Independent Thinking and Hard Working, or Caring and Well Behaved? Short- and Long-Term Impacts of Gender Identity Norms

Núria Rodríguez-Planas
City University of New York (CUNY), Queens College

Anna Sanz-de-Galdeano
University of Alicante and IZA

Anastasia Terskaya
University of Alicante

This version: January 2019

Abstract

Using the National Longitudinal Study of Adolescent to Adult Health, we explore the causal effect of gender-identity norms on female teenagers' engagement in risky behaviors relative to males in the US. To do so, we exploit idiosyncratic variation across adjacent grades within schools in the proportion of high-school peers' mothers who think that important skills for *both* boys and girls to possess are traditionally masculine ones, such as to think for him or herself or work hard, as opposed to traditionally feminine ones, namely to be well-behaved, popular or help others. We find that a higher proportion of mothers who believe that independent thinking and working hard matter for either gender reduces the gender gap in risky behaviors, traditionally more prevalent among males, both in the short and medium run. We also find evidence of convergence in the labor market in early adulthood. Short- and medium-run results are driven by a reduction in males' engagement in risky behaviors; long-run results are driven by females' higher annual earnings and lower welfare dependency.

Keywords: Gender-identity norms, short-, medium- and long-run effects, risky behaviors and labor market outcomes, and Add Health.

JEL Codes: I10, I12, J15, J16, J22, Z13

Authors' contact: Núria Rodríguez-Planas, Queens College - CUNY, Economics Department, Powdermaker Hall, 65-30 Kissena Blvd., Queens, New York 11367, USA. Email: nuria.rodriguezplanas@qc.cuny.edu. Anna Sanz-de-Galdeano and Anastasia Terskaya, FAE, Universidad de Alicante, Carretera de San Vicente s/n, 03080 San Vicente – Alicante, Spain. Emails: anna.sanzdegaldeano@gmail.com and anastasiaterskaya@gmail.com. Sanz-de-Galdeano is also affiliated with CRES-UPF. She acknowledges financial support from the Spanish Ministry of Economy and Competitiveness Grants ECO2014-58434-P and ECO2017-87069-P.

The authors would like to thank Ana Balsa, Eleonora Patacchini, Stephen Ross, and Marcos Vera-Hernández for their comments on our work as well as participants at the LACEA's Health Economics Network 1st Annual Workshop, 2017 UAB/MOVE Workshop on Health Economics; the 2018 Applied Economics Seminar at the University Massachusetts, Amherst; the 2nd IZA/Barnard Workshop on Gender and Family Economics; the 2018 meetings of the Society of Economics of the Household (SEHO) in Paris; and the Applied Economics Seminar at the University of Alicante for comments that helped us improve the paper.

"Female trial lawyer, male nurse, woman Marine—all conjecture contradiction. Why? Because trial lawyers are viewed as masculine, nurses as feminine, and a Marine as the ultimate man."

Akerlof and Kranton (2000).

Interviewer:	*Are there any groups in society that should not use drugs?*
Female Informant:	*Yes, well that is the group of women who should raise and take care of our children and demonstrate the female sex role, they should really be chained to the kitchen….*
Male Informant:	*I have female friends who smoke (cannabis), but they smoke a bit less often.*
Interviewer:	*How come?*
Male Informant:	*I think it may have to do with gender role expectations, that it is not as OK for girls to use drugs. Well, in general it is not as, well, OK that girls use drugs as it is for guys to do it, so I think it has to do with that a lot. I have difficulties thinking about a biological reason for why girls would use less.*

Sznitman (2007).

1. Introduction

Men's and women's lives have converged considerably in the last century in the US, as in many other developed countries. Importantly, the labor-force participation rates of men and women (Goldin 2006) have converged over time and the gender wage gap has narrowed (Goldin, 2014; Blau and Kahn, 2000). Men and women have also converged in their human capital investment decisions. In fact, the gender gap in educational attainment has often reversed with girls outperforming boys in high-school performance (Fortin et al., 2015) and graduation (Murnane, 2013), years of completed schooling (Charles and Luoh, 2003), and college enrollment and graduation (Goldin, Katz, and Kuziemko, 2006). Last but not least, while men are known to engage in risky behaviors more than women, this gap is also narrowing (Keyes et al. 2007; Warren et al. 2006).

The converging roles of men and women is undoubtedly a multifaceted phenomenon, explained by a combination of factors.[1] One of such factors is the evolution of gender identity (Akerlof and Kranton, 2000; Bertrand, 2010; Goldin 2006). From the theoretical viewpoint, in several influential papers, Akerlof and Kranton (2000, 2002, 2005) use the term identity to refer to a person's self-image and his or her sense of belonging to a social category. Akerlof and Kranton (2000) propose a model where one's identity enters the utility function and, since norms as to how individuals should behave depend on their social category, deviating from such norms decreases utility. Within the gender context, there are two social categories, "men" and "women", which differ in their prescriptions of appropriate behaviors. In particular, women are traditionally thought of as "*generally weak, careful, obedient, socially responsible and sensible, well-behaved, and anxious about and responsive to others' opinion*", whereas

[1] See for instance Goldin (2006) and the references therein.

men, in contrast, are perceived as "*independent, daring, and fearless, inherently curious, and holders of relaxed attitudes*" (Sznitman, 2007). Indeed, traditionally feminine traits have subordinated women to childrearing and domestic tasks, enhancing their identity by working inside the home and establishing invisible boundaries of respectable behavior for them (Abrahamson, 2004). In contrast, men's role as breadwinner highlight their persistency, hard work, strong will, and independent thinking, but also their strength, fearlessness and willingness to take risks (Sznitman, 2007). Hence, as the framework in Akerlof and Kranton (2000) predicts, a societal change that would remove or attenuate gender associations from behaviors both inside and outside the household would decrease women's gains (men's losses) in identity from focusing on traditionally women's activities, and decrease the identity loss of women (men) engaging in traditionally men's (women's) behaviors.

From the empirical viewpoint, devising causal tests of the impact of gender identity is difficult because of both identification challenges, and measurement and data availability issues. A recent strand of the economics literature has empirically estimated the effects of different proxies for gender-identity norms or gender-related cultural dimension on women's human capital, labor-market and fertility decisions. In this context, gender-identity norms tend to be described as "*differences in preferences regarding family structure and women's role in market versus home work"* (Antecol, 2001). This literature has emphasized the relevance of one's cultural background by uncovering positive effects of source-country labor force participation (Fernandez and Fogli, 2006; Blau et al., 2013), education (Blau et al., 2013) and fertility (Fernandez and Fogli, 2006 and 2009; Blau et al., 2013) on these outcomes for second-generation immigrant women living in the US. Instead of focusing only on women, other studies have emphasized the effects of the source-country gender gaps in wages (Antecol, 2001), labor force participation (Antecol, 2000), and smoking (Rodríguez-Planas and Sanz-de-Galdeano, forthcoming) on the same gaps for immigrants living in the same host country.[2] The underlying logic for isolating cultural effects in these studies is that immigrants living in a given host country experience the same institutional environment but differ in their country-of-ancestry preferences towards women's role in society (proxied by the source-country variables), which may have been transmitted to them by their parents.

[2] Rodríguez-Planas and Sanz-de-Galdeano (forthcoming) use data on second-generation immigrant teenagers living in Spain, while Antecol (2000, 2001) uses data on adult immigrants living in the US.

Other papers have instead relied on more direct measures of gender identity norms to study their association with several socioeconomic outcomes. Fortin (2005) shows that gender identity norms (as measured by statements such as "being a housewife is just as fulfilling as working for pay" and "when jobs are scarce, men should have more right to a job than women") are indeed strong predictors of women's labor market outcomes across 25 OECD countries. In the education area, Pope and Sydnor (2010) show that the gender gap in high achievement on test scores[3] is larger in US states where there is more agreement with statements such as "women are better suited for the home" and "math is for boys". More recently, Bertrand, Kamenika and Pan (2015) present evidence that the social norm "a man should earn more than his wife" affects the distribution of relative income within households, women's labor supply and income conditional on working, the patterns of marriage and divorce, and the division of home production.

In contrast with previous studies, the current article studies whether exposure to norms that value traditionally masculine skills (to work hard and think independently) versus traditionally feminine ones (to be well behaved, popular and take care of others) *both* for girls and boys during high school affects female adolescents' engagement in risky behaviors relative to males. Additionally, we explore how the impact of these gender-identity norms evolves as male and female teenagers grow up, and whether they also affect their subsequent labor market involvement as adults.

Based on the aforementioned evidence, it is plausible that norms enhancing masculine skills *both* for girls and boys may reduce the gender gap in risky behaviors, which have been traditionally more prevalent among males.[4] In other words, the relaxation of gender stereotypes may result in girls behaving "more like boys" or boys behaving "more like girls". Because traditional gender roles exert more rigorous social control over women than men (Kaplan et al. 1990, and Waldron et al., 1988), when they are relaxed, women may well increase their experimentation with tobacco, alcohol and illicit drugs, whereas men may well experience less pressure to "act as boys" and be fearless, bold and mischievous. Kaplan et al. (1990) and Waldron et al. (1988) explain that traditional female norms define smoking as unfeminine and inappropriate, and women whose actions do not correspond to the gender norms of behavior face sanctions against smoking under the close social monitoring of traditional societies. These

[3] Males are disproportionally represented at the high end of the math and science test score distributions, while females are disproportionally represented at the top of reading test scores.
[4] Keyes et al. (2007) and Warren et al. (2006) report that adult men are generally more likely to engage in risky behaviors than women, but the gap is getting smaller all around the world.

authors find that men do not perceive such negative connotations as smoking is socially accepted for them and related to their greater social power. Similarly, Abrahamson (2004) notes that women are aware of gender-specific norms establishing invisible boundaries of respectable behavior for them when drinking, while men do not articulate such concerns. Instead, men describe alcohol intoxication as an experience of pure abandonment of self-control where there is no tomorrow, they dare to be themselves and they are freed from every demand. Sznitman (2007) argues that traditional female norms also prevent women from losing control and being selfish, which tends to be associated with the consumption of illicit drugs. Again, such restrictions do not apply to men. Instead, men are considered mentally stronger than women and thus able to control drug use better than women.

Whether non-traditional gender-identity norms push girls to behave "like boys" or boys to behave "like girls" in terms of risky behaviors is an empirical question that we aim at investigating in this paper. Our second objective is to analyse the medium- to long-term consequences of growing up in an environment with less traditional gender norms. In this case, we explore whether non-traditional gender norms improve young adult females' wages and labor supply and decrease their welfare dependence relative to their male counterparts.

To study the causal effect of gender-identity norms during high school on different short-, medium- and long-term outcomes, we exploit idiosyncratic variation across adjacent grades within schools in the proportion of high-school grade-mates'[5] mothers who think that important skills for both girls and boys to learn are traditionally masculine skills (namely, to think for him or herself or work hard) as opposed to traditionally feminine skills (namely, to be well-behaved, popular or help others) using the National Longitudinal Study of Adolescent Health (hereafter, Add Health).[6] The Add Health dataset is ideal for our purposes because, on top of containing detailed information on youths' outcomes, characteristics, and family background including mothers' beliefs on relevant skills for boys and girls, it includes students from multiple grades in a nationally representative sample of high schools and follows them over time. For each student, a school-grade gender identity norm indicator is constructed using *only* information on other students' mothers, that is, we exclude the respondent's own mother. We use school- and grade-fixed effects, as well as school-specific time trends to control for unobserved factors that might confound the non-traditional gender-norm effect in schools. To

[5] "Grade-mates" refer to students in an individual's school-specific grade.

[6] Our identification strategy is drawn from the education literature and exploits variation in student composition across cohorts, within schools, to avoid the endogeneity of friendship networks (Angrist and Lang, 2004; Friesen and Krauth, 2007; Hanushek et al. 2002; Hoxby, 2000; Lavy and Schlosser, 2011; Lavy et al., 2012; Bifulco, Fletcher and Ross, 2011; and Olivetti, Patacchini, and Zenou, 2018).

support the validity of our identification strategy, we follow Lavy and Schlosser (2011) and use Monte-Carlo simulations to show that our actual within-school variation in the proportion of non-traditional mothers is "as good as random" or, more specifically, empirically undistinguishable to the one we obtain when the grade composition of maternal gender-identity norms is randomly generated. Furthermore, we do not find that this within-school variation is related to students' or their parents' predetermined characteristics.

We find that having a greater share of grade-mates' whose mothers gender social norms are non-traditional reduces the high-school gender gap in risky behaviors. This reduction appears to be driven by boys curbing risky behaviors and girls' increasing their engagement in risky behaviors (regular smoking, getting drunk, ever tried marijuana and ever expelled) relative to their male counterparts. While these results persist when youths are in their early twenties, they fade away when they are close to their thirties. Interestingly, at that point in time we also find evidence that a greater share of high-school grade-mates' mothers with non-traditional gender norms increases gender convergence in the labor market by raising women's annual earnings and decreasing their welfare dependency relative to their male counterparts. There is no evidence that our measure of gender-identity norms during high school affects adult males' labor-market outcomes. Overall, our findings suggest that the relaxation of gender stereotypes results in boys behaving "more like girls" in terms of their short- and medium-run engagement in risky behaviors and girls behaving "more like boys" in terms of long-term labor-market outcomes, underscoring an impact on *both* genders. Crucially, we find that gender-identity norms experienced during adolescence improve women's economic outcomes as adults.

In addition to contributing to the gender-identity literature described above, our research also speaks to recent work using Add Health and aiming at identifying factors that shape women's gender-role identity. Using the same identification strategy, Olivetti, Patacchini, and Zenou (forthcoming) find that a woman's labor supply as a young adult is shaped by the work behavior of her adolescent peers' mothers. While these authors' findings "emphasize the influence of (high-school peers') mothers as (women's) role models", our work underscores the influence of high-school peers' mothers' *beliefs* regarding gender norms on *both* males' and females' engagement in risky behaviors during their teens and twenties, and on females' wages and welfare dependency in their thirties. Our work delivers a broader picture of the role played by gender-identity norms in shaping the gap in relevant indicators beyond labor-market outcomes, as we show that their effects already appear during

adolescence, and, crucially, that they affect men too by influencing their engagement in risky behaviors as teenagers and as they transition into adulthood. [7]

The remainder of this paper is organized as follows. The next section presents the data, sample selection and descriptive statistics. Section 3 describes the identification strategy used and its validity. Section 4 presents the main results. Prior to concluding in Section 6, Section 5 presents several robustness checks.

2. Data and Descriptive Statistics

2.1. The Add Health Dataset

The National Longitudinal Study of Adolescent to Adult Health (Add Health) is a school-based longitudinal survey, nationally representative of the United States population of 7th to 12th graders during the school year 1994/95. A stratified sample of 80 high schools and their main feeder school (typically a middle school) was selected, and within each high school (middle school) up to four (two) different grades were generally sampled in Wave I.[8] Next, within each school and grade, a random sample of approximately 17 males and 17 females were selected in 1994/95 (hereafter Wave I). These randomly selected students constitute the so called *core sample*, which is nationally representative of US adolescents in grades 7th to 12th, and they were subsequently interviewed in 1996 (hereafter Wave II), in 2001/02 (hereafter Wave III) and in 2008 (hereafter Wave IV). In addition, students from specific minorities were oversampled in Wave I and followed over time. These students, together with the *core sample* students, constitute the so called *in-home* sample.

The *in-home* survey of Add Health, which was mostly conducted at the respondents' homes, collects comprehensive information on their health-related behaviors during adolescence and early adulthood, as well as other post-secondary outcomes.[9] Additionally, a parent (preferably, and usually, the resident mother) of each adolescent who completed the *in-home* questionnaire was also asked to complete a questionnaire.[10] Our outcome variables

[7] Several papers have also studied whether the impact of school peers' characteristics persists into adulthood (Bifulco et al. 2014; Bifulco, Fletcher and Ross, 2011; Black, Devereux and Salvanes, 2013; Anelli and Peri, 2017; and Carrell, Hoekstra and Kuka, 2016), albeit their focus is not on gender-identity norms.

[8] Because Add Health only covers grades 7th to 12th, middle schools at Wave I only covered two grades (grades 7th and 8th), even though most middle schools begin in 6th grade in the US.

[9] Add Health conducts computer-assisted self-interviews (CASI) in which the respondent listens through earphones to pre-recorded questions and enters the answers directly on a laptop to minimize misreporting or non-response. Prior to responding to the questionnaire, the interviewee is assured that his or her responses will be and remain confidential.

[10] In particular, 90.8% of the responding parents were the adolescents' mothers.

(described in detail below) and individual control variables come from the *in-home* survey. In particular, we use the following individual controls in our analysis: student's sex, age, race, ethnicity, the quality of the residential building they live in,[11] and whether they live with both parents, as well as their parents' age and education. We also control for the student's test scores on the Peabody Picture Vocabulary test (PPVT), which is an age-specific test used to assess verbal ability and receptive vocabulary. The PPVT is often considered a measure of verbal IQ and is strongly correlated with the Wechsler Intelligence Test and the Armed Forces Qualifying Test (Dunn and Dunn 2007).

During Wave I, Add Health also administered a shorter survey (the *in-school* survey) to *all* students at the sampled schools who were present in a given interview date. While we used the *in-school* survey to estimate most of the grade-level variables included in our analysis, we were unable to use it more broadly because: (1) it lacks the detailed information on risky behaviors contained in the *in-home* survey, and (2) it is not longitudinal as it was only administered in Wave I. Nonetheless, it is worth noting that we were able to use the *in-school* census of students (as opposed to the *in-home* random sample) to construct the following grade-level variables that are used as controls in our analysis: grade size and grade-mates' average age, share of females, and share of minorities.

2.2. Gender Identity Norms

Our measure of traditional gender identity norms is constructed from the following two separate questions in the parent questionnaire administered in Wave I: "*Of the following, which do you think is the most important thing for a girl to learn? (1) to be well-behaved; (2) to be popular; (3) to think for herself; (4) to work hard; or (5) to help others*" and "*Of the following, which do you think is the most important thing for a boy to learn? (1) to be well-behaved; (2) to be popular; (3) to think for himself; (4) to work hard; or (5) to help others*", where respondents had to select one of the possible 5 choices in each case. These two questions were asked to all responding parents regardless of whether they had a daughter or a son. We classify mothers' gender-identity beliefs as non-traditional if they answered "*to think for herself*" or "*to work hard*" for *both* girls and boys, while we classify them as traditional if they answered otherwise. Using this binomial variable, we calculated, for each student in our sample, the

[11] Residential building quality is reported by the interviewer as opposed to the child or the parent and is one of the variables Add Health users sometimes rely on to capture family socio-economic background (see, for instance, Olivetti, Patacchini and Zenou, 2018).

proportion of non-traditional mothers in his or her grade and school (excluding his or her own mother's answer and regardless of whether mothers had a daughter or a son).[12] It is noteworthy that our key explanatory variable was constructed with the a sample of over 10,000 mothers of high-school students.

To the extent that this measure indeed captures "non-traditional" maternal gender-identity, we would expect it to be correlated with actual behaviors and outcomes traditionally less prevalent among women. We provide evidence that this is indeed the case both at the individual and county levels (shown in Panels A and B in Appendix Table 1). In Panel A of Appendix Table 1, we use information on several choices and behaviors of all the mothers available in Add Health (about 15,000 mothers) and show that mothers with non-traditional gender-related beliefs are more likely to report working for pay than traditional mothers, and doing so an average of 2.2 more hours per week. They are also less likely to live in an only-male-breadwinner household. In addition, they are more likely to be more educated (with, on average, close to one more year of education) and to be more educated than their spouse than traditional mothers. While these estimates only capture correlations, they provide evidence that our measure of non-traditional gender-identity norms relates to maternal choices commonly linked to non-traditional gender roles.

Panel B in Appendix Table A.1 presents complementary evidence that our main gender norms explanatory variable is also correlated with indicators typically associated with gender equality at the county level. In particular, we use the following county-level variables available in Add Health: a female labor force opportunity index[13] (which reflect economic and social opportunities for women) and age-specific child/woman ratios[14] (which serve as county-level fertility measures for different cohorts of women). In line with our expectations, we find that the female labor force opportunity index is higher in counties with a higher proportion of non-traditional mothers, while fertility is lower instead. Again, while we are only capturing correlations, our measure of non-traditional gender-identity norms seems to be related with county-level variables commonly used to reflect greater gender equality.

[12] We base our computation of the proportion of non-traditional mothers only on the answers given by mothers. Nonetheless, our findings are unaltered if instead we also consider the responses given by fathers.

[13] The female labor force opportunity index included in Add Health indicates the expected number of jobs for female workers relative to the potential supply of female workers, taking into account the sex-segregated nature of the labor market.

[14] Age-specific child/woman ratios are calculated by dividing the number of children ever born (not counting still births) to women in the specified age group by the mid-year population of women in that age group.

2.3. Outcome Variables

We are interested in two broad types of outcomes: risky behaviors and post-secondary socioeconomic outcomes. Among the former, we consider the following six binary outcomes: smoking regularly, getting drunk in the past year, ever smoking marijuana, ever using illicit drugs (other than marijuana), ever being expelled from school, and having intercourse before age 16. Among the latter, we consider three post-secondary outcomes, namely whether the individual has ever worked for pay at a full-time job, personal annual earnings, and whether the individual is a welfare recipient.[15] Appendix Table A.2 displays the Add Health definition of the variables from the *in-home* survey used to construct all these outcomes.

Because we are interested in analyzing whether the effects of gender-identity norms change as youths grow up, we measure most of these behaviors at three different points in time: during adolescence (Wave I), as youths transition into adulthood (Wave III), and in young adulthood (Wave IV). There are a few exceptions: having intercourse before age 16 is not measured in Wave I because most respondents were younger than 16 at that point in time, and being expelled from school is not measured in Wave IV as it is no longer relevant. Similarly, the three post-secondary outcomes are only measured during young adulthood, in Wave IV. [16]

2.4. Sample Restrictions

Because we study youths' engagement in behaviors during high school and follow them as they transition into and subsequently settle into young adulthood, we restrict our analysis to students in the *in-home* sample who were in high school in Wave I. This leaves us with a sample of 14,406 students. In addition, we restrict our sample to those students who were successfully tracked from waves I to IV, further reducing our sample to 8,547 students.[17] We also dropped 5 students for whom age or race was missing. Given our focus on the proportion of grade-mates whose mothers have non-traditional gender-related beliefs during Wave I, we dropped 366 additional students because some grades had less than 10 students to eliminate grades in which parental preferences would be imprecisely estimated, leaving us with a longitudinal

[15] Personal earnings include wages or salaries, tips, bonuses, overtime pay, and income from self-employment.
[16] Wave II data were collected in 1996. Because we are interested in analyzing the short-, medium- and long-run behavioral effects of high-school gender-identity norms, we preferred to focus our attention on waves I, III and IV as they were each 6 years apart. Nonetheless, results using wave II are similar to those from wave I and available from the authors upon request.
[17] This restriction was applied *after* the peer variable was constructed from the Wave I in-home survey.

sample of 8,181 students from 72 schools and 283 school-grade cells.[18] In Section 5, we address potential concerns for attrition bias.

2.5. Descriptive Statistics

Table 1 presents descriptive statistics for the longitudinal sample used in our analyses including our benchmark set of control variables both at the individual level (Table 1.A) and at the grade level (Table 1.B). Table 2 provides summary statistics for our grade-level measure of maternal gender identity and for our outcome variables. The first two columns display means and standard deviations for girls and boys, respectively, while the last column presents the gender gap. Consistent with previous evidence on boys' underperformance (relative to girls') in many non-cognitive outcomes (see Bertrand and Pan, 2013, and Autor et al., 2016, and the references therein), we find that male teenagers are generally more likely to engage in risky and disruptive behaviors than female teenagers. Interestingly, this difference appears to widen as time goes by. In contrast, men tend to outperform women in labor market outcomes in early adulthood. Table 2 also presents descriptive statistics of our key explanatory variable, the share of grade-mates' mothers with non-traditional gender beliefs. On average, about 68% of grade-mates have mothers reporting non-traditional gender-related beliefs. This variable barely varies by gender.

3. Identification Strategy and Validity

3.1. Identification Strategy

Our goal is to estimate the causal effect of gender-identity norms during high-school on the gender gap in risky behaviors from adolescence into early adulthood and in labor market outcomes in early adulthood. This effect may be confounded by the effects of unobserved correlated factors if gender-identity norms affect students' sorting across schools, or if they are correlated with other characteristics of the school that may also affect students' outcomes. To account for these sources of confounding factors we follow a quasi-experimental research design, first developed by Hoxby (2000), which is based on across grades comparisons within a school. In our context, the basic idea of this approach is to exploit within-school variation in gender-identity norms across adjacent grades. Hence, we assume that, conditional on school,

[18] This restriction is common in papers analyzing the effects of high school grade-mates' characteristics and using Add Health data (see Bifulco, Fletcher and Ross, 2011; and Olivetti, Patacchini and Zenou, 2018).

the variation in grade-mates' exposure to different gender-identity norms is quasi-randomly assigned. In particular, we estimate the following model:

$$Y_{igs,w} = \beta_0 + \beta_1 Female_{igs} + \beta_2 NonTraditionalMothers_{-igs,1} + \beta_3 NonTraditionalMothers_{-igs,1} * Female_{igs} + \delta_g + \rho_s + X'_{igs,1}\alpha + G'_{gs,1} + \varphi\pi_s(Grade_g) + \varepsilon_{igs,w} \quad (1)$$

, where $Y_{igs,w}$ is the outcome of interest in wave *w* for an individual *i* who attended high school *s* and grade *g* in Wave I. $Female_{igs}$ takes the value 1 if individual *i* is female and 0 otherwise and accounts for behavioral differences across genders in the outcome variable. $NonTraditionalMothers_{-igs,1}$ is the proportion of students (excluding individual *i*) in grade *g* and school *s* whose mothers have non-traditional gender-related beliefs in Wave I. For each student *i*, we construct this variable using *only* information on his/her grade-mates, that is, excluding each student's own mother's gender-identity norms. $X'_{igs,1}$ is a vector of student-specific characteristics and $G'_{gs,1}$ is a vector of grade-specific characteristics in a particular school *s*. Both vectors measure characteristics at Wave I and control for student- and grade-specific characteristics that may be related to youths' engagement in risky behaviors, such as age, race or IQ (their own and those of their grade-mates). In addition, to account for the most obvious issues that may arise at the school and grade level such as the static selection of students into schools or the fact that some grades may differ from others within a school, we include both school (ρ_s) and grade (δ_g) fixed effects. One may still be concerned that time-varying unobserved factors correlated with the changes in grade composition within schools may be biasing our results. For example, let's suppose that the proportion of non-traditional mothers is increasing over time in some schools more than in others. To the extent that parents may be able to detect this change and act upon it based on their preferences related to gender-identity norms, students in higher and lower grades may differ in unobserved ways that may in turn affect boys' and girls' relative propensities to engage in risky behaviors. To address this concern we include a full set of school-grade trends, $\pi_s(Grade_g)$. Hence, identification is based on the deviation in the proportion of grade-mates' non-traditional mothers across grades from its school trend.

Since our goal is to examine whether gender-identity norms affect the gender gap in individual outcomes, our main coefficient of interest is that of the interaction between

$NonTraditionalMothers_{-igs,1}$ and the female indicator, that is, β_3. For instance, if *Y* denotes regular smoking, a positive and significant estimate of β_3 would suggest that a higher proportion of non-traditional mothers in grade *g* and school *s* is associated with a higher prevalence of smoking among girls relative to boys from the same grade and school, and thus a *smaller* male-female gender gap in smoking. Note also that the coefficient β_2 captures the effect of the proportion of non-traditional mothers on the outcomes of interest for boys, while (β_2+ β_3) is the (absolute) effect of the proportion of non-traditional mothers on the outcomes of interest for girls.

Since we analyze multiple outcomes, we must address the concern that an increase in the number of tests increases the likelihood of rejecting the null hypothesis using traditional inferential techniques. We do so using two alternative and complementary strategies. First, we apply the Romano and Wolf (2005) stepwise multiple testing procedure that asymptotically controls the familywise-error rate to estimate adjusted p-values. Following Heckman *et al.* (2010), we group hypotheses into economic and substantially meaningful categories by survey waves. Thus, the analysis focuses on indicators from two key families of outcomes: risky behaviors and labor market outcomes, the former measured at three different points in time (Waves I, II and IV) and the latter measured at Wave IV. Second, to address concerns that methods that adjust individual p-values for multiple testing to control for familywise-error rate may be overly conservative in terms of power, we also construct several summary indices (using the same families of outcomes as explained above).[19] Summary indices are a common method to adjust for multiple hypothesis testing,[20] and, in addition, their use offers a broad snapshot of our results' overall patterns. Each summary index variable, *Y**, is constructed as the unweighted average of all standardized outcomes within a family:

$$Y^* = \frac{\sum_k Y_k^*}{k} \qquad \text{where} \qquad Y_k^* = \frac{Y_k - \mu_k}{\sigma_k}$$

where Y_k is the k^{th} of *K* outcomes within each family, μ_k denotes its mean and σ_k its standard deviation. Y_k^* is the standardized version of Y_k. Because the labor market summary index in Wave IV contains both adverse and beneficial outcomes, we switch the sign for the adverse outcome (welfare receipt), so that a higher value of this normalized measure represents a more

[19] Anderson (2008) explains that summary indices are preferred to alternative methods that adjust individual p-values for multiple testing when there is a priori no reason to believe that a group of outcomes will be affected in a consistent direction.

[20] See Kling, Liebman, and Katz, 2007, and Rodríguez-Planas, 2012 and 2017, among others.

"beneficial" outcome.[21] We describe the summary indices used in the paper in Appendix A. Appendix Table A.3 presents summary statistics of these summary indices by gender.

Finally, regarding our set of covariates, we use the following controls in our benchmark analyses: i) at the individual level: the student's sex, age, race, ethnicity, the student's scores on the PPVT, the quality of the residential building they live in, whether they live with both parents, parental age and education; ii) at the grade level: grade size and grade-mates' average age, share of females, and share of minorities; iii) grade fixed effects, school fixed effects, and school-grade trends. Importantly, we also report sensitivity analyses (which confirm the robustness of our results) to the inclusion of each student's own mother's gender-identity beliefs and its interaction with the female dummy; school-grade fixed effects as well as other parental and grade-level controls such as: maternal work and smoking indicators and grade-level shares of working and smoking mothers, and their interactions with the female dummy.

3.2. Validity of the Identification Strategy

Our key identifying assumption postulates that changes across grades in the proportion of non-traditional mothers within a school result from random fluctuations, and hence are uncorrelated with unobserved differences across grades in students' characteristics that may in turn affect their outcomes. In order to effectively rely on this identification strategy two things must happen. First, the data need to display enough variation in the proportion of mothers with non-traditional beliefs within schools and across grades so as to estimate their effects with precision. And second, changes across grades in the proportion of non-traditional mothers within a school must result from as good as random rather than systematic fluctuations. This section assesses both aspects.

Table 3 examines the extent of variation in the proportion of non-traditional mothers that is left after removing grade and school fixed effects (Table 3, Panel B), and after removing school/grade trends on top of grade fixed effects and school fixed effects (Table 3, Panel C). Removing school and grade fixed effects reduces the standard deviation in the proportion of students with non-traditional mothers by 40%, and additionally removing school-grade trends reduces this variation by an additional 10%. There seems to be sufficient variation in the data to estimate the effects of interest even after one controls for grade fixed effects, school fixed effects and school-grade trends. This assessment is indeed reinforced by the fact that we do

[21] For the risky behaviors' summary indices, a higher value represents greater engagement in risky behaviors and hence a "detrimental" outcome.

estimate statistically significant impacts for many of our outcomes, as we will discuss in the next section.

Because in most schools in the US, the grade a student attends is a function of his or her birth date and a cut-off date, it ought to be beyond the influence of the student, parents or school administrator (as argued by Elsner and Isphording 2017, among others). To rule out sorting across grades, Table 4 presents balance tests for our non-traditional gender-identity norms variable. More specifically, we explore whether variation in the share of non-traditional mothers across grades is indeed unrelated to a number of socio-demographic characteristics net of grade and school fixed effects. In Table 4, we report estimates of the coefficients of the share of non-traditional mothers and of their interactions with the female dummy. These tests reveal that only one of our 28 coefficients is statistically significantly different from zero at the 10% level, and none are statistically significantly different from zero at the 5% level, which is less than what we would expect by chance, suggesting that controlling for grade and school fixed effects is likely to be sufficient to isolate variation in grade composition that is not systematically related to students' socio-demographic composition within schools.

Additionally, we follow Lavy and Schlosser (2011), and perform Monte-Carlo simulations to show that our actual within-school variation in the share of grade-mates' mothers with non-traditional gender beliefs is "as good as random". To do so, for each student in each school, we randomly draw a dummy variable indicating beliefs in non-traditional gender roles using a binomial distribution with population mean equal to the actual school-specific mean of this measure. Based on this random draw, we compute the simulated proportion of non-traditional mothers in each grade for each school. Then, we calculate the standard deviation of this proportion in each school. We repeat this procedure 1,000 times to obtain a 90% confidence interval for our simulated within-school standard deviations, and we finally check if the actual within-school standard deviations of our variable of interest are within this confidence interval. Indeed, we find that more than 90% of our schools have an actual standard deviation that falls within the 90% confidence interval based on simulated data, which suggests that the within-school variation of our actual measure of grade-mates mothers' gender identity is as good as random.

4. Results

4.1. Summary Indices

Table 5 presents estimates by domain or summary index, with the outcomes included in each summary index indicated in Appendix Table A.3.[22] The third row in Table 5 displays estimates of β_3, which capture the effect of high-school gender-identity norms on the gender gap in risky behaviors in the short-, medium- and long-run (columns 1 to 3). In Waves I and III (shown in columns 1 and 2), this coefficient is positive and statistically significant at the 1% level, revealing that having a higher proportion of grade-mates with mothers with non-traditional beliefs during high school increases girls' engagement in risky behaviors *relative* to their male counterparts, hence reducing the gender gap. This effect fades away as $\hat{\beta}_3$ is smaller in size and no longer statistically significant in Wave IV (shown in column 3 in Table 5).

This gender convergence in risky behaviors when there is exposure to non-traditional gender-identity norms may be the result of a reduction in males' engagement in risky behaviors, an increase in females' engagement in risky behaviors, or both. The evidence indicates that a greater proportion of peers with non-traditional mothers during high school curves males' risky behaviors in the short- and medium-run as $\hat{\beta}_2$ (which captures the effect of gender-identity norms on boys' engagement in risky behaviors) is negative and statistically significant at the 10% level in Wave I and 5% level in Wave III. While exposure to non-traditional gender-identity norms increases girls' (absolute) engagement in risky behaviors during high school ($\hat{\beta}_2+\hat{\beta}_3$=+0.186), there is not enough precision to reject the null of no (absolute) effect among girls. Moreover, the size of ($\hat{\beta}_2+\hat{\beta}_3$) decreases in Waves III and IV, suggesting that any potential short-term detrimental effect on girls vanishes by the time females are in their early twenties.

Column 4 in Table 5 explores the effect of gender-identity norms during high school on labor market outcomes in adulthood. We find a positive and persistent effect of high-school gender-identity norms on females' labor market outcomes in Wave IV both in absolute terms ($\hat{\beta}_2+\hat{\beta}_3$) and relative to their male counterparts ($\hat{\beta}_3$). This effect is statistically significant at the 5% level and consistent with earlier findings that a relaxation of traditional gender norms reduces the gender gap in the labor market as shown by Bertrand, Kamenica and Pan (2015) for US married couples. We find no effect of high-school grade-mates' mothers with non-traditional beliefs on adult males' labor market outcomes.

[22] Standard errors are clustered at the school level. Following Add Health protocols, our analyses use longitudinal sample weights so that our estimates are nationally representative of the US high-school student population in school year 1994/95.

It is worth noting that reverse causality is difficult to argue as it would imply that students' behavior affects the beliefs of his or her peers' mothers, and that it does so differentially by gender. Even if this were possible during high school, hence affecting our Wave I estimates, it is even more difficult to argue that adults' behaviors (say females' welfare receipt and earnings in their early 30s) could have affected the beliefs of their mothers' high-school peers 12 years earlier.

Appendix Table A.4. replicates our analysis in Table 5 with a different non-traditional gender norms variable. In this case, we define as non-traditional mothers those who answered "*to think for herself*" or "*to work hard*" to the question on the most important thing for a ***girl*** to learn. In contrast with our main definition, this one also includes as non-traditional mothers those who aspire for their daughters to be independent even if they aspire for their sons to be caring, popular or nurturing. Estimates with this alternative variable reveal that our main results remain unaltered.

The first five columns of Appendix Table A.5 show the sensitivity of coefficients in Table 5 to sequentially adding fixed effects, school-grade trends and individual, parental and grade controls. Overall, the size and precision of $\hat{\beta}_3$ is relatively stable across specifications, suggesting that school-grade trends matter little and that omitted individual-level variable bias is unlikely to be a problem. The estimated coefficient $\hat{\beta}_2$ is a bit more sensitive to the specification, in particular to the inclusion of school-grade trends.

Since our main coefficient of interest is $\hat{\beta}_3$, another robustness check is to include school-grade fixed effects, instead of including school-grade trends. This is shown in column 6 of Appendix Table A.5. While this specification does not allow us to identify $\hat{\beta}_2$, it provides a convincing robustness check of the main results, as both the magnitude and significance of the estimates of β_3 are mostly unchanged.

4.2. Potential Mechanisms

Concerns that our findings may be picking up individuals' own mother's beliefs or behaviors are addressed in columns 7 and 8 of Appendix Table A.5. Here, we add to our preferred specification (shown in column 5) a control for each individual's own mother's gender-identity beliefs and its interaction with the female dummy (column 7), and maternal smoking behavior and labor force participation, as well as their interactions with the female dummy (column 8). Importantly, estimates in both specifications shown in columns 7 and 8 go in line with our main finding that being exposed to less traditional gender norms reduces the gender gap in risky

behaviors. In our preferred specification, we excluded these own mother's gender-identity beliefs, smoking behavior and labor force participation because, unlike our grade-level measure $NonTraditionalMothers_{-igs,1}$ (whose within-school and across grades variation is argued to be idiosyncratic), they are potentially endogenous as a mother's beliefs on girls' or boys' most relevant skills, as well as her smoking behavior and her labor force participation could be correlated with unobserved heterogeneity shaping her child's engagement in risky behaviors during adolescence and adulthood as well as her labor market choices. Appendix Table A.6 displays coefficient estimates for each student's own mother's gender-identity beliefs[23] and its interaction with the female dummy (as well as our main coefficients of interest) when using the same specification as in Appendix Table A.5 Column 7 so as to provide some sense of the conditional correlations. As expected, having a non-traditional mother reduces the gender gap in risky behaviors and labor market outcomes.

Columns 9 to 11 of Appendix Table A.5 explore the extent to which our findings are driven by gender role models in female employment or female smoking habits observed during high school. To do so, we add to our preferred specification (shown in column 5) the share of working mothers (column 9), the share of smoking mothers (column 10) or both variables (column 11), all interacted with the female dummy. Adding these variables has a negligible effect on our estimated coefficients of interest, suggesting that our findings on gender-identity norms are not mediated via grade-mates' mothers' behavior, but instead via their gender-identity norms being transmitted from them to their children, who in turn interact with their grade-mates.[24] Importantly, estimates from columns 7 to 11 show that our coefficients of interest $\hat{\beta}_2$ and $\hat{\beta}_3$ are robust to the inclusion of these controls, suggesting that they are not picking up behavioral role models from the teenagers' own mothers or their peers' mothers. Moreover, findings in columns 9 to 11 suggest that the relevance of mothers' beliefs during high school affects behaviors above and beyond the contextual effect of the employment status or smoking habits of the teenager peers' mothers. To put it differently, these findings suggest that gender roles in female employment or female smoking observed during high school are not driving our findings, underscoring the relevance of gender-identity beliefs and

[23] Note however that the magnitudes of the estimated coefficients of the variables "Mother is non-traditional" and "Share of non-traditional mothers" (as well as their interactions with the female dummy) are not comparable, since the former is a dummy and the latter is a percentage.

[24] To the extent that these students are in high school, where parents' presence is relatively rare when compared to elementary school, the transmission of gender-identity norms is unlikely to happen between students' mothers or from mothers to their children's grade-mates.

distinguishing our results on labor market outcomes from those of Olivetti, Patacchini and Zenou (forthcoming).

Data limitations prevent us from precisely disentangling all the possible channels through which the exposure to less traditional gender norms during high school operates. Having said that, Appendix Table A.7 presents some suggestive evidence that, indeed, when individuals are exposed to less traditional norms the association between gender and certain gender traits and behavioral prescriptions weakens among them. For instance, we study individuals' responses to a standard question that is used to measure impulse control:[25] "When making decisions, you usually go with your 'gut feeling' without thinking too much about the consequences of each alternative?" The answers are coded 1 to 5, where 1 means "strongly agree" and 5 means "strongly disagree", such that the higher this variable is, the higher the level of self-control. As expected, boys are on average more impulsive than girls, and, interestingly, we find that an increase in the share of non-traditional grade-mates' mothers of 10 percentage points would significantly reduce the observed gender gap by about 22%.[26] Additionally, we consider two indicators that can proxy for the social stigma attached to sex, its association with gender and, most importantly, how this association is affected when gender identity is less traditional. In particular, individuals were asked while still in high school whether they thought their mother would be upset if they had sex and whether a pregnancy (getting pregnant for girls and getting someone pregnant for boys) would be an embarrassment for their families. Again, there is a baseline gender gap in both indicators (with girls expressing significantly more concern regarding these issues) that is, as expected, in turn significantly reduced among individuals who are exposed to less traditional gender norms during high school. In sum, while tentative, we interpret this evidence as suggestive that indeed the attenuation of the traditional association between gender and certain traits and prescriptions linked to gender identity may be one crucial channel for our findings.

4.3. Individual Outcomes

Panels A and B in Table 6 show the effects of gender-identity norms on individual outcomes using our preferred specification (using the same controls as those in the summary index specifications in Table 5). Panel A focuses on risky behaviors in the short, medium and long

[25] See for instance Battaglini, Díaz and Patacchini (2017) and the references therein.
[26] The average value of the self-control variable is 3.19 for boys and 3.00 for girls, so the raw gender gap is 0.19.

run, while Panel B presents labor-market outcomes in the long run. Each table presents estimates of β_2 and β_3.

Below we only discuss estimates when the coefficient on the summary index (shown in Table 5) is statistically significantly different from zero. Even though summary indices are frequently preferred to alternative methods that adjust individual p-values for multiple testing (Anderson 2008), in Table 6 we also adjusted p-values using stepwise multiple testing procedure that asymptotically controls the familywise error rate proposed by Romano and Wolf (2005). Coefficients in bold are those that are statistically significant at the 10% level or lower using this alternative correction method.

Short-Run Effects on Risky Behaviors

Focusing on the short-run effects first, we find that being exposed to non-traditional gender norms during high school increases girls' regular smoking, getting drunk in the past year, ever smoking marijuana, and ever being expelled from school *relative* to their male counterparts. Indeed, estimates of β_3 (shown in column 2) indicate that a 10 percentage-point increase in the share of grade-mates' mothers with non-traditional beliefs results in an increase in girls' likelihood of getting drunk during the past year of 2.6 percentage points relative to their male counterparts (that is, a 7% increase of girls' "getting drunk" prevalence in Wave I). Since the raw male-female gender gap in Wave I is 3.8 percentage points, this represents a 68% decrease in such gap.[27] Similarly, a 10 percentage-point increase in the share of grade-mates' mothers with non-traditional beliefs increases high-school girls' likelihood of: (1) regular smoking by 1.48 percentage points relative to their male counterparts (or a 6.8% increase of girls' smoking prevalence); (2) ever smoking marijuana by 2.52 percentage points relative to their male counterparts (or a 57% reduction of the raw gender gap); (3) ever being expelled from school by 1.40 percentage points relative to their male counterparts (or a 32% reduction of the raw gender gap). All of these estimates are statistically significant at the 5% level or lower, even after adjusting for multiple hypotheses testing with the Romano and Wolf procedure. The effect of being exposed to non-traditional gender norms during high school on using other illegal drugs goes in the same direction, although $\hat{\beta}_3$ is only statistically significant at the 10% level and it further loses precision after adjusting for multiple-hypothesis testing. Even though the short-run estimate for the summary index, $\hat{\beta}_2$, shown in Table 5 and discussed in the

[27] In Wave I, the prevalence of boys and girls reporting getting drunk in the last 12 months is 39.8% and 36.1%, respectively.

previous sub-section, was statistically significant for males, none of the individual outcomes are statistically significantly different from zero. Nonetheless, all of them are negative and their size is far from zero.

Medium- and Long-Run Effects on Risky Behaviors

Moving now to Wave III, we explore whether high-school grade-mates mothers' gender-identity norms continue to affect the gender gap in risky behaviors by the time youths were on their early twenties and out of high school. The summary-index estimate of β_3 in Table 5 was positive and statistically significant indicating that high-school gender-identity norms still affected the gender gap in risky behaviors in the medium run. Even though most estimates of β_3 in Column 4 in Table 6 remain positive and sizeable, only three of them are statistically significantly different from zero and of these, only one remains statistically significant once we adjust the standard errors for multiple hypothesis testing. Hence, it appears that some of the short-run effects on the gender gap in risky behaviors fade away as youths grow up (this is, for instance, the case for getting drunk). Nonetheless, other effects persist as we find that a 10 percentage-point increase in the share of grade-mates' mothers with non-traditional beliefs during high school results in an increase in the likelihood of smoking marijuana among young female adults of 2.4 percentage points (the equivalent of a 28,5% decrease in the gender gap). In addition, it also increases females' likelihood of having ever been expelled from school by 2.23 percentage points relative to their male counterparts (the equivalent of a 25,6% decrease in the gender gap), or of having sex before age 16 by 1.93 percentage points (the equivalent of a 68% widening of the gender gap). The estimated effects for smoking marijuana and having sex lack precision when the Romano-Wolf adjustment is applied.

During Wave III, when youths were in their early twenties, the summary-index estimate of β_2 in Table 5 was negative and statistically significant at the 5% level, indicating that a higher exposure to non-traditional norms during high school continued to curb males' risky behaviors 6 years later. Table 6 shows that this effect is driven by boys reducing regular smoking and marijuana use (marginally) as well as the probability of being expelled from school. While these effects for males are statistically significant at the 10% level or lower, precision is lost after adjusting for multiple-hypothesis testing using the Romano-Wolf procedure.

The effects of exposure to non-traditional gender-identity norms on risky behaviors fade away by the time youths are in their late twenties/early thirties (Wave IV). Indeed, none

of the coefficients on the summary indices are statistically significantly different from zero, and the only statistically significant coefficient we estimate when analyzing individual outcomes (β_3 for ever trying marijuana) is no longer significant when we adjust for multiple hypotheses testing.

Long-Run Labor Market Outcomes

Interestingly, we find that exposure to non-traditional gender-identity norms during high school benefits adult women's labor market outcomes relative to their male counterparts. We find that women who were exposed to less traditional gender identity norms during high school earned higher annual income 12 years later, when they were in their late twenties, and they were also less likely to receive welfare relative to their male counterparts. More specifically, Panel B in Table 6 shows that a 10 percentage-point increase in the share of grade-mates' mothers with non-traditional beliefs during high school increases women's annual personal income by 14% relative to their male counterparts, reducing the earnings gap from $12,801 to $11,01 US dollars (or from 42% to 36% of females' average earnings). Similarly, we find that a 10 percentage-point increase in the share of grade-mates' mothers with non-traditional beliefs during high school results in a decline of 1.9 percentage points in the likelihood that women ever receive welfare since adulthood relative to their male counterparts, the equivalent to a 20% decline in the raw gender gap. Both estimates are statistically significant at the 5% level and survive the Romano-Wolf adjustment.

5. Robustness Checks

Gender Norms Based on the Full Sample versus Gender Norms Based on the Core Sample

Because Add Health over-samples minorities, one may be concerned that our results may be affected by the proportion of non-traditional mothers in each grade and school being measured with error if the mothers of minority students and non-minority students systematically differ in terms of their gender identity norms. To assess whether this is the case, we first re-estimated our gender-identity norms variable using only the core sample, which is a random sample from each school and grade and does not over-sample minorities. This newly constructed variable is highly correlated with the one estimated using the full sample that we use for our main analysis (ρ=0.88). Second, we re-estimated equation (1) using this alternative gender-identity norms variable on the full sample, first, and the core sample, second (both estimates of β_3 for the three waves and the different summary indices are show in Appendix Table A.8 columns 2 and 3, respectively). Overall, these estimates are quite similar to those from our preferred

specification (also shown in Appendix Table A.8 column 1), suggesting that it is unlikely that over-sampling of minorities in Add Health is driving our results.

Attrition

Due to the longitudinal nature of Add Health, attrition is inevitable. Between Waves I and IV, Add Health loses about 40% of its sample, which is a potential problem for the validity of our medium- and long-term estimates (those from Waves III and IV) if attrition were systematically correlated with gender norms. To explore whether this is the case, we regress an attrition dummy on a female dummy, the proportion of non-traditional mothers, their interaction, and school and cohort fixed effects. The estimated coefficients $\hat{\beta}_2$and $\hat{\beta}_3$ are neither individually, nor jointly statistically significant, suggesting that attrition is independent on gender norms.[28]

Strategic Delay and Anticipation of School Entry

Our estimation strategy requires that there is no systematic selection into grades within the school. As noted by Black, Devereux and Salvanes (2013), endogeneity may arise if parents strategically choose the school starting age of their children in light of the peer composition of a particular cohort in a given school. In particular, our key explanatory variable (the share of non-traditional mothers) might be potentially endogenous in equation (1) if parents' decision to delay or anticipate their children's school entry is systematically linked to the gender related beliefs of mothers of children of the same grade. To deal with this issue we follow Bertoni, Brunello and Cappellari (2017) and define as school peers the students who belong to the same birth cohort and are enrolled in the same school. [29] More specifically, we re-compute all grade-level variables using this new peer group definition, and we replace the grade dummies with birth-year dummies and the share of non-traditional mothers in the grade with the share of non-traditional mothers in the birth cohort. We re-estimate equation (1) using this new peer group definition and, as it is shown in Appendix Table A.7 column 4, these estimates are very similar to those from our preferred specification. This is suggestive that parents' strategic choice of school starting age is not a particular issue for us and that it is unlikely that parents systematically delay or anticipate their children's school entry depending on their gender-identity beliefs.

[28] For the proportion of non-traditional mothers, the coefficient $\hat{\beta}_2$ is 0.043, the standard error is 0.067, and the t-statistic is 0.65. The coefficient for the proportion of non-traditional mothers interacted with the female dummy ($\hat{\beta}_3$) is -0.099, the standard error is 0.075, and the t-statistic is -1.32. The F-statistic for joint significance is 0.88.
[29] We define a birth cohort X as all children born between the 1st of September of year X and the 1st of September of year X+1, since the majority of schools in the US start the academic year on the 1st of September.

Falsification Tests

Finally, we check whether our results are spuriously picking up the effect of unobserved confounders or merely due to chance by performing falsification tests. In particular, we generate random data on maternal gender identity and, for each school and grade, we construct "fake" shares of non-traditional mothers using these 1000 random draws as we did in Section 3.2. This time, however, we use these randomly generated grade indicators of gender identity to re-estimate our benchmark model, which otherwise includes the same set of covariates as before. In line with the idea that our results are indeed genuine, these placebo regressions only deliver statistically significant results for $\hat{\beta}_3$ in less than 5% of the cases, which is less than what one would expect by chance. These results are summarized in Figure 1, which displays the placebo t-value distributions of the test $\beta_3 = 0$ for our four summary indices.

7. Conclusion

Using the National Longitudinal Study of Adolescent to Adult Health, we explore the causal effect of gender-identity norms on female teenagers' engagement in risky behaviors in the US, relative to their male counterparts. To do so, we exploit idiosyncratic variation across adjacent grades within schools in the proportion of high-school grade-mates' mothers who think that important skills for *both* girls and boys to learn are traditionally masculine skills (such as to think for herself or himself or work hard) as opposed to traditionally feminine skills (such as to be well-behaved, popular or help others). We find that a higher proportion of non-traditional mothers (that is, being exposed to less traditional gender-identity norms during high school) reduces the gender gap in risky behaviors, traditionally more prevalent among males, both in the short run, when individuals are still in high school, but also in the medium run, when they are transitioning into adulthood. This effect is driven by a decrease in boys' engagement in risky behaviors. In the long run, we also find that a greater proportion of non-traditional mothers improves adult females' annual earnings and reduces adult females' welfare dependency relative to their male counterparts. No labor-market effects are found among adult males.

Our work contributes to a recent literature studying the role of gender-identity norms on women's behavioral choices (Bertrand, Kamenica and Pan, 2015; Olivetti, Patacchini and Zenou, forthcoming). While these studies find evidence that gender-identity norms affect women's outcomes (such as labor force participation, hours worked, divorce and tenure), we

are the first to find evidence that gender-identity norms also shape males' behavioral choices and to look into risky behaviors. In particular, exposure to less traditional gender stereotypes during high school appears to weaken the perceived association between masculinity and traits such as fearlessness and boldness, reducing in turn males' engagement in risky behaviors. Overall, our findings suggest that gender-identity norms lead males to behave "more like females" in terms of their engagement in risky behaviors and females to behave "more like males" in the labor market.

References

Abrahamson, M. (2004). Alcohol in courtship contexts: focus-group interviews with young Swedish women and men. *Contemp. Drug Probs.*, 31: 3-29.

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

Akerlof, G. A., and Kranton, R. E. (2002). Identity and schooling: Some lessons for the economics of education. *Journal of Economic Literature*, 40(4): 1167-1201.

Akerlof, G. A., and Kranton, R. E. (2005). Identity and the Economics of Organizations. *Journal of Economic Perspectives*, 19(1): 9-32.

Anderson, M. (2008). Multiple Inference and Gender Differences in the Effects of Early Intervention: A Reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects. *Journal of the American Statistical Association*, vol. 103 (484): 1481-1495.

Anelli, M., and Peri, G. (2017). Peers' Composition Effects in the Short and in the Long Run: College Major, College Performance and Income. *Economic Journal* (October 2017). https://doi.org/10.1111/ecoj.12556

Angrist, J., and K. Lang. (2004). Does School Integration Generate Peer Effects? Evidence from Boston's Metco Program. *American Economic Review*, 94(5): 1613–34.

Antecol, H. (2000). An Examination of Cross-Country Differences in the Gender Gap in Labor Force Participation Rates. *Labour Economics*, 7(4): 409–26

Antecol, H. (2001). Why is there interethnic variation in the gender wage gap?: The role of cultural factors. *Journal of Human Resources*, 36(1): 119-143.

Autor, D., Figlio, D., Karbownik, K., Roth, J. and Wasserman, M., (2016), Family Disadvantage and the Gender Gap in Behavioral and Educational Outcomes, No. 22267, NBER Working Papers, National Bureau of Economic Research, Inc.

Battaglini, M., Díaz, C. and Patacchini, E. (2017). Self-control and peer groups: An empirical analysis. *Journal of Economic Behavior and Organization*, 134: 240-254.

Bertoni, M., Brunello, G. and Cappellari, L. (2017). Parents, Siblings and Schoolmates: the Effects of Family-School Interactions on Educational Achievement and Long-Term Labor Market Outcomes. *IZA Discussion Paper No. 11200.*

Bertrand, M. (2010). New Perspectives on Gender. In O. Ashenfelter, and D. Card (Eds.), Handbook of Labor Economics, Volume 4B (pp. 1545-1592). Amsterdam: Elsevier.

Bertrand, M. and Pan, J. (2013). The trouble with boys: Social influences and the gender gap in disruptive behavior. *American Economic Journal: Applied Economics*, 5(1): 32-64.

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

Bifulco, R., Fletcher, J. M., and Ross, S. L. (2011). The effect of classmate characteristics on post-secondary outcomes: Evidence from the Add Health. *American Economic Journal: Economic Policy*, 3(1): 25-53.

Bifulco, R., Fletcher, J. M., Oh, S. J., and Ross, S. L. (2014). Do high school peers have persistent effects on college attainment and other life outcomes? *Labour Economics*, 29: 83-90.

Black, S. E., Devereux, P. J., and Salvanes, K. G. (2013). Under pressure? The effect of peers on outcomes of young adults. *Journal of Labor Economics*, 31(1): 119-153.

Blau, F. D., and Kahn, L. M. (2000). Gender differences in pay. *Journal of Economic Perspectives*, 14(4): 75-99.

Blau, Francine D., Lawrence M. Kahn, Albert Yung-Hsu Liu, and Kerry L. Papps. 2013. "The Transmission of Women's Fertility, Human Capital, and Work Orientation across Immigrant Generations." *Journal of Population Economics,* 26(2): 405–35.

Carrell, S. E., Hoekstra, M., and Kuka, E. (2016). The long-run effects of disruptive peers. National Bureau of Economic Research, Working Paper w22042.

Charles, K. K., and Luoh, M. C. (2003). Gender differences in completed schooling. *The Review of Economics and Statistics*, 85(3): 559-577.

Elsner, B., and Isphording, I. E. (2017). Rank, Sex, Drugs and Crime. *Journal of Human Resources*, 0716-8080R.

Fernández, Raquel, and Alessandra Fogli. 2006. Fertility: The Role of Culture and Family Experience. *Journal of the European Economic Association*, 4(2-3): 552–61.

Fernández, Raquel, and Alessandra Fogli. 2009. Culture: An Empirical Investigation of Beliefs, Work, and Fertility. *American Economic Journal: Macroeconomics*, 1(1): 146–77.

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

Fortin, N. M., Oreopoulos, P., and Phipps, S. (2015). Leaving boys behind gender disparities in high academic achievement. *Journal of Human Resources*, 50(3): 549-579

Friesen, J., and Krauth, B. (2007). Sorting and inequality in Canadian schools. *Journal of Public Economics*, 91(11): 2185-2212.

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

Goldin, C., Katz, L. F., and Kuziemko, I. (2006). The homecoming of American college women: The reversal of the college gender gap. *The Journal of Economic Perspectives*, 20(4): 133-13.

Goldin, Claudia. 2006. The Quiet Revolution That Transformed Women's Employment, Education, and Family. *American Economic Review*, 96 (2): 1–21.

Hanushek, E. A., Kain, J. F., Markman, J. M., and Rivkin, S. G. (2003). Does peer ability affect student achievement? *Journal of Applied Econometrics*, 18(5): 527-544.

Heckman, J., Moon, S. H., Pinto, R., Savelyev, P., and Yavitz, A. (2010). Analyzing social experiments as implemented: A reexamination of the evidence from the High Scope Perry Preschool Program. *Quantitative Economics*, 1(1): 1-46.

Hoxby, Caroline. 2000. "Peer Effects in the Classroom: Learning from Gender and Race Variation." National Bureau of Economic Research, Working Paper 7867

Kaplan, M., Carriker, L., and Waldron, I. (1990). Gender differences in tobacco use in Kenya. *Social Science & Medicine*, 30(3): 305-310.

Keyes, K. M., Grant, B. F., and Hasin, D. S. (2007). Evidence for a closing gender gap in alcohol use, abuse, and dependence in the United States population. *Drug & Alcohol Dependence*, 93(1): 21-29.

Kling, J. R., Liebman, J. B., and Katz, L. F. (2007). Experimental analysis of neighborhood effects. *Econometrica*, 75(1): 83-119.

Lavy, V., and Schlosser, A. (2011). Mechanisms and impacts of gender peer effects at school. *American Economic Journal: Applied Economics*, 3(2): 1-33.

Lavy, V., M. D. Paserman, and A. Schlosser (2012). Inside the Black Box of Ability Peer Effects: Evidence from Variation in the Proportion of Low Achievers in the Classroom. *Economic Journal*, 122 (559): 208–237.

Murnane, R. J. (2013). US high school graduation rates: Patterns and explanations. *Journal of Economic Literature*, 51(2): 370-422.

Olivetti, C., Patacchini, E., and Zenou, Y. (2018). Mothers, Peers and Gender Identity.

Pope, D. G., and Sydnor, J. R. (2010). Geographic variation in the gender differences in test scores. *The Journal of Economic Perspectives*, 24(2): 95-108.

Rodríguez-Planas, N. (2017). School, Drugs, Mentoring and Peers: Evidence from a Randomized Trial in the U.S. *Journal of Economic Behavior and Organization*. July 2017, 139: 166-181.

Rodríguez-Planas, N. (2012). Longer-Term Impacts of Mentoring, Educational Services, and Learning Incentives: Evidence from a Randomized Trial in the U.S. *American Economic Journal: Applied Economics*, 4 (4): 121–139.

Rodríguez-Planas, N., and A. Sanz-de-Galdeano. (Forthcoming). Intergenerational Transmission of Gender Social Norms and Teenage Smoking. *Social Science & Medicine*.

Romano, J. P., and Wolf, M. (2005). Stepwise multiple testing as formalized data snooping. *Econometrica*, 73(4): 1237-1282.

Sznitman, S. R. (2007). Drugs and gender. *Nordic Studies on Alcohol and Drugs*, 24(2): 107-126.

Waldron, I., Bratelli, G., Carriker, L., Sung, W. C., Vogeli, C., and Waldman, E. (1988). Gender differences in tobacco use in Africa, Asia, the Pacific, and Latin America. *Social Science & Medicine*, 27(11): 1269-1275.

Warren CW, Jones NR, Eriksen MP, Asma S. Global Tobacco Surveillance System (GTSS) collaborative group. 2006. "Patterns of global tobacco use in young people and implications for future chronic disease burden in adults." *Lancet*, 367: 749-53.

Table 1. Panel A: Longitudinal Sample Description: Individual Chacteristics

	Mean	Std. Dev.
Grade 9	0.250	0.433
Grade 10	0.254	0.435
Grade 11	0.235	0.424
Grade 12	0.261	0.439
Age	16.949	1.222
White	0.654	0.476
Black	0.168	0.374
Hispanic	0.112	0.316
PPVT	102.671	13.484
High quality residential building	0.572	0.495
Resident parent college graduate	0.236	0.424
Parents live together	0.750	0.433
Parental age	42.545	6.123
Observations	8,181	

Note: All variables are measured using the in-home survey of Add Health for Wave I- Parental variables are meaured using Add Health's parental survey (Wave I). Observations are weighted using longitudinal weights.

Table 1. Panel B: Longitudinal Sample Description: Grade Characteristics

	Mean	Std. Dev.
Grade size	38.756	54.547
Average age	15.951	1.101
Share of female students	0.498	0.079
Share of minorities	0.400	0.297
Average PPVT	101.813	6.314
Observations	8,181	

Note: Grade size, average age, the share of female students and the share of minorities are constructed using Add Health in-school survey (Wave I); the share of non-traditional mothers is measured using the parental survey (Wave I); and average PPVT is measured using the in-home survey from Wave I. Observations are weighted using longitudinal weights.

Table 2: Summary Statistics for Main Outcomes. Longitudinal Sample

	Female mean/sd	Male mean/sd	Female-male
Share of non-traditional mothers	0.679	0.684	-0.005
	(0.136)	(0.132)	(0.004)
Regular smoker. W1	0.217	0.218	-0.002
	(0.412)	(0.413)	(0.015)
Got drunk during the past year.W1	0.361	0.398	-0.038**
	(0.480)	(0.490)	(0.017)
Ever tried marijuana. W1	0.329	0.373	-0.044**
	(0.470)	(0.484)	(0.018)
Ever tried other illegal drugs. W1	0.139	0.147	-0.008
	(0.346)	(0.354)	(0.014)
Expelled from school. W1	0.024	0.068	-0.044***
	(0.153)	(0.252)	(0.008)
Regular smoker. W3	0.290	0.323	-0.034**
	(0.454)	(0.468)	(0.016)
Got drunk during the past year.W3	0.488	0.601	-0.113***
	(0.500)	(0.490)	(0.018)
Ever tried marijuana. W3	0.551	0.635	-0.084***
	(0.497)	(0.481)	(0.017)
Ever tried other illegal drugs. W3	0.260	0.336	-0.076***
	(0.439)	(0.472)	(0.018)
Expelled from school. W3	0.046	0.133	-0.087***
	(0.210)	(0.340)	(0.011)
Had sex before 16. W3	0.305	0.277	0.028*
	(0.460)	(0.448)	(0.015)
Regular smoker. W4	0.255	0.315	-0.060***
	(0.436)	(0.465)	(0.016)
Got drunk during the past year.W4	0.411	0.569	-0.157***
	(0.492)	(0.495)	(0.017)
Ever tried marijuana W4	0.617	0.717	-0.101***
	(0.486)	(0.450)	(0.016)
Ever tried other illegal drugs. W4	0.317	0.438	-0.121***
	(0.465)	(0.496)	(0.016)
Ever worked for pay full time. W4	0.953	0.967	-0.014**
	(0.212)	(0.178)	(0.006)
Annual personal income (in thousand US dollars). W4	30.764	43.566	-12.801***
	(37.117)	(41.453)	(1.415)
Welfare recipient. W4	0.259	0.164	0.095***
	(0.438)	(0.371)	(0.015)
Observations	4,404	3,777	

*Note: In columns (1-2) standard deviations are displayed in parentheses, while in column 3 standard errors clustered at the school level are shown in parentheses. Observations are weighted using longitudinal weights. *** p<0.01, ** p<0.05, * p<0.1*

Table 3: Variation in Grade Composition Measures After Removing School Fixed Effect and Trends

	Panel A. Raw grade variables			
	Mean	SD	Min	Max
% of non-traditional mothers	0.682	0.134	0.235	1.000
	Panel B. Residuals after removing grade and school fixed effects			
	Mean	SD	Min	Max
% of non-traditional mothers	-0.000	0.081	-0.404	0.284
	Panel C. Residuals after removing grade fixed effects, school fixed effects and school trends			
	Mean	SD	Min	Max
% of non-traditional mothers	-0.000	0.068	-0.224	0.328
Observations	8,181			

Note: Observations are weighted using longitudinal weights.

Table 4: Balance Tests

Variable	Share of non-traditional mothers	Share of non-traditional mothers*Female
White	-0.039	0.082
	(0.057)	(0.061)
Black	0.053	-0.037
	(0.045)	(0.054)
Hispanic	-0.012	-0.049
	(0.052)	(0.049)
PPVT	-4.823*	2.075
	(2.757)	(2.673)
High quality residential building	-0.063	0.087
	(0.111)	(0.115)
Live with both parents	0.141	-0.121
	(0.101)	(0.095)
Parental age	1.160	0.071
	(1.399)	(1.584)
Total family income before tax 1994.	-0.048	0.174
(in hundred thousand US dollars)	(0.108)	(0.109)
Number of siblings	0.061	-0.448
	(0.274)	(0.343)
Mother born in the US	-0.030	0.091
	(0.061)	(0.064)
Mother smokes	-0.010	-0.068
	(0.101)	(0.118)
Father smokes	0.020	-0.146
	(0.101)	(0.148)
Mother is a college graduate	0.058	-0.010
	(0.070)	(0.074)
Father is a college graduate	0.038	0.027
	(0.079)	(0.088)
Observations	8,181	

*Note: The figures in each row are coefficients from regressions that include, in addition to the share of non-traditional mothers and its interaction with the female dummy, grade fixed effects and school fixed effects. All variables are measured using Add Health's Wave I. Standard errors (in parentheses) are clustered at the school level. *** p<0.01, ** p<0.05, * p<0.1*

Table 5: The Effect of Mothers of Grademates' Gender Identity on the Gender Gap in Risky Behaviors and Labor Maket Outcomes. Summary Indices.

	Risky behavior. W1	Risky behavior. W3	Risky behavior. W4	Labor market. W4
Female	-0.834	-2.124	-2.374	1.326
	(3.726)	(2.872)	(4.069)	(3.009)
Share of non-traditional mothers	-0.300*	-0.328**	-0.162	-0.0244
	(0.172)	(0.137)	(0.162)	(0.135)
Share of non-traditional mothers*Female	0.486***	0.361***	0.222	0.317**
	(0.138)	(0.128)	(0.166)	(0.149)
B2+B3 (effect for females)	0.186	0.0324	0.0591	0.292**
	(0.130)	(0.123)	(0.155)	(0.136)
Observations	8,181	8,181	8,181	8,181
R-squared	0.123	0.123	0.154	0.141

*Note: All regressions include the female dummy, the share of non-traditional mothers and its interaction with the female dummy, school and grade fixed effects, and school-specific time trends as well as the individual student covariates and grade-level characteristics listed in Table 1. Standard errors (in parentheses) are clustered at the school level. *** p<0.01, ** p<0.05, * p<0.1 ***

Table 6. Panel A: The Effect of Mothers of Grademates' Gender Identity on the Gender Gap in Risky Behaviors

Dependent variable:	Wave 1		Wave 3		Wave 4	
	Share of non-traditional mothers	Share of non-traditional mothers*Female	Share of non-traditional mothers:	Share of non-traditional mothers*Female	Share of non-traditional mothers:	Share of non-traditional mothers*Female
Regular smoker	-0.073	**0.148****	-0.182*	0.119	-0.164	0.162
	(0.082)	(0.074)	(0.102)	(0.093)	(0.106)	(0.105)
Got drunk during the past 12 months	-0.145	**0.260****	-0.108	0.104	-0.094	0.092
	(0.122)	(0.106)	(0.136)	(0.111)	(0.120)	(0.109)
Ever tried marijuana	-0.136	**0.252****	-0.207*	0.240**	-0.129	0.224**
	(0.129)	(0.121)	(0.124)	(0.113)	(0.109)	(0.109)
Ever tried other illegal drugs	-0.083	0.090*	-0.041	-0.006	0.089	-0.070
	(0.081)	(0.052)	(0.114)	(0.116)	(0.108)	(0.109)
Ever expelled from school	-0.094	**0.140*****	-0.156**	**0.223*****		
	(0.065)	(0.052)	(0.072)	(0.063)		
Had sex before 16			-0.129	0.193*		
			(0.125)	(0.112)		

*Note: All regressions include the female dummy, the share of non-traditional mothers and its interaction with the female dummy; school and grade fixed effects, and school-specific time trends as well as the individual student covariates and grade-level characteristics listed in Table 1. Standard errors (in parentheses) are clustered at the school level. No. of observations: 8,181. *** p<0.01, ** p<0.05, * p<0.1, in bold if Romano-Wolf p<0,1*

Table 6. Panel B: The Effect of Grademates Mothers' Gender Identity on the Gender Gap in Labor Market Outcomes

	Wave 4	
	Share of non-traditional mothers	Share of non-traditional mothers*Female
Ever worked for pay >35	0.039	0.001
hours per week	(0.038)	(0.044)
Log of personal income	-0.543	**1.405****
	(0.608)	(0.691)
Welfare recipient	0.037	**-0.189****
	(0.111)	(0.082)

*Note: All regressions include the female dummy, the share of non-traditional mothers and its interaction with the female dummy; school and grade fixed effects, and school-specific time trends as well as the individual student covariates and grade-level characteristics listed in Table 1. Standard errors (in parentheses) are clustered at the school level. No. of observations: 8,181. *** p<0.01, ** p<0.05, * p<0.1, in bold if Romano-Wolf p<0,1*

Appendix A. Definition of Summary Indexes.

Below, we define each of the different standardized indexes used in the main paper. As explained in section 3.1 of the paper, the standardized indexes are obtained as the average of the variables listed below, after each has been standardized by subtracting the control group's mean and dividing by the control group's standard deviation. Because the labor market summary index in Wave IV contains both adverse and beneficial outcomes, we switch the sign for the adverse outcome (welfare receipt), so that a higher value of this normalized measure represents a more "beneficial" outcome. Appendix Table A.3 presents summary statistics of these summary indexes by gender.

When youths are in their teens (Wave I):

Risky behaviors = (Regular smoker + Got drunk during the past year + Ever tried marijuana + Ever tried other illegal drugs + Expelled from school)/5

When youths are transitioning into adulthood (Wave III):

Risky behaviors = (Regular smoker + Got drunk during the past year + Ever tried marijuana + Ever tried other illegal drugs + Expelled from school + Had sex before 16)/6

When youths have become young adults (Wave IV):

Risky behaviors = (Regular smoker + Got drunk during the past year + Ever tried marijuana + Ever tried other illegal drugs)/4

Labor market = (Ever worked for pay full time + Log(Annual personal income) - Welfare recipient)/3

Table A1. Panel A: Relationship Between Women's Non-Traditional Beliefs and Labor Market and Educational Outcomes

	Coefficient	Standard Error
Works	0.0645***	(0.00707)
Hours worked	2.185***	(0.325)
Completed college	0.112***	(0.00776)
Years of education	0.887***	(0.0398)
Better educated than the spouse	0.0181**	(0.00922)
Only male works in the couple	-0.0450***	(0.00814)
Observations	15,686	

*Note: OLS coefficients from regressions of mothers' labor market and educational outcomes (reported in each row) on her own non-traditional gender beliefs indicator. Standard errors (in parentheses) are reported. *** p<0.01, ** p<0.05, * p<0.1*

Table A1. Panel B: Relationship Between Women's Non-Traditional Beliefs and Variables Related to the Gender Equality at County Level

	Coefficient	Standard Error
FLF opportunity index	0.0147**	(0.00617)
child women ratio (age 15-24)	-0.212***	(0.0518)
child women ratio (age 25-34)	-0.155	(0.118)
Observations	237	

*Note: OLS coefficients from regressions of country-level characteristics (reported in each row) on county level share of non-traditional mothers and standard errors (in parentheses) are reported. *** p<0.01, ** p<0.05, * p<0.1*

Table A2: Definition of Outcome Variables and Add Health Questions Used

	Using Waves I, III and IV of AddHealth:
Smoking	Youths who answered "*at least 10 out of 30 days*" to the question: "*During the past 30 days, on how many days did you smoke cigarettes?*"
Drinking	Youths who answered "*one or more days*" to the question "*Over the past 12 months, on how many days have you gotten drunk or "very, very high" on alcohol?*"
Smoking marijuana	Youths who reported an age when asked the question "*How old were you when you tried marijuana for the first time?*" in wave I; and youths who answered "*yes*" to the questions: "*Since June 1995, have you used marijuana?*", and "*Have you ever used any of the following drugs: marijuana?*" in Waves III and IV, respectively.
Illicit drugs (other than marijuana)	Youths who reported an age when asked at least one of the following questions: "*How old were you when you tried any kind of cocaine— including powder, freebase, or crack cocaine—for the first time?*", and "*How old were you when you first tried any other type of illegal drug such as LSD, PCP, ecstasy, mushrooms, speed, ice, heroin, or pills, without a doctor's prescription?"* in Wave I Youths who answered "*yes*" to at least one of the following questions : "*Since June 1995, have you used any kind of cocaine—including crack, freebase, or powder?*", "*Since June 1995, have you used crystal meth?*", and "*Since June 1995, have you used any other types of illegal drugs, such as LSD, PCP, ecstasy, mushrooms, inhalants, ice, heroin, or prescription medicines not prescribed for you?*" in Wave III. Youths who answered "yes" to the question "*Have you ever used any of the following drugs: cocaine, crystal meth or other types of illegal drugs, such as LSD, PCP, ecstasy, heroin, or mushrooms; or inhalants?"* in Wave IV.
Expelled from school	We coded as being expelled from school, youths who answered "*yes*" to the question "*Have you ever been expelled from school?*". This question was not asked in wave IV, so this indicator was constructed using only Waves I and III.
Having sex prior to age 16	Youths who responded "*16 years old or younger*" to the question "*How old were you the first time you had vaginal intercourse?*". This indicator was only constructed for wave III because many respondents were younger than 16 at Wave I.
	Using Wave IV of Add Health
Working full time	Individuals who answered "*yes*" to the question "*Have you ever worked full time at least 35 hours a week at a paying job while you were not primarily a student? Do not include summer work."*
Average yearly earnings	"*Now think about your personal earnings. In {2006/2007/2008}, how much income did you receive from personal earnings before taxes—that is, wages or salaries, including tips, bonuses, and overtime pay, and income from self-employment?"*
Welfare receipt	Individuals who answered "*yes*" to the question "*Between {1995/2002} and {2006/2007/2008}, did you or others in your household receive any public assistance, welfare payments, or food stamps?*"

Table A3: Summary Indexes

	Female mean/sd	Male mean/sd	Female-male
Wave I Risky behavior index	0.010	0.094	-0.085***
	(0.641)	(0.689)	(0.015)
Wave 3 Risky behavior index	-0.028	0.121	-0.149***
	(0.560)	(0.577)	(0.013)
Wave 4 Risky behavior index	-0.060	0.169	-0.229***
	(0.683)	(0.685)	(0.015)
Wave 4 Labor market index	-0.127	0.106	-0.233***
	(0.722)	(0.543)	(0.014)
Observations	4,404	3,777	

*Note: For columns (1-2) standard deviations are in parentheses and for column 3 standard errors clustered at the school level are in parentheses. Observations are weighted using longitudinal weights. *** p<0.01, ** p<0.05, * p<0.1*

Table A4: The Effect of Mothers of Grademates' Gender Identity on the Gender Gap in Risky Behaviors and Labor Maket Outcomes. Alternative Definition of Maternal Gender Norms.

	Risky behavior. W1	Risky behavior. W3	Risky behavior. W4	Labor market. W4
Female	-1.180	-2.369	-2.551	1.118
	(3.760)	(2.891)	(4.085)	(2.983)
Share of non-traditional mothers	-0.343*	-0.398***	-0.241	-0.0434
	(0.181)	(0.147)	(0.180)	(0.144)
Share of non-traditional mothers*Female	0.541***	0.374**	0.273	0.347**
	(0.143)	(0.143)	(0.194)	(0.159)
B2+B3 (effect for females)	0.198	-0.0237	0.0319	0.304**
	(0.141)	(0.140)	(0.168)	(0.142)
Observations	8,181	8,181	8,181	8,181
R-squared	0.123	0.123	0.154	0.141

*Note: The mother is definded as non-traditional if she believes that the most important thing to learn for a girl is "to think for herself" or "to work hard". All regressions include the female dummy, the share of non-traditional mothers and its interaction with the female dummy, school and grade fixed effects, and school-specific time trends as well as the individual student covariates and grade-level characteristics listed in Table 1. Standard errors (in parentheses) are clustered at the school level. *** p<0.01, ** p<0.05, * p<0.1 ****

Table A5: The Effect of Mothers' of Grademates Gender Identity on the Gender Gap in Risky Behaviors and Labor Market Outcomes.

		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Risky behavior. W1	Share of non-traditional mothers	-0.234	-0.325*	-0.334**	-0.302*	-0.300*		-0.259	-0.268	-0.345**	-0.299*	-0.351**
		(0.181)	(0.168)	(0.163)	(0.169)	(0.172)		(0.168)	(0.161)	(0.173)	(0.173)	(0.173)
	Share of non-traditional mothers*Female	0.507***	0.497***	0.506***	0.488***	0.486***	0.474***	0.422***	0.422***	0.575***	0.486***	0.583***
		(0.138)	(0.136)	(0.132)	(0.137)	(0.138)	(0.143)	(0.140)	(0.135)	(0.138)	(0.138)	(0.136)
Risky behavior. W3	Share of non-traditional mothers	-0.194	-0.352***	-0.341**	-0.315**	-0.328**		-0.300**	-0.307**	-0.360***	-0.322**	-0.350**
		(0.120)	(0.132)	(0.131)	(0.133)	(0.137)		(0.136)	(0.134)	(0.135)	(0.138)	(0.137)
	Share of non-traditional mothers*Female	0.364***	0.373***	0.379***	0.364***	0.361***	0.344**	0.327**	0.322**	0.389***	0.355***	0.373***
		(0.119)	(0.122)	(0.122)	(0.127)	(0.128)	(0.131)	(0.132)	(0.128)	(0.126)	(0.128)	(0.127)
Risky behavior. W4	Share of non-traditional mothers	-0.197	-0.228	-0.182	-0.155	-0.162		-0.111	-0.124	-0.188	-0.153	-0.167
		(0.164)	(0.162)	(0.160)	(0.161)	(0.162)		(0.162)	(0.161)	(0.165)	(0.162)	(0.166)
	Share of non-traditional mothers*Female	0.227	0.239	0.239	0.222	0.222	0.196	0.156	0.154	0.243	0.212	0.213
		(0.166)	(0.168)	(0.164)	(0.165)	(0.166)	(0.170)	(0.168)	(0.159)	(0.169)	(0.163)	(0.166)
Labor market. W4	Share of non-traditional mothers	-0.117	-0.0718	-0.00541	-0.0152	-0.0244		-0.0208	-0.0193	-0.0108	-0.0406	-0.0423
		(0.113)	(0.117)	(0.127)	(0.130)	(0.135)		(0.137)	(0.138)	(0.139)	(0.131)	(0.136)
	Share of non-traditional mothers*Female	0.342**	0.324**	0.322**	0.324**	0.317**	0.338**	0.301**	0.281*	0.249*	0.331**	0.294**
		(0.140)	(0.144)	(0.148)	(0.149)	(0.149)	(0.148)	(0.151)	(0.154)	(0.148)	(0.138)	(0.140)
Grade and school FE		Yes	Yes	Yes	Yes	Yes		Yes	Yes	Yes	Yes	Yes
School specific trend			Yes	Yes	Yes	Yes		Yes	Yes	Yes	Yes	Yes
Individual controls				Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Parental controls					Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Grade characteristics						Yes		Yes	Yes	Yes	Yes	Yes
Own mother is non-traditional*Female								Yes	Yes			
Own mother's behaviors*Female									Yes			
Share of working mothers*Female										Yes		Yes
Share of smoking mothers*Female											Yes	Yes
School-grade FE							Yes					

*Note: All regressions include the female dummy, the share of non-traditional mothers and its interaction with the female dummy, as well as school and grade fixed effects. The regressions in columns (2)-(10) control for school-specific trends. The regressions in columns (3)-(5) sequentially add individual-, parental- and grade-level controls. Individual and parental controls are listed in Table 1 Panel A; grade characteristics are listed in Table 1 Panel B. The regression in column (6) controls for an indicator that each student's own mother is non-traditional and its interaction with the female dummy; the regression in column (7) controls for indicators of maternal behavior, such as working and smoking, and their interactions with the female dummy. The regression in column (8) controls for the share of working mothers in the grade and its interaction with the female dummy; the regression in column (9) controls for the share of smoking mothers in the grade and its interaction with the female dummy; both are constructed at the school-grade level using Wave I parental survey of the Add Health. The regression in column (11) controls for school-grade fixed effects instead of school fixed effects and grade fixed effects. Standard errors (in parentheses) are clustered at the school level. No. of observations: 8,181. *** p<0.01, ** p<0.05, * p<0.1*

Table A6: The Effect of Mothers of Grademates' Gender Identity on the Gender Gap in Risky Behaviors and Labor Maket Outcomes. Table A4 Column 6 Specification, Own Mothers' Gender Identity Coefficients Displayed

	Risky behavior. W1	Risky behavior. W3	Risky behavior. W4	Labor market. W4
Female	-1.052	-2.262	-2.624	1.311
	(3.720)	(2.869)	(4.056)	(3.007)
Share of non-traditional mothers	-0.259	-0.300**	-0.111	-0.0208
	(0.168)	(0.136)	(0.162)	(0.137)
Share of non-traditional mothers*Female	0.422***	0.327**	0.156	0.301**
	(0.140)	(0.132)	(0.168)	(0.151)
Mother is non-traditional	-0.0145	0.0135	0.0169	-0.0244
	(0.0384)	(0.0296)	(0.0352)	(0.0267)
Mother is non-traditional*Female	0.0967**	0.0456	0.0913**	0.0344
	(0.0462)	(0.0410)	(0.0451)	(0.0483)
Observations	8,181	8,181	8,181	8,181
R-squared	0.126	0.124	0.157	0.141

*Note: This specification is the same as in Column 6 Table A.4, but this table also shows coefficients on own mothers' gender identity and its interaction with the female dummy. Standard errors (in parentheses) are clustered at the school level. *** p<0.01, ** p<0.05, * p<0.1 ****

Table A7: The Effect of Mothers' of Grademates Gender Identity on the Link Between Gender and Several Traits/Prescriptions

	Goes with gut feeling: 1 (strongly agree) to 5 (strongly disagree). W1	Mother would be upset if had sex. W1	Getting (someone) pregnant would be an embarrassment for
Female	-6.567	4.529	1.209
	(5.070)	(3.163)	(3.213)
Share of non-traditional mothers	0.263	0.115	0.176
	(0.285)	(0.128)	(0.123)
Share of non-traditional mothers*Female	-0.442**	-0.251**	-0.193**
	(0.206)	(0.100)	(0.0921)
B2+B3 (effect for females)	-0.179	-0.136	-0.0173
	(0.270)	(0.0904)	(0.105)
Observations	8,143	7,243	7,679
R-squared	0.110	0.164	0.172

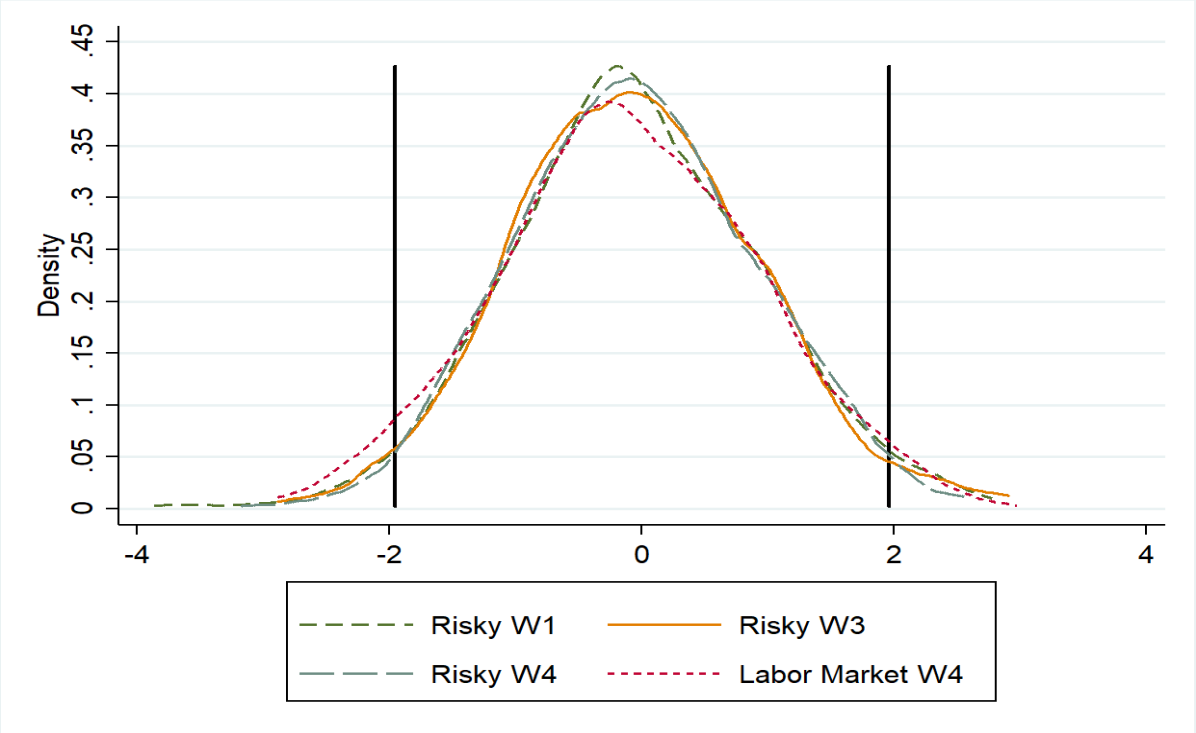
*Note: All regressions include the female dummy, the share of non-traditional mothers and its interaction with the female dummy, school and grade fixed effects, and school-specific time trends as well as the individual student covariates and grade-level characteristics listed in Table 1. Standard errors (in parentheses) are clustered at the school level. *** p<0.01, ** p<0.05, * p<0.1 ****

Table A8: Robustness of results. Coefficient for the share of non-traditional mothers*Female

	Full sample	Share of non-traditional mothers based on the core sample	Only observations from the core sample	No selective delay
Risky behavior. W1	0.486***	0.428***	0.487***	0.496***
	(0.138)	(0.126)	(0.151)	(0.150)
Risky behavior. W3	0.361***	0.273**	0.377***	0.376***
	(0.128)	(0.122)	(0.127)	(0.139)
Risky behavior. W4	0.222	0.174	0.180	0.239
	(0.166)	(0.138)	(0.173)	(0.162)
Labor market. W4	0.317**	0.277**	0.284*	0.381**
	(0.149)	(0.114)	(0.147)	(0.180)
Observations	8,181	8,181	4,725	7,893

*Note: Table reports the OLS coefficients for the share of non-traditional mothers interacted with the female dummy. In addition, all regressions include the female dummy, the share of non-traditional mothers, school and grade fixed effects, and school-specific time trends as well as the individual student covariates and grade level characteristics listed in Table 1. Standard errors (in parentheses) are clustered at the school level. *** p<0.01, ** p<0.05, * p<0.1*

Figure 1. Distribution of placebo t-values



Note: This graph shows the distributions of the t-values of the test $\beta_3 = 0$ obtained when estimating 1000 placebo regressions of equation (1) for our four indices. Instead of using the actual values of $NonTraditionalMothers_{-igs,1}$ we replace them with randomly generated grade-indicators of gender identity.