incomes listed.

Career Variables

- Full time in 1980 - This binary variable indicates if the individual in the study worked full time. The SBC data had 1980 census data on the amount of hours that an individual in the study worked per week in 1980. If an individual worked 35 hours per week or more in 1980, then we classified them as "full time in 1980."

- Part time in 1980 - This binary variable indicates if the individual in the study worked part time. The SBC data had 1980 census data on the amount of hours that an individual in the study worked per week in 1980. If an individual worked less than 35 hours per week in 1980, then we classified them as "part time in 1980."

- Not employed in 1980 - This binary variable indicates if the individual in the study was not working in 1980. The SBC data had 1980 census data on whether the individual was employed in 1980. If an individual was not employed in 1980, then we classified them as "not employed in 1980."

- Professional - This binary variable indicates if the individual in the study worked as a professional in their career. The SBC data had census data on the socio-economical index of the individual. If the individual had a socio-economical index that listed them as "professional or other higher non-manual posts" or "self-employed professionals" , then we categorized them as being a "professional" laborer.

Occupational Variables

- Legal or business - This binary variable indicated whether an individual worked in a legal or business industry between 1970 and 1980. The SBC data had census data variables for the years 1970, 1975, and 1980 that indicated the three-digit industry code of a given individual's occupation, assuming they worked. An individual was assigned as having worked in the "legal or business" industry if they have a legal or business industry code listed in the given years.

- STEM - This binary variable indicated whether an individual worked in a STEM industry between 1970 and 1980. The SBC data had census data variables for the years 1970, 1975, and 1980 that indicated the three-digit industry code of a given individual's occupation, assuming they worked. An individual was assigned as having worked in the "STEM" industry if they have a STEM industry code listed in the given years.

- Blue collar - This binary variable indicated whether an individual worked in a blue collar industry between 1970 and 1980. The SBC data had census data variables for the years 1970, 1975, and 1980 that indicated the three-digit industry code of a given individual's occupation, assuming they worked. An individual was assigned as having worked in the "blue collar" industry if they have a blue collar industry code listed in the given years.

- Clerical support - This binary variable indicated whether an individual worked in a clerical support industry between 1970 and 1980. The SBC data had census data variables for the years 1970, 1975, and 1980 that indicated the three-digit industry code of a given individual's occupation, assuming they worked. An individual was assigned as having worked in the "clerical support" industry if they have a clerical support industry code listed in the given years.

- Teacher-other health - This binary variable indicated whether an individual worked in a educational or non-medical doctor health related position between 1970 and 1980. The SBC data had census data variables for the years 1970, 1975, and 1980 that indicated the three-digit industry code of a given individual's occupation, assuming they worked. An individual was assigned as having worked in the "teacher-other health" industry if they have a educational or non-medical doctor health related position industry code listed in the given years.

- Service and sales - This binary variable indicated whether an individual worked in a service and sales industry between 1970 and 1980. The SBC data had census data variables for the years 1970, 1975, and 1980 that indicated the three-digit industry code of a given individual's occupation, assuming they worked. An individual was assigned as having worked in the "service and sales" industry if they have a service and sales industry code listed in the given years.

- Did not work - This binary variable indicated whether an individual did not work between 1970 and 1980. The SBC data had census data variables for the years 1970, 1975, and 1980 that indicated if the individual worked. An individual was assigned "did not work" if the census data indicated they did not work in the given years.

Marriage Variables

- Married by 1980 - This binary variable indicates whether an individual in the study was married by 1980. The SBC data contains 1970, 1975, and 1980 census data on marriage outcomes. If an individual is listed as married, divorced, or widowed in the 1970, 1975, or 1980 census data, then the individual is classified as having been "married by 1980."

- Divorced by 1980 - This binary variable indicates whether an individual in the study was divorced by 1980. The SBC data contains 1970, 1975, and 1980 census data on marriage outcomes. If an individual is listed as divorced in the 1970, 1975, or 1980 census data, then the individual is classified as having been "divorced by 1980."

Mental Health Variables

- Mental health - This binary variable indicates whether an individual in the study was hospitalized for any mental health disorders between 1981-2008. The data come from inpatient records.
- Addiction - This binary variable indicates whether an individual in the study was hospitalized for any alcohol and/or drug related disease between 1981-2008. The data come from inpatient records.

Birth Variables The data come from the (Swedish) Medical Birth Register of 1984 and contain birth records for each of the female individuals up until 1983, but not male individual in the study.

- Childless - This variable is a binary variable that states whether the female individual in the study had children by age 40, close to completed fertility. If the individual was female, then she was assigned "childless" if she had zero births listed in the birth records.
- Total births - This variable states the number of children that female individuals in the study had. If the individual was female, then the number of births was tallied up based on summing the number of births listed in the birth records.
- Teenage mother - This is a binary variable that states whether the individual had a child in her teen years. These birth records listed the year of the birth. The age of the mother during her first birth was estimated by subtracting the year she was born from the year of her earliest delivery. Conditional on the individual having at least one child and her age at her first birth being less than 20, then we assigned her as a "teen mother."

- Age at first birth - This variable provides the age at which a female individual in this study was when she first gave birth conditional on her ever having had a child. These birth records listed the year of the birth. The age of the mother during her first birth was estimated by subtracting the year she was born from the year of her earliest delivery.

B.2 Household Variables

Sibling Variables

- Older brother - This variable is a binary variable that tells us whether the individual had at least one older brother. The SBC data contained a variable with the number of older brothers that an individual in the study had. We constructed the this variable by checking if the older brother variable was greater than 0.
- Older sister This variable is a binary variable that tells us whether the individual had at least one older sister. The SBC data contained a variable with the number of older sisters that an individual in the study had. We constructed the this variable by checking if the older sister variable was greater than 0.
- Multiple older brothers - This variable is a binary variable that tells us whether the individual had multiple older brothers. The SBC data contained a variable with the number of older brothers that an individual in the study had. We constructed the this variable by checking if the older brother variable was greater than 1.
- Multiple older sisters - This variable is a binary variable that tells us whether the individual had multiple older sisters. The SBC data contained a variable with the number of older sisters that an individual in the study had. We constructed the this variable by checking if the older sister variable was greater than 1.

Father Variables

- Fatherless - This variable is a binary variable that tells us whether there was a father present in the household. The SBC data provided 1960 census data with a variable that stated the "sex of the head of the household" as well as the "marital status of the head of the household." We assigned the variable as "fatherless" if the head of house was female and not married.
- More educated father - The SBC data provided 1960 census data on the amount of schooling that the individual's parents had completed at that point in time. This variable was constructed by checking if the father had completed more school than the mother.

- Father less than HS - This variable is a binary variable that tells us whether the individual's father did not finish high school. The SBC data provided 1960 census data on the amount of schooling that the individual's parents had completed at that point in time. This variable was constructed by checking if the father was listed as having completed high school.

- Father any college - This variable is a binary variable that tells us whether the individual's father attended college at all. The SBC data provided 1960 census data on the amount of schooling that the individual's parents had completed at that point in time. This variable was constructed by checking if the father was listed as having attended college.

Mother Variables

- Motherless - This variable is a binary variable that tells us whether there was a father present in the household. The SBC data provided 1960 census data with a variable that stated the "sex of the head of the household" as well as the "marital status of the head of the household." We assigned the variable as "motherless" if the head of house was male and not married.

- Professional mother - This variable is a binary variable that tells us whether the mother worked outside the home in a professional position. The SBC data contained 1960 census data on the "type of occupation of wife of head of household." We assigned the variable as "professional mother" if the wife was coded as a "professional" conditional on her being being present in the household and being "economically active."

- Working mother - This variable is a binary variable that tells us whether the mother worked outside the home. The SBC data contained 1960 census data on the "economic activity of the wife of the head of household." We assigned the variable as "working mother" sif the wife was "active" conditional on there being a wife in the household.

- Female head of household - This variable is a binary variable that tells us whether the head of the household that the individual in the study lived in was female. The SBC data provided 1960 census data with a variable that stated the "sex of the head of the household." This variable was constructed by checking if the head of the household was listed as female.

- More educated mother - The SBC data provided 1960 census data on the amount of schooling that the individual's parents had completed at that point in time. This variable was constructed by checking if the mother had completed more school than the father.

- Mother less than HS - This variable is a binary variable that tells us whether the individual's mother did not finish high school. The SBC data provided 1960 census data on the amount of schooling that the individual's parents had completed at that point in time. This variable was constructed by checking if the mother was listed as having completed high school.
- Mother any college - This variable is a binary variable that tells us whether the individual's mother attended college at all. The SBC data provided 1960 census data on the amount of schooling that the individual's parents had completed at that point in time. This variable was constructed by checking if the mother was listed as having attended college.

C Identification of Latent Factors

The type of models we use in this paper can be described as a set of measurement systems that are linked by a factor structure.[13] We start with a measurement system which we use to identify the distributional parameters of the unobserved factors. The measurement system would have the following form:

$$\mathbf{T} = \mathbf{X}_T \beta^T + \alpha^{\mathbf{T},\mathbf{A}} \theta^A + \mathbf{e}^{\mathbf{T}} \tag{7}$$

where $\mathbf{T}$ is a $L \times 1$ vector of measurements (e.g., test scores, measures of behaviors), $\mathbf{X}_T$ is a matrix with all observable controls for each measurement and $\alpha^{T,A}$ is a vector containing the loadings of unobserved factor A in each mean measurement T. We assume that $\left(\theta^A, \mathbf{X}_T\right) \perp \mathbf{e}^{\mathbf{T}}$, that all the elements of the $L \times 1$ vector $\mathbf{e}^{\mathbf{T}}$ are mutually independent and have associated distributions $f_{e^h}(\cdot)$. To explain how the parameters of the measurement system (7) are identified, let us focus on the matrix $COV\left(\mathbf{T} \,|\mathbf{X}_T\right)$ whose elements in the diagonal are of the form:

$$COV\left(T_i, T_i \,|\mathbf{X}_T\right) = \left(\alpha^{T_i,A}\right)^2 \sigma^2_{\theta^A} + \sigma^2_{e^{T_i}}$$

and the off-diagonal elements are of the form:

$$COV\left(T_i, T_j \,|\mathbf{X}_T\right) = \alpha^{T_i,A} \alpha^{T_j,A} \sigma^2_{\theta^A}$$

As it is, the model is underidentified. To see this, note that the is no way to use the off-diagonal elements of $COV\left(\mathbf{T} \,|\mathbf{X}_T\right)$ to come up with unique values for the parameters we intend to estimate. More precisely, note that

$$\frac{COV\left(T_2, T_3 \,|\mathbf{X}_T\right)}{COV\left(T_1, T_2 \,|\mathbf{X}_T\right)} = \frac{\alpha^{T_3,A}}{\alpha^{T_1,A}}, \quad \frac{COV\left(T_2, T_3 \,|\mathbf{X}_T\right)}{COV\left(T_1, T_3 \,|\mathbf{X}_T\right)} = \frac{\alpha^{T_2,A}}{\alpha^{T_1,A}}, \quad \frac{COV\left(T_1, T_3 \,|\mathbf{X}_T\right)}{COV\left(T_1, T_2 \,|\mathbf{X}_T\right)} = \frac{\alpha^{T_3,A}}{\alpha^{T_2,A}}$$

Therefore, identification requires some assumptions. First, we acknowledge that latent factors have no metric or scale of their own [Bartholomew et al., 2011]. Hence, we need to normalize to unity one loading, and the remaining loadings should be interpreted as relative to the one used as nummeraire. If, with out loss of generality, we normalize $\alpha^{T_3,A} = 1$, the remaining loadings are identified from the quotients of the off-diagonal elements of $COV\left(\mathbf{T} \,|\mathbf{X}_T\right)$. This also shows that identification requires $L \geq 3$. That

[13] This Appendix follows closely the argument put forth in [Sarzosa and Urzua, 2016].

is, we need at least three scores per factor.[14] Having identified the loadings, we can further use the off-diagonal elements of $COV\left(\mathbf{T}\,|\mathbf{X}_T\right)$ to identify the factor variance σ_{θ^A} and the diagonal elements of $COV\left(\mathbf{T}\,|\mathbf{X}_T\right)$ to identify σ_{θ^A}.

Now that we have identified all the loadings, factor variances and measurement residual variances, together with the fact that the means of θ^A, θ^B and $\mathbf{e^T}$ are finite—in fact, equal to zero because we allow the measurement system (7) to have intercepts—we can invoke the Kotlarski Theorem to use the manifest variables $\mathbf{T}$ to non-parametrically identify the distributions of $f_{\theta^A}\left(\cdot\right)$ [Kotlarski, 1967].[15]

With the distribution of the factors in hand, we can consider a model linking the outcome variables we observe with the factor structure. That is,

$$y = \mathbf{x_Y}\beta^Y + \alpha^Y\theta + e^Y \tag{8}$$

where e^Y is a vector of error terms with distributions $f_{e^y}\left(\cdot\right)$. We assume that $e^Y \perp (\theta, \mathbf{x_Y})$. This is the type of models considered, for instance, by Heckman et al. [2006] and Urzua [2008]. Note that the econometrician does not observe the actual value of θ for each observation. Instead, in the first stage, s/he estimates the distributions they are drawn from and uses it to integrate it out θ in (8).

[14] Analogously, we can also infer the minimum number of scores by acknowledging that we can use the $\frac{L(L-1)}{2}$ off-diagonal elements to identify $L-1$ loadings—taking into account we normalized one loading—and the factor variance [Carneiro et al., 2003] . We then need that $\frac{L(L-1)}{2} \geq (L-1)+1$. Thus $L \geq 3$.

[15] The basic idea of the Kotlarski Theorem is that if there are three independent random variables e_{T_1}, e_{T_2} and θ and define $T_1 = \theta + e_{T_1}$ and $T_2 = \theta + e_{T_2}$, the joint distribution of (T_1, T_2) determines the distributions of e_{T_1}, e_{T_2} and θ, up to one normalization. Note that, given that we have already identified all the loadings, we can write (7) in terms of $T_\tau = \theta + e_{T_\tau}$ by dividing both sides by the loading. See more details in Carneiro et al. [2003].