Introduction
A mutual fund pools money from many investors to purchase tradable assets. While firms’ management is advised and monitored by a board of directors who are elected by shareholders, funds have trustees who are elected by investors. A large literature has found that funds with effective boards have higher performance. Our current work is preparing data for two studies.

1. Investigate the impact of overlapping boards on funds’ performance
2. Examine the impact of connections between fund trustees and firm directors on the funds’ and firms’ operations.

Boards of Directors
- Each mutual fund is required to have a board.
- Funds belong to families, which tend to appoint both types of boards include inside directors, corporations also have boards of directors.
- Both types of boards include inside directors, who are usually top management, and outside directors, who are assumed to be independent.
- Individuals can serve simultaneously or subsequently on boards of mutual funds and corporations.
- Directors with firm-fund connections may have private information; this information asymmetry could affect the actions of the directors.

Data
- Mutual Funds are required to publicly release N-CSR filings (i.e., a combined reporting form) that discloses information about board membership.
- The Securities and Exchange Commission’s (SEC) EDGAR system archives N-CSR filings.
- Using a text scraper/web-crawler, data was collected from the SEC EDGAR system to create a sample of the 601 actively managed equity funds sponsored by the 25 largest fund families in 2008-2020.

Cleaning
- The scraped data presents many different issues. These issues were observed within individual funds over time and within families at common points of time. The most common issues include:
  - Inconsistent formatting of data
  - Duplicate observations
  - Paragraphs of extraneous information around names (e.g., full bibliographic entries covering an individual’s professional history to date)
  - Prefixes and suffixes are recorded inconsistently
  - “Year of birth” is often reported as current age
  - Last name changes during sampling period (e.g., some women changed names upon marriage)
  - Some entries are missing first names and age or year of birth
  - Spelling errors (e.g., typos such as John or Jon

Future Work
Study #1: Mutual Fund Board Overlaps
- Construct measures of overlapping boards within fund families.
- Estimate models such as $\text{OUTCOME} = \beta_0 + \beta_1 \text{Overlaps} + \beta_2 X_{\text{id}} + \epsilon_{\text{id}}$.
- Outcomes include expenses, return gap, performance transfer, and window dressing.

Study #2: Corporate and Fund Director Overlaps
- Fuzzy matching of fund director names with firm director names as reported in BoardEx database.
- The fund directors and BoardEx dataset may have non-standardized presentations of names.

Acknowledgements
We thank Professors Hornstein, Kaparakis, Kabacoff, and Oleinikov for all their help working on this project, and for the QAC workshops and classes. We also thank Harrison Karp ’21 and James Bayer ’22 for developing code for this project over the last two summers.