Data files for "Mapping the Landscape of Transactions: the Governance of Business Relations in Latin America" by David C. Francis, Nona Karalashvili, and Peter Murrell

August 17, 2018

We use a new set of survey questions – posed as part of the Enterprise Surveys (ES) implemented in six South American countries in 2017 and 2018 – to map governance structures that firms employ to support the successful implementation of transactions. Using latent class analysis (LCA), we discover economically meaningful patterns of governance structures that readily match well-known constructs in the literature. For each firm, we produce estimates of the posterior probability that the firm uses each of the discovered governance structures. The datafiles included in this zipfile provide all information necessary for other researchers to take our analysis further and test their own hypotheses relevant to Williamson's (1991) discriminating alignment agenda using additional data from the ES or elsewhere. Section VI of the paper is an example of a simple version of such analysis.

The core question that was adapted from Hendley and Murrell (2003) specially for the paper asks about the effectiveness of six different individual mechanisms in resolving or preventing problems in transactions. The mechanisms are trust, mutual interest, private dispute resolution, private third parties, governments, and the legal system. The question was posed separately concerning agreements for trade with suppliers and with customers. The responses were analyzed separately, and placed in pairs of datasets, one for supplier relations and another for customer relations.

The zipfile contains two such pairs of datafiles (plus one xlsx file of Latent GOLD output). One pair of datafiles contains: (i) the survey responses to the corresponding series of questions, (ii) estimates produced by Latent GOLD of the posterior probability that each firm uses each of the discovered governance structures, and (iii) the unique identifier that enables one-to-one matching of each firm with the full ES data (variable idstd), along with the country identifier, sampling weights, and the stratum of each firm. These files are called "Landscape of Transactions_Data on Supplier Relations.dta" and "Landscape of Transactions_Data on Customer Relations.dta". The variable labels within the data files define each variable.

The second pair of datafiles contains estimates of the posterior probability of using each of the discovered governance structures for the full spectrum of possible response patterns. As noted in the paper, with six questions, each with five categories of responses, there are 15,625 possible response patterns (5^6). The ES respondents use far fewer than this. For relations with suppliers, we observe 711 distinct combinations of answers. For relations with customers, we observe 631. We use estimated model parameters – provided in the xlsx file included in this zipfile – to estimate posterior probabilities for all possible response patterns, including the ones that we do not observe in the six South American countries that we analyze. This pair of datafiles thus enables other researchers to use their own survey data and obtain posterior probabilities for their observed

patterns of responses. These files are called "Estimated Posteriors for Each Possible Response Pattern_Supplier Relations.dta" and " Estimated Posteriors for Each Possible Response Pattern_Customer Relations.dta". The variable labels within the data files define each variable.

Notably, the model parameters that we use to estimate posterior probabilities are obtained from the software Latent GOLD (Vermunt and Magidson 2016), which does not provide exact parameters and applies some rounding (see the xlsx file with estimated model parameters). As a result, our estimates of posterior probabilities calculated from the estimated model parameters differ somewhat from the estimates that are obtained directly from the Latent GOLD output. We have included both sets of posterior probabilities, ones based on the estimated model parameters and ones obtained directly from the Latent GOLD, even though the latter is available only for the response patterns that we do observe. The first pair of datafiles with survey responses contains the posterior probabilities obtained directly from the Latent GOLD.

Any questions about the paper or data should be directed to Nona Karalashvili at nkaralashvili@worldbank.org, or Peter Murrell at murrell@econ.umd.edu.