**Readme.docx**

**Documentation for "**[**How Much Energy Do Building Energy Codes Save? Evidence from California Houses**](http://faculty.georgetown.edu/aml6/pdfs&zips/BuildingCodes.pdf)**" American Economic Review by Arik Levinson.**

There are three parts to the paper, corresponding to the three empirical approaches. Part 1 studies California using the annual RASS data from 2003 and 2009. Part 2 uses the monthly RASS billing data. And Part 3 compares California to other states using the RECS.

For parts 1 and 2 there are two programs for each part. One is called part#data.do, which assembles the data used in the analysis. The dataset is called Part#.dta. And the other program is called Part#results.do, which generates the tables and figures in the paper. For part 3 there’s just a part3results.do program, which does both the data assembly (which is minor) and the analysis.

Tables generated by the programs go into the “Tables” folder, with names corresponding to the table numbers in the paper. Figures generated by the programs go into the “Figures” folder.

Part 1 notes

* First merges the two RASS datasets (2003 and 2009), and saves the house characteristics, without the weather or energy data in a dataset called “Part1\_NoEnergy\_1".
* Then merges in the weather and energy data created using the *actual* energy data, which is compiled in Part2data.do.
* But Part2data.do uses the compiled house characteristics from Part1data.do.
* So the data programs need to be run in the following way:
  + Part1data.do up to line 491, which reads   
    “\* Run up to here first, then run Part2data.do before continuing;”
  + Then run Part2data.do, return to Part1data.do and run the rest of the program.

Part 2 notes

* Weather:
  + Gets the weather data for each weather station in California, from NOAA.
  + Cross with zip code population centroids.
  + Gets the weighted avg weather at each population centroid, within 20, 30, and 40 miles, weighted by inverse distance.
* Billing data:
  + Billing cycles don’t correspond to calendar months. The 2003 RASS data I acquired had already been “calendarized”, which means the energy used by each home are assigned proportionally to calendar months. See the 2003 RASS documentation, pp 99ff. This program calendarizes the 2009 RASS billing data in the same way.

|  |  |
| --- | --- |
| **Location of Figures and Tables in Programs** | |
| **Table or Figure** | **Program** |
| Table 1 | Horn et al. (1980) |
| Table 2 | Part1results.do andPart3results.do |
| Table 3 | Part1results.do |
| Table 4 | Part2results.do |
| Table 5 | Part2results.do |
| Table 6 | Part3results.do |
|  |  |
| Figure 1 | Part1results.do |
| Figure 2 | Part3results.do |
| Figure 3 | Part1results.do |
| Figure 4 | Part1results.do |
| Figure 5 | Part2results.do |
| Figure 6 | Part2results.do |
| Figure 7 | Part2results.do |
| Figure 8 | Part3results.do |
| Figure 9 | Part3results.do |
|  |  |
| Appendix Figure A1 | Part3results.do |
| Appendix Figure A2 | Part1results.do |
| Appendix Figure A3 | Part1results.do |
|  |  |
| Appendix Table A1 | Various |
| Appendix Table A2 | Part1results.do |
| Appendix Table A3 | Part1results.do |
| Appendix Table A4 | Part1results.do |
| Appendix Table A5 | Part2results.do |
| Appendix Table A6 | Part2results.do |