

The data used are taken from

Guryan, Jonathan, Kory Kroft, and Matthew J. Notowidigdo. 2009. "Peer Effects in the Workplace: Evidence from Random Groupings in Professional Golf Tournaments." *American Economic Journal: Applied Economics*, 1 (4): 34-68

and can be downloaded in Stata format from the publisher's website at

<https://www.aeaweb.org/articles?id=10.1257/app.1.4.34>

The data concern professional golf tournaments. They contain 17,942 data points on 81 tournaments spanning three PGA seasons (2002, 2005, 2006). A detailed description on the data and variable construction is given in Guryan et al. (2009). Table 1 in Jochmans (2023) provides descriptive statistics for the variables relevant to the analysis performed there. These variables are

handicap: a golf player's handicap

hand_i: the average handicap of a golfer's opponents

tourn: tournament dummies (there are 81 tournaments in the data)

round: round dummies (a tournament consists of two rounds)

cat: player category dummies (there are three categories based on player 'quality')

grouping_id: player identities for variance clustering

To replicate the test results in Table 2 of Jochmans (2023) first create tournament-by-category dummies as

```
gen tourncat = tourn + "_" + cat
```

Then the three different test statistics can be obtained by running

```
areg handicap hand_i if round==1, vce(cluster grouping_id) absorb(tourncat)
```

```
rassign handicap hand_i if round==1, group(tourncat)
```

```
areg handicap hand_i mean_handicap if round==1, vce(cluster group_id) absorb(tourncat)
```