# FINAL TAKE-HOME EXAM: REPLICATION OF COOK & LUDWIG (2006)

---

**This version:** January 24, 2018

**Goal:** In this final exam you will mainly discuss, replicate and extend results and methods of the article by Cook & Ludwig (2006) and discuss the findings. In contrast to the midterm, this exam is less on data construction. Now you will go a step further and draw statistical conclusions. The focus is not so much on the exact implementation (as you have shown already in the mid-term), but rather on analyzing questions and investigating economic effects. For parts of the exercise it will be up to you to decide which tests or plots to do in order to answer the question.

**Rules:** By completing the full exercise, you can collect 28 points (out of the total of 60 points, achievable for the course). The points in the brackets will give you a basic orientation. Use Stata to generate the answers for all the results. Turn to Stata help-files and then to the internet when searching for a solution to a problem.

Organize your work nicely: mark clearly each of the exercises within the do-file, *label all newly generated variables* and comment on less obvious procedures or choices. Carefully double check your codes.

Feel free to discuss the solutions with your colleagues. You will have to hand in distinct do-, log-, and pdf-files and formulate your answers/comments on your own. In case of identical wordings and other cases of obvious copy-pasting, the copied parts will not be evaluated in any of the affected submissions (and will hence receive 0 points).

**Preliminaries:** Before starting with the exam you should read the article "The social cost of gun ownership" by Philip J. Cook and Jens Ludwig, published in the Journal of Public Economics in 2006.
Note that the provided dataset differs slightly. So for the replication parts you should not expect to find the exact same effects as in the published paper.

**Deliverables:** (1) A do-file and (2) the corresponding log-file (saved in .log-format), saved under the name of `familyname_final.do/log`, containing all following steps (numbered in the do-file). (3) A pdf-file, containing concise written answers to those of the following questions where an explanation is required explicitly, saved under the name of `familyname_final.pdf`.

**Modalities:** Upload the three files (.do, .log, .pdf with your family name) until March, 19, 2018, 23:55 to Ilias. You cannot submit your solution after the submission deadline has expired, so please make sure that you can hand in the final in time.

**Work plan:**

1. Start your do-file the usual way. Open the dataset `final.dta` and make yourself familiar with the variables and data structure. Think of a small test to check whether the variables `e95`, `e955` have obvious measurement problems.

   **PDF-Report:** Briefly describe how you investigate the measurement of the variables and interpret your results. Is there anything else about the dataset that you discovered?

   [1]

2. Compute the main variable of interest, namely $FSS_{i,t}$, as the share of suicides with firearms of total suicides in a given county and given period of time. Then, investigate how stable $FSS_{i,t}$ is at the county level over time, i.e. is the measure quite jumpy or rather slowly moving over time in a given county.

   **PDF-Report:** Use your results on the stability of $FSS_{i,t}$ to argue whether it is a reasonable proxy for gun ownership over time.

   [2]

3. Replicate the descriptive statistics of table 1 of the paper. Generate new variables if necessary.
   **PDF-Report:** Set up a table with descriptive statistics according to table 1 of the paper
   How close are you relative to the original dataset. Do you expect any differences due to deviations of the descriptive statistics? Given the results, how important is a weighting of the data?

Table 1.1: Descriptive statistics for county data

| | CL (2006) | unweighted | weighted |
|---|---|---|---|
| *Full period (1980–1999)* | | | |
| FSS in selected years | 49.9 | | |
| Homicide rate | 11.0 | | |
| Gun homicide rate | 7.3 | | |
| %Urban | 92.6 | | |
| %Percent black | 14.0 | | |
| %Female household head | 18.0 | | |
| # Suicides | 195.8 | | |
| | | | |
| *FSS in selected years* | | | |
| 1980 | 48.0 | | |
| 1990 | 52.8 | | |
| 1999 | 48.0 | | |

[3]

4. Use `xtset` to set the panel dimensions with `fips` and `year`. Then, replicate the results of table 2 as close as possible using `xtreg` to estimate specification (1.1) from below. Generate new variables if necessary. The full model of the last column of table 2 is given by:

$$\begin{aligned}\log\left(\frac{homicides_{i,t}}{population_{i,t}}\right) = \; & \alpha + \beta \log\left(\frac{firearmsuicides_{i,t-1}}{suicides_{i,t-1}}\right) + \gamma_1 \log\left(\frac{robberies_{i,t}}{population_{i,t}}\right) \\ & + \gamma_2 \log\left(\frac{burglaries_{i,t}}{population_{i,t}}\right) + \gamma_3 \log\left(\frac{blacks_{i,t}}{population_{i,t}}\right) \\ & + \gamma_4 \log\left(\frac{urbanresidents_{i,t}}{population_{i,t}}\right) + \gamma_5 \log\left(\frac{samehouse_{i,t}}{population_{i,t}}\right) \\ & + \sigma \log\left(\frac{femaleheadedhouseholds_{i,t}}{households_{i,t}}\right) + d_i + d_t + \varepsilon_{i,t},\end{aligned} \tag{1.1}$$

   where $i$ corresponds to the county and $t$ describes the year.
   When estimating the model, make sure that you constrain your sample so that you estimate all regressions on the same sample.
   **PDF-Report:** Set up a table in which you compare your results with the results of the original paper. Interpret your findings with a particular emphasis on potential differences. Are the results qualitatively the same? What happens if you cluster the standard errors on the county level instead of using robust standard errors?

[6]

5. Is the model robust to adding $\log(population_{i,t})$ as additional regressor to the specification?
   **PDF-Report:** Discuss your findings.

[2]

6. Estimate a transformation of the original model of the following form:

$$\begin{aligned}\log(homicides_{i,t}) = \; & \alpha + \beta \log\left(\frac{firearmsuicides_{i,t-1}}{suicides_{i,t-1}}\right) + \tilde{\gamma_1} \log(robberies_{i,t}) \\ & + \tilde{\gamma_2} \log(burglaries_{i,t}) + \tilde{\gamma_3} \log(blacks_{i,t}) + \tilde{\gamma_4} \log(urbanresidents_{i,t}) \\ & + \tilde{\gamma_5} \log(samehouse_{i,t}) + \sigma \log\left(\frac{femaleheadedhouseholds_{i,t}}{households_{i,t}}\right) \\ & + \delta \log(population_{i,t}) + d_i + d_t + \varepsilon_{i,t}\end{aligned} \tag{1.2}$$

**PDF-Report:** Add a table in which you compare the estimates of both models for the full specification of the model. Discuss your findings.

[3]

7. If you look careful you can see that the original model is actually a restricted version of the general model from the previous question. The transformation relies on a linear restrictions imposed on $\delta$. Use the post-estimation command `test` to test the validity of the relevant restriction.

   **PDF-Report:** Write about half a page in which you discuss the linear restrictions and the relation between the models. Interpret the outcome of your test. Briefly discuss also what restrictions you would impose on an even more general model, which does not include fractions, i.e. contains $firearmsuicides_{i,t-1}$, $suicides_{i,t-1}$, $femaleheadedhouseholds_{i,t}$, and $households_{i,t}$ as separate regressors.
   Bonus-Point: Estimate this more general model and test for the linear restrictions that you have to impose.

[3(+1)]

8. Estimate the model under the explicit imposition of the constraint to verify your constraint. To do so, estimate the model with `cnsreg` after explicitly modeling the constraint, using the `constraint` command.

   **PDF-Report:** Set up a table in which you compare the coefficient of $\gamma$ in the four different estimations from exercise 4, 5, 6, and 8. Briefly discuss the findings.

[3]

9. Estimate the model of question 6 in first differences, i.e.

$$\begin{aligned}\Delta\log(homicides_{i,t}) = \; & \beta\Delta\log\left(\frac{firearmsuicides_{i,t-1}}{suicides_{i,t-1}}\right) + \tilde{\gamma_1}\Delta\log(robberies_{i,t}) \\ & + \tilde{\gamma_2}\Delta\log(burglaries_{i,t}) + \tilde{\gamma_3}\Delta\log(blacks_{i,t}) + \tilde{\gamma_4}\Delta\log(urbanresidents_{i,t}) \\ & + \tilde{\gamma_5}\Delta\log(samehouse_{i,t}) + \sigma\Delta\log\left(\frac{femaleheadedhouseholds_{i,t}}{households_{i,t}}\right) \\ & + \delta\Delta\log(population_{i,t}) + d_t + \varepsilon_{i,t}\end{aligned} \tag{1.3}$$

**PDF-Report:** Interpret your findings. Then, write about half a page and argue whether there is a robust relation between homicides and gun ownership in light of your results. Think about a problem that the initial problem might face in which you divided almost all variables by population. Which model would you choose and why?

[3]

10. As last exercise, implement an additional empirical analysis that you consider interesting.
    **PDF-Report:** Describe what you want to test/analyze and why it is interesting. Interpret your findings.

    [2]

11. Save your dataset and close your log-file.