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

DISCUSSION PAPER SERIES

IZA DP No. 13470

Gender Stereotyping in Sports

Miriam Marcén
Marina Morales
Almudena Sevilla

JULY 2020

I Z A Institute
of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 13470

Gender Stereotyping in Sports

Miriam Marcén
University of Zaragoza

Marina Morales
University of Zaragoza

Almudena Sevilla
UCL and IZA

JULY 2020

ISSN: 2365-9793

IZA – Institute of Labor Economics

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

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

www.iza.org

ABSTRACT

Gender Stereotyping in Sports*

This paper contributes to the literature of gender differences in academic attainment by putting together several sources of data going back several decades to investigate how gender stereotypes and parental time investments shape sport choices of boys and girls during high school. Using data from the 2002-2019 National Federation of State High School Association, which provides information for every state on the total number of high school participants by gender in each sport, we document that states with more gender-equal norms are also states where boys and girls tend to break stereotypes when making sport choices in high school. We also identify parental time investments as being an important cultural-transmission mechanism.

JEL Classification: J10, J16, J18

Keywords: gender, stereotypes, bias, sports

Corresponding author:
Almudena Sevilla
University College London
55-59 Gordon Square
London WC1H 0NT
United Kingdom
E-mail: A.Sevilla@ucl.ac.uk

* The authors declare that they have no relevant or material financial interests that relate to the research described in this paper.

Gender differences in academic achievement have dramatically reversed in the last decades. In the United States whereas in the 1960s there were 1.6 men for every woman graduating from four-year colleges, there are now 1.35 women for every man (Goldin et al., 2006). Yet there still remain important gender differences in educational attainment that seem to be persistent over time. For example, girls continue to perform relatively worse than boys in math tests, particularly at the top of the ability distribution (Guiso et al., 2008; Fryer and Levitt, 2010; Pope and Sydnor, 2010). A much lesser understood phenomenon is how the practice of sports while in high school differs by gender. The economic literature has documented positive causal effects on later-life economic outcomes from the participation in sports during high school (Stevenson, 2010). Beyond the direct physiological benefits, sports can foster the acquisition of important skills such as the ability to cooperate, compete, and team work, which are likely to be valued in the market later on. This paper contributes to the literature of gender differences in academic attainment by putting together several sources of data going back several decades to investigate how gender stereotypes and parental time investments shape sport choices of boys and girls during high school.

Following the passage of Title IX in 1972, which required schools to provide equal access to all sport activities by 1978, the number of high-school girls participating in sports as a percentage of female high-school enrollment increased ten-fold from close to 3 in 100 girls in 1972 to almost 30 in 100 girls in 1978 (Stevenson, 2007). However, the increase in female participation in sports was not homogeneous across all sports. Although the legislation did not make any stipulation as to the type of sports to be taken on by girls, girls stayed away from highly popular male-dominated sports such as football and baseball, and instead new sports emerged such as softball and volleyball that rapidly became female-dominated. At the same time, the number of boys participating in less popular sports such as field hockey and

gymnastic dramatically dropped following the sharp rise in participation by girls (Stevenson, 2007).

Using data from the 2002-2019 National Federation of State High School Association, which provides information for every state on the total number of high school participants by gender in each sport, we construct a Gender Stereotype Defier (GSD) sports index to capture the share of boys and girls practicing sports dominated by the opposite sex by state. Whereas there is a constant ratio of 7 girls for every 10 boys playing sports over this period, the GSD sports index reveals a high degree of specialization in the choice of sport by sex. Athletes are 37 times more likely to play a sport dominated by their own sex than are athletes from the opposite sex.

We also document large cross-state differences in the GSD sports index. In the state with the largest GSD sports index boys (girls) are 6 times more likely than girls (boys) to play a male (female)-dominated sport, whereas in the states with the lowest GSD index (Alabama and South Carolina) hardly any athletes play sports dominated by the opposite sex. This cross-state variation in the rates at which boys and girls participate in a sport that is dominated by the opposite sex have remained quite persistent over time. It is possible that differences in physical capabilities between boys and girls, which have been shown to emerge at the age of 12, could explain the lack of convergence across states over this period (McKay et al., 2017). However, physical differences between boys and girls should be the same regardless of the state, and thus seem unlikely to be able to explain the large cross-state variation in the GSD sports index documented here. Additionally, given the fact that sports competitions are usually single-sex, it is very unlikely that comparative advantage in physical abilities can drive the gendered pattern in sports choice documented here.

Using several questions from the 1972-2018 General Social Survey (GSS) on attitudes towards women, indicators on the status of women in society from the Institute for

Women's Policy Research (IWPR) for the period 1989-2006, as well as objective measures of labor and non-labor market outcomes from the 2002-2018 American Community Survey (ACS) and the 2002-2018 Current Population Survey (CPS), we document that states with more gender-equal norms and where the relative position of women is relatively better are also states where boys and girls tend to break stereotypes when making sport choices in high school.

We also identify parental time investments as being an important cultural-transmission mechanism through which gender stereotypical patterns in the choice of sports across US states are maintained. We use the 2003-2018 America Time Use survey, which records detailed information on individuals' activities for 24-hour of the previous day. A particular advantage of the ATUS over other time diary surveys is that parents record the time they spend with each child in the household. We implement a siblings fixed-effect estimation strategy and document that whereas fathers spend more time with sons, the gender gap in father's time is halved for parents living in states with a higher GSD sports index. This is particularly so for time spent in recreational child care, which includes playing sports with children as well as attending events. This type of child care is particularly important during middle childhood as children's lives extend beyond the family to include peers, when parent's role in arranging for human-capital enhancing extracurricular academic, recreational, and social activities become more important (Kalil, 2012).

This paper is organized as followed. Section I presents variation in the GSD sports index across states. Section II presents the correlation between the GSD sports index with subjective and objective indicators of the position of women in society. Section III looks at parental time investments as a driving cultural-transmission mechanism in sport choice. Section IV concludes.

I Gender Stereotype Defier (GSD) Sports Index

We use publicly available data from the 2002/2003 to 2018/2019 academic years from the National Federation of State High School Associations (NFHS), which collects and publishes on-line information on the number of players in each sport by gender for each state over time.[2] Each of their 51 member state associations (50 states plus District of Columbia) is responsible for gathering information on high school sports from individual schools, covering about 80 per cent of the total students enrolled in high school in the U.S. Our sample consists of 128,294,593 high school students, about three million girls and four and a half million boys playing 91 sports across 19,500 schools over this period.[3]

For each state, we construct a GSD sports index that captures the relative share of girls doing male-dominated sports and the relative share of boys doing female-dominated sports, as follows:

$$GSD_j = \left[\sum_k^{m,f} GSD_j^k\right]/2 \quad (1)$$

where *j* refers to state and, *k* refers to either male-dominated (*m*) or female-dominated sport (*f*). We consider a sport to be female-dominated if over the analyzed period the proportion of girls playing a sport (over all the players in that sport) is over 80 per cent, and male-dominated if the proportion of boys playing a sport (over all the players in that sport) is over

[2] https://members.nfhs.org/participation_statistics

[3] Sports listed by the NFHS are Adapted Basketball, Adapted Bocce (Indoor), Adapted Bowling, Adapted Floor Hockey, Adapted Football, Adapted Soccer, Adapted Softball, Adapted Track, Adapted Volleyball, Adaptive Corn Toss, Adaptive Golf, Adaptive Handball, Adaptive Strength Training, Adaptive Tennis, Air Riflery, Archery, Badminton, Baseball, Basketball, Bass Fishing, Beach Volleyball, Bocce (Outdoor), Bowling, Canoe Paddling, Canoeing, Competitive Spirit Squad (Boys who cheer/Girls who cheer), Crew, Cross Country, Cycling, Dance, High Kick, Jazz, Dance/Drill, Decathlon, Drill Team, Equestrian, Fencing, Field Hockey, Figure Skating, Flag Football, Football (11 player), Football (6 player), Football (8 player), Football (9 player), Golf, Gymnastics, Heptathlon, Ice Hockey, Judo, Kayaking, Lacrosse, Martial Arts, Mixed 6-Coed Volleyball, Mt. Biking, Native Youth Olympics, None, Outrigger Canoe Paddling LL, Pentathlon, Rugby, Riflery, Rock, Climbing, Rodeo, Roller Hockey, Rhythmic Gymnastics, Sand Volleyball, Skiing (Alpine), Skiing (Cross Country), Snowboarding, Soccer, Soft Tennis, Softball (Fast Pitch), Softball (Slow Pitch), Squash, Surfing, Swimming and Diving, Synchronized Swimming, Team Tennis, Tennis, Track and Field (Indoor), Track and Field (Outdoor), Trap Shooting, Ultimate Frisbee, Unified Basketball, Unified Flag Football, Unified Track and Field (Outdoor), Volleyball, Water Polo, Weight Lifting, Wrestling, Sailing, Other.

80 per cent. We choose a national cut-off of 80 per cent as a conservative threshold, and our results are robust to thresholds of 70 per cent and 60 per cent (See Tables A.1 and A.2). Considering a national-level cut-off, as opposed to state-level thresholds makes sense since professional leagues are national labor markets. Out of the 91 sports listed by the NFHS over this period, there are 15 female-dominated sports (dance, dance team (high kick), dance team (jazz), dance/drill, field hockey, cheer leader, drill team, equestrian, figure skating, flag football, gymnastics, synchronized swimming, volleyball, heptathlon and softball), and 11 male-dominated sports (American football (6, 8, 9 or 11 players), baseball, rugby, bass fishing, ice hockey, adaptive golf, native youth Olympics and wrestling).

For each state, GSD_j^m is constructed as follows:

$$GSD_j^m = \left(\frac{\frac{\sum_{i=1}^{NF_j} I_{i,j}^m}{NF_j}}{\frac{\sum_{i=1}^{NM_j} I_{i,j}^m}{NM_j}} \right) \quad (2)$$

where NF_j and NM_j are the number of girls and boys in our sample who play sports in high school in state j over this period. $I_{i,j}^m$ takes value 1 if an individual i plays a male-dominated sport and 0 otherwise. The numerator is the share of girls who play a male-dominated sport (relative to the total number of girls playing sports). The denominator is the share of boys who play a male-dominated sport (relative to the total number of boys playing sports). Higher values of GSD_j^m represent breaking with stereotypes in the choice of sports either as a result of more girls playing male-dominated sports, or fewer boys playing male-dominated sports.

Similarly, GSD_j^f is constructed as follows:

$$GSD_j^f = \left(\frac{\frac{\sum_{i=1}^{NM_j} I_{i,j}^f}{NM_j}}{\frac{\sum_{i=1}^{NF_j} I_{i,j}^f}{NF_j}} \right) \quad (3)$$

where $I_{i,j}^{f}$ takes value 1 if an individual i plays a female-dominated sport and 0 otherwise. The numerator is now the share of boys who play a female-dominated sport (relative to the total number of boys playing sports), and the denominator is the share of girls who play a female-dominated sport (relative to the total number of girls playing sports). Higher values of GSD_j^f represent breaking with stereotypes in the choice of sports either as a result of more boys playing female-dominated sports, or fewer girls playing female-dominated sports.

We construct the GSD_j sports index as an average of GSD_j^m and GSD_j^f. Values of the GSD_j sports index closer to 1 indicate a higher probability that girls and boys break stereotypical gender patterns in the choice of sport. Closer values to 1 may either result from the share of girls playing male-dominated sports being similar to the share of boys playing male-dominated sports, i.e. $GSD_j^m = 1$, or from the share of boys playing female-dominated sports being similar to the share of girls playing female-dominated sports, i.e. $GSD_j^f = 1$.[4] The values of the GSD sports index range from 0 to 0.17, with an average of 0.027 and standard deviation of 0.032. The average of the GSD sports index is far from 1 at a value of 0.027, indicating that the share of boys(girls) playing a male(female)-dominated sport is 37 times the share of girls(boys) playing a male(female)-dominated sport.[5]

II Gender Norms and Stereotyping in Sport

Figure 1 shows that the national average of the GSD sports index marks high level of heterogeneity across states. An *F*-test rejects the null hypothesis that the rates at which boys and girls participate in sports dominated by the opposite gender are the same across states,

[4] GSD_j^m and GSD_j^f are highly correlated with the GSD_j sports index with a Pearson correlation coefficients of 0.70 and 0.88 respectively.

[5] For example, if we focus on male-dominated sports, then $GSD_j^m = 0.027$ translates into $\frac{\sum_{i=1}^{NM_j} I_{i,j}^m}{NM_j} = 37 \frac{\sum_{i=1}^{NF_j} I_{i,j}^m}{NF_j}$, where 37=1/0.027.

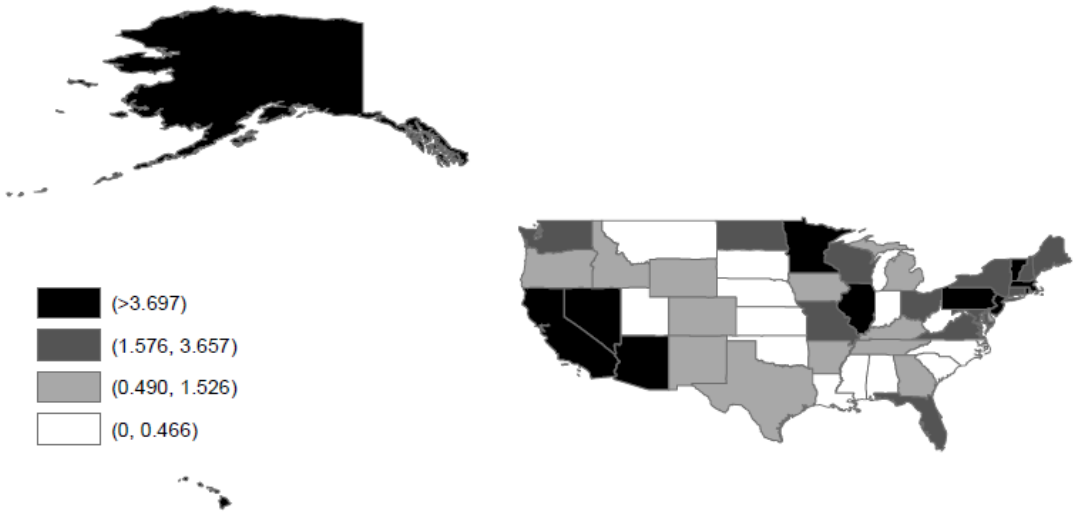
with *p*-values below 0.05 in every case.[6] At the 95th percentile, the state with the largest GSD sports index is Hawaii (0.17), where boys (girls) are 6 times more likely than girls (boys) to play a male (female)-dominated sport. The two states with the lowest value GSD sports index take are Alabama and South Carolina, followed closely by West Virginia and Indiana, where hardly any children play sports dominated by the opposite sex. There also seem to be geographical clusters of states that are more likely to break with stereotypical gender choices in sports, as shown by the darker areas in Figure 1. These regions are the West, Southwest and Northeast. In contrast, the lighter areas in Figure 1 coincide with the South and Mountain West regions, where high school children are less likely to break stereotypical gender patterns when practicing sports.

Despite large cross-state differences in the GSD sports index, the rates at which boys and girls participate in a sport that is dominated by the opposite gender remain quite persistent over time, with no sign of convergence across states over this period. Formally, analyses of the R^2 resulting from regressions that relate the GSD sports index to state and year fixed effects shows that additionally controlling by the interaction of state and year dummies can account for about an additional 2 per cent of the variation over time in state level variation in the GSD sports index. [7]

Figure 1: Gender Stereotype Defier (GSD) Sports Index across US States

[6] The F-test for the equality test among the GSD sports index (by year) across states is 111.84 with a p-value below 0.01.

[7] The R^2 of regressions that relate the state-level GSD sports index to state and year fixed-effects only yield an R^2 of 0.887, and adding the interaction of state and year fixed effects increases the R^2 to 0.908.



Notes: Labels represents four GSD sports index quartiles. Darker shades indicate a higher GSD sports index. The values of the GSD sports index are multiplied by 100 for ease of exposition.

Alabama	0	Iowa	0.491	Washington	1.576	Pennsylvania	3.698
South Carolina	0	Idaho	0.674	Missouri	1.992	Minnesota	3.838
North Carolina	0.033	Wyoming	0.679	Maryland	2.143	Arizona	4.100
West Virginia	0.048	Michigan	0.814	Virginia	2.199	New Jersey	4.142
Indiana	0.068	Kentucky	0.932	North Dakota	2.401	Nevada	5.713
Louisiana	0.134	New Mexico	0.945	DC	2.558	Illinois	5.879
Mississippi	0.157	Arkansas	0.953	Wisconsin	2.660	Massachusetts	6.525
Utah	0.166	Texas	1.025	Florida	2.721	Alaska	6.778
Montana	0.234	Tennessee	1.081	New Hampshire	3.087	California	8.397
Kansas	0.369	Georgia	1.086	Connecticut	3.129	Rhode Island	9.351
South Dakota	0.391	Colorado	1.124	Ohio	3.196	Vermont	9.854
Oklahoma	0.412	Oregon	1.392	Maine	3.643	Hawaii	17.143
Nebraska	0.466	Delaware	1.526	New York	3.657		

It is possible that comparative advantage considerations in physical abilities that differ between boys and girls, which have been shown to emerge at the age of 12, could explain the lack of convergence across states over this period (McKay et al., 2017). However, physical innate abilities between boys and girls should be the same regardless of the state, and thus seem unlikely to be able to explain the large cross-state variation in the GSD sports index documented here. Additionally, given the fact that girls and boys compete against athletes of the same sex, makes an explanation based on comparative advantage considerations less likely and suggests the presence of relatively constant state-level factors behind the state variation in GSD. The GSD sport index is highest in the state where boys and girls are most likely to break gender stereotypes in the choice of sport (Hawaii, 0.17). If gender equality in

sport choice is captured by a value of the GSD sports index of 1, then at least 17 per cent (1-0.17)-(1-0.027))/(1-0.027) of the gender stereotypical sports choices can be explained by these cross-state factors.

We next investigate how gender norms about the position of women in society may relate to gender stereotyping in sports choice. Using several questions from the 1972-2018 General Social Survey (GSS) we construct the share of individuals that strongly agree with a gender-equal statement (or strongly disagree with a non gender-equal statement) in each of the nine US regions for which the GSS is publicly available by calculating.[8] On average 50 per cent percent of respondents display gender-equal attitudes in the US over the 1972-2018 period, consistent with findings in the literature (Charles et al., 2019). Second, we use indicators on women's social and economic autonomy, political participation, women's reproductive rights, and health and well-being from the Institute for Women's Policy Research (IWPR) for the period 1989-2006.[9]

[8] https://gss.norc.org/get-the-data is publicly available for 9 US regions New England, Middle Atlantic, East north Central, West north Central, South Atlantic, East south Central, West. south Central, Mountain, and Pacific. We use the following GSS Questions: (1) Do you approve of a married woman earning money in business or industry if she has a husband capable of supporting her? (Answer Approve: coded as gender-equal attitudes=1); (2) If your party nominated a woman for president, would you vote for her if she were qualified for the job? (Answer Yes: coded as gender-equal attitudes=1); (3) Do you agree or disagree with this statement? Women should take care of running their home and leave running the country up to men. (Answer disagree: coded as gender-equal attitudes=1) (4) Most men are better suited emotionally for politics than are most women. (Answer disagree: coded as gender-equal attitudes=1) (5) A working mother can establish just as warm and secure a relationship with her children as a mother who does not work. (Answer strongly agree: coded as gender-equal attitudes=1) (6) A preschool child is likely to suffer if his or her mother works (Answer strongly disagree: coded as gender-equal attitudes=1); (7) It is more important for a wife to help her husband's career than to have one herself (answer strongly disagree: coded as gender-equal attitudes=1); (8) It is much better for everyone involved if the man is the achiever outside the home and the women takes care of the home and family (answer strongly disagree: coded as gender-equal attitudes=1). Results do not change when we consider the dummy to take value 1 if a respondent strongly agrees/agrees with a gender-equal statement (or strongly disagrees/disagrees with a non-gender-equal statement) in those questions in which both alternatives are available.

[9] These indicators can be downloaded from https://iwpr.org/tools-data/data-for-researchers/status-women-states-data/. A detailed description of how these indicators are constructed can be found at https://iwpr.org/wp-content/uploads/wpallimport/files/iwpr-export/publications/appendices.pdf. We average across the years in which the information is available for each index as follows: social and economic autonomy (1989-2005), political participation (1992-2004), reproductive rights (1996-2004) and health and well-being (1991-2002). Indicators capture how far a state is from reaching equality. Equality in women's status in the political participation area is achieved in a state: when women's voter registration and voter turnout are set at the value of the highest state for these components; when 50 percent of elected positions are held by women; and when a

Panel A in Table 1 provides the coeffificient estimates from a simple linear regression of the state-level gender norms indicators on the state-level GSD sports index. Women appear to do better in states with a higher GSD sports index. Column 1 shows the estimated effect of the share of individuals with gender-equal attitudes on the GSD variable. The coefficient of 0.005 indicates that the difference the proportion of individuals with equal-gender attitudes between two states where the GSD sports index varies by one standard deviation (representing approximately the difference between living in Illinois rather than Florida, or in Wisconsin rather than Alabama) is 1.6 percentage points. Similarly, columns 2-5 show that a one standard deviation increase in the GSD sports index is positively related to the status of women in society for all the indicators considered, explaining between 26 percent and 60 percent of the standard deviations of the indicators of the status of women.[10] Looking at the R^2s the state-level GSD sports index accounts between 7 percent and 37 percent of the variation in gender norms. We check the robustness of our estimates to outliers such as Hawaii. Results do not change (see Appendix A.3).

We next look at how the GSD sports index is associated with objective measures of the position of women in society. To that end we use information from the 2002-2018 American Community Survey (ACS) and the 2002-2018 Current Population Survey (CPS) to construct state-level variables of labor force participation gender gaps, the (log) wage gender

---

state has both a commission for women and a women's legislative caucus in each house of the state legislature. In the case of the social and economic autonomy, equality is considered: when a state achieves the highest value for all states in the percentage of women with health insurance; when the percentage of women with higher education achieves that of men at the national level; when the percentage of businesses owned by women are set as if 50 percent of businesses were owned by women; and when the percentage of women in poverty are equal to that of men at the national level. For the reproductive rights index equality takes place when a state assumes to have: no notification/consent or waiting period policies; public funding for abortion, prochoice government, 100 percent of women living in counties with an abortion provider, insurance mandates for contraceptive coverage and infertility coverage, maximum legal guarantees of second-parent adoption, and mandatory sex education for students. The health and well-being index consider equality in a state when: mortality rates (from heart disease, lung cancer, breast cancer, and suicide), the incidence of some diseases (diabetes, chlamydia, and AIDS), and the mean days of poor mental health and mean days of activity limitations are equal to the national goal, and in the absence of goals to the level of the best state among all states.

[10] For example for Political participation: 3.2 (GSD sdx100) x 0.327 (coef.)=1.0464; 1.0464/4.029 (sd Political participation indicator)=0.259 (approx.26%)

gap, the share of females never married, and the average female age at first child, and to see whether they are related to the level of gender stereotypes in sport participation.[11] These variables have been shown to be negatively correlated with the level of sexism in a state (Charles et al., 2019). Labor market outcomes are estimated using a sample of natives aged 25-64 and non-labor market outcomes are for native women aged 20 to 40.

Panel B of table 1 presents the result from an OLS regression. States with higher values of the GSD sports index have lower gender wage, albeit non-statistically significant, and labor force participation gaps, and women marry and have a first child at a later age. In particular, an increase of one standard deviation in the GSD sports index in a state is associated with a 6.4 percent decrease in the gender gap in labor force participation. Similarly, columns 3-4 show that comparing two adult women living in two states where the GSD sports index varies in one standard deviation (representing approximately the difference between living in Illinois rather than Florida, or in Wisconsin rather than Alabama), a woman living in a state with the highest GSD sports index is 3.2 percentage points less likely to be married, and bears her first child more than a quarter year later.

III Gender Stereotypes in Sport Choice and Parental Time Investments

This section looks at whether differences in parental time investments are related to stereotypical gender patterns in sports choice. To that end we pool data from the 2003-2018 America Time Use survey, which records detailed information on individuals' activities for 24-hour of the previous day. A particular advantage of the ATUS over other time diary

[11] The labor force participation gap is constructed as the difference between the percentage of females in labor force and the percentage of males in labor force in each state. The wage gender gap is constructed as the difference between the average female wages and the average male wages (conditional on working). The share of females never married is calculated as the percentage of females never married by state and the average age at first birth by state is obtained using information on how old a woman was when her first child was born from the reported age of her eldest child living in the same household. We use ACS weighting.

surveys is that parents record the time they spend with every child living in the household, and the activity they engage with. Together with the information on the child's gender, we can construct the time that boys and girls receive from parents as the sum of all minutes per day spent in parental activities with the child as primary activity.[12] Our main sample includes parents between 21 and 55 years old with at least one child aged 6 to 11 living in the household. We focus on children before the high school years because two main reasons. First, parental time investments are more important during this period than during adolescence, when children become autonomous and child's own investments matter more than that of the parents (Del Boca et al., 2017). Second, we want to make sure that parental time does not capture parent's reactions to the differential rates of physical development for boys and girls after the child is 12 (McKay et al., 2017).

As in Guryan et al. (2008) we define "child care" as the sum of three primary time use components. *Basic child care* is time spent on the basic needs of children, *Educational child care* includes reading to/with children and helping children with homework, and *Recreational child care* involves playing with children and attending children's events.[13]

Panel C and D present the results from a siblings-FE model of the time that parents spent with a child on the GSD sports index for fathers and mothers separately.[14] Results from

[12] The information on the gender of the child is limited to children who are classified as household members. We cannot use information on child care of non-household members. Results are maintained when we use a sample of married individuals who are supposed to be less likely to have non-household children than those divorced or separated individuals.

[13] Categories of the time use survey are described as follows, where children refer to household children only. Basic child care: Physical care for hh children, Organization and planning for hh children, Looking after hh children (as a primary activity), Waiting for/with hh children, Picking up/dropping off hh children, Caring for and helping hh children, n.e.c, Activities Related to Household Children's Health, Providing medical care to hh children, Obtaining medical care for hh children, Waiting associated with hh children's health, Activities related to hh child's health, n.e.c.. Recreational child care is defined incorporating: Playing with hh children, not sports, Arts and crafts with hh children, Playing sports with hh children, and Attending hh children's events. Educational child care includes: Reading to/with hh children, Talking with/listening to hh children, Activities Related to Household Children's Education, Homework (hh children), Meetings and school conferences (hh children), Home schooling of hh children, Helping or teaching hh children, Waiting associated with hh children's education, and Activities related to hh child's education, n.e.c.

[14] In particular, we estimate: $Y_{ijs} = \alpha_1 \text{female}_{j,s} + \alpha_2 \text{female}_{j,s} * \text{GSD}_s + x_{j,s} + U_{i,s} + \varepsilon_{ij}$

an OLS regression model are qualitatively the same (see Table A.4). The coefficient on the female dummy in the first row of Panel C shows that fathers spend around 9 minutes less per day with daughters than with sons. The gender difference in father's time is economically meaningful representing a 18 per cent decrease in father's time for daughters with respect to sons, and it holds for the three kinds of parental time investments considered here. Yet, fathers living in states with a higher GSD sports index spend more time with daughters (relative to sons) than fathers living in states with a lower GSD sports index. This is particularly so for time spent in basic care and recreational activities. In particular, a standard deviation increase (approximately the difference between living in Illinois rather than Florida, or in Wisconsin rather than Alabama) increases the time fathers spend with their daughters (relative to sons) by 3.5 minutes per day, reducing the gender gap in paternal time by almost half. A big proportion of fathers increase in time with daughters relative to sons is concentrated in recreational child care, which includes playing sports with children as well as attending events. This type of child care is particularly important during middle childhood as children's lives extend beyond the family to include peers, when parent's role in arranging for human-capital enhancing extracurricular academic, recreational, and social activities become more important (Kalil, 2012).

IV Conclusion

where *i* denotes father (mother), *j* denotes child and s indicates state. Y_{ijs} are minutes per day that a father (mother) spends with child *j*. $female_{j,s}$ is an indicator equal to one if the child *j* is a girl and zero otherwise. GSD_s is the gender stereotypical defier sports index in state s, $x_{j,s}$ captures child characteristics such as age, and $U_{i,s}$ captures household invariant characteristics.

This paper documents that whereas there is a large heterogeneity in stereotypical gender choices of sports during high school across states, the rates at which boys and girls participate in a sport that is dominated by the opposite gender remains quite persistent over time. Using several sources of data over long periods of time, we present correlational evidence suggesting that the extent to which boys and girls break stereotype when choosing what sports to practice during high school depends on how women are viewed in society. We also identify parental time investments as being an important cultural-transmission mechanism through which gender stereotypical patterns in the choice of sports across US states may be maintained.

Establishing causal effects for these state-level variations is beyond the scope of this paper. We cannot rule out that the degree of gender specialization in sports documented here may reflect that resources for these sports may simply be allocated toward a particular gender. Given the importance of practicing sports for later labor market outcomes, understanding these associations can point towards future research on gender differences in sport choice during high school, and inform a public policy issue of first-order importance.

TABLE 1—GENDER STEREOTYPING IN SPORTS, GENDER NORMS, AND PARENTAL TIME INVESTMENTS

Panel A					
	(1)	(2)	(3)	(4)	(5)
Dependent variable:	Gender-equal	Political Participation	Social and Economic Autonomy	Reproductive Rights	Health and Well-being
GSD sports index	0.005***	0.327**	0.051***	0.269***	0.031**
	(0.002)	(0.152)	(0.013)	(0.035)	(0.012)
Observations	51	50	51	51	51
R-squared	0.225	0.070	0.219	0.367	0.071
Mean	0.499	0.890	7.002	2.362	2.038

Panel B				
	(1)	(2)	(3)	(4)
Dependent variable:	LFP Gap (Female - Male) (%)	Log Wage Gaps, conditional on working (Female-Male)	Share of Females Never married (%)	Average Female Age at First Birth
GSD sports index	0.200**	0.009	0.997***	0.098**
	(0.089)	(0.009)	(0.344)	(0.042)
Observations	51	51	51	51
R-squared	0.079	0.047	0.178	0.177
Mean	-9.335	-0.139	45.412	23.56

Panel C: Parental Time Investments - Fathers with children 6-11				
	(1)	(2)	(3)	(4)
Dependent variable	Total Time (minutes per day)	Time Spent in Basic Care (minutes per day)	Time Spent in Recreational Activities (minutes per day)	Time Spent in Educational activities (minutes per day)
Female	-8.635***	-4.425***	-3.044***	-1.166***
	(1.479)	(1.243)	(0.667)	(0.408)
GSD x Female	1.090***	0.525***	0.437***	0.128
	(0.282)	(0.203)	(0.159)	(0.088)
Observations	18,716	18,716	18,716	18,716
R-squared	0.015	0.008	0.006	0.003
N of households	13,609	13,609	13,609	13,609
Mean	48.31	21.023	18.568	8.72

Panel D: Parental Time Investments - Mothers with children 6-11				
	(1)	(2)	(3)	(4)
Dependent variable	Total Time (minutes per day)	Time Spent in Basic Care (minutes per day)	Time Spent in Recreational Activities (minutes per day)	Time Spent in Educational activities (minutes per day)
Female	2.901	1.343	1.201**	0.357
	(1.827)	(1.196)	(0.535)	(1.005)
GSD x Female	0.096	-0.105	0.079	0.122
	(0.451)	(0.322)	(0.155)	(0.181)
Observations	27,575	27,575	27,575	27,575
R-squared	0.025	0.030	0.005	0.003
N of households	20,278	20,278	20,278	20,278
Mean	74.404	41.417	15.615	17.371

Notes: Panel A shows a state-level OLS regression of gender equality and women's status on the GSD sports index (multiplied by 100). The dependent variable in column 1 is the proportion of individuals reporting gender-equal attitudes from the 1972-2018 General Social Survey. Columns 2-5 includes average state-level variables

on the status of women from the Institute for Women's policy research. There is no availability of information on the Political Participation Index for the District of Columbia (Panel A, column 2). Panel B shows a state-level OLS regressions of labor and non-labor market outcomes for women relative to men on the GSD sports index: The labor force participation gender gap, the share of females never married, and the average female age at first child are constructed from the 2002-2018 American Community Survey (ACS), and the (log) wage gender gap is constructed from the 2002-2018 Current Population Survey (CPS) on the hourly wage. Labor market outcomes are estimated on a sample of natives aged 25-64 and non-labor market outcomes are estimated for a sample of native women aged 20 to 40. Panels C and D present siblings fixed effects model of parental time (minutes per day) from the 2003-2018 America Time Use survey. The sample includes parents aged 21 to 55 with at least one child between 6 and 11 years in the household. Estimations are obtained using survey-specific weights and include controls for age of children. Robust standard errors are in parentheses.
*** Significant at the 1 percent level.
** Significant at the 5 percent level.
* Significant at the 10 percent level.

References

Charles, Kerwin Kofi, Jonathan Guryan, and Jessica Pan. 2018. “The effects of sexism on American women: The role of norms vs. discrimination.” *National Bureau of Economic Research* No. w24904.

Del Boca, Daniela, Chiara Monfardini, and Cheti Nicoletti. 2017. “Parental and child time investments and the cognitive development of adolescents.” *Journal of Labor Economics* 35(2): 565-608.

Fryer, Roland Gerhard, and David Steven Levitt. 2010. “An empirical analysis of the gender gap in mathematics.” *American Economic Journal: Applied Economics* 2(2): 210-40.

Goldin, Claudia, Larry Katz, and Ilyana Kuziemko. 2006. “The homecoming of American college women: The reversal of the college gender gap” *Journal of Economic Perspectives* 20(4): 133-156.

Guiso, Luigi, Paola Sapienza Monte Ferdinando, and Luigi Zingales. 2008. “Culture, gender, and math.” *Science* 320(5880): 1164-1165.

Guryan, Jonathan, Erik Hurst, and Melissa Kearney. 2008. “Parental education and parental time with children.” *Journal of Economic Perspectives* 22(3): 23-46.

Kalil, Ariel, Rebecca Ryan, and Michael Corey. 2012. "Diverging destinies: Maternal education and the developmental gradient in time with children." *Demography* 49(4): 1361-1383.

McKay, Marnee, Jennifer Baldwin, Paulo Ferreira, Milena Simic, Natalie Vanicek, Joshua Burns, and 1000 Norms Project Consortium. 2017. “Reference values for developing responsive functional outcome measures across the lifespan.” *Neurology* 88(16): 1512-1519.

Pope, Devin, and Justin Sydnor. 2010. “Geographic variation in the gender differences in test scores.” *Journal of Economic Perspectives* 24(2): 95-108.

Stevenson, Betsey. 2007. “Title IX and the evolution of high school sports.” *Contemporary Economic Policy* 25(4): 486-505.

Stevenson, Betsey. 2010. “Beyond the classroom: Using Title IX to measure the return to high school sports.” *The Review of Economics and Statistics* 92(2): 284-301.

Data Sources

Flood, Sarah, Miriam King, Renae Rodgers, Steven Ruggles, and J. Robert Warren. Integrated Public Use Microdata Series, Current Population Survey: Version 7.0 [dataset]. Minneapolis, MN: IPUMS, 2019. https://doi.org/10.18128/D030.V7.0

General Social Survey https://gss.norc.org/Get-The-Data

Hofferth, Sandra L., Sarah Flood, and Matthew Sobek. American Time Use Survey Data Extract Builder: Version 2.7 [dataset]. College Park, MD: University of Maryland and Minneapolis, MN: IPUMS, 2018. https://doi.org/10.18128/D060.V2.7

National Federation of State High School Associations
https://members.nfhs.org/participation_statistics

Status of Women in the States Data, Institute for Women’s Policy Research https://iwpr.org/tools-data/data-for-researchers/status-women-states-data/

Ruggles, Steven, Sarah Flood, Ronald Goeken, Josiah Grover, Erin Meyer, Jose Pacas and Matthew Sobek. IPUMS USA: Version 9.0 [dataset]. Minneapolis, MN: IPUMS, 2019. https://doi.org/10.18128/D010.V9.0

TABLE A1—GENDER STEREOTYPING IN SPORTS, GENDER NORMS, AND PARENTAL TIME INVESTMENTS (70 PER CENT THRESHOLD)

Panel A

	(1)	(2)	(3)	(4)	(5)
Dependent variable:	Gender-equal	Political Participation	Social and Economic Autonomy	Reproductive Rights	Health and Well-being
GSD sports index	0.005***	0.339**	0.050***	0.268***	0.029**
	(0.002)	(0.155)	(0.013)	(0.034)	(0.013)
Observations	51	50	51	51	51
R-squared	0.221	0.077	0.217	0.374	0.065
Mean	0.499	0.890	7.002	2.362	2.038

Panel B

	(1)	(2)	(3)	(4)
Dependent variable:	LFP Gap (Female – Male) (%)	Log Wage Gaps, conditional on working (Female-Male)	Share of Females Nevermarried (%)	Average Female Age at First Birth
GSD sports index	0.197**	0.009	1.000***	0.096***
	(0.088)	(0.009)	(0.301)	(0.041)
Observations	51	51	51	51
R-squared	0.080	0.046	0.183	0.174
Mean	-9.335	-0.139	45.421	23.564

Panel C: Parental Time Investments - Fathers with children 6-11

	(1)	(2)	(3)	(4)
Dependent variable	Total Time (minutes per day)	Time Spent in Basic Care (minutes per day)	Time Spent in Recreational Activities (minutes per day)	Time Spent in Educational activities (minutes per day)
Female	-8.536***	-4.371***	-2.984***	-1.182***
	(1.450)	(1.229)	(0.648)	(0.397)
GSD sports index x Female	1.003***	0.481***	0.395***	0.127
	(0.250)	(0.179)	(0.142)	(0.078)
Observations	18,716	18,716	18,716	18,716
R-squared	0.015	0.008	0.006	0.003
N of households	13,609	13,609	13,609	13,609
Mean	48.310	21.023	18.568	8.720

Panel D: Parental Time Investments - Mothers with children 6-11

	(1)	(2)	(3)	(4)
Dependent variable	Total Time (minutes per day)	Time Spent in Basic Care (minutes per day)	Time Spent in Recreational Activities (minutes per day)	Time Spent in Educational activities (minutes per day)
Female	3.150*	1.459	1.197**	0.494
	(1.785)	(1.169)	(0.535)	(0.976)
GSD sports index x Female	0.013	-0.137	0.076	0.074
	(0.413)	(0.302)	(0.145)	(0.155)
Observations	27,575	27,575	27,575	27,575
R-squared	0.025	0.030	0.005	0.003
N of households	20,278	20,278	20,278	20,278
Mean	74.404	41.417	15.615	17.371

Notes: See Table 1.

*** Significant at the 1 percent level.

** Significant at the 5 percent level.

* Significant at the 10 percent level.

TABLE A2—GENDER STEREOTYPING IN SPORTS, GENDER NORMS, AND PARENTAL TIME INVESTMENTS (60 PER CENT THRESHOLD)

Panel A					
	(1)	(2)	(3)	(4)	(5)
Dependent variable:	Gender-equal	Political Participation	Social and Economic Autonomy	Reproductive Rights	Health and Well-being
GSD sports index	0.004***	0.270*	0.037***	0.179***	0.036**
	(0.001)	(0.142)	(0.011)	(0.031)	(0.009)
Observations	51	50	51	51	51
R-squared	0.168	0.060	0.144	0.203	0.118
Mean	0.499	0.890	7.002	2.362	2.038

Panel B				
	(1)	(2)	(3)	(4)
Dependent variable:	LFP Gap (Female - Male) (%)	Log Wage Gaps, conditional on working (Female-Male)	Share of Females Nevermarried (%)	Average Female Age at First Birth
GSD sports index	0.153**	0.003	0.616**	0.058*
	(0.070)	(0.007)	(0.237)	(0.030)
Observations	51	51	51	51
R-squared	0.059	0.007	0.085	0.079
Mean	-9.335	-0.139	45.412	23.564

Panel C: Parental Time Investments - Fathers with children 6-11				
	(1)	(2)	(3)	(4)
Dependent variable	Total Time (minutes per day)	Time Spent in Basic Care (minutes per day)	Time Spent in Recreational Activities (minutes per day)	Time Spent in Educational activities (minutes per day)
Female	-9.943***	-5.071***	-3.746***	-1.126**
	(1.726)	(1.310)	(0.917)	(0.547)
GSD sports index x Female	0.778***	0.381***	0.342***	0.059
	(0.219)	(0.143)	(0.128)	(0.080)
Observations	18,716	18,716	18,716	18,716
R-squared	0.014	0.008	0.006	0.003
N of households	13,609	13,609	13,609	13,609
Mean	48.310	21.023	18.568	8.720

Panel D: Parental Time Investments - Mothers with children 6-11				
	(1)	(2)	(3)	(4)
Dependent variable	Total Time (minutes per day)	Time Spent in Basic Care (minutes per day)	Time Spent in Recreational Activities (in minutes per day)	Time Spent in Educational activities (in minutes per day)
Female	2.367	0.987	1.536*	-0.156
	(2.592)	(1.724)	(0.812)	(1.313)
GSD sports index x Female	0.140	0.006	-0.016	0.150
	(0.383)	(0.274)	(0.131)	(0.154)
Observations	27,575	27,575	27,575	27,575
R-squared	0.025	0.030	0.005	0.003
N of households	20,278	20,278	20,278	20,278
Mean	74.404	41.417	15.615	17.371

Notes: See Table 1.
*** Significant at the 1 percent level.
** Significant at the 5 percent level.
* Significant at the 10 percent level.

TABLE A3—GENDER STEREOTYPING IN SPORTS, GENDER NORMS, AND PARENTAL TIME INVESTMENTS (WITHOUT HAWAII)

Panel A					
	(1)	(2)	(3)	(4)	(5)
Dependent variable:	Gender-equal	Political Participation	Social and Economic Autonomy	Reproductive Rights	Health and Well-being
GSD sports index	0.008***	0.485**	0.067***	0.294***	0.021
	(0.001)	(0.208)	(0.015)	(0.054)	(0.017)
Observations	50	49	50	50	50
R-squared	0.283	0.091	0.230	0.291	0.020
Mean	0.498	0.861	6.994	2.295	2.025

Panel B				
	(1)	(2)	(3)	(4)
Dependent variable:	LFP Gap (Female - Male) (%)	Log Wage Gaps, conditional on working (Female-Male)	Share of Females Nevermarried (%)	Average Female Age at First Birth
GSD sports index	0.299***	0.023***	1.500***	0.161***
	(0.104)	(0.005)	(0.329)	(0.039)
Observations	50	50	50	50
R-squared	0.105	0.174	0.238	0.280
Mean	-9.352	-0.136	45.331	23.563

Panel C: Parental Time Investments - Fathers with children 6-11				
	(1)	(2)	(3)	(4)
Dependent variable	Total Time (minutes per day)	Time Spent in Basic Care (minutes per day)	Time Spent in Recreational Activities (minutes per day)	Time Spent in Educational activities (minutes per day)
Female	-8.749***	-4.483***	-3.091***	-1.174***
	(1.494)	(1.257)	(0.675)	(0.414)
GSD sports index x Female	1.136***	0.548***	0.455***	0.132
	(0.291)	(0.211)	(0.165)	(0.092)
Observations	18,670	18,670	18,670	18,670
R-squared	0.015	0.008	0.006	0.003
N of households	13,574	13,574	13,574	13,574
Mean	48.317	21.093	18.599	8.695

Panel D: Parental Time Investments - Mothers with children 6-11				
	(1)	(2)	(3)	(4)
Dependent variable	Total Time (minutes per day)	Time Spent in Basic Care (minutes per day)	Time Spent in Recreational Activities (minutes per day)	Time Spent in Educational activities (minutes per day)
Female	3.089	1.561	1.127**	0.401
	(1.905)	(1.248)	(0.563)	(1.036)
GSD sports index x Female	0.021	-0.192	0.108	0.104
	(0.498)	(0.352)	(0.175)	(0.197)
Observations	27,492	27,492	27,492	27,492
R-squared	0.025	0.030	0.005	0.003
N of households	20,216	20,216	20,216	20,216
Mean	74.375	41.670	15.584	17.484

Notes: See Table 1.

*** Significant at the 1 percent level.

** Significant at the 5 percent level.

* Significant at the 10 percent level.

TABLE A4—PARENTAL TIME INVESTMENTS (OLS REGRESSIONS)

Panel C: Parental Time Investments - Fathers with children 6-11				
	(1)	(2)	(3)	(4)
Dependent variable	Total Time (minutes per day)	Time Spent in Basic Care (minutes per day)	Time Spent in Recreational Activities (minutes per day)	Time Spent in Educational activities (minutes per day)
Female	-8.440***	-1.960*	-4.908***	-1.572**
	(1.742)	(1.003)	(1.092)	(0.687)
GSD Index x Female	1.016**	0.205	0.533**	0.277
	(0.460)	(0.287)	(0.262)	(0.195)
State FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Observations	18,716	18,716	18,716	18,716
R-squared	0.030	0.025	0.016	0.018
Mean	48.310	21.023	18.568	8.720
Panel D: Parental Time Investments - Mothers with children 6-11				
	(5)	(6)	(7)	(8)
Dependent variable	Total Time (minutes per day)	Time Spent in Basic Care (minutes per day)	Time Spent in Recreational Activities (minutes per day)	Time Spent in Educational activities (minutes per day)
Female	3.394*	3.225**	-0.778	0.947
	(1.938)	(1.298)	(0.877)	(1.005)
GSD Index x Female	-0.322	-0.135	0.070	-0.258
	(0.511)	(0.350)	(0.234)	(0.247)
State FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Observations	27,575	27,575	27,575	27,575
R-squared	0.042	0.044	0.018	0.014
Mean	74.404	41.417	15.615	17.371

Notes: Panels C and D present OLS regression model of parental time (minutes per day) from the 2003-2018 America Time Use survey on GSD sports index (multiplied by 100). The sample includes parents aged 21 to 55 with at least one child between 6 and 11 years in the household. Estimations are obtained using survey-specific weights and include controls for age of children, education of parents, race of parents, state and year fixed effects. Race is included as a set of two dummies (white, black, other(omitted)). Education is included as a set of three dummies indicating whether the father/mother has completed high school, 3 years of college, or 4 or more years of college. Robust standard errors are in parentheses.

*** Significant at the 1 percent level.

** Significant at the 5 percent level.

* Significant at the 10 percent level.