Explaining Taxpayer Compliance
Evidence from UK Administrative Data

Nigar Hashimzade
Durham University and Institute for Fiscal Studies

Gareth D. Myles
University of Exeter and Institute for Fiscal Studies

Hana Yousefi
University of Exeter

Abstract
A considerable academic literature has developed that provides theoretical models of the compliance decision (Hashimzade et al., Journal of Economic Surveys, 2013). In contrast, there is a relative paucity of empirical research that confronts these models with data. The notable exception to this is the work of Feinstein (Rand Journal, 1991) who uses a sample of IRS data to determine the taxpayer characteristics that explain non-compliance. This paper develops the Feinstein analysis and applies it to anonymised administrative data from the UK HMRC that includes tax records and audit outcomes for the past decade. The basis of the Feinstein method is that the undeclared income discovered by an audit is only a fraction of the actual hidden income. The modeling strategy is to jointly model the individual choice of undeclared income and the proportion of undeclared income discovered. Our dataset contains a very wide range of individual characteristics including sex, age, occupation, and location that are used as conditioning variables for modeling individual choice. Hence, we model the proportion discovered using as explanatory variables (among many in our dataset) the time length of audit, whether full or aspect, whether random or targeted, geographical location, and whether a simple or complex case. The estimated relationship is then used to test success in predicting compliance behavior and, hence, to assess the potential benefits of predictive analytics in targeting tax audits.

E-mail
gdmyles@ex.ac.uk

Phone
0044 1392 724487