Homophily and Transmission of Behavioral traits in Friendship networks

Palaash Bhargava [1] Daniel Chen [2] Matthias Sutter [3] Camille Terrier [4]

[1]Columbia University

[2]Toulouse School of Economics

[3]Max Planck Institute, University of Bonn

[4]Queen Mary University of London

February 24th, 2023

Motivation

- **What determines the personality and socio-economic outcomes of individuals?**
  - Intra family characteristics
  - Education
  - **Social Networks** Further Discussion
  - **Behavioral Traits** Further Discussion

- **How do social networks and behavioral traits interact?**
  - Do social networks exhibit homophily on behavioral traits?
  - Are these behavioral traits malleable? Can they be influenced by your peers?
  - Does malleability of social actions imply malleability of behavioral traits?

Motivation

- **What determines the personality and socio-economic outcomes of individuals?**
  - Intra family characteristics
  - Education
  - **Social Networks** Further Discussion
  - **Behavioral Traits** Further Discussion

- **How do social networks and behavioral traits interact?**
  - Do social networks exhibit homophily on behavioral traits?
  - Are these behavioral traits malleable? Can they be influenced by your peers?
  - Does malleability of social actions imply malleability of behavioral traits?

Motivation

- **What determines the personality and socio-economic outcomes of individuals?**
  - Intra family characteristics
  - Education
  - **Social Networks** Further Discussion
  - **Behavioral Traits** Further Discussion
- **How do social networks and behavioral traits interact?**
  - Do social networks exhibit homophily on behavioral traits?
  - Are these behavioral traits malleable? Can they be influenced by your peers?
  - Does malleability of social actions imply malleability of behavioral traits?

Motivation

- **What determines the personality and socio-economic outcomes of individuals?**
  - Intra family characteristics
  - Education
  - **Social Networks** Further Discussion
  - **Behavioral Traits** Further Discussion
- **How do social networks and behavioral traits interact?**
  - Do social networks exhibit homophily on behavioral traits?
  - Are these behavioral traits malleable? Can they be influenced by your peers?
  - Does malleability of social actions imply malleability of behavioral traits?

Research questions

1. **What patterns of homophily do we find w.r.t. behavioral traits?**
2. **Does this homophily exist due to transmission of behavioral traits or do similar individuals simply seek out each other?**
    1. What are the potential mechanisms behind this transmission?
    2. Are these transmissions similar across demographic characteristics?

Research questions

1. **What patterns of homophily do we find w.r.t. behavioral traits?**
2. **Does this homophily exist due to transmission of behavioral traits or do similar individuals simply seek out each other?**
    1. What are the potential mechanisms behind this transmission?
    2. Are these transmissions similar across demographic characteristics?

Research questions

1. **What patterns of homophily do we find w.r.t. behavioral traits?**
2. **Does this homophily exist due to transmission of behavioral traits or do similar individuals simply seek out each other?**
    1. What are the potential mechanisms behind this transmission?
    2. Are these transmissions similar across demographic characteristics?

Overview - Setup

1. We partner with 67 French high schools in 3 regions: Créteil, Nantes and Montpellier
   - **Lab in class:** Measures of behavioral traits from incentivized games.
   - Match this data with Administrative data from the French Ministry of Education.
   - Additional information on college / occupational aspirations, gender biases, policy preferences, self reported news exposure and SEL outcomes collected.
2. Peer group **(Endogenous)** - upto 5 closest friends reported from the class. Links are directed.
3. Descriptive OLS regressions to document homophily controlling for a host of fixed effects and shared characteristics.
4. Estimation stategy for peer effects: 3SLS IV approach motivated by Bramoullé et al. [2009] / Case and Katz [1991] and Konig et al. [2019].

Overview - Setup

1. We partner with 67 French high schools in 3 regions: Créteil, Nantes and Montpellier
   - **Lab in class:** Measures of behavioral traits from incentivized games.
   - Match this data with Administrative data from the French Ministry of Education.
   - Additional information on college / occupational aspirations, gender biases, policy preferences, self reported news exposure and SEL outcomes collected.
2. Peer group **(Endogenous)** - upto 5 closest friends reported from the class. Links are directed.
3. Descriptive OLS regressions to document homophily controlling for a host of fixed effects and shared characteristics.
4. Estimation stategy for peer effects: 3SLS IV approach motivated by Bramoullé et al. [2009] / Case and Katz [1991] and Konig et al. [2019].

Overview - Setup

1. We partner with 67 French high schools in 3 regions: Créteil, Nantes and Montpellier
   - **Lab in class:** Measures of behavioral traits from incentivized games.
   - Match this data with Administrative data from the French Ministry of Education.
   - Additional information on college / occupational aspirations, gender biases, policy preferences, self reported news exposure and SEL outcomes collected.
2. Peer group **(Endogenous)** - upto 5 closest friends reported from the class. Links are directed.
3. Descriptive OLS regressions to document homophily controlling for a host of fixed effects and shared characteristics.
4. Estimation stategy for peer effects: 3SLS IV approach motivated by Bramoullé et al. [2009] / Case and Katz [1991] and Konig et al. [2019].

Overview - Setup

1. We partner with 67 French high schools in 3 regions: Créteil, Nantes and Montpellier
   - **Lab in class:** Measures of behavioral traits from incentivized games.
   - Match this data with Administrative data from the French Ministry of Education.
   - Additional information on college / occupational aspirations, gender biases, policy preferences, self reported news exposure and SEL outcomes collected.
2. Peer group **(Endogenous)** - upto 5 closest friends reported from the class. Links are directed.
3. Descriptive OLS regressions to document homophily controlling for a host of fixed effects and shared characteristics.
4. Estimation stategy for peer effects: 3SLS IV approach motivated by Bramoullé et al. [2009] / Case and Katz [1991] and Konig et al. [2019].

Overview - Setup

1. We partner with 67 French high schools in 3 regions: Créteil, Nantes and Montpellier
   - **Lab in class:** Measures of behavioral traits from incentivized games.
   - Match this data with Administrative data from the French Ministry of Education.
   - Additional information on college / occupational aspirations, gender biases, policy preferences, self reported news exposure and SEL outcomes collected.
2. Peer group **(Endogenous)** - upto 5 closest friends reported from the class. Links are directed.
3. Descriptive OLS regressions to document homophily controlling for a host of fixed effects and shared characteristics.
4. Estimation stategy for peer effects: 3SLS IV approach motivated by Bramoullé et al. [2009] / Case and Katz [1991] and Konig et al. [2019].

Overview - Results

- **Homophily results:**
  - Large degree of homophily on demographic characteristics (not explained by homophily on behavioral traits)
  - Large degree of homophily on behavioral traits, above and beyond homophily on demographic characteristics
  - Similarity in demographic characteristics (w.r.t. gender and middle school) strongly amplifies homophily on behavioral traits
  - Higher the number of similar traits, higher the homophily.
- **Peer effects results:**
  - Students are influenced considerably by the behavioral traits of their peers.
    - Positive peer effects w.r.t. depth of reasoning, risk tolerance, coordination, cooperation
    - Negative peer effects w.r.t. competition.
  - Similarity in demographic characteristics (gender), longevity of friendship and popularity of peers strongly amplify peer effects in behavioral traits.

Overview - Results

- **Homophily results:**
  - Large degree of homophily on demographic characteristics (not explained by homophily on behavioral traits)
  - Large degree of homophily on behavioral traits, above and beyond homophily on demographic characteristics
  - Similarity in demographic characteristics (w.r.t. gender and middle school) strongly amplifies homophily on behavioral traits
  - Higher the number of similar traits, higher the homophily.
- **Peer effects results:**
  - Students are influenced considerably by the behavioral traits of their peers.
    - Positive peer effects w.r.t. depth of reasoning, risk tolerance, coordination, cooperation
    - Negative peer effects w.r.t. competition.
  - Similarity in demographic characteristics (gender), longevity of friendship and popularity of peers strongly amplify peer effects in behavioral traits.

Overview - Results

- **Homophily results:**
  - Large degree of homophily on demographic characteristics (not explained by homophily on behavioral traits)
  - Large degree of homophily on behavioral traits, above and beyond homophily on demographic characteristics
  - Similarity in demographic characteristics (w.r.t. gender and middle school) strongly amplifies homophily on behavioral traits
  - Higher the number of similar traits, higher the homophily.
- **Peer effects results:**
  - Students are influenced considerably by the behavioral traits of their peers.
    - Positive peer effects w.r.t. depth of reasoning, risk tolerance, coordination, cooperation
    - Negative peer effects w.r.t. competition.
  - Similarity in demographic characteristics (gender), longevity of friendship and popularity of peers strongly amplify peer effects in behavioral traits.

Overview - Results

- **Homophily results:**
  - Large degree of homophily on demographic characteristics (not explained by homophily on behavioral traits)
  - Large degree of homophily on behavioral traits, above and beyond homophily on demographic characteristics
  - Similarity in demographic characteristics (w.r.t. gender and middle school) strongly amplifies homophily on behavioral traits
  - Higher the number of similar traits, higher the homophily.
- **Peer effects results:**
  - Students are influenced considerably by the behavioral traits of their peers.
    - Positive peer effects w.r.t. depth of reasoning, risk tolerance, coordination, cooperation
    - Negative peer effects w.r.t. competition.
  - Similarity in demographic characteristics (gender), longevity of friendship and popularity of peers strongly amplify peer effects in behavioral traits.

Overview - Results

- **Homophily results:**
  - Large degree of homophily on demographic characteristics (not explained by homophily on behavioral traits)
  - Large degree of homophily on behavioral traits, above and beyond homophily on demographic characteristics
  - Similarity in demographic characteristics (w.r.t. gender and middle school) strongly amplifies homophily on behavioral traits
  - Higher the number of similar traits, higher the homophily.
- **Peer effects results:**
  - Students are influenced considerably by the behavioral traits of their peers.
    - Positive peer effects w.r.t. depth of reasoning, risk tolerance, coordination, cooperation
    - Negative peer effects w.r.t. competition.
  - Similarity in demographic characteristics (gender), longevity of friendship and popularity of peers strongly amplify peer effects in behavioral traits.

Overview - Results

- **Homophily results:**
  - Large degree of homophily on demographic characteristics (not explained by homophily on behavioral traits)
  - Large degree of homophily on behavioral traits, above and beyond homophily on demographic characteristics
  - Similarity in demographic characteristics (w.r.t. gender and middle school) strongly amplifies homophily on behavioral traits
  - Higher the number of similar traits, higher the homophily.
- **Peer effects results:**
  - Students are influenced considerably by the behavioral traits of their peers.
    - Positive peer effects w.r.t. depth of reasoning, risk tolerance, coordination, cooperation
    - Negative peer effects w.r.t. competition.
  - Similarity in demographic characteristics (gender), longevity of friendship and popularity of peers strongly amplify peer effects in behavioral traits.

Overview - Results

- **Homophily results:**
  - Large degree of homophily on demographic characteristics (not explained by homophily on behavioral traits)
  - Large degree of homophily on behavioral traits, above and beyond homophily on demographic characteristics
  - Similarity in demographic characteristics (w.r.t. gender and middle school) strongly amplifies homophily on behavioral traits
  - Higher the number of similar traits, higher the homophily.
- **Peer effects results:**
  - Students are influenced considerably by the behavioral traits of their peers.
    - Positive peer effects w.r.t. depth of reasoning, risk tolerance, coordination, cooperation
    - Negative peer effects w.r.t. competition.
  - Similarity in demographic characteristics (gender), longevity of friendship and popularity of peers strongly amplify peer effects in behavioral traits.

Related literature

1. **Behavioral traits and socio-economic outcomes:** Luhrmann et al. [2018], Sutter et al. [2013], Caliendo et al. [2010, 2014], Buser et al. [2014]

2. **Peer effects:** Shan and Zolitz [2022], Rao [2019], Alan et al. [2020], Santavarita and Sarzosa [2019], Case and Katz [1991], Gaviria and Raphael [2001], Lee et al [2020], Patacchini et al. [2017], Epple and Romano [2011], Sacerdote [2014]

3. **Homophily on behavioral traits:** Girard et al. [2015], Jackson et al. [2022], Charroin et al. [2022]

Experimental setup

- Lab in class - 3 regions of France - Créteil, Nantes and Montpellier, 67 schools
- Economic measures of behavioral traits ellicited from high school students (Seconde, Premier, Terminale) through standard incentivised games.
- Students asked to report **upto 5 close friends** from their class.
- Data then matched to administrative data from French Education Ministry.
- Additional information on college / occupational aspirations, gender biases, policy preferences, self reported news exposure and SEL outcomes collected.

Measures of behavioral traits

- **Tolerance for inequality:** Redistribution game, Cappelan et al. [2007]
- **Trust:** Berge et al. [1995]
- **Morality:** Donation to UNICEF, Kirchler et al. [2016]
- **Generosity:** Donate part of earnings to a charity
- **Cooperation:** Angerer et al. [2016]
- **Coordination:** Cooper et al. [1990]
- **Risk tolerance:** Bomb task, Crosetto and Filipin [2013]
- **Competitiveness:** Piece rate vs Competitive pay for performance, Niederle and Vesterlind [2007]
- **Depth of Reasoning:** Beauty Contest game, Nagel [1995]
- **Educational aspirations:** Highest academic degree one wishes to obtain

Preview of some games

Administrative data

- Gender
- Ethnicity generated by ethnicolor race predictor (Python)
- Nationality
- Age
- Socio-economic status: Parent's PCS occupation category (upto second digit categorization)
  - Agriculteurs = 1, Employès = 5, Ouvriers = 6, Retraitès = 7, Autres inactifs = 8, inconnu = 9 (low skill or out of market)
  - Chefs d'entreprise = 2, Cadres = 3, Prof. Intermèdiaires = 4 (high and intermediate skill)
- Commune of residence
- Country of birth
- Number of children from each parent.

Summary Statistics Typical Classroom Network

Documenting Homophily

- We document homophily by investigating how the probability of friendship between two students in classroom depend on the similarity in behavioral traits.
- We run a separate OLS regression for each trait.

$$d_{ij} = \beta_0 + \beta_1 \left(-|y_i - y_j|\right) + \beta_2 \mathbf{1}[\mathbf{x}_i = \mathbf{x}_j] + \zeta_i + \psi_j + \nu_{ij}$$

  - d_{ij} is a potential friendship pair. Potential friendship pair restricted to within the classroom.
  - y_i captures the normalised measure of the trait.
  - $\mathbf{x}_i$ captures the demographic characteristics: age, ethnicity, nationality, country of birth, parental occupation, age, number of siblings, dummy for single child and commune of residence.
  - ζ_i, ψ_j: sender and receiver fixed effects.
- Friendship links are directed, i.e. $d_{ij} = 1$ doesn't necessary imply $d_{ji} = 1$.

Documenting Homophily

- We document homophily by investigating how the probability of friendship between two students in classroom depend on the similarity in behavioral traits.
- We run a separate OLS regression for each trait.

$$d_{ij} = \beta_0 + \beta_1 \left(-|y_i - y_j|\right) + \beta_2 \mathbf{1}[\mathbf{x}_i = \mathbf{x}_j] + \zeta_i + \psi_j + \nu_{ij}$$

  - d_{ij} is a potential friendship pair. Potential friendship pair restricted to within the classroom.
  - y_i captures the normalised measure of the trait.
  - $\mathbf{x}_i$ captures the demographic characteristics: age, ethnicity, nationality, country of birth, parental occupation, age, number of siblings, dummy for single child and commune of residence.
  - ζ_i, ψ_j: sender and receiver fixed effects.

- Friendship links are directed, i.e. $d_{ij} = 1$ doesn't necessary imply $d_{ji} = 1$.

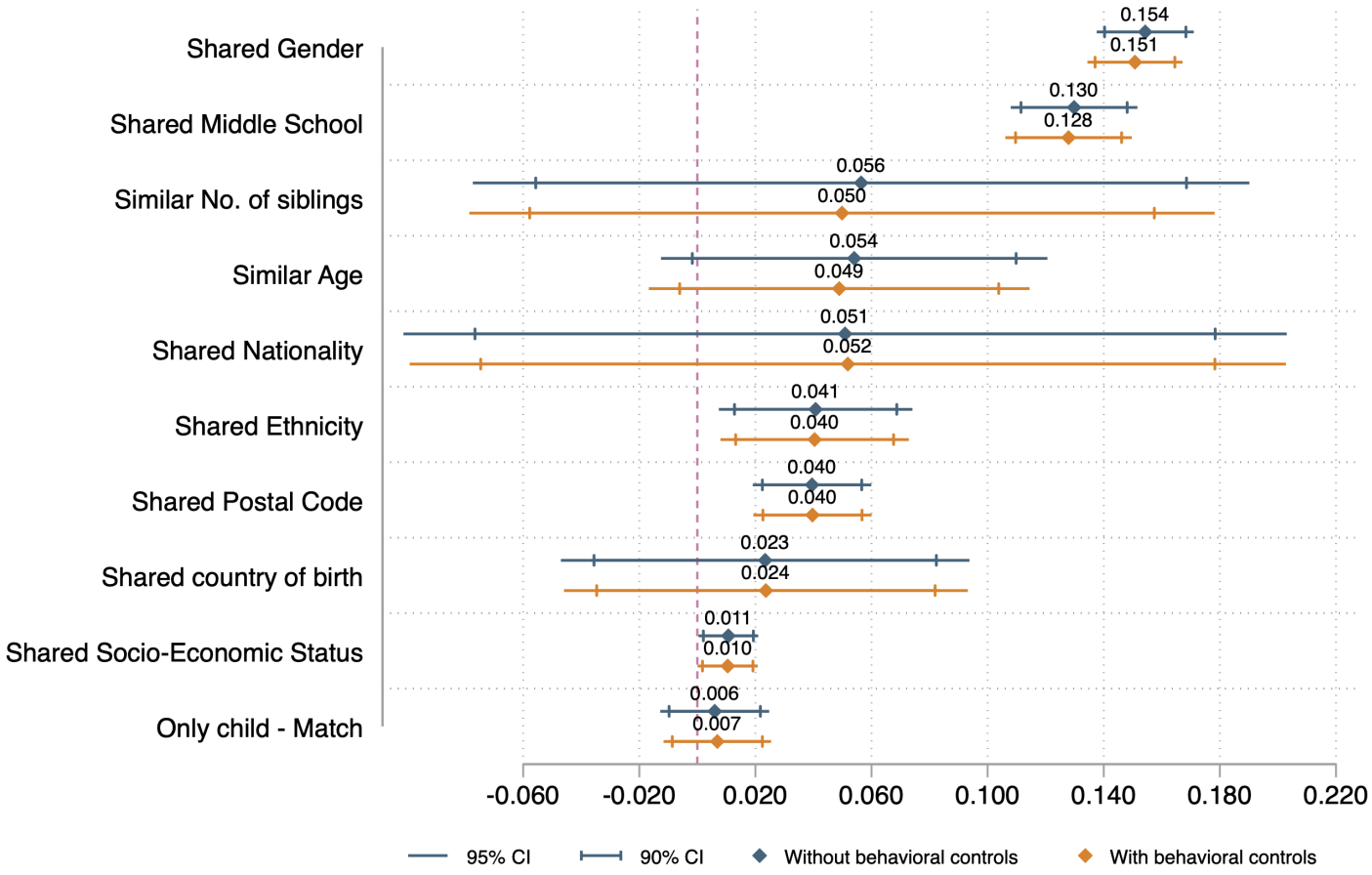
Documenting Homophily

- We document homophily by investigating how the probability of friendship between two students in classroom depend on the similarity in behavioral traits.
- We run a separate OLS regression for each trait.

$$d_{ij} = \beta_0 + \beta_1 \left(-|y_i - y_j|\right) + \beta_2 \mathbf{1}[\mathbf{x}_i = \mathbf{x}_j] + \zeta_i + \psi_j + \nu_{ij}$$

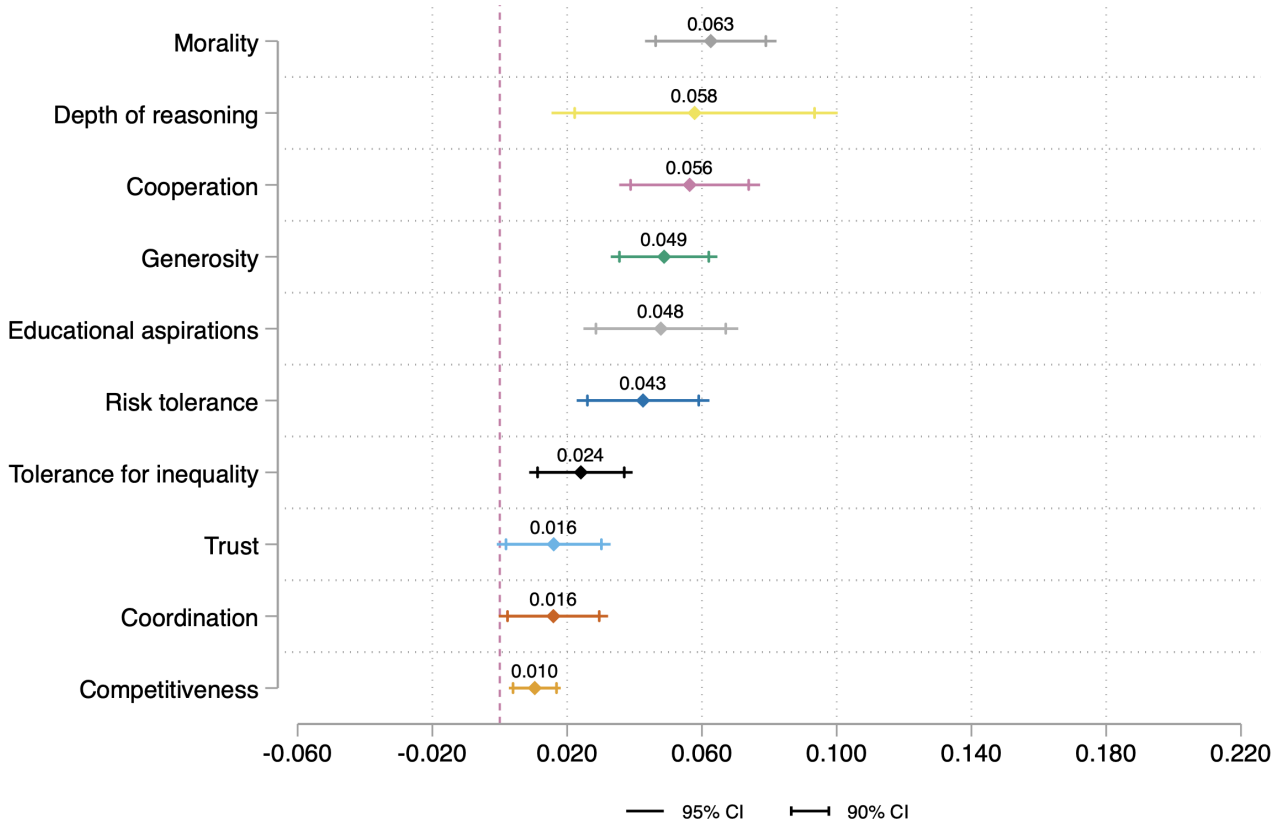
  - d_{ij} is a potential friendship pair. Potential friendship pair restricted to within the classroom.
  - y_i captures the normalised measure of the trait.
  - $\mathbf{x}_i$ captures the demographic characteristics: age, ethnicity, nationality, country of birth, parental occupation, age, number of siblings, dummy for single child and commune of residence.
  - ζ_i, ψ_j: sender and receiver fixed effects.
- Friendship links are directed, i.e. $d_{ij} = 1$ doesn't necessary imply $d_{ji} = 1$.

Homophily Fact 1



High school students exhibit large homophily based on demographic characteristics.

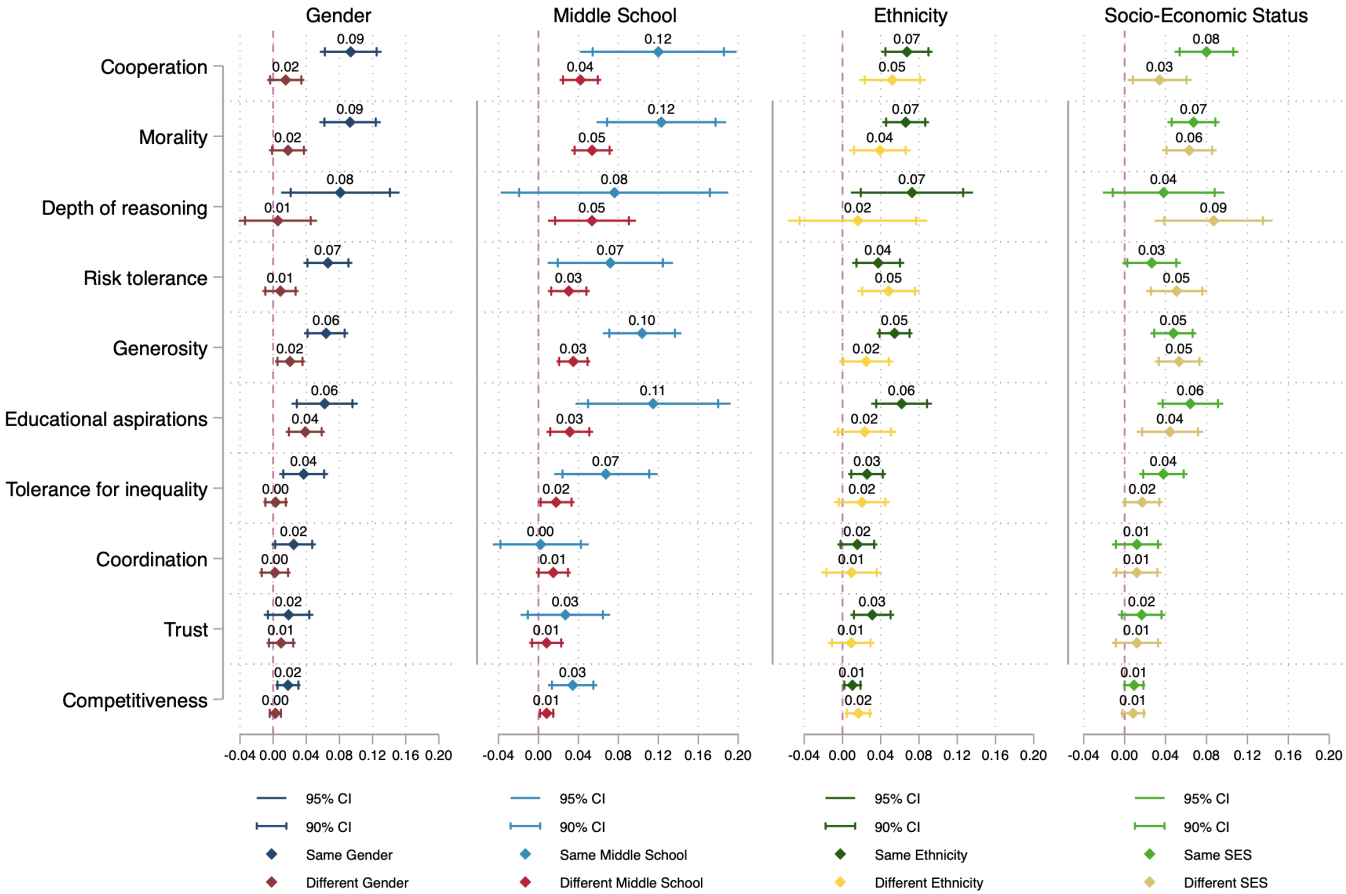
Homophily Fact 2



High school students exhibit large homophily based on behavioral traits, above and beyond the well-documented homophily on demographic characteristics.

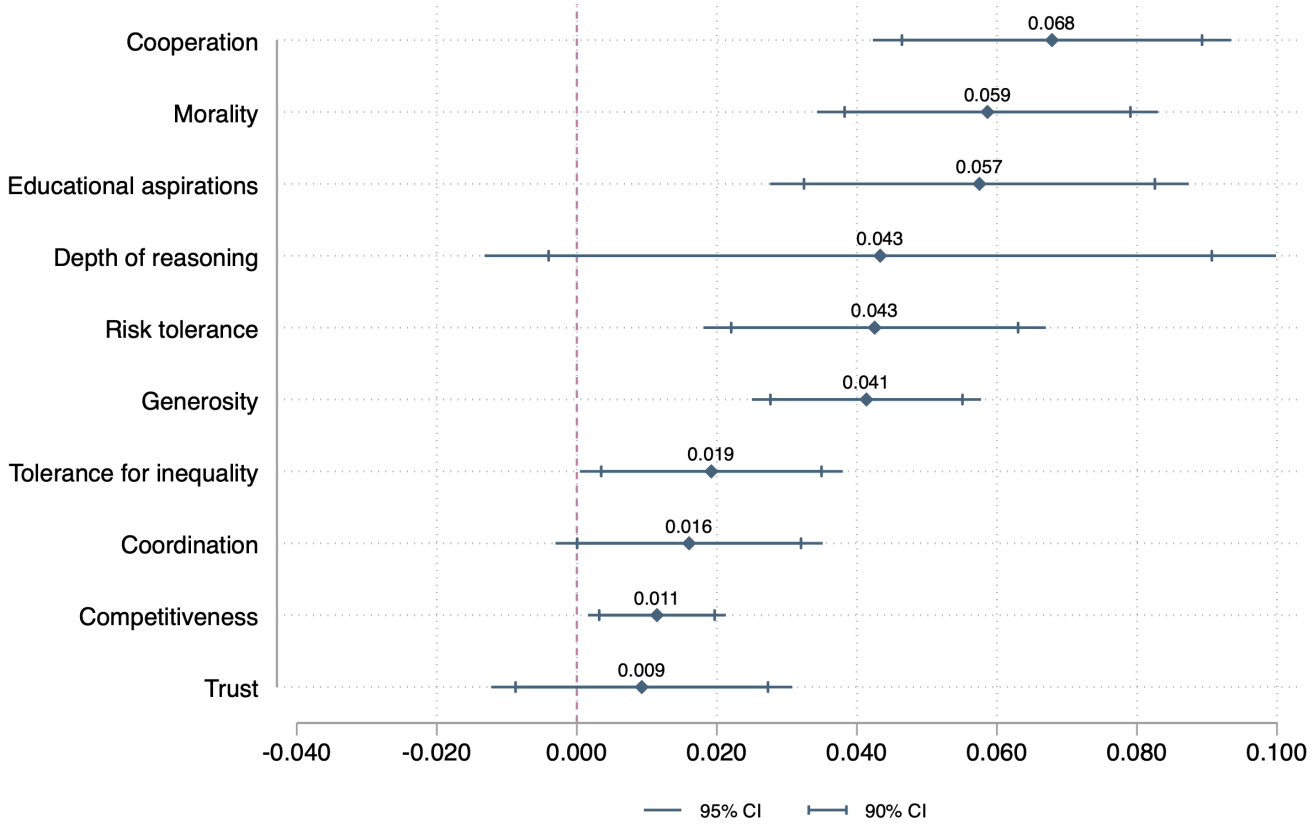
Alternate Network Specifications

Homophily Fact 3



Similarity in demographic characteristics (such as gender or middle school attended) strongly amplifies homophily based on behavioral traits.

Homophily Fact 4



The larger the number of behavioral traits that students share, the higher friendship chances are. In other words, similarity in one behavioral trait does not substitute well for similarity in another trait when it comes to determining friendships.

Homophily $\Longrightarrow$ Peer effects?

- Natural question which pops up: do we see this homophily because similar people are attracted to each other or because individuals become friends and then behavioral traits transmit?
- Are peer effects higher when groups are more homogenous?
- To answer these questions, we now turn to our peer effects analysis.

Peer effects - Basic Equation

- To identify peer effects, basic equation to estimate:

$$y_{li} = \beta \frac{\sum_{j \in P_{li}} y_{lj}}{n_{li}} + \gamma \mathbf{x}_{li} + \delta \frac{\sum_{j \in P_{li}} \mathbf{x}_{lj}}{n_{li}} + \eta_l + \epsilon_{li} \quad (1)$$

where i indexes individual, j indexes friend, y_{li}: behavioral trait, $\mathbf{x}_{li}$: demographic characteristics, η_l: classroom fixed effect

- Let $\mathbf{G}$ denote the interaction (adjacency) matrix such that $G_{ij} = \frac{1}{n_{li}}$ if j is a friend of i and 0 otherwise. (Row normalised)
- Eq. 1 can be rewritten in the matrix form:

$$\mathbf{y} = \beta \mathbf{G}\mathbf{y} + \gamma \mathbf{x} + \delta \mathbf{G}\mathbf{x} + \eta + \epsilon \quad (2)$$

- The above empirical equation can be rationalised through a simple behavior adjustment and social cohesion model.

Social Cohesion Game

Peer effects - Basic Equation

- To identify peer effects, basic equation to estimate:

$$y_{li} = \beta \frac{\sum_{j \in P_{li}} y_{lj}}{n_{li}} + \gamma \mathbf{x}_{li} + \delta \frac{\sum_{j \in P_{li}} \mathbf{x}_{lj}}{n_{li}} + \eta_l + \epsilon_{li} \quad (1)$$

  where i indexes individual, j indexes friend, y_{li}: behavioral trait, $\mathbf{x}_{li}$: demographic characteristics, η_l: classroom fixed effect

- Let $\mathbf{G}$ denote the interaction (adjacency) matrix such that $G_{ij} = \frac{1}{n_{li}}$ if j is a friend of i and 0 otherwise. (Row normalised)

- Eq. 1 can be rewritten in the matrix form:

$$\mathbf{y} = \beta \mathbf{G}\mathbf{y} + \gamma \mathbf{x} + \delta \mathbf{G}\mathbf{x} + \eta + \epsilon \quad (2)$$

- The above empirical equation can be rationalised through a simple behavior adjustment and social cohesion model.

Social Cohesion Game

Peer effects - Basic Equation

- To identify peer effects, basic equation to estimate:

$$y_{li} = \beta \frac{\sum_{j \in P_{li}} y_{lj}}{n_{li}} + \gamma \mathbf{x}_{li} + \delta \frac{\sum_{j \in P_{li}} \mathbf{x}_{lj}}{n_{li}} + \eta_l + \epsilon_{li} \quad (1)$$

  where i indexes individual, j indexes friend, y_{li}: behavioral trait, $\mathbf{x}_{li}$: demographic characteristics, η_l: classroom fixed effect

- Let $\mathbf{G}$ denote the interaction (adjacency) matrix such that $G_{ij} = \frac{1}{n_{li}}$ if j is a friend of i and 0 otherwise. (Row normalised)

- Eq. 1 can be rewritten in the matrix form:

$$\mathbf{y} = \beta \mathbf{G}\mathbf{y} + \gamma \mathbf{x} + \delta \mathbf{G}\mathbf{x} + \eta + \epsilon \quad (2)$$

- The above empirical equation can be rationalised through a simple behavior adjustment and social cohesion model.

Social Cohesion Game

Identification of peer effects - I

- Issues to tackle to identify β:

1. **Manski's reflection problem:** Do I impact my friends or my friends impact me? In real life, probably both. How do we disentangle?
2. **Correlated effects:** Unobservable components in our shared environment impact us similarly, same school, same class, same teacher, presence of bars around our school etc.
3. **Endogenous link formation:** Who am I friends with is not random? Friendships based on shared experiences, traits, characteristics of self and family.
    - **G** can be independently correlated with **y**.

Identification of peer effects - I

- Issues to tackle to identify β:

1. **Manski's reflection problem:** Do I impact my friends or my friends impact me? In real life, probably both. How do we disentangle?

2. **Correlated effects:** Unobservable components in our shared environment impact us similarly, same school, same class, same teacher, presence of bars around our school etc.

3. **Endogenous link formation:** Who am I friends with is not random? Friendships based on shared experiences, traits, characteristics of self and family.
    - **G** can be independently correlated with **y**.

Identification of peer effects - I

- Issues to tackle to identify β:

1. **Manski's reflection problem:** Do I impact my friends or my friends impact me? In real life, probably both. How do we disentangle?

2. **Correlated effects:** Unobservable components in our shared environment impact us similarly, same school, same class, same teacher, presence of bars around our school etc.

3. **Endogenous link formation:** Who am I friends with is not random? Friendships based on shared experiences, traits, characteristics of self and family.
   - **G** can be independently correlated with **y**.

Identification of peer effects - I

- Issues to tackle to identify β:

1. **Manski's reflection problem:** Do I impact my friends or my friends impact me? In real life, probably both. How do we disentangle?

2. **Correlated effects:** Unobservable components in our shared environment impact us similarly, same school, same class, same teacher, presence of bars around our school etc.

3. **Endogenous link formation:** Who am I friends with is not random? Friendships based on shared experiences, traits, characteristics of self and family.
    - **G** can be independently correlated with **y**.

Identification of peer effects - II

- Solution: Use a 3SLS strategy
  - Tackle **correlated effects** by differencing out classroom fixed effects
  - Tackle **reflection problem** by instrumenting the endogenous behavioral trait of friends ($\mathbf{Gy}$) with the exogenous characteristics (gender, race, nationality etc.) of friends of friends and friends of friends of friends ($\mathbf{G^2x}, \mathbf{G^3x}$) (Bramoullé et al. [2009], Case and Katz [1991]) Reflection problem
  - Tackle **endogeneity in link formation** by replacing the adjacency matrix $\mathbf{G}$ with a predicted adjacency matrix $\hat{\mathbf{G}}$, i.e. instruments for $\mathbf{Gy}$ are $Z = [\hat{\mathbf{G}}\mathbf{x}, \hat{\mathbf{G}}^2\mathbf{x}, \hat{\mathbf{G}}^3\mathbf{x}]$ Endogenous links
- Instruments are valid under the assumption that peers within the class who the individual doesn't interact with as friends do not differentially impact the behavioral traits of the individual apart from the impact already captured by classroom fixed effects.

Identification of peer effects - II

- Solution: Use a 3SLS strategy
  - Tackle **correlated effects** by differencing out classroom fixed effects
  - Tackle **reflection problem** by instrumenting the endogenous behavioral trait of friends ($\mathbf{Gy}$) with the exogenous characteristics (gender, race, nationality etc.) of friends of friends and friends of friends of friends ($\mathbf{G^2x}, \mathbf{G^3x}$) (Bramoullé et al. [2009], Case and Katz [1991]) Reflection problem
  - Tackle **endogeneity in link formation** by replacing the adjacency matrix $\mathbf{G}$ with a predicted adjacency matrix $\hat{\mathbf{G}}$, i.e. instruments for $\mathbf{Gy}$ are $Z = [\hat{\mathbf{G}}\mathbf{x}, \hat{\mathbf{G}}^2\mathbf{x}, \hat{\mathbf{G}}^3\mathbf{x}]$ Endogenous links
- Instruments are valid under the assumption that peers within the class who the individual doesn't interact with as friends do not differentially impact the behavioral traits of the individual apart from the impact already captured by classroom fixed effects.

Identification of peer effects - II

- Solution: Use a 3SLS strategy
  - Tackle **correlated effects** by differencing out classroom fixed effects
  - Tackle **reflection problem** by instrumenting the endogenous behavioral trait of friends ($\mathbf{Gy}$) with the exogenous characteristics (gender, race, nationality etc.) of friends of friends and friends of friends of friends ($\mathbf{G^2x}, \mathbf{G^3x}$) (Bramoullé et al. [2009], Case and Katz [1991]) Reflection problem
  - Tackle **endogeneity in link formation** by replacing the adjacency matrix $\mathbf{G}$ with a predicted adjacency matrix $\hat{\mathbf{G}}$, i.e. instruments for $\mathbf{Gy}$ are $Z = [\hat{\mathbf{G}}\mathbf{x}, \hat{\mathbf{G}}^2\mathbf{x}, \hat{\mathbf{G}}^3\mathbf{x}]$ Endogenous links
- Instruments are valid under the assumption that peers within the class who the individual doesn't interact with as friends do not differentially impact the behavioral traits of the individual apart from the impact already captured by classroom fixed effects.

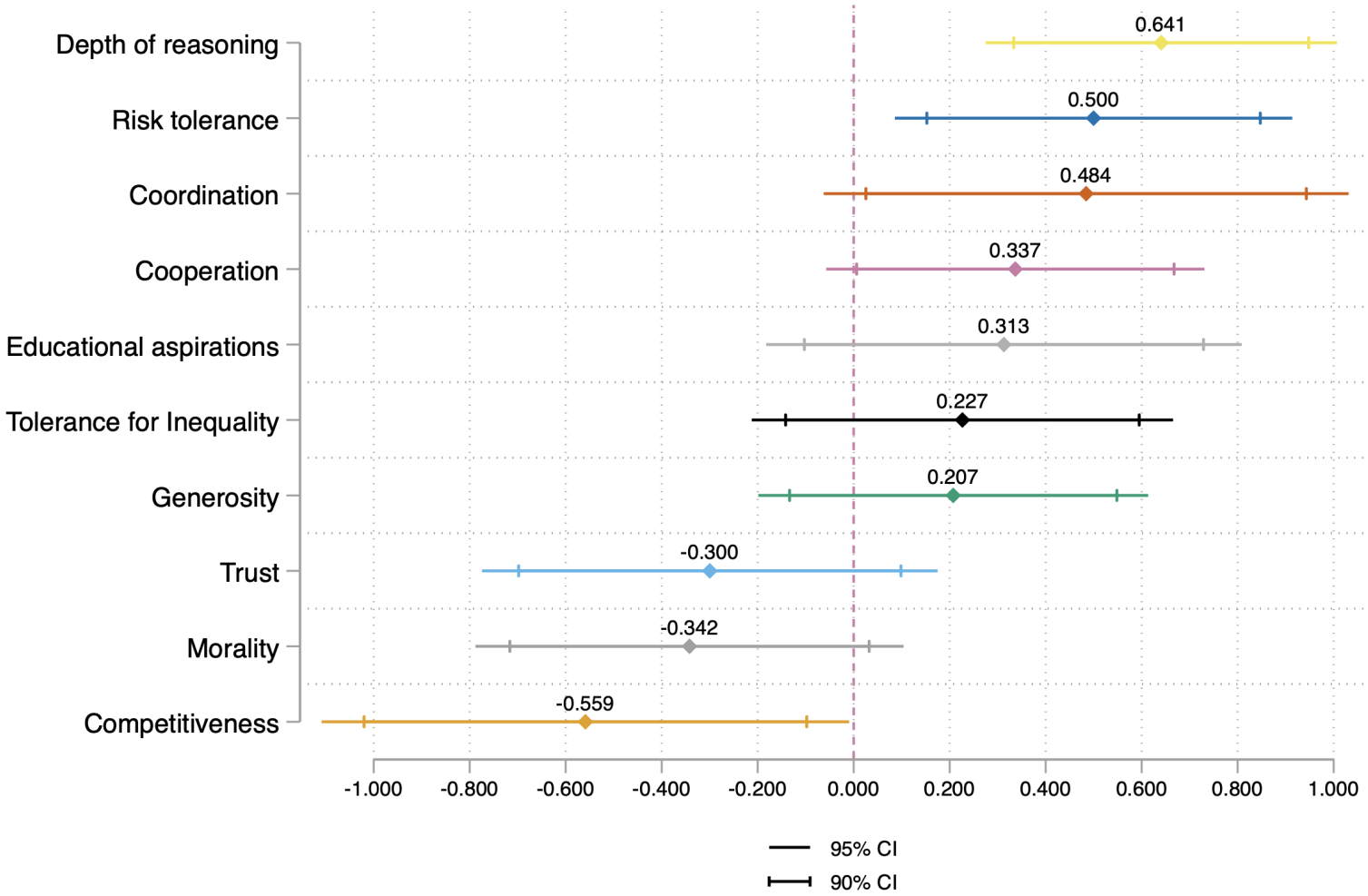
Identification of peer effects - II

- Solution: Use a 3SLS strategy
  - Tackle **correlated effects** by differencing out classroom fixed effects
  - Tackle **reflection problem** by instrumenting the endogenous behavioral trait of friends ($\mathbf{Gy}$) with the exogenous characteristics (gender, race, nationality etc.) of friends of friends and friends of friends of friends ($\mathbf{G^2x}, \mathbf{G^3x}$) (Bramoullé et al. [2009], Case and Katz [1991]) Reflection problem
  - Tackle **endogeneity in link formation** by replacing the adjacency matrix $\mathbf{G}$ with a predicted adjacency matrix $\hat{\mathbf{G}}$, i.e. instruments for $\mathbf{Gy}$ are $Z = [\hat{\mathbf{G}}\mathbf{x}, \hat{\mathbf{G}}^2\mathbf{x}, \hat{\mathbf{G}}^3\mathbf{x}]$ Endogenous links
- Instruments are valid under the assumption that peers within the class who the individual doesn't interact with as friends do not differentially impact the behavioral traits of the individual apart from the impact already captured by classroom fixed effects.

Identification of peer effects - II

- Solution: Use a 3SLS strategy
  - Tackle **correlated effects** by differencing out classroom fixed effects
  - Tackle **reflection problem** by instrumenting the endogenous behavioral trait of friends ($\mathbf{Gy}$) with the exogenous characteristics (gender, race, nationality etc.) of friends of friends and friends of friends of friends ($\mathbf{G^2x}, \mathbf{G^3x}$) (Bramoullé et al. [2009], Case and Katz [1991]) Reflection problem
  - Tackle **endogeneity in link formation** by replacing the adjacency matrix $\mathbf{G}$ with a predicted adjacency matrix $\hat{\mathbf{G}}$, i.e. instruments for $\mathbf{Gy}$ are $Z = [\hat{\mathbf{G}}\mathbf{x}, \hat{\mathbf{G}}^2\mathbf{x}, \hat{\mathbf{G}}^3\mathbf{x}]$ Endogenous links
- Instruments are valid under the assumption that peers within the class who the individual doesn't interact with as friends do not differentially impact the behavioral traits of the individual apart from the impact already captured by classroom fixed effects.

Peer Effects - Main findings

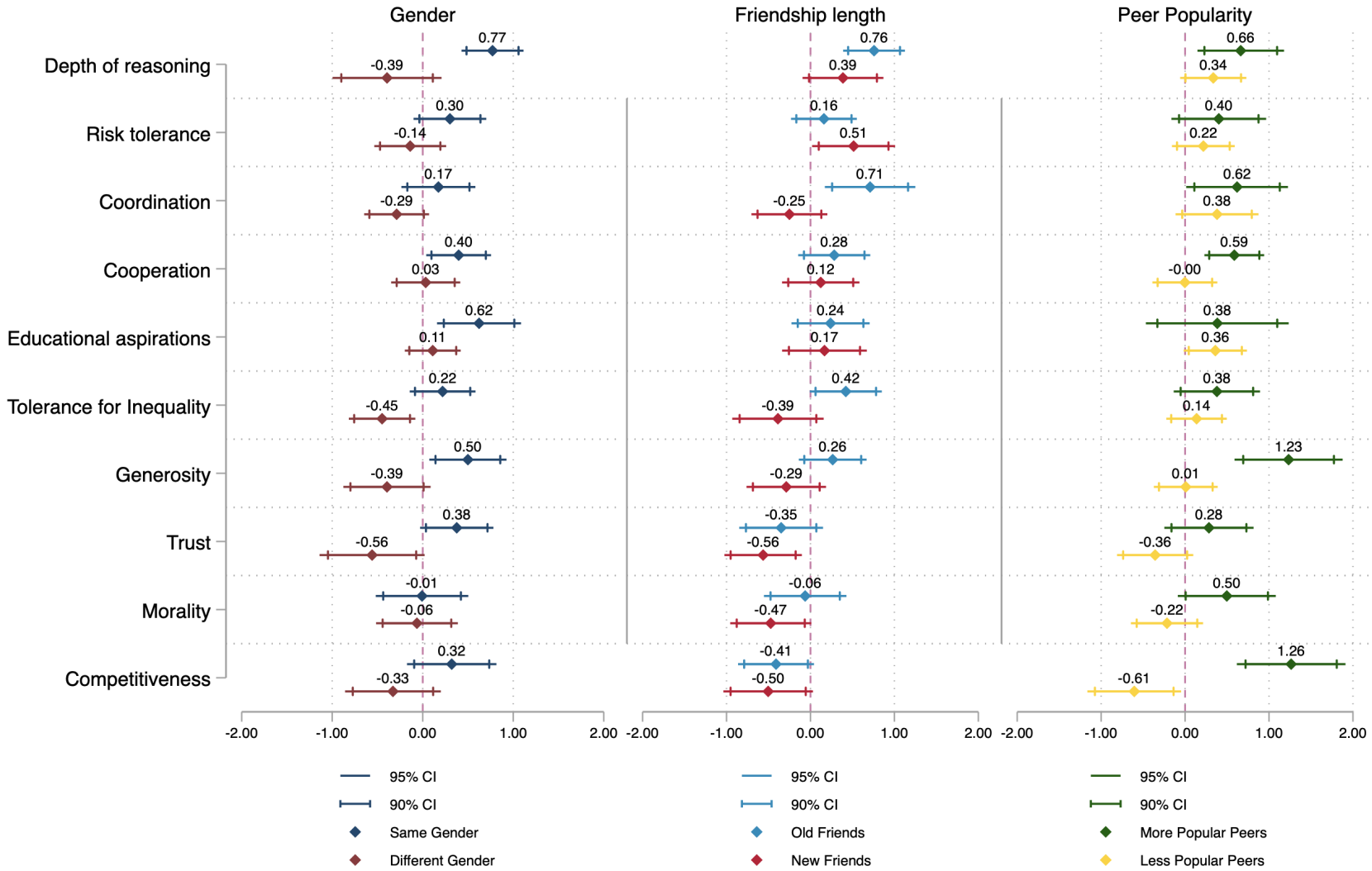


First Stage
Without Contextual variables
Alternate Network Specifications
Robustness to first stage variations

Who matters and how do they matter?

- Social interactions have different intensities.
- Individuals look up to popular peers and try to be like them and try to differentiate themselves from less popular peers.
- Traits have social value. Who endows these traits with the social value?
- Behavioral spillovers can be strong if networks last for a longer time.

Heterogeneity in peer effects



Conclusion

Friends matter! You choose your friends but they may influence you too.

Questions and Suggestions?

Feel free to reach out!

palaash.bhargava@columbia.edu

@palaashbhargava

https://sites.google.com/view/palaashbhargava

Appendix

Social Networks and Homophily

- Societies extremely segregated by ethnicity, income, gender, age, profession, religion and caste.
  - US 2020: 56% of African Americans have social networks composed entirely of African Americans. 53% (55%) of Republicans (Democrats) have social circles composed entirely of Republicans (Democrats) (Cox et al. [2020]).
- Social networks and homophily can lead to reproduction of inequalities, echo chambers, spillovers from interventions (Jackson [2021], Banerjee et al. [2019], Kearney and Levine [2015], Lobel and Sadler [2016]).
- Uncovering different aspects on which social groups exhibit homophily is important, specifically for young adults / students. Evidence currently is limited w.r.t. traits.

Back

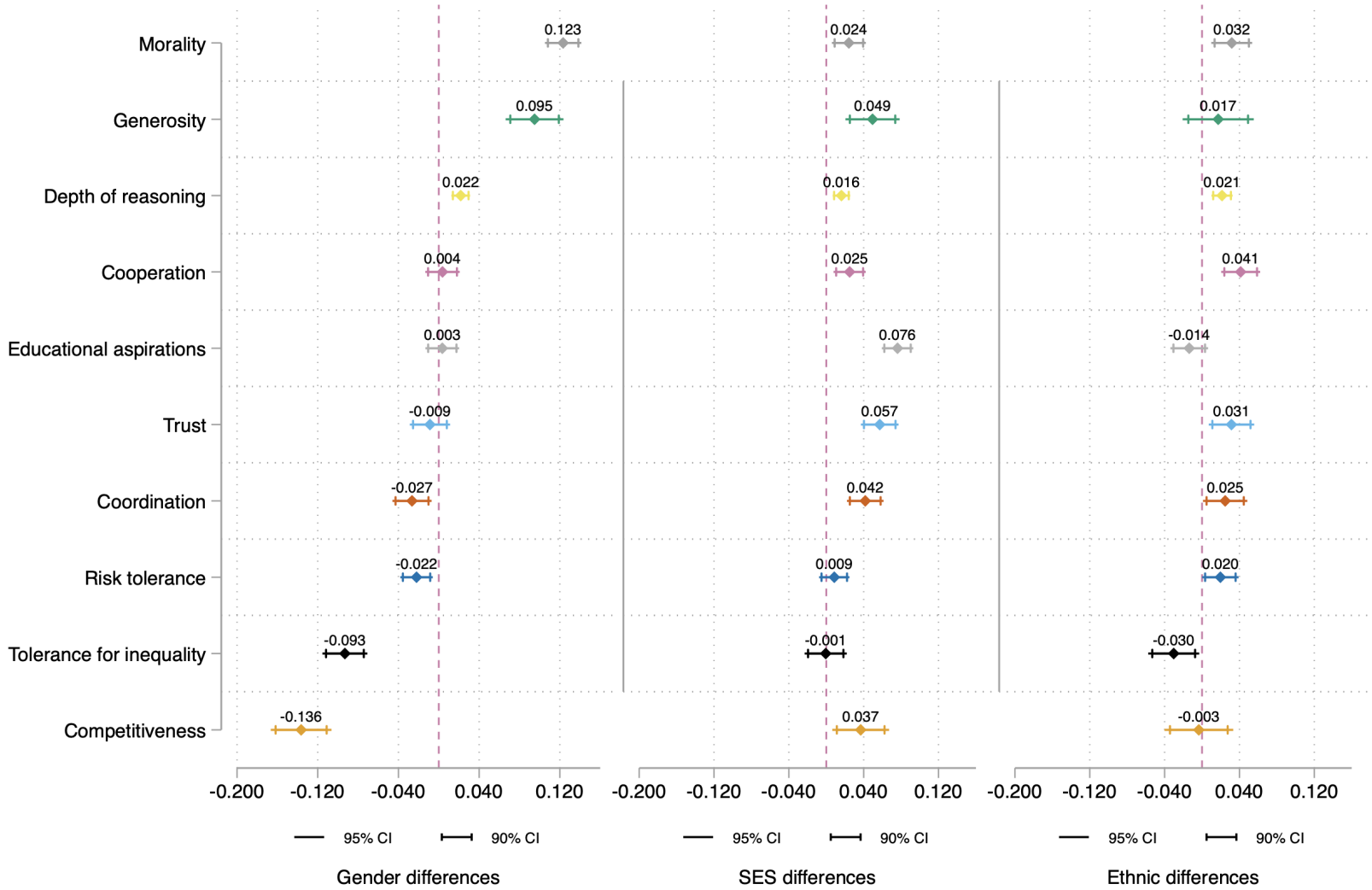
Value of Behavioral Traits

- Behavioral traits influence a broad array of life-time outcomes of individuals:
  - Wages and Choice of occupation (Buser et al. [2014], Flory et al. [2015])
  - Labour market success (Deming [2017], Kosse et al. [2020])
  - Health (Sutter et al. [2013])
  - Financial success (Meier and Sprenger [2010], [2012], Dohmen et al. [2011])
  - Educational Achievement and School dropout rates (Cadena and Keys [2015], Castillo et al. [2011], [2018])

- Systemic differences in behavioral traits between individuals from different gender, socio-economic backgrounds and ethnicity (Falk et al. [2020]) Evidence

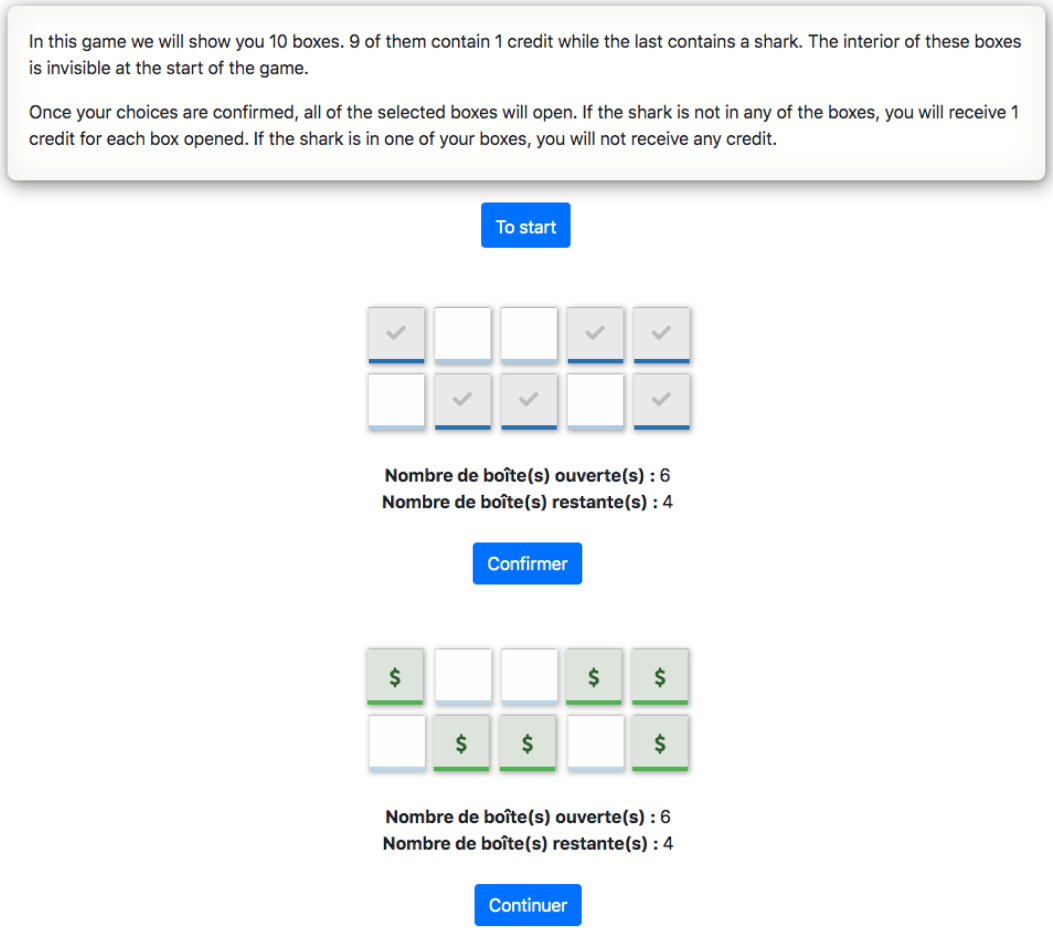
Back

Behavioral Correlations - Demographic Characteristics



Back

Preview of incentivised games - Risk Tolerance



Back

Preview of incentivised games - Competition

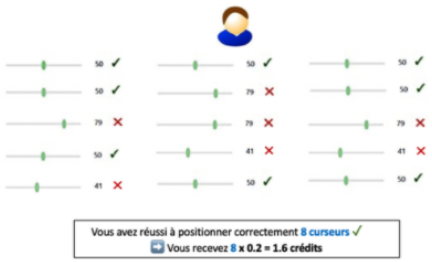
In this game, we suggest you position a cursor in the middle of a horizontal line ranging from 0 to 100. As in the example below, when you move the cursor along the axis, its positioning will be displayed. to the right of the axis. The objective is to position it on 50.

79

The next page will contain 48 of these axes. You will have 2 minutes to correctly place the greatest number of cursors out of 50.

Each correct positioning will earn you credits and we offer you to choose between two options to receive credits.

Option A: You receive **0.2 credits** for each correctly positioned cursor over 50.



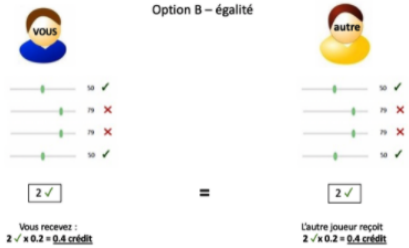
Vous avez réussi à positionner correctement 8 curseurs ✓
Vous recevez 8 x 0.2 = 1.6 crédits

Option B: You play against a partner (randomly selected).

The second participant is also in your class.
If the number of sliders you position correctly is **greater** than the number of the other participant, you will receive **0.5 credits** for each correctly positioned slider.

If the number of sliders you position correctly is **less than** the number of the other participant, **you receive nothing** .

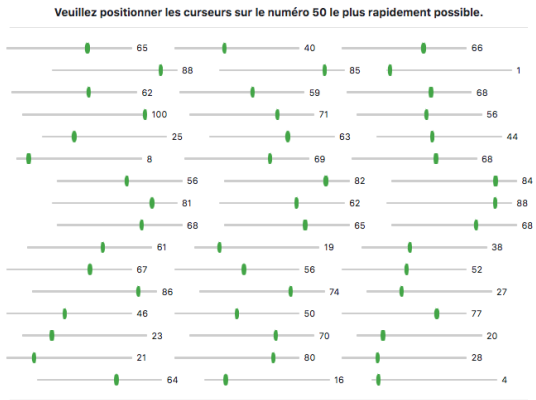
If you position the **same number** of cursors correctly, you receive 0.2 credits for each correctly positioned cursor.



Which option do you prefer to receive the credits?

- ○ Option A: 0.20 credit for each correctly positioned cursor
- ○ Option B: 0.50 credit for each correctly positioned cursor if my number is greater than the number of the other participant. If my number is lower, I get nothing.

To confirm



Back

Preview of incentivised games - Trust

In this game, we offer you to interact with another participant.

You will each have to make 2 decisions.

Decision 1

We allocate you 5 credits.

1 / You can send between 0 and 5 of these credits to the other participant. The quantity you send will be tripled.

2 / Among this tripled quantity, the other participant must decide how much he wishes to return to you in turn. The quantity returned to you by the other participant is not tripled.

Your final credits are calculated as follows: Your initial 5 credits **minus** your transfer to the other participant **plus** what the other participant returned to you.

Decision 2

Now the other participant receives an endowment of 5 credits.

1 / He chooses the quantity he wishes to send you (between 0 and 5 credits). This quantity is tripled and then sent to you.

2 / You then decide how much you want to return to the other participant. This returned quantity is not tripled.

Your final credits are calculated as follows: credits sent by the other participant tripled **minus** what you returned to the other participant.

When you are ready to start with decision 1, press the start button.

To start

Please indicate how much (between 0 and 5 credits) you wish to transfer to the other participant. Remember that this transferred quantity is tripled and that the other participant can return part of it to you afterwards.

I am transferring:
(Please use multiples of 0.50)

credit (s)

To confirm

Back

Summary Statistics

Variable	Mean	Std. Dev.	Min.	Max.	N
Panel A: Friendship Information					
No. of friends reported	4.640	0.856	1	5	2565
No. of friends matched	3.395	1.298	1	5	2565
No. of times reported as a friend	2.287	1.681	0	10	2565
Panel B: Behavioral Traits					
Tolerance for inequality	1.629	2.672	0	8	2332
Morality	7.550	3.028	0	10	2450
Trust	2.433	1.619	0	5	2332
Generosity	0.452	0.404	0	1	2061
Cooperation	0.488	0.278	0	1	2332
Coordination	0.479	0.310	0	1	2332
Risk Tolerance	5.712	2.750	0	10	2565
Competitiveness	0.477	0.500	0	1	2332
Depth of reasoning	33.496	14.198	0	100	2332
Educational aspirations	2.852	0.816	1	4	2565

Back

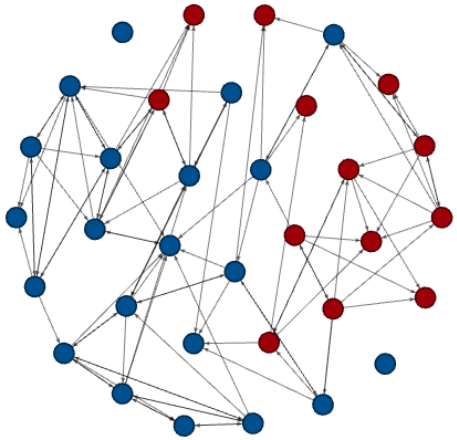
Summary Statistics

Variable	Mean	Std. Dev.	Min.	Max.	N
Panel C: Demographic characteristics					
Female	0.557	0.497	0	1	2565
French	0.961	0.193	0	1	2565
White	0.792	0.406	0	1	2565
Arab	0.053	0.223	0	1	2565
Hispanic	0.062	0.242	0	1	2565
Black	0.061	0.239	0	1	2565
Asian	0.032	0.177	0	1	2565
Primary parent occupation: low skill	0.419	0.493	0	1	2565
No. of siblings from primary parent	1.073	1.047	0	11	2565
Single Child	0.329	0.470	0	1	2565
Born in France	0.950	0.218	0	1	2565
Age (in years)	15.766	0.942	13	19	2565
From Créteil	0.170	0.376	0	1	2565
From Montpellier	0.306	0.461	0	1	2565
From Nantes	0.524	0.500	0	1	2565
Grade 10	0.498	0.500	0	1	2565
Grade 11	0.274	0.446	0	1	2565
Grade 12	0.228	0.420	0	1	2565

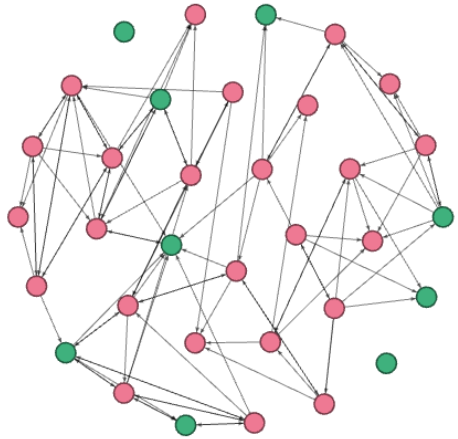
Back

Typical network

Figure: Network visualization for a select classroom



(a) Nodes colored by gender

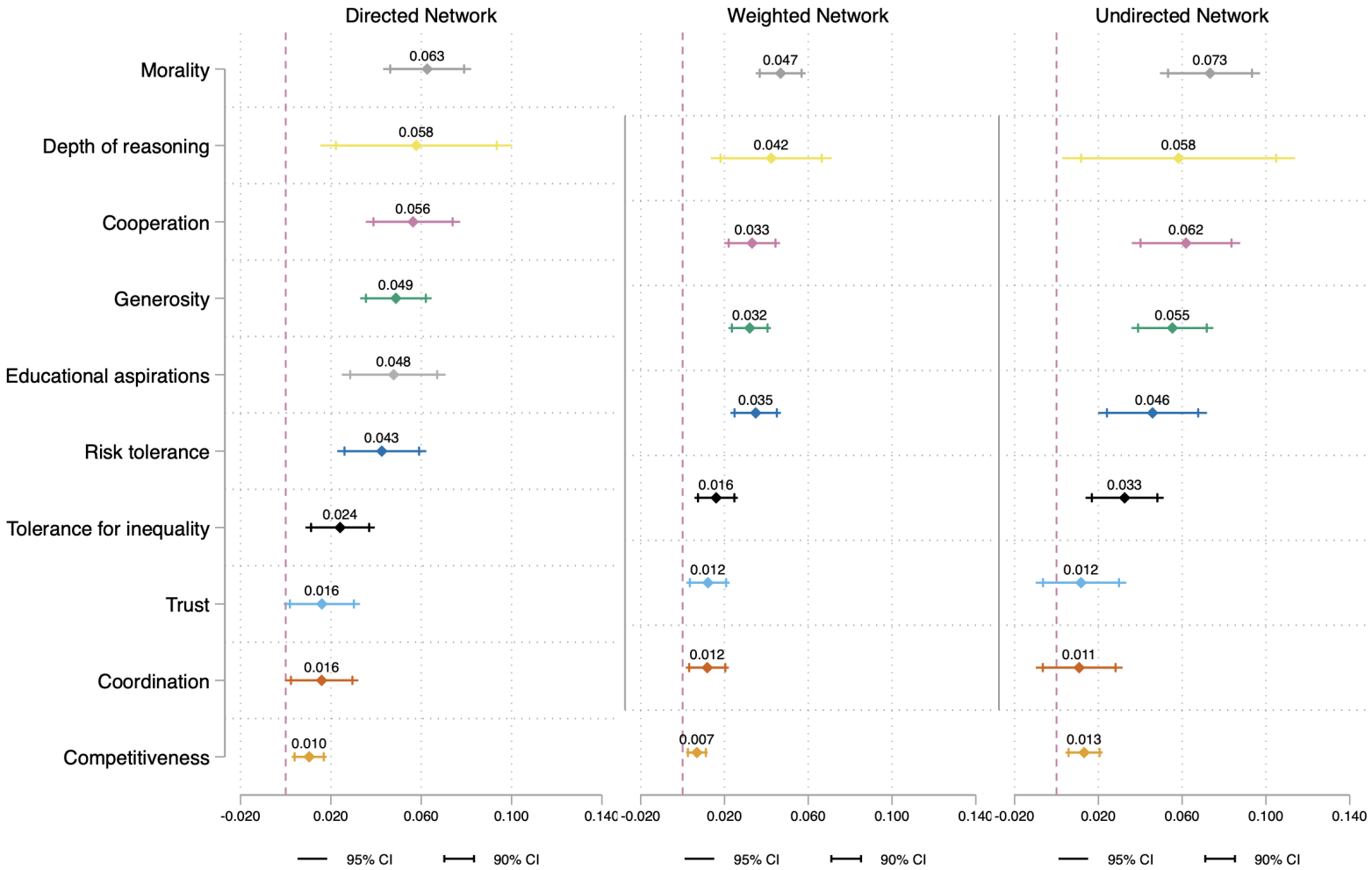


(b) Nodes colored by SES

Note: The figure above graphs the social network for a given classroom. The plot on the left colors the nodes by gender where blue nodes are males and red nodes are females. The plot on the right colors the nodes by the socio-economic status (SES) of the individual. Pink nodes correspond to High SES individuals (individuals whose primary parent in engaged in a non-low skill job) and green nodes correspond to low SES individuals (individuals whose primary parent is engaged in a low skill job). Considerable homophily exists on both gender and socio-economic status. However, homophily is more pronounced on the gender dimension.

Back

Homophily - Alternate Specifications



Back

Microfoundation - I

- Finite set of agents: $\mathcal{N} = \{1, 2, ..., n\}$
- Social connections in the network represented in the adjacency matrix: **G** where $g_{ij} = 1/n_i$ if individual i sends a link to individual j, 0 otherwise. n_i represents the out-degree of individual i.
- Let P_i represent the neighborhood of individual i
- Individuals have an intrinsic behavioral type a_i which is a linear function of his own exogenous characteristics and social environment (exogenous characteristics of friends) i.e.

$$a_i = \gamma_1 x_i + \delta_1 \sum_{j \in P_i} g_{ij} x_j + \eta$$

- y_i is the observed behavioral type (in control of the agent).

Microfoundation - II

- Agents incurs a cost for deviating from his intrinsic type and also a cost for deviating from the social norm represented by the average type of his friends.
- Agent tries to minimize this cost, i.e. the objective function of the agent is:

$$\max_{y_i} \mathcal{U}_i = \max_{y_i} \left(-(a_i - y_i)^2 - \zeta \left(\sum_{j \in P_i} g_{ij} y_j - y_i \right)^2 \right)$$

- Quadratic utility $\implies$ linear best replies: $y_i^* = \frac{a_i}{1+\zeta} + \frac{\zeta}{1+\zeta} \sum_{j \in P_i} g_{ij} y_j^*$
- If the spectral radius of $\mathbf{G} < \frac{1+\zeta}{\zeta}$, unique Nash equilibrium.
- Observed behavior types function of the Katz-Bonacich centrality of the player (weighted with respect to $\mathbf{a}$). (Ballester et al. [2006])

Back

Reflection problem - I

- Recall Eq. 2 reads as:

$$\mathbf{y} = \beta \mathbf{G}\mathbf{y} + \gamma \mathbf{x} + \delta \mathbf{G}\mathbf{x} + \eta + \epsilon$$

- Rearranging and using a series expansion for $(\mathbf{I} - \beta \mathbf{G})^{-1} = \sum_{k=0}^{\infty} \beta^k \mathbf{G}^k$, we get:

$$\mathbf{y} = \eta/(1-\beta)\iota + \gamma \mathbf{x} + (\gamma\beta + \delta) \sum_{k=0}^{\infty} \beta^k \mathbf{G}^{k+1}\mathbf{x} + \sum_{k=0}^{\infty} \beta^k \mathbf{G}^k \boldsymbol{\epsilon}$$

- Pre-multiplying with $\mathbf{G}$ and taking conditional expectation gives us:

$$\mathbb{E}(\mathbf{G}\mathbf{y} \mid \mathbf{x}) = \eta/(1-\beta)\iota + \gamma \mathbf{G}\mathbf{x} + (\gamma\beta + \delta) \sum_{k=0}^{\infty} \beta^k \mathbf{G}^{k+2}\mathbf{x}$$

Back

Reflection problem - II

- Instruments for $\mathbb{E}(\mathbf{Gy} \mid \mathbf{x})$: $\mathbf{G}^{k+2}\mathbf{x}$, $k = 0, 1, \ldots$
- **Intuition?:** Let $\mathbf{G}$ take the following form:

$$G_{ij} = \begin{cases} 1, & j = i - 1 \\ 0, & j \neq i - 1 \end{cases}$$

- Eq. 2 boils down to:

$$y_i = \eta + \beta y_{i-1} + \gamma x_i + \delta x_{i-1} + \epsilon_i$$

Panel data structure! Use $x_{i-k}, \ldots$ where $k = 2, 3, \ldots$ as an instrument for y_{i-1}

Back

Endogenous link formation

- Friendships aren't formed at random, i.e. $\mathbf{G}$ and ϵ are correlated. Moreover, homophily on behavioral traits may exist, i.e. $\mathbf{G}$ and $\mathbf{y}$ maybe independently correlated. Therefore, just using $\mathbf{G}^2\mathbf{x}, \mathbf{G}^3\mathbf{x}, ...$ as instruments for $\mathbf{Gy}$, will still produce biased estimates.

- Rely on patterns of homophily based on exogenous characteristics (König et al 2019). Well documented homophily patterns based on shared gender, race, nationality status, parents occupation / social status etc.

- Estimate $\hat{\mathbf{G}}$ based on pre-determined shared characteristics.

- Specifically, estimate the probability of person i sending a friendship link to person j within a classroom based on shared gender, race, nationality status, parental occupation etc. while controlling for the characteristics of link sender and receiver.

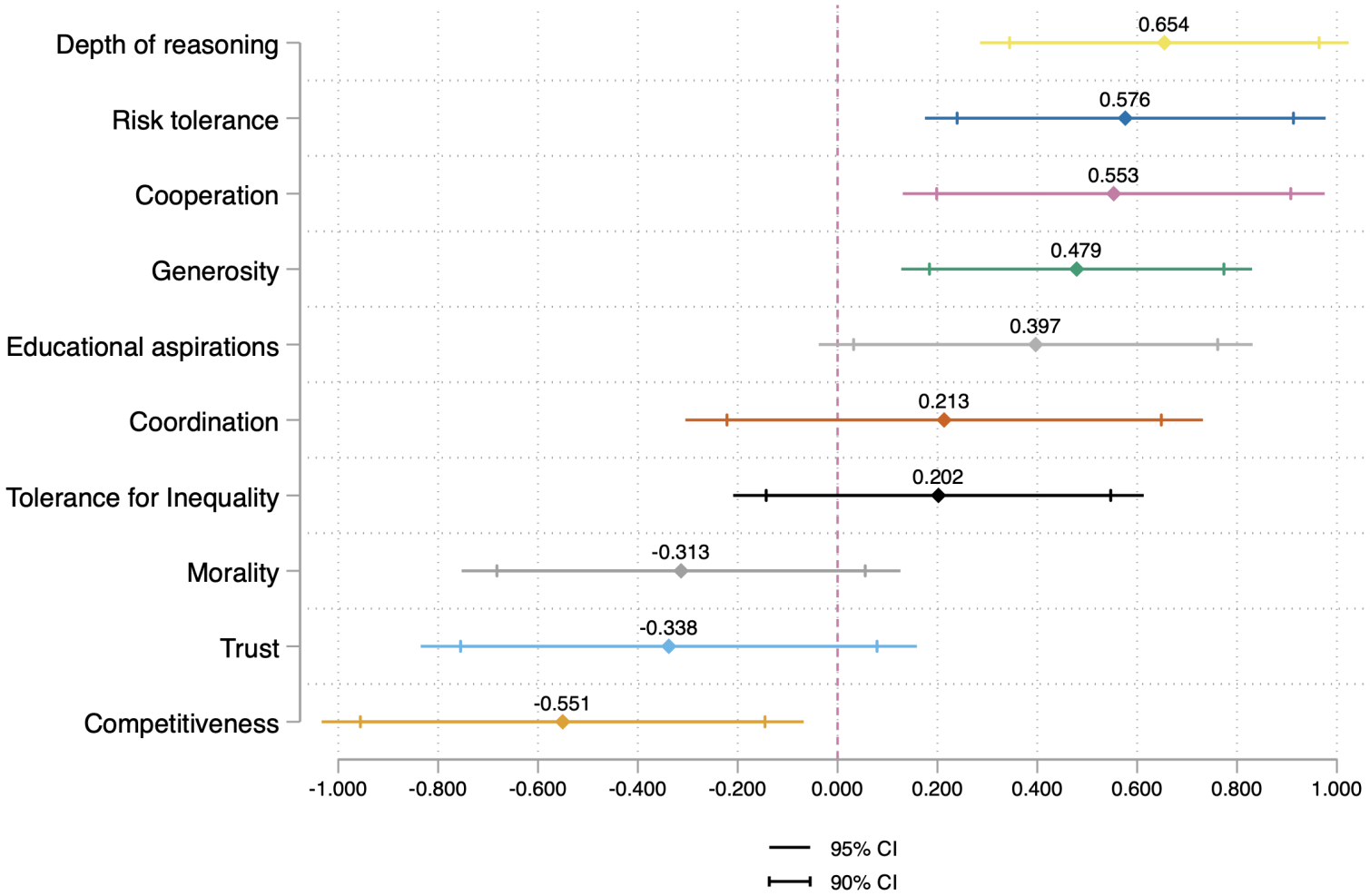
Back

First stage - Results

Variable	Logit	Logit	Logit	Logit
Shared postal code	0.275***	0.293***	0.373***	0.390***
	(0.063)	(0.065)	(0.082)	(0.083)
Shared gender	1.009***	1.014***	1.010***	1.016***
	(0.082)	(0.083)	(0.083)	(0.084)
Shared nationality	0.937*	0.884*	0.912+	0.858+
	(0.549)	(0.532)	(0.573)	(0.550)
Shared ethnicity	0.284**	0.284**	0.298**	0.298**
	(0.125)	(0.127)	(0.129)	(0.131)
Similar age (in months)	0.023***	0.024***	0.025***	0.024***
	(0.006)	(0.007)	(0.006)	(0.008)
Similar no. of siblings	0.074	0.082	0.080	0.087
	(0.065)	(0.070)	(0.064)	(0.071)
Only child - Match	0.026	0.025	0.038	0.026
	(0.097)	(0.100)	(0.099)	(0.103)
Shared primary parent occu. cat.	-0.002	0.000	0.000	0.003
	(0.043)	(0.043)	(0.044)	(0.044)
Shared country of birth	0.073	0.171	0.074	0.146
	(0.368)	(0.330)	(0.381)	(0.350)
Sender and Receiver Characteristics	Y	Y	Y	Y
Classroom Fixed Effects	N	Y	N	Y
Interaction terms	N	Y	N	Y
Mc. Fadden R-sq	0.041	0.045	0.055	0.059
Mc. Fadden Adj. R-sq	0.038	0.035	0.052	0.049
N	21521	21521	21476	21476
*** p<0.01, **p<0.05, * p<0.1, + p<0.15				

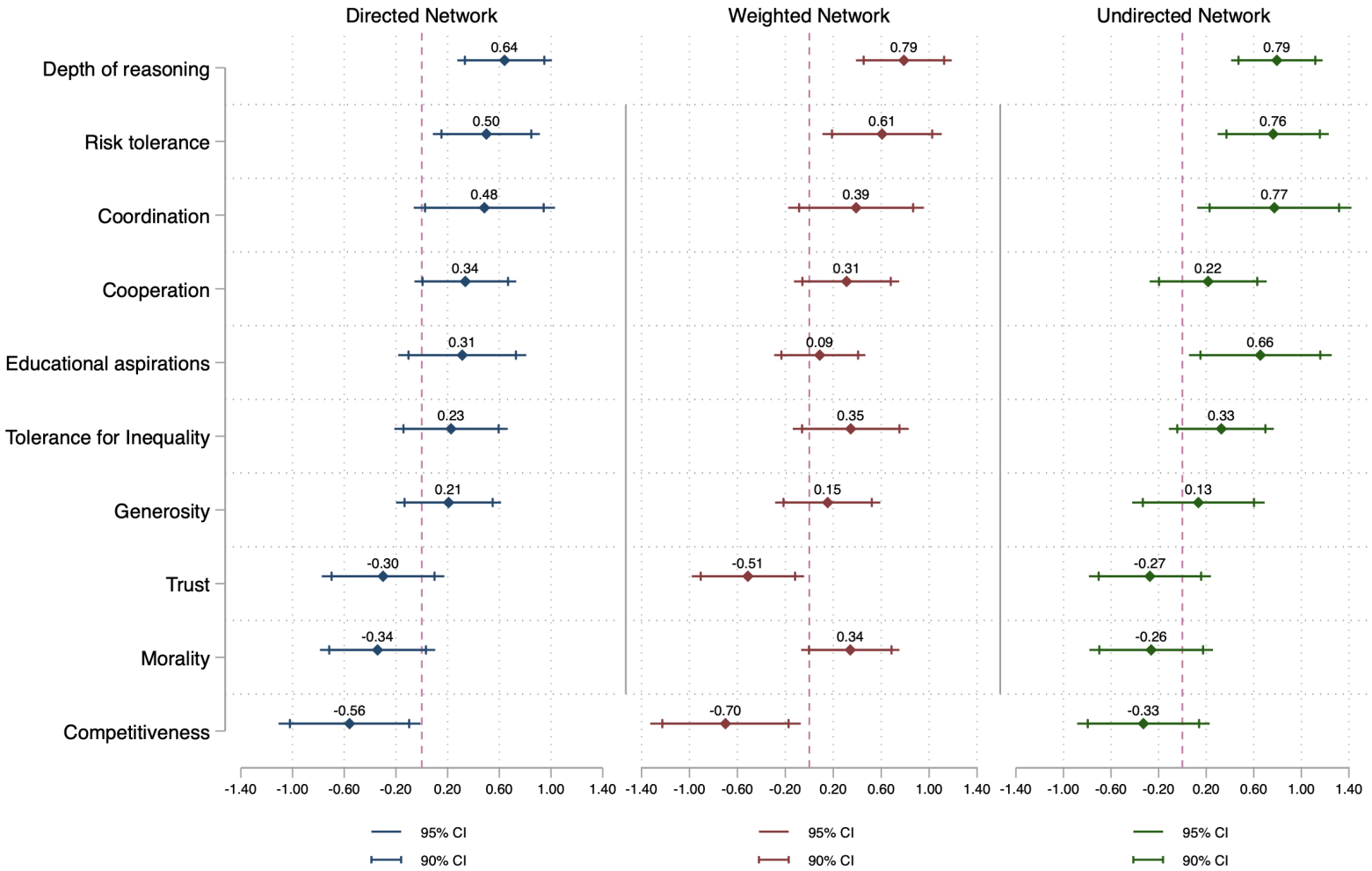
Back

Peer Effects - Without Contextual Variables



Back

Peer Effects - Alternate Network Specifications



Back

Robustness - First Stage variations

	Dem Char		Dem Char + Int		Dem Char + FE		Dem Char + Int + FE	
Cont. var.	Coeff. (1)	F stat. (2)	Coeff. (3)	F stat. (4)	Coeff. (5)	F stat. (6)	Coeff. (7)	F stat (8)
				Outcome: Depth of reasoning				
Y	0.549***	2.566	0.557***	2.646	0.631***	2.444	0.641***	2.487
	(0.191)		(0.191)		(0.188)		(0.187)	
N	0.577***	2.464	0.587***	2.550	0.651***	2.263	0.654***	2.332
	(0.192)		(0.193)		(0.189)		(0.188)	
				Outcome: Risk tolerance				
Y	0.566***	2.328	0.552***	2.294	0.541***	2.413	0.500**	2.479
	(0.209)		(0.210)		(0.212)		(0.211)	
N	0.628***	2.743	0.623***	2.642	0.609***	2.853	0.576***	2.939
	(0.203)		(0.203)		(0.205)		(0.205)	
				Outcome: Coordination				
Y	0.359	4.989	0.386	4.311	0.456*	3.378	0.484*	3.093
	(0.289)		(0.293)		(0.276)		(0.279)	
N	0.091	8.070	0.091	6.517	0.198	4.149	0.213	3.788
	(0.266)		(0.268)		(0.261)		(0.264)	
				Outcome: Cooperation				
Y	0.228	2.108	0.274	2.281	0.287+	2.018	0.337*	2.188
	(0.199)		(0.201)		(0.198)		(0.201)	
N	0.423*	1.692	0.465**	1.879	0.500**	1.674	0.553***	1.822
	(0.219)		(0.215)		(0.218)		(0.216)	

Back

Robustness - First Stage variations

	Dem Char		Dem Char + Int		Dem Char + FE		Dem Char + Int + FE	
Cont. var.	Coeff. (1)	F stat. (2)	Coeff. (3)	F stat. (4)	Coeff. (5)	F stat. (6)	Coeff. (7)	F stat (8)
				Outcome: Educational aspirations				
Y	0.186	2.017	0.215	2.041	0.264	1.910	0.313	1.849
	(0.224)		(0.227)		(0.245)		(0.253)	
N	0.273	2.210	0.286+	2.228	0.363*	2.101	0.397*	2.045
	(0.194)		(0.196)		(0.218)		(0.222)	
				Outcome: Tolerance for inequality				
Y	0.301	2.929	0.287	2.946	0.250	2.924	0.227	2.863
	(0.225)		(0.226)		(0.225)		(0.224)	
N	0.260	5.580	0.253	5.941	0.219	5.310	0.202	5.366
	(0.208)		(0.209)		(0.211)		(0.210)	
				Outcome: Generosity				
Y	0.152	2.154	0.193	1.940	0.189	1.763	0.207	1.648
	(0.207)		(0.200)		(0.209)		(0.207)	
N	0.417**	2.061	0.437***	1.790	0.463***	1.682	0.479***	1.573
	(0.165)		(0.167)		(0.178)		(0.179)	
				Outcome: Trust				
Y	-0.243	2.024	-0.248	2.020	-0.298	2.151	-0.300	2.045
	(0.246)		(0.246)		(0.239)		(0.242)	
N	-0.280	2.066	-0.307	2.103	-0.315	2.272	-0.338	2.099
	(0.263)		(0.257)		(0.256)		(0.254)	

Back

Robustness - First Stage variations

	Dem Char		Dem Char + Int		Dem Char + FE		Dem Char + Int + FE	
Cont. var.	Coeff.	F stat.	Coeff.	F stat.	Coeff.	F stat.	Coeff.	F stat
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
				Outcome: Morality				
Y	-0.278	1.605	-0.292	1.457	-0.286	1.499	-0.342+	1.384
	(0.218)		(0.222)		(0.225)		(0.228)	
N	-0.201	2.344	-0.233	2.145	-0.247	2.146	-0.313	1.973
	(0.213)		(0.217)		(0.218)		(0.224)	
				Outcome: Competitiveness				
Y	-0.785***	4.650	-0.586**	4.767	-0.700***	4.191	-0.559**	4.248
	(0.256)		(0.284)		(0.262)		(0.280)	
N	-0.715***	7.536	-0.612**	8.043	-0.640***	6.495	-0.551**	6.741
	(0.239)		(0.248)		(0.240)		(0.246)	

Back