README file for

"Paying on the margin for medical care: Evidence from breast cancer treatments"

By Liran Einav, Amy Finkelstein and Heidi Williams

The two data sets used in the paper cannot be posted online:

1. The first data set is collected by a private company (IMV; http://www.imvinfo.com) and is therefore proprietary, but it is available for purchase.

2. The second source of data is the California Cancer Registry, which records information about cancer patients diagnosed in the state of California (http://www.ccrcal.org/). We obtained a research version of this database with confidential variables as part of a research agreement with PI Professor Joseph Doyle at MIT. The research agreement was approved by the California Committee for the Protection of Human Subjects and the MIT COUHES (Committee on the Use of Humans as Experimental Subjects). The state of California has a standardized application process for obtaining these data (described here: http://www.ccrcal.org/Data_and_Statistics/ Cancer_Data_for_Research.shtml).

For further details about obtaining access to data files associated with this program or questions regarding the code, please contact the authors.

The Stata program code.do documents our data construction and generates all of the empirical estimates presented in the paper, with the following exceptions: Table 5 and standard errors for column 6 of Tables 3, A1, A2, and A3.

The excel file "Table5_calibration" includes data and calculations for producing Table 5 in the paper. The "data" worksheet includes a list of simulated draws underlying the demand curve in Figure 3(a). The "calculations" worksheet uses this data as an input and computes the numbers that are presented in Table 5. Cells M3-M9 include parameter values (including the risk aversion estimates). The numbers in M30-Q33 correspond to Panel A in Table 5, M37-Q40 correspond to Panel B, and M44-47 correspond to Panel C.

The Stata programs Bootstrap_Table3_trvtime.do, Bootstrap_TableA1_trvtime_fullsample.do, Bootstrap_TableA2_sphdist.do, and Bootstrap_TableA3_trvdist.do generate bootstrapped standard errors for column 6 of Tables 3, A1, A2, and A3. Each bootstrap program is implemented as parallelized Stata jobs as described at the beginning of each file. The main program (code.do, described above) creates the intermediate data files called in these programs.