



# **PREDICTING COLLEGE RETENTION RATES: AN APPLICATION OF THE CRITICAL MASS THEORY**

Kevin Sun

Wednesday, April 4, 2018

MA Computational Social Science

Research Proposal

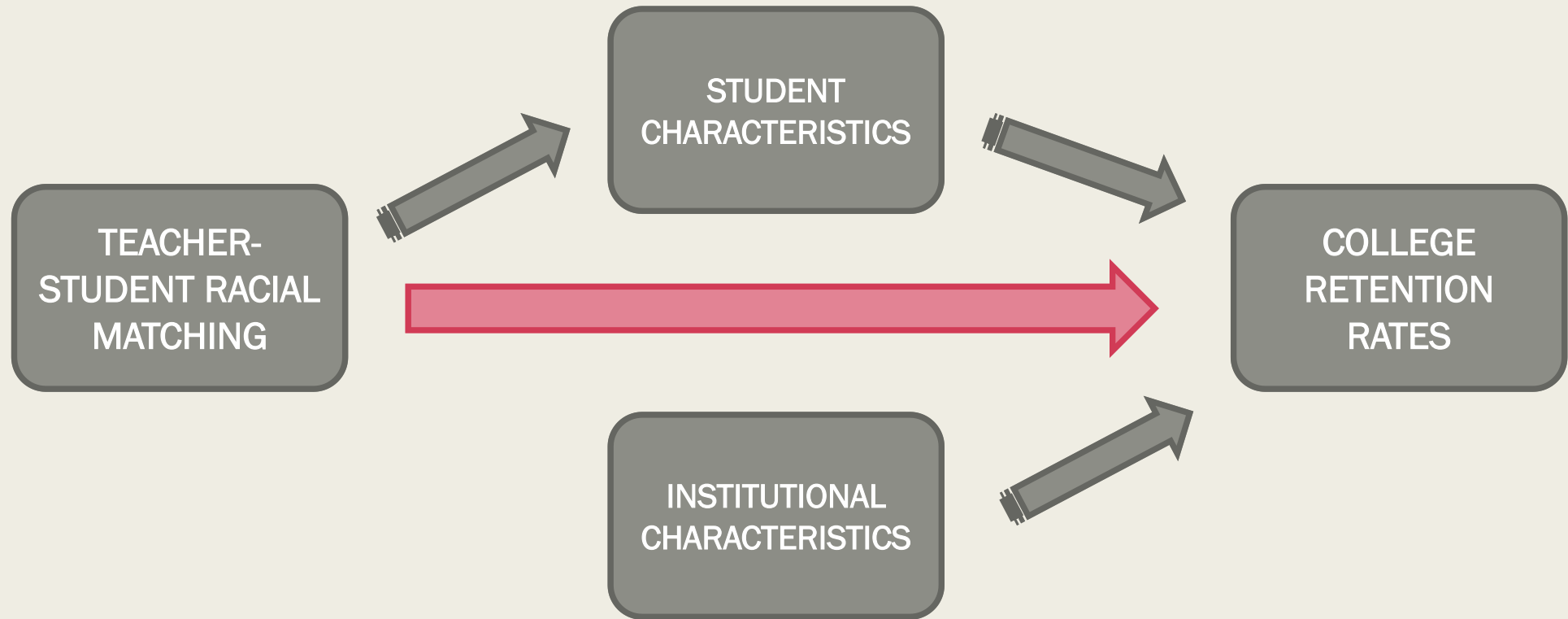


# RESEARCH QUESTION:

To what extent can college retention-rates be predicted by bureaucratic representation at the K-12 level?

- **College retention:** students who return to college their second year
- **Bureaucratic representation:** Teachers & administrators sharing demographic similarities with students

# THE LAY OF THE LAND:



# THE LITERATURE: On Retention

- **High School Achievement** (Astin, Korn, Green 1987)
- **Integration into academic and social community** (Tinto 1975, 1987)
- **Demographic Traits** (Astin 1975; St. John, Hu, Simmons, Musoba 2001)
- **Psychological Traits** (Trapmann, Hell, Hirn, Schuler 2007; Oswald, Schmitt, Kim, Ramsay, Gillepsie 2004)

# THE LITERATURE: Racial Matching

- **Assessment of student behavior & disciplinary outcomes**  
(Bates & Glick 2013; Lindsay & Hart 2017)
- **Expectation of student potential** (Gershenson, Holt, Papageorge 2016; McGrady and Reynolds 2012; Dee 2005)
- **Representation of non-white students in gifted programs**  
(Grissom, Rodriguez, Kern 2017)
- **Math & reading achievement** (Dee 2004; Clotfelter, Ladd, and Vigdor 2007)
- **Student perception of non-white teachers** (Cherng & Halpin 2016)

# THE LITERATURE: Racial Matching

- **Assessment of student behavior & disciplinary outcomes**  
(Bates & Glick 2013; Lindsay & Hart 2017)
- **Expectation of student potential** (Gershenson, Holt, Papageorge 2016; McGrady and Reynolds 2012; Dee 2005)
- **Representation of non-white students in gifted programs**  
(Grissom, Rodriguez, Kern 2017)
- **Math & reading achievement** (Dee 2004; Clotfelter, Ladd, and Vigdor 2007)
- **Student perception of non-white teachers** (Cherng & Halpin 2016)

# THE DATA:

- **Teacher & Administration Demographics: Chicago Public Schools**
  - *Individual teachers at each school*
    - Impute race/ethnicity of each teacher
    - *NamSor*
    - *ethnicolr*
- **Other Demographics & School-Level Data: Chicago Public Schools**
- **College Attendance & Persistence Rates: National Student Clearinghouse**

# METHODS & MODELS

$$Retention = \beta_0 + \beta_1(BureacraticRepresentation) + \beta_2(DemographicControls)$$



# METHODS & MODELS

- OLS
- Decision Tree
- Random Forest

# ANTICIPATED CHALLENGES & EXTENSIONS:

- Imputing race based on names
- Analysis on every school district in the U.S.
- Ideal: student-teacher racial matching
- Hypothesis: A “critical mass” of teachers/administrators of color at a school will be associated with higher college retention/persistence levels in that school’s graduates