# Gender Attitudes and Later Life Outcomes

Sreevidya Ayyar[1], Uta Bolt[2], Eric French[3], Cormac O'Dea[4]

October 2023

[1]LSE, [2]University of Bristol, [3]University of Cambridge, [4]Yale University

# Gender attitudes are a potentially important driver of women's decisions

**Key questions:**

**Approach:**

# Gender attitudes are a potentially important driver of women's decisions

**Key questions:**

1. What are the determinants of gender attitudes?
2. How are traditional gender attitudes related to later life outcomes?

**Approach:**

# Gender attitudes are a potentially important driver of women's decisions

**Key questions:**

1. What are the determinants of gender attitudes?
2. How are traditional gender attitudes related to later life outcomes?

**Approach:**

Unique data set of essays written by girls at age 11 +

Information on their outcomes over the lifecycle

- Family formation
- Education
- Labour market outcomes

# Literature & Contribution

**Formation of gender attitudes:** Fernandez et al (2004), Alesina et al (2013), Giuliano (2021), Vella(1994), Johnston et al (2013), Akerlof and Kraton(2000), Dhar et al (2019, 2022)

**Effect of gender attitudes:** Fernandez & Fogli (2009), Boelman et al (2021), Bursztyn & Yanagizawa-Drott (2020), Bertrand(2011), Fortin(2005), Blau et al (2011), Johnston et al (2013), Vella(1994)

# Literature & Contribution

**Formation of gender attitudes:** Fernandez et al (2004), Alesina et al (2013), Giuliano (2021), Vella(1994), Johnston et al (2013), Akerlof and Kraton(2000), Dhar et al (2019, 2022)

**Effect of gender attitudes:** Fernandez & Fogli (2009), Boelman et al (2021), Bursztyn & Yanagizawa-Drott (2020), Bertrand(2011), Fortin(2005), Blau et al (2011), Johnston et al (2013), Vella(1994)

**Our contribution:**

1. Use of text data (age 11 essays) to elicit gender attitudes
    - Measured at individual level
    - Measured early in life
    - Can elicit underlying gender attitudes
2. Gender attitudes linked to outcomes over the whole lifecycle

# Data: National Child Development Study

- Population born in one week in Britain in 1958
- Followed at ages 0, 7, 11, 16, 23, 26, 33, 37, 42, 49, 55, (62)

# Data: National Child Development Study

- Population born in one week in Britain in 1958
- Followed at ages 0, 7, 11, 16, 23, 26, 33, 37, 42, 49, 55, (62)
- Detailed information:
  - Childhood: Family background, cognitive skills, non-cognitive skills
  - Adulthood: Educational attainment, hours worked, earnings, marital status, fertility
  - Age 11 essays: "Imagine you are now 25 years old. Write about the life you are leading, your interests, your home life and your work at the age of 25. (You have 30 minutes to do this)."

    → 3,514 essays. Extract underlying gender attitudes using natural language processing

# Extracting Gender Attitudes from Text Data

**Definition of Gender Attitudes:**

Girls have traditional gender attitudes if the roles and behaviours they support and want to take on in adulthood are stereotypically associated with females (Davis and Greenstein 2009).

# Extracting Gender Attitudes from Text Data

**Definition of Gender Attitudes:**

Girls have traditional gender attitudes if the roles and behaviours they support and want to take on in adulthood are stereotypically associated with females (Davis and Greenstein 2009).

**Identifying Assumption:**

Girls have more traditional gender attitudes if they *write* about typically feminine gender roles/norms or about engaging in a traditionally female-dominated activity.

# Extracting Gender Attitudes from Text Data

**Estimation Roadmap:**

1. Correct spelling mistakes and convert each essay into a vector of word counts

# Extracting Gender Attitudes from Text Data

**Estimation Roadmap:**

1. Correct spelling mistakes and convert each essay into a vector of word counts
2. Train a word-embedding model (WEM)
3. Construct a gender dimension by averaging across vector values of of gender word pairs
4. Project the words in an essay onto the gender dimension vector
5. Aggregate projection-weighted word counts for each essay, residualise on essay length, standardise.

# Word Embedding Model (WEM)

Represents words as real-valued vectors

$\rightarrow$ captures meanings and associations.

Train our WEM on 1 million books written between 1958 & 1978

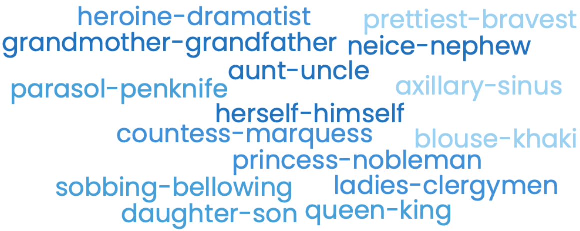
# Word Embedding Model (WEM)

Represents words as real-valued vectors

$\rightarrow$ captures meanings and associations.

Train our WEM on 1 million books written between 1958 & 1978

**Validity Check**: Which pairs of words are "closest" in the WEM to $\overrightarrow{\text{female}} - \overrightarrow{\text{male}}$?



# Extracting Gender Attitudes from Text Data

**Estimation Roadmap:**

1. Correct spelling mistakes and convert each essay into a vector of word counts
2. Train a word-embedding model (WEM)
3. Construct a gender dimension by averaging across vector values of gender word pairs
4. Project the words in an essay onto the gender dimension vector
5. Aggregate projection-weighted word counts, residualise on essay length, standardise.

# Gender Dimension Vector

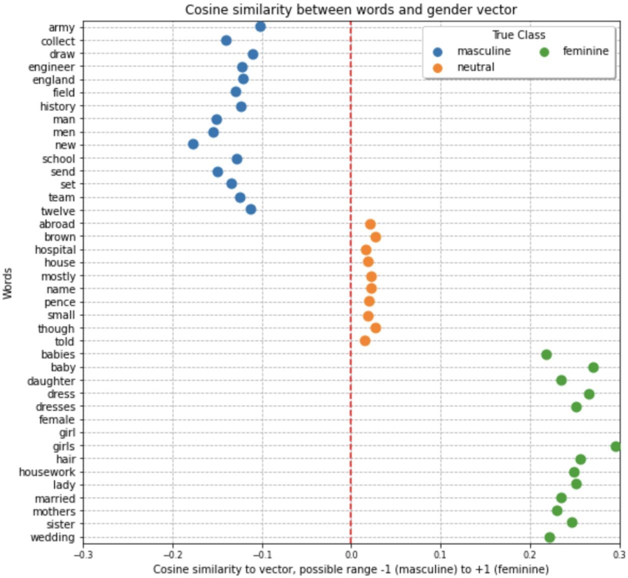
$$\vec{GD} \equiv \frac{1}{10}\{(\vec{woman} - \vec{man}) + (\vec{women} - \vec{men}) + (\vec{she} - \vec{he}) + (\vec{her} - \vec{him}) + (\vec{her} - \vec{his}) + (\vec{hers} - \vec{his}) + (\vec{girl} - \vec{boy}) + (\vec{girls} - \vec{boys}) + (\vec{female} - \vec{male}) + (\vec{feminine} - \vec{masculine})\}$$

- Vector's direction is from male to female
- Pairs of words adopted from Kozlowski et al. (2019)

**Estimation Roadmap:**

1. Correct spelling mistakes and convert each essay into a vector of word counts
2. Train a word-embedding model (WEM) on 1 million books written between 1958 & 1978
3. Construct a gender dimension by averaging across vector values of of gender word pairs
4. Project the words in an essay onto the gender dimension vector
5. Aggregate projection-weighted word counts, residualise on essay length, standardise.

# 5. Projecting Words onto the Gender Vector



# Extracting Gender Attitudes from Text Data

**Estimation Roadmap:**

1. Correct spelling mistakes and convert each essay into a vector of word counts
2. Train a word-embedding model (WEM) on 1 million books written between 1958 & 1978
3. Construct a gender dimension by averaging across vector values of of gender word pairs
4. Project the words in an essay onto the gender dimension vector
5. Create gender score

# Creating the Gender Score

1. For each essay: Sum over projection-weighted word counts
2. Partial out polynomial in essay length
3. Standardize (mean 0, variance 1)

$\rightarrow$ Traditional Attitude to Gender Score (TAGS)

# Results

1. **What determines gender attitudes?**
2. **How do traditional gender attitudes relate to** ...

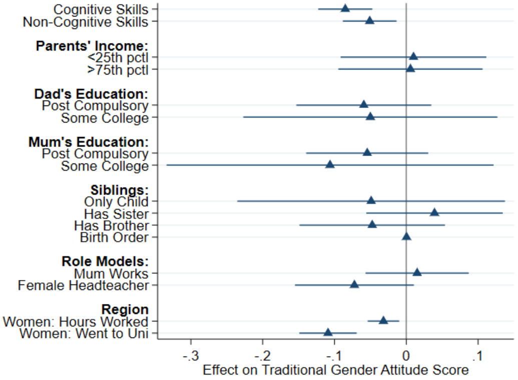
    ... family formation?

    ... education?

    ... labour supply?

    ... earnings & wages?

# What determines gender attitudes?




Notes: Additional controls include birth order, number of siblings and gender of sibling, as well as quadratic terms for cognitive and non-cognitive skills. Sample size is 3,514.

# Results

1. **What determines gender attitudes?**
2. **How do traditional gender attitudes relate to ...**

   ... **family formation?**

   ... education?

   ... labour supply?

   ... earnings & wages?

# How do traditional gender attitudes relate to family formation?

| | Married 23 | Ever Married | #kids 23 | #kids 50 |
|---|---|---|---|---|
| TAGS | 0.019** | 0.005 | 0.021* | 0.019 |
| | (0.009) | (0.006) | (0.013) | (0.021) |
| Log Parental Income | -0.013 | 0.023 | -0.046 | 0.013 |
| | (0.020) | (0.014) | (0.029) | (0.050) |
| Cognitive Skills | -0.046*** | -0.008 | -0.147*** | -0.096*** |
| | (0.010) | (0.007) | (0.015) | (0.025) |
| Non-Cognitive Skills | -0.008 | -0.010 | 0.021 | 0.007 |
| | (0.010) | (0.007) | (0.015) | (0.025) |

Notes: Additional controls include birth order, number of siblings, parental education (mother and father), parental aspirations, ethnicity, divorce / family difficulties, and region. All columns include dummies for missings. Sample size is 3,514.

# Results

1. **What determines gender attitudes?**
2. **How do traditional gender attitudes relate to ...**

   ... **family formation?**

   - Small effect on being married early and early fertility

   ... education?

   ... labour supply?

   ... earnings & wages?

# How do traditional gender attitudes relate to education?

| | Years of Education | Attend Uni |
|---|---|---|
| TAGS | -0.074*** | -0.014*** |
| | (0.025) | (0.005) |
| Log Parental Income | 0.147** | 0.018 |
| | (0.058) | (0.012) |
| Cognitive Skills | 0.781*** | 0.121*** |
| | (0.029) | (0.007) |
| Non-Cognitive Skills | -0.079*** | -0.014** |
| | (0.028) | (0.005) |

Additional controls: birth order, number of siblings, parental education (mother and father), parental aspirations, ethnicity, divorce / family difficulties, and region. All columns include dummies for missings. Sample size is 3,514.

# Results

1. **What determines gender attitudes?**
2. **How do traditional gender attitudes relate to ...**

   ... family formation?

   - No effect on early or completed fertility
   - Slightly more likely to be married early & continuously

   ... **education?**

   - 1 SD increase in TAGS decreases education by **0.9 months**.

   ... labour supply?

   ... earnings & wages?

# How do traditional gender attitudes relate to labour supply?

| | Employed | Employed | Employed | Employed |
|---|---|---|---|---|
| TAGS | -0.008** | -0.004 | -0.003 | -0.002 |
| | (0.004) | (0.004) | (0.004) | (0.004) |
| LogParentalIncome | | -0.013 | -0.016* | -0.015* |
| | | (0.009) | (0.009) | (0.008) |
| CognitiveSkills | | 0.031*** | 0.010 | 0.007 |
| | | (0.005) | (0.006) | (0.006) |
| Non-CognitiveSkills | | -0.008* | -0.006 | -0.006 |
| | | (0.004) | (0.004) | (0.004) |
| Cog16 | | | 0.019*** | 0.019*** |
| | | | (0.007) | (0.007) |
| YrsEd | | | 0.015*** | 0.016*** |
| | | | (0.003) | (0.003) |
| NumberofChildren | | | | -0.044*** |
| | | | | (0.004) |

Additional controls: squared cognition, birth order, number of siblings, parental education (mother and father), parental aspirations, ethnicity, divorce / family difficulties, and region. All columns include dummies for missings.

# How do traditional gender attitudes relate to labour supply?

| | AvgHours | AvgHours | AvgHours | AvgHours |
|---|---|---|---|---|
| TAGS | -0.387*** | -0.250* | -0.220* | -0.193 |
| | (0.128) | (0.130) | (0.130) | (0.126) |
| LogParentalIncome | | 0.174 | 0.137 | 0.166 |
| | | (0.290) | (0.290) | (0.275) |
| CognitiveSkills | | 1.059*** | 0.890*** | 0.758*** |
| | | (0.150) | (0.208) | (0.203) |
| Non-CognitiveSkills | | 0.285* | 0.296** | 0.309** |
| | | (0.148) | (0.149) | (0.142) |
| Cog16 | | | -0.213 | -0.231 |
| | | | (0.236) | (0.229) |
| YrsEd | | | 0.300*** | 0.307*** |
| | | | (0.087) | (0.084) |
| NumberofChildren | | | | -2.079*** |
| | | | | (0.118) |

Additional controls: squared cognition, birth order, number of siblings, parental education (mother and father), parental aspirations, ethnicity, divorce / family difficulties, and region. All columns include dummies for missings Disaggregated

# Results

1. **What determines gender attitudes?**
2. **How do traditional gender attitudes relate to ...**

   ... family formation?

   - No effect on early or completed fertility
   - Slightly more likely to be married early & continuously

   ... education?

   - 1 SD increase in TAGS decreases education by 0.9 months

   ... **labour supply?**

   - No effect on extensive margin
   - intensive margin: 1 SD increase in TAGS $\downarrow$ hours by approx 15 mins per week

   ... earnings & wages?

# How do traditional gender attitudes relate to earnings & wages

| | LTE | LTE | LTE | LTE |
|---|---|---|---|---|
| TAGS | -0.042*** | -0.027** | -0.020 | -0.018 |
| | (0.013) | (0.013) | (0.013) | (0.012) |
| LogParentalIncome | | 0.003 | -0.010 | -0.007 |
| | | (0.029) | (0.029) | (0.028) |
| CognitiveSkills | | 0.127*** | 0.055*** | 0.044** |
| | | (0.015) | (0.020) | (0.019) |
| Non-CognitiveSkills | | -0.007 | -0.001 | 0.001 |
| | | (0.014) | (0.014) | (0.014) |
| Cog16 | | | 0.011 | 0.009 |
| | | | (0.022) | (0.022) |
| YrsEd | | | 0.080*** | 0.080*** |
| | | | (0.009) | (0.008) |
| NumberofChildren | | | | -0.186*** |
| | | | | (0.013) |

Additional controls: squared cognition, birth order, number of siblings, parental education (mother and father), parental aspirations, ethnicity, divorce / family difficulties, and region. All columns include dummies for missings.

# How do traditional gender attitudes relate to earnings & wages

| | ln avg wage | ln avg wage | ln avg wage | ln avg wage |
|---|---|---|---|---|
| TAGS | -0.034*** | -0.020* | -0.013 | -0.011 |
| | (0.011) | (0.011) | (0.011) | (0.010) |
| LogParentalIncome | | 0.001 | -0.012 | -0.010 |
| | | (0.024) | (0.024) | (0.024) |
| CognitiveSkills | | 0.105*** | 0.030* | 0.022 |
| | | (0.013) | (0.016) | (0.016) |
| Non-CognitiveSkills | | -0.016 | -0.009 | -0.008 |
| | | (0.011) | (0.011) | (0.011) |
| Cog16 | | | 0.015 | 0.014 |
| | | | (0.019) | (0.018) |
| YrsEd | | | 0.079*** | 0.079*** |
| | | | (0.007) | (0.007) |
| NumberofChildren | | | | -0.123*** |
| | | | | (0.011) |

Additional controls: squared cognition, birth order, number of siblings, parental education (mother and father), parental aspirations, ethnicity, divorce / family difficulties, and region. All columns include dummies for missings.

# Results

1. **What determines gender attitudes?**
2. **How do traditional gender attitudes relate to ...**

... family formation?

- No effect on early or completed fertility
- Slightly more likely to be married early & continuously

... education?

- 1 SD increase in TAGS decreases education by **0.9 months**.

... labour supply?

- No effect on extensive margin
- intensive margin: 1 SD increase in TAGS $\downarrow$ hours by approx 15 mins per week

... **earnings & wages?**

- Effects on earnings and wages, driven by education

# Results

1. **What determines gender attitudes?**
2. **How do traditional gender attitudes relate to …**

   … family formation?

   - No effect on early or completed fertility
   - Slightly more likely to be married early & continuously

   … education?

   - 1 SD increase in TAGS decreases education by **0.9 months**.

   … labour supply?

   - No effect on extensive margin
   - intensive margin: 1 SD increase in TAGS $\downarrow$ hours by approx 15 mins per week

   … earnings & wages?

   - Effects on earnings and wages, driven by education

# More Results

- Other dimensions of interest
- A model

# Conclusion

Use novel text data set to construct index of gender attitudes for large sample of girls at age 11

**1. What determines attitudes?**

Hard to say - skills, characteristics of the area, parental education

**2. How are traditional gender attitudes related to later life outcomes?**

Girls with more traditional gender attitudes ...

... attain less education

... work fewer hours

... earn less

Thank you!

# Other dimensions?

| | Years of Education | Log Lifetime Earnings | Employed | Average Hours | Log Average Wages |
|---|---|---|---|---|---|
| Gender Attitude Index | -0.063** | -0.024* | -0.004 | -0.265** | -0.017 |
| | (0.026) | (0.013) | (0.004) | (0.133) | (0.011) |
| Education Score | 0.019 | 0.020 | 0.002 | -0.104 | 0.023* |
| | (0.028) | (0.015) | (0.005) | (0.145) | (0.012) |
| Cultivation Score | 0.040 | 0.001 | -0.002 | 0.034 | -0.001 |
| | (0.033) | (0.017) | (0.005) | (0.171) | (0.014) |
| Affluence Score | -0.022 | 0.003 | 0.005 | -0.025 | 0.005 |
| | (0.027) | (0.013) | (0.004) | (0.135) | (0.011) |
| Status Score | -0.072*** | -0.016 | -0.002 | 0.227* | -0.017 |
| | (0.026) | (0.014) | (0.004) | (0.133) | (0.011) |
| Morality Score | -0.047 | -0.009 | -0.000 | 0.049 | -0.007 |
| | (0.032) | (0.016) | (0.005) | (0.163) | (0.013) |
| Employment Score | 0.036 | -0.009 | -0.006 | 0.074 | -0.008 |
| | (0.029) | (0.015) | (0.004) | (0.149) | (0.012) |

Additional controls: Birth order, number of siblings, cognitive skills, log parental income, cognitive skills, non-cognitive skills Back

# A Model

In adulthood, women maximise

$$
\begin{aligned}
V_{adult}(ed) &= max_{c_m, c_h, l, hp}\ u(c_m, c_h, l, hp) \\
&= max_{c_m, c_h, l, hp}\ ln[(1-\gamma)c_m^{\tau} + \gamma c_h^{\tau}]^{\frac{1}{\tau}} + \delta ln(l)
\end{aligned}
$$

subject to:

$$
\begin{aligned}
T &= hp + l + l_m \\
c_m &= l_m \cdot w(ed) \\
c_h &= \alpha \cdot hp
\end{aligned}
$$

# Model

In childhood, i.e. at school-leaving age, girls assume they will behave optimally in adulthood. Moreover, assume there is no uncertainty. In this period, girls choose how much education to attain and how much to work. They maximise

$$\begin{aligned} V_{child} &= max_{ed,l_{16}} u(c_{16}, l_{16}) + \beta V_{adult}(ed) \\ &= max_{ed,l_{16}} ln(c_{16}) + \delta ln(l_{16}) + \beta V_{adult}(ed) \end{aligned}$$

subject to:

$$\begin{aligned} T &= l_{16} + ed + l_{m,16} \\ c_{16} &= l_{m,16} \cdot w_1 \end{aligned}$$

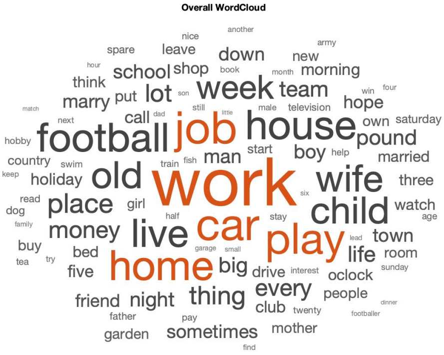
Women with higher $\gamma$, that is more feminine gender attitudes, will...

1. ...attain less education () have lower wages, only if w(ed) is increasing in education)
2. ...work more at age 16 (for contemporaneous consumption)
3. ...works less in adulthood relative to time spent in home production

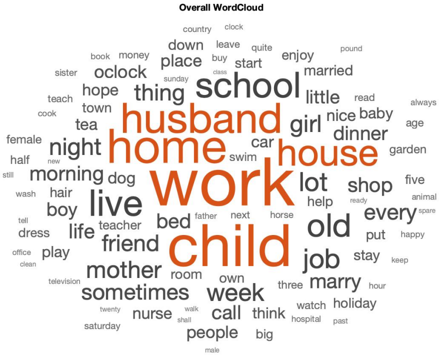
Back

# Word Clouds

Back




(a) Male




(b) Female

# Traditional Attitude to Gender Score

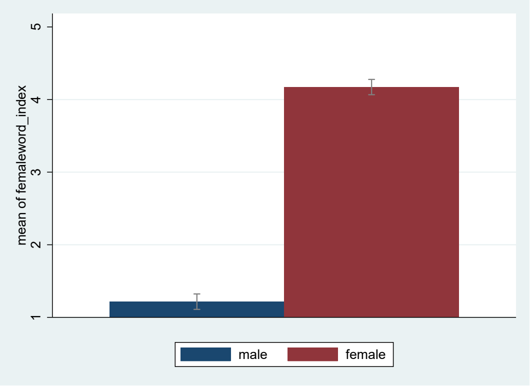
1. Select words traditionally associated with women

| | |
|---|---|
| Domestic chores: | cook, wash, clean, breakfast, knit, tidy, kitchen, housework, sew, washing, cake |
| Childcare: | boy, girl, little, baby |
| Other words: | husband, wedding, house, stay, hostess, part, let |

# Traditional Attitude to Gender Score

1. Select words traditionally associated with women

2. Count how often each essay mentions a word in the list



# Traditional Attitude to Gender Score

1. Select words traditionally associated with women

2. Count how often each essay mentions a word in the list

3. Partial out polynomial in essay length

# Traditional Attitude to Gender Score

1. Select words traditionally associated with women
2. Count how often each essay mentions a word in the list
3. Partial out polynomial in essay length
4. Standardize: Mean 0, Variance 1

# Traditional Attitude to Gender Score

1. Select words traditionally associated with women
2. Count how often each essay mentions a word in the list
3. Partial out polynomial in essay length
4. Standardize: Mean 0, Variance 1.

$\Rightarrow$ **Traditional Attitudes to Gender Score (TAGS)**

**Essay 1: TAGS = 2.5 (very traditional)**

**Essay 2: TAGS = -0.75 (somewhat non-traditional)**

# Word Embedding

Word embeddings:

1. Represent words as real-valued vectors
2. $\rightarrow$ detailed analysis of word meanings and associations.
3. Example: $\overrightarrow{\text{king}} - \overrightarrow{\text{man}} + \overrightarrow{\text{woman}} = \overrightarrow{\text{queen}}$
4. We use Word2vec by Google to create word embeddings - trained on large corpus of text
5. Word2vec by Google is the most common approach to creating word embeddings

   Word2vec algorithm learns a representation of words in high-dimensional vector space by training a shallow, two-layer neural network on a large corpus of text split into n-gram phrases. By embedding words in a numerical space, we can apply linear algebra to words.
6. word2vec takes as its input a large corpus of text and produces a vector space, typically of several hundred dimensions, with each unique word in the corpus being assigned a corresponding vector in the space.

# Hours: Disaggregated

| | Hours23 | Hours33 | Hours42 | Hours50 | Hours55 |
|---|---|---|---|---|---|
| TAGS | 0.293* | -0.539** | -0.004 | -0.301 | -0.362 |
| | (0.172) | (0.246) | (0.248) | (0.220) | (0.273) |
| Log Parental Income | 0.309 | 1.119** | -0.130 | -0.071 | -0.483 |
| | (0.379) | (0.562) | (0.564) | (0.546) | (0.652) |
| Cognitive Skills | 0.457** | 1.723*** | 1.304*** | 0.394 | 2.162*** |
| | (0.216) | (0.293) | (0.299) | (0.277) | (0.346) |
| Non-Cognitive Skills | 0.196 | 0.726** | 0.158 | 0.189 | 0.153 |
| | (0.194) | (0.283) | (0.273) | (0.265) | (0.329) |
| Number of Siblings | -0.033 | 0.127 | 0.549*** | 0.421** | 0.361 |
| | (0.172) | (0.189) | (0.202) | (0.185) | (0.234) |
| Birth Order | 0.370* | 0.011 | 0.136 | -0.254 | -0.143 |
| | (0.200) | (0.268) | (0.264) | (0.251) | (0.314) |

Additional controls: squared cognition, birth order, number of siblings, parental education (mother and father), parental aspirations, ethnicity, divorce / family difficulties, and region. All columns include dummies for missings. Back