

guide to data and replication for

Evans, Carolyn L. and James Harrigan, “Distance, Time, and Specialization: Lean Retailing in General Equilibrium”, *American Economic Review* March 2005.

We have provided a zip file that contains the following three files:

```
AER 20030497 guide to data.pdf      (the document you are reading)
AER 20030497 data.dta
AER 20030497 regression log files.pdf
```

The dataset `AER 20030497 data.dta` is extensively self-documented using Stata’s `notes` feature (see below for discussion). If you don’t have access to Stata you use the freeware program “StatTransfer” to convert it to other formats, or you can e-mail us and we’ll do it for you.

The file `AER 20030497 regression log files.pdf` shows the exact commands and results that were used to generate the results reported in the tables in the paper. The dataset used for the regressions was exactly the same as the one included in this package, `AER 20030497 data.dta`, with one key difference: the variables related to replenishment have been deleted due to confidentiality requirements.

We hope these files are useful. If you have any questions, please contact James Harrigan at [<james.harrigan@ny.frb.org>](mailto:james.harrigan@ny.frb.org).

A note on notes

The dataset “AER 20030497 data.dta” is extensively self-documented using Stata’s `notes` feature. To see the definition of any variable x , simply open the dataset in Stata and type

```
. notes x
```

Most variables were created using an ado file called “gg”. The `gg` command is the same as `generate`, except that it adds a note to the variable being created which records how it was created. For example, suppose we wanted to define a variable $z1$ as the sum of $x1$ and $x2$. We would type

```
. gg z1 = x1 + x2
```

which is equivalent to the two commands

```
. generate z1 = x1 + x2
. notes z1 : gg z1 = x1 + x2
```

As a consequence, the variable $z1$ will have a note attached to it giving its definition.