

Codebook for GAIN Data

Description

These datasets report treatment assignment and earnings outcomes for participants in the GAIN job training program.

The data used in the paper are from Hotz, Imbens and Klerman (2006). We have created a simulated dataset which produces similar results, to demonstrate our method.

The file ***Simulated All Locations GAIN Data*** contains data for all sites. The file ***Simulated Riverside GAIN Data*** subsets the data to the Riverside site.

The data are organized long on participant identifier and wide on outcomes, so that there is exactly one row per individual.

Codebook

Variable	Type	Description
id	Num	Unique identifier for each participant.
site	String	Name of GAIN site (Riverside, Alameda, San Diego or Los Angeles).
treatment	Num	Indicator variable for assignment to the treatment group: control = 0, treatment = 1.
emp[q]	Num	Indicator variable for employment [q] quarters after the treatment: unemployed = 0, employed = 1. Employment outcomes are provided for the first 36 quarters after treatment.