

## **README file for**

### **“The Health Effects of Prison”**

Randi Hjalmarsson  
University of Gothenburg and CEPR

Matthew J. Lindquist  
Swedish Institute for Social Research, Stockholm University

May 28, 2021

#### **1. Introduction**

We are not permitted to place our microdata on the *Journal's* homepage. In Section 2, we provide instructions on how researchers can purchase similar data directly from Statistics Sweden (SCB), Sweden's National Council for Crime Prevention, the Swedish Prison and Probation Service, and the Swedish National Board of Health and Welfare. These data can be accessed through SCB's secure server system MONA. In Section 3, we provide a comprehensive list of variables that need to be ordered as well as their sources. In Section 4, we provide a list of all of the STATA do files needed to replicate our study as well as instructions on how to implement these do files. Section 5 includes specific instructions and details that enable the *Journal's* Data Editor to access our analysis data sets and to replicate the tables and figures presented in our paper.

#### **2. Obtaining Access to Swedish Microdata**

To obtain access to Swedish register data, researchers must first have their projects approved by Sweden's Central Ethical Review Board. Instructions can be found on the review board's homepage at <https://etikprovningssmyndigheten.se/>. Please note that part of this application must be done in Swedish. Once a project has been approved, then a researcher can turn to Statistics Sweden (SCB) in order to obtain access to microdata. Contact information (in English) can be found on their homepage at <http://www.scb.se>. SCB assigns a contact person to each new researcher. The contact person's job is to guide researchers through the data ordering process and to coordinate the delivery of the data via SCB's own secure server system called [MONA](#).

Please note that SCB will also do its own review of your project plan. Also, they will only give access to researchers who have an affiliation with a Swedish university, research institute, or other organization that has been granted permission to work with Swedish microdata. One does not have to be a Swedish citizen, and researchers are allowed to access the MONA system from other countries within the European Union. However, one is not allowed to work with the data while located outside of the European Union.

The sampling frame used in this paper was constructed by the [Swedish Prison and Probation Service](#) and consists of all individuals who entered a Swedish prison between January

1992 and December 2013. The sampling frame also includes a snapshot of all individuals who were in prison in January 1992. This list of individuals was then given to SCB along with the prison data used in this paper (more on this below).

Crime conviction and court decision data are supplied by Sweden's National Council for Crime Prevention. Information can be found at <http://www.bra.se>. In practice, SCB sends a list of all individuals in the sample to BRÅ, who match on the necessary crime variables and then send these data back to SCB. Mortality and hospitalization data can be obtained from the Swedish National Board of Health and Welfare. They also cooperate with SCB to deliver data. See, e.g., <https://www.socialstyrelsen.se/en/statistics-and-data/register/>.

### **3. Variable List and Source**

We provide a list of variable names in English. In our experience, the given Swedish name of the variable can – in some instances – change from one data order to the next (although the content is the same). Your contact person will help extract the correct variable. If needed, the exact variable names and labels that were given to us can be seen in our do files before we give them more user friendly variable names in English.

The following variables were ordered from each of the following government authorities.

#### **1. The Swedish Prison and Probation Authority provided data on:**

- a. Personal identifiers for all individuals who entered a Swedish prison between January 1992 and December 2013, as well as personal identifiers for all individuals who were in prison in January 1992.
- b. For this list of individuals, the prison authority provided data on:
  - i. Start date
  - ii. End date
  - iii. Facility placement
    1. Many inmates have more than one placement.
    2. Start and end date for each placement
  - iv. Main crime type
  - v. Age
  - vi. Gender
  - vii. Citizenship
  - viii. Died in prison
  - ix. Transferred from pre-trial detention to prison
  - x. Transferred from prison to pre-trial detention
  - xi. Transferred to foreign prison
  - xii. Transferred from abroad to Swedish prison
  - xiii. Rehabilitative treatment program participation, including start date, end date (if not completed), completion date, and program name. (only available from 2002)
  - xiv. In prison visits with a doctor, including dates. (only available from 2009)

- xv. In prison visits with a nurse, including dates. (only available from 2009)
  - xvi. In prison visits with a psychologist, including dates. (only available from 2009)
  - xvii. Received medication. (only available from 2009)
- 2. Statistics Sweden provide data from various official data registers that they maintain. You need only provide a list of variables and your contact person at SCB will extract them from the appropriate register. SCB provided us with data on:
  - a. Marital status
  - b. Children in the home
  - c. Annual pre-tax total factor income
  - d. Annual pre-tax earnings
  - e. Employment status in November
  - f. Education level
  - g. Born in Sweden (yes or no)
  - h. Immigration date
    - i. A person may have more than 1
  - i. Emigration date
    - i. A person may have more than 1
  - j. Death date
- 3. Sweden's National Council for Crime Prevention provide data from Sweden's national convictions register (*lagföringsregistret* in Swedish).
  - a. For those individuals in our sampling frame we have:
    - i. All convictions between 1973 and 2016, including information on:
      - 1. Conviction date
      - 2. Sentence received (fine, prison, other)
  - b. Appendix Figure 2 was created using the universe of all convictions and court decisions between July 1991 and July 1995 and between January 1997 and January 2001. In these data we observe
    - i. Decision date
    - ii. Decision type (conviction, waiver of prosecution, summary sanction order)
    - iii. Sentence received (fine, prison, other)
- 4. Sweden's National Board of Health and Welfare provided data from their
  - a. Inpatient hospitalization register (*patientregister* in Swedish)
    - i. Start date
    - ii. End date
    - iii. Hospital ward type (MVO code)
  - b. Cause of death register (*dödsorsaksregister* in Swedish)
    - i. ICD10 code for main cause of death

- ii. Coroner's note for alcohol or drug related death

## 4. STATA do Files

### 4.1. Preparing the Register Data

Our primary data set (named *Prison Data 1.dta*) is created directly from the register data mentioned above. This data is created using the STATA do file named *create\_primary\_data\_set\_AEJ.do*.

Death date and cause of death are included in this data set and are used to create or main outcome variables for mortality. Mortality outcome variables are created in the analysis data set do files discussed in section 4.2. below.

Outcome variables that are created directly from the register data mentioned above include: crime outcomes (recidivism), hospital outcomes, and employment outcomes (earnings, income, and employment in November). These outcome variables are included in our data sets named:

*Crime Outcome Variables.dta*,  
*Hospital Outcome Variables.dta*, and  
*Labor Market Outcome Variables.dta*.

These outcome variables were created using the following two STATA do files:

*create\_labor\_market\_outcomes\_AEJ.do*, and  
*create\_crime\_and\_hospital\_outcome\_variables\_AEJ.do*.

Importantly, the three STATA do files discussed above must be run before one can create the analysis data sets. The data sets that they generate are used as input when creating our analysis data sets.

Other data sets that were created directly from the register data available to us include:

The *Data For Figures 8 and 9.dta* used in Figures 8 and 9 were created using *create\_data\_for\_figures\_8\_and\_9\_AEJ.do*.

The *Data For Appendix Figure 2.dta* were created using *create\_data\_for\_appendix\_figure\_2\_AEJ.do*.

The *Data For Appendix Figure 7.dta* were created using *create\_data\_for\_appendix\_figure\_7\_AEJ.do*.

### 4.2. Creating the Analysis Data Sets

The following STATA do files create the data sets used in the analysis. We list them here, and summarize their role.

Please note that you must set your own correct “path” at the top of each file listed below in order for them to run properly.

1. *create\_analysis\_data\_set\_AEJ.do*: This do file creates the main analysis data set called *cleaned\_sample.dta*. It imposes all of the necessary baseline sample restrictions.
2. *create\_analysis\_data\_set\_startdate\_analyses\_AEJ.do*: This do file generates a robustness check data set called *cleaned\_startdate\_sample.dta*. This data set differs from the main analysis sample in two ways. It includes the straddle sample of individuals convicted before the reforms but who start sentences after; these individuals need to be manually dropped if you want to use the main analysis sample. It also measures all outcomes relative to the start date of prison rather than the release date from prison.
3. *create\_analysis\_data\_set\_for\_sensitivity\_AEJ.do*: This do file generates the sensitivity analysis sample that allows us to test the robustness of our results to various sample creation restrictions. It allows us to (i) include the straddle sample, (ii) include the trimmed sample of sentences with share time served longer than 110% and less than 10%, and (iii) those who died in prison. One must remember to drop the appropriate sub-samples when conducting sensitivity checks. The resulting data set is called *cleaned\_sample\_sensitivity.dta*

#### 4.3 List of STATA do Files That Generate Figures and Tables

Below, we list each figure and table in the paper or appendix, and indicate the name of the STATA do file that creates that table or figure. The do file titled *main\_analysis\_file\_AEJ.do* calls on three data sets to conduct the analyses: *cleaned\_sample*, *cleaned\_startdate\_sample*, and *cleaned\_sample\_sensitivity*. The STATA do file that creates Figures 8 and 9 calls on *Data For Figures 8 and 9.dta*. The STATA do file that creates Appendix Figures 2 calls on *Data For Appendix Figure 2.dta*. The STATA do file that creates Appendix Table 7 calls on *Data For Appendix Figure 7.dta*.

Figure/ Table	#	Title	STATA do file name
Figure	1	Sweden's Early Release Policies 1990-2002	main_analysis_file_AEJ.do
Figure	2	Descriptive Statistics: Dynamic Paths of Mortality, Recidivism, Hospitalization	main_analysis_file_AEJ.do
Figure	3	Comparison of Mean 10-Year Mortality (All-Cause and Suicide) for Inmates Affected and Unaffected by the Reforms.	main_analysis_file_AEJ.do

Figure	4	Kernal Densities: Sentence and Share Time Served Distributions for Pre-reform, Post Reform and Straddle Samples	main_analysis_file_AEJ.do
Figure	5	Implementation of 1993 and 1999 Reforms: Effect on Share and Days Served by Sentence Length Bins	main_analysis_file_AEJ.do
Figure	6	Dynamics of Mortality Effects	main_analysis_file_AEJ.do
Figure	7	Heterogeneity Analysis of Mortality Effects for High Risk Sub-Groups	main_analysis_file_AEJ.do
Figure	8	In-Prison Medical Data for All Inmates with 4-48 month sentences 2009-2013	figures_8_and_9_AEJ.do
Figure	9	In-Prison Treatment Program Participation of Inmates with 4-48 Month Sentences Entering Prison Between 2009 and 2013.	figures_8_and_9 AEJ.do
Table	1	Summary Statistics	main_analysis_file_AEJ.do
Table	2	Relevance: The Effect of the Two-Thirds Sentencing Reform on Days Served	main_analysis_file_AEJ.do
Table	3	The Effect of Exposure to the Two-Thirds Reform on Mortality Overall and by Cause	main_analysis_file_AEJ.do
Table	4	Robustness Checks of the Effect of the Two-Thirds Reform on Mortality	main_analysis_file_AEJ.do
Table	5	Heterogeneous Effects of the Two-Thirds Reform on Overall Mortality	main_analysis_file_AEJ.do
Table	6	The Effects of Exposure to the Two-Thirds Reform on Recidivism and Labor Market Outcomes	main_analysis_file_AEJ.do
Appendix Figure	1	Swedish Prisons in International Perspective	main_analysis_file_AEJ.do
Appendix Figure	2	Implementation: No manipulation of prosecutor or judge decisions	appendix_figure_2_AEJ.do
Appendix Figure	3	Dynamics of Mortality Effects for Other Causes of Death	main_analysis_file_AEJ.do
Appendix Figure	4	Parallel trends in pre-reform hospitalization rates	main_analysis_file_AEJ.do
Appendix Figure	5	Dynamics of Mortality Effects Using the Prison Start Date as the At-Risk Date	main_analysis_file_AEJ.do
Appendix Figure	6	Average Annual Stock of Inmates and Average Annual Prison Capacity	main_analysis_file_AEJ.do
Appendix Figure	7	Monthly Prison Level Descriptive Statistics, 1992-20004	appendix_figure_7_AEJ.do

Appendix Table	1	Share Time Served Prescribed by the Law	main_analysis_file_AEJ.do
Appendix Table	2	Sample Restrictions (A log file sample_creation is created as you run the do file)	create_analysis_data_set_AEJ.do
Appendix Table	3	Relevance across Sub-samples	main_analysis_file_AEJ.do
Appendix Table	4	Do Observables Change with Exposure to the reform?	main_analysis_file_AEJ.do
Appendix Table	5	Sensitivity of Mortality Results to Sample Creation Decisions (Straddle Sample and Trimming)and Sample Window	main_analysis_file_AEJ.do
Appendix Table	6	heterogeneity for High-risk subsamples	main_analysis_file_AEJ.do
Appendix Table	7	Heterogeneous Effects by Sentence Bin	main_analysis_file_AEJ.do
Appendix Table	8	Heterogeneity in recidivism and labor market results	main_analysis_file_AEJ.do
Footnote 24		To address question on changing samples in Table 3	main_analysis_file_AEJ.do
In text p. 19-20		Culling discussion	main_analysis_file_AEJ.do
Footnote 26		Parallel trends seen in pre-reform hospitalization rates separately for each reform.	main_analysis_file_AEJ.do

## 5. Additional Instructions for the Data Editor

**Step 1:** Please write and sign a data access and confidentiality statement:

“I (... Insert your name and work title here ...) understand that I am being granted server and data access for the sole purpose of replicating the results reported in Hjalmarsson and Lindquist’s (2021) paper entitled “The Health Effects of Prison” in order to fulfill the replication requirements of the AEJ: Applied Economics. I will not use the access granted for any other purpose. I promise that I will not attempt to download any data from the server, nor will I attempt to download any other outputs from which data could be reconstructed. I will also not upload additional data to the server. Even within the server environment, I will not attempt to identify specific individuals or specific groups of individuals in the data. I furthermore promise that I will use the servers in good faith, and will not knowingly endanger either the security of the server or the confidentiality of the data stored on it. I understand that login credentials are personal, and should not be shared with others, nor should they be stored in a way in which they could easily be observed by others.

Please add a signature to this statement along with your work address, email, and telephone number. A pdf copy of this signed statement can be sent to [matthew.lindquist@sofi.su.se](mailto:matthew.lindquist@sofi.su.se).

Please note that server use is logged, and that logs are monitored by the server administrators.

**Step 2:** After receiving the data access and confidentiality statement, Matthew Lindquist will put you in contact with a server administrator.

**Step 3:** The server administrator will provide instructions on how to access the server account and help configure the necessary client software.

**Step 4:** In the home folder of this account, there is a folder entitled “AEJ\_Replication\_Folder”.

In this folder, there are 7 STATA data sets (see details in Section 4):

*Prison Data 1.dta*

*Labor Market Outcome Variables.dta*

*Crime Outcome Variables.dta*

*Hospital Outcome Variables.dta*

*Data For Figures 8 and 9.dta*

*Data For Appendix Figure 2.dta*

*Data For Appendix Figure 7.dta*

These data are drawn directly from the full set of register data that we have access to.

In this folder, there are 6 STATA do files that were used to create these data sets. These do files cannot be run, since they access data outside of the folder that the replicator has access to.

*create\_primary\_data\_set\_AEJ.do.*

*create\_labor\_market\_outcomes\_AEJ.do*

*create\_crime\_and\_hospital\_outcome\_variables\_AEJ.do.*

*create\_data\_for\_figures\_8\_and\_9\_AEJ.do.*

*create\_data\_for\_appendix\_figure\_2\_AEJ.do.*

*create\_data\_for\_appendix\_figure\_7\_AEJ.do.*

**Step 5:** The replicator should now run the STATA program *create\_analysis\_data\_set\_AEJ.do*. This do file creates the main analysis data set called *cleaned\_sample.dta*. It imposes all of the necessary baseline sample restrictions.

**Step 6:** The replicator should now run the STATA program code *create\_analysis\_data\_set\_startdate\_analyses\_AEJ.do*. This do file generates a robustness check data set called *cleaned\_startdate\_sample.dta*. This data set differs from the main analysis sample in two ways. It includes the straddle sample of individuals convicted before the reforms but who start sentences after; these individuals need to be manually dropped if you want to use



the main analysis sample. It also measures all outcomes relative to the start date of prison rather than the release date from prison.

**Step 7:** The replicator should now run the STATA program *create\_analysis\_data\_set\_for\_sensitivity\_AEJ.do*: This do file generates the sensitivity analysis sample that allows us to test the robustness of our results to various sample creation restrictions. It allows us to (i) include the straddle sample, (ii) include the trimmed sample of sentences with share time served longer than 110% and less than 10%, and (iii) those who died in prison. One must remember to drop the appropriate sub-samples when conducting sensitivity checks. The resulting data set is called *cleaned\_sample\_sensitivity.dta*

Please note that all data sets needed for replicating the paper now exist in the folder.

**Step 8:** The STATA program *main\_analysis\_file\_AEJ.do* can now be used to replicate the tables and figures in our paper. At the top of this program file there is a table of contents. Here the replicator can turn each part of the program on or off. Then that particular table or figure is produced. For example, to create Figure 1 “set local figure1 = 1”. Changing the number from 0 to 1 allows the replicator to run the code that produces that figure and only that figure.

**Step 9:** To create Figures 8 and 9 run *figures\_8\_and\_9 AEJ.do*.

**Step 10:** To create Appendix Figure 2 run *appendix\_figure\_2\_AEJ.do*.

**Step 11:** To create Appendix Figure 7 run *appendix\_figure\_7\_AEJ.do*.

**End:** All figures and tables have now been successfully replicated.