README

Replication Package (version 3) for
"Measuring the Incentive to Collude: The Vitamin Cartels, 1990–1999"
(*The Review of Economic Studies*, forthcoming)

Mitsuru Igami Takuo Sugaya

July 14, 2021

1. Overview

This replication package contains the data and the code to generate 9 figures and 4 tables in the paper. The replicator should expect the code to run for *about 12 minutes* to replicate the baseline results, which are Table 2 (column 1), Figure 5, and Figure 6.

Additionally, the same amount of time is needed for each of the three "other vitamin" markets, the alternative models, and the counterfactual simulations in the paper.

2. Data Availability Statements

The paper uses three main sources of data: (i) the expert witness report by Dr. B. Douglas Bernheim, (ii) the report by the European Commission (EC), and (iii) the data from the World Bank. The paper refers to these sources as Bernheim (2002a), EC (2003), and World Bank (2020), respectively.

Source (i) is publicly available for copying at the United States District Court for the District of Columbia. Sources (ii) and (iii) are publicly available for download on the EC's and the World Bank's websites, respectively.

Thus, *all data are publicly available*, and *we certify that the authors of the manuscript have legitimate access to and permission to use the data used in this manuscript*. The code is licensed under a Creative Commons/CC-BY-NC license.

3. Dataset List

The following data files are included in the replication package.

Data file	Source	Notes	Provided
`data.mat`	Bernheim (2002a), EC (2003)	Includes most of the variables	Yes
`worldbank_hic2.mat`	World Bank (2020)	Includes world GDP and population	Yes

4. Computational Requirements

Software: MATLAB (code was run with Release R2013b). One of the main programs (‘`bootstrapICC.m`’) uses the ‘`parfor`’ command, which requires an additional package called Parallel Computing Toolbox. Nevertheless, the replicator should be able to run this program without it by replacing ‘`parfor`’ with regular ‘`for`’.

Memory and Runtime: Approximate time needed to reproduce the main/baseline analyses on a standard 2014 desktop machine is 12 minutes. The code was last run on a 4-core Intel-based desktop PC with Windows 8.1 Pro and 32 GB of RAM. If the replicator replaces ‘`parfor`’ with regular ‘`for`’ (see above), the runtime is expected to double.

5. Description of Programs

‘`gmm_bootstrap.m`’ and ‘`gmm_obj.m`’ will conduct the GMM estimation of the demand and cost parameters in the vitamin C market reported in Table 2 (column 1). The former calls the latter as a function file, and hence the replicator needs to run only the former.

‘`calculateICC.m`’ and ‘`bootstrapICC.m`’ will generate the collusive-incentive estimates in the vitamin C market reported in Figure 6. The latter calls an additional function file ‘`prctile.m`’, which computes percentiles.

The same analyses for the three other vitamin markets (vitamins A and E, and beta carotene), various counterfactual simulations, and alternative models (including columns 2 and 3 of Table 2, and many sensitivity analyses in section 6 and the Online Appendix) will be conducted by the same/similar sets of files in 13 subfolders.

6. Instructions to Replicators (Overview)

The replicator can obtain a single set of results (the demand-and-cost parameter estimates and their dynamic implications in terms of the incentive to collude under a specific model setup) by sequentially running the three programs as follows.

A) Run '`gmm_bootstrap.m`', which takes less than a minute.

B) Run '`calculateICC.m`', which takes less than a minute.

C) Run '`bootstrapICC.m`', which takes about 10 minutes (or longer without `parfor`).

7. List of Exhibits and Programs

This package reproduces all numbers used in the main text, tables, and figures in the paper. The TEX files for the tables, and the EPS files for the figures, are included as well.

Exhibit	Data/Programs	Line #	Output file	Note
Table 1	`data.mat`	—	—	Summary statistics based on price, cost, & quantity data
Table 2				
Column 1	`gmm_bootstrap.m` `(& gmm_obj.m)`	175 (57)	`result_theta.mat`	Main parameter estimates under three alternative specifications of the fringe supply
Column 2	`00a_lambda¥gmm_bootstrap.m` `(& 00a_lambda¥gmm_obj.m)`	175 (57)	`result_theta.mat`	
Column 3	`00b_lambda_kappa¥gmm_bootstrap.m` `(& 00b_lambda_kappa ¥gmm_obj.m)`	175 (57)	`result_theta.mat`	
Table 3	`02_merger¥calculateICC.m` `02_merger¥bootstrapICC.m`	266 369	`result_ICC.mat` `result_ICC_bs.mat`	(See note on Figures 5 & 6 for the location of the main results.)
Table 4	`03_other_mergers¥calculateICC.m` `03_other_mergers¥bootstrapICC.m`	266 369	`result_ICC.mat` `result_ICC_bs.mat`	(See note on Figures 5 & 6 for the location of the main results.)

Figure 1	`data.mat`
	`08_vitaminA¥data.mat`
	`09_vitaminE¥data.mat`
	`10_betacarotene¥data.mat`
		—	—	Line graph based on the prices of vitamins C, A, E, and beta carotene

Figure 2	`data.mat`	—	—	Line graph based on the data on fringe supply
Figure 3	`data.mat`	—	—	Line graph and area graph based on the data on price, cost, and output quantities

Figure 4	08_vitaminA¥data.mat 09_vitaminE¥data.mat 10_betacarotene¥data.mat	—	—	Line graph and area graph based on the data on price, cost, and output quantities
Figure 5	calculateICC.m	266	result_ICC.mat	The main result is stored in two variables named: 'ICC_indiv_est' (a 228x4 matrix) and 'ICC_collect_est' (a 228x1 matrix).
Figure 6	bootstrapICC.m	369	result_ICC_bs.mat	The main result is stored in a variable named: 'ICC_collect_bscomp' (a 228x10 matrix).
Figure 7	08_vitaminA¥calculateICC.m 08_vitaminA¥bootstrapICC.m 09_vitaminE¥calculateICC.m 09_vitaminE¥bootstrapICC.m 10_betacarotene¥calculateICC.m 10_betacarotene¥bootstrapICC.m	266 369 266 369 266 369	result_ICC.mat result_ICC_bs.mat result_ICC.mat result_ICC_bs.mat result_ICC.mat result_ICC_bs.mat	(See note on Figures 5 & 6 for the location of the main results.)
Figure 8	01_decomposition¥calculateICC.m 01_decomposition¥bootstrapICC.m	266 369	result_ICC.mat result_ICC_bs.mat	(See note on Figures 5 & 6 for the location of the main results.)
Figure 9	02_merger¥calculateICC.m 02_merger¥bootstrapICC.m	266 369	result_ICC.mat result_ICC_bs.mat	(See note on Figures 5 & 6 for the location of the main results.)

Note: Other subfolders (not listed here) contain sets of similar programs for various sensitivity analyses in the Online Appendix.

8. How to Replicate Figures (Details)

Each figure in the table above is a manually customized graphic representation of the computational results via Microsoft Excel 2007. The replicator can use this or any other software to create them, as follows.

Figure 1 is a line graph of the four vitamin prices.

A) In each data file (e.g., 'data.mat' in the top folder for vitamin C), the price time-series is stored in 'P', a 228-by-1 vector, each row of which corresponds to a month from January 1980 to December 1998. Hence, rows 121–228 correspond to January 1990–December 1998.[1]

[1] Even though Figure 1 shows the prices for January 1999–December 2001 as well, we do not use this period in the analysis because none of the other variables are recorded.

B) Rescale each of the four prices by (i) dividing all elements of ‘P’ by the element in row 181 (January 1995) and then (ii) multiplying the result by 100.

Figure 2 is a line graph of the fringe output.

A) The fringe output is stored in ‘q_fri’, another 228-by-1 vector.

B) Each of the five dashed lines reflects the “forecast by the cartel,” which is simply a sequence of 12-month-lagged values of the actual fringe output in ‘q_fri’.

C) These quantities are measured in *million* kilograms in the data files, whereas our figures display them in *thousand* kilograms. Hence, the former should be multiplied by 1,000 to match the scale in Figure 2.

D) The cartel period is stored in ‘I_cartel’, a 228-by-1 vector. We plot this variable as a bar graph on a separate scale.

Figure 3 (left) is a filled-area graph of output volumes in which all firms’ time-series are stacked vertically.

A) Each firm’s output volume is stored in ‘q_(*firm name*)’.

B) The same unit adjustment as in Figure 2 (i.e., multiplication by 1,000) applies.

Figure 3 (right) is a line graph of price, cost, and cartel operation.

A) The price is stored in ‘P’.

B) The cost is stored in ‘c_roche’.

C) The cartel indicator is stored in ‘I_cartel’.

Figure 4 is the same as Figure 3 except that the graphs are based on the data from the three other vitamin markets.

Figure 5 is a line graph of the implied monopoly price.

A) The point estimates (“If Monopoly”) are stored in ‘P_mono_est’ in ‘`result_ICC.mat`’.

B) The “standard error” series are stored in ‘P_mono_estcomp’ in ‘`result_ICC_bs.mat`’.

C) The “Data” series is the same as the price in Figure 3 (right).

D) The “Actual Cartel Period” is the same as the cartel period in Figures 2 and 3 (right).

Figure 6 (top) is a line graph of the individual firms' collusive incentives.

A) These point estimates are stored in rows 133–228 of 'ICC_indiv' (a 228-by-4 matrix) in '`result_ICC.mat`'.

B) The "Actual Cartel Period" is the same as in the previous graphs.

Figure 6 (bottom) is a line graph of the cartel's collective incentive.

A) The point estimates are stored in (rows 133–228, column 6) of 'ICC_collect_bscomp' (a 228-by-10 matrix) in '`result_ICC_bs.mat`'.

B) The "95% confidence interval" series are stored in columns 3 and 9 of the same matrix.

C) The "99% confidence interval" series are stored in columns 2 and 10 of the same matrix.

Figure 7 is the same as Figure 6 (bottom) except that the graphs are based on the results from the three other vitamin markets.

Figure 8 is a line graph of the incentive estimates in "decomposition" counter-factual scenarios.

A) As such, each of the four series corresponds to the "point estimates" series in Figure 6 (bottom) or its scenario 1–3 counterparts.

B) The "Actual" series is the same as the point estimates in Figure 6 (bottom).

Figure 9 is a bar graph of the individual firms' incentives under different "synergy" counter-factual scenarios.

A) As such, each set of four bars corresponds to the point estimates in Figure 6 (top) or their counterparts, in row 188 (August 1995).

References

Bernheim, B. Douglas. 2002a. "Expert Report of B. Douglas Bernheim," *In Re: Vitamins Antitrust Litigation*, MDL No. 1285, Misc 99-0197.

European Commission. 2003. "Case COMP/E-1/37.512 -- Vitamins," *Official Journal of the European Communities*.

World Bank. 2020. Data on "Real GDP", "Population", and "Livestock Production" in *DataBank*. Accessed at https://databank.worldbank.org/home.aspx on February 24, 2020.