Self-Selected or Regulatory: Which SIC Code is Accurate?*

Allen Hu†

This Draft: November 15, 2019

Abstract

Firms deviate from the regulatory provision when self-reporting industry classifications on the public information platform. This leads to a pervasive (more than 20%) discrepancy between firms’ SIC codes presented on EDGAR (self-selected by firms) and the SIC codes recorded in Compustat (designated according to the regulatory rule). This information inconsistency is due neither to higher information quality of EDGAR nor to the mismatch of time horizons but rather to opportunistic actions of firms who seek to categorize themselves into industries with higher valuations. This firm-induced friction doubles the equity return comovement between firms and their self-selected industries and generates return predictability at the firm and industry levels. After taking the opportunistic actions, firms are 40% more likely to restate their financial reports and invest aggressively in the following year.

Keywords: Information Friction, Industry Classification, Investor Bias

---

*I sincerely thank Nick Barberis, Jason Chen, Lauren Cohen, Stefano Giglio, Frank Li, Song Ma, Chris Malloy, Trung Nguyen, Kelly Shue, Terry Wu and Ning Zhu for many helpful comments.
†Yale School of Management. Email: allen.hu@yale.edu