REPLICATION PACKAGE FOR ARTICLE: *Optimal Pricing of a New Utility Service: The Case of Piped Water in Vietnam* (Do and Jacoby)

Review of Economic Studies, *forthcoming*

Overview

There are 3 datasets and one main Stata do file in this replication package. The Stata do file runs all of the code needed to generate 5 figures and 2 tables in the paper. The replicator should expect the code to run for about 5 days.

Data Availability and Provenance Statements

Statement about Rights

- I certify that the author(s) of the manuscript have legitimate access to and permission to use the data used in this manuscript.

Summary of Availability

- All data **are** publicly available.

Data Provenance

The data comes from a field experiment and associated survey conducted by the authors combined with administrative data from An Thinh water utility, Vietnam.

Cleaned versions of the raw data stripped of all individual identifiers and processed for analysis have been uploaded to the Review of Economic Studies repository as part of the replication package.

Dataset list

- section0_end2016.dta: raw data – household survey
- admin2016.dta: raw data – administrative data
- Usage_July15_May18.dta: Monthly water usage and price during experiment
- Usage_Sep2012_July2015.dta: Monthly water usage before experiment

Software Requirements

- Stata (code was last run with version 18/MP with a minimum of 4 processors)
- Copy file …\adofile\F_eff.ado into your \ado folder
- In Stata command window, type "search **grc1leg**" and click on *install all*
- In Stata command window, type "search **grc1leg2**" and click on *install all*
- In Stata command window, type "search **st0594**", and click on *install all*

- Install the following ado files in stata by typing in the command window:
    - ssc install **outreg2**, replace
    - ssc install **pdslasso**, replace
    - ssc install **ranktest**, replace
    - ssc install **avar**, replace
- For **ivreghdfe** to be properly installed (which includes **ivreg2**, **ftools**, **reghdfe**, and **ivreghdfe**), follow instructions from https://github.com/sergiocorreia/ivreghdfe#installation

Controlled Randomness

- Random seed is set at line 143 of program Table2.do
- Random seed is set at line 94 & 103 of program Figure6.do
- Random seed is set at line 94 & 103 of program Figure7.do

Memory and Runtime Requirements

Approximate time needed to reproduce the analyses on a standard desktop machine:

- 5 days (with 1000 bootstrap replications)
- 5 hours (with 10 bootstrap replications—for debugging)

Instructions

- Install software (see software requirements)
- Place all contents of replication package in a folder
- Open file "....\analysis code\MAIN.do" for editing and set path as needed in line 18, by indicating the location of the folder where replication package has been placed as per previous bullet point.
- Run "...\analysis code\ MAIN.do" to generate all tables and figures. Output files are given names (`Table1.xls`, `Figure5.pdf`) corresponding to objects in the manuscript and are written to the "...\analysis results\" subdirectory.