README file for

"The Effects of Parental and Sibling Incarceration: Evidence from Ohio"

By Samuel Norris, Matthew Pecenco, and Jeffrey Weaver

Openicpsr-132762

This readme file contains instructions on how to replicate the results of our paper. The underlying data used for the project is not publicly shareable, and so we cannot post that data or share it with other researchers. However, we provide details on accessing all of the data sets used in the project on the next page. We also provide all of the cleaning code that can be used to transform the raw data for analysis as well as all of our analysis code.

Changes Made (2/10/2022): The code in this deposit have been updated after publication of the article. The repository contained an old version of the code for creation of the UJIVE instrument. This revision updates that code to the newest version in the folder cleaning/helper.[1] All the results are the same since they were based on the updated version of the code.

Project Folders

The following diagram shows the levels of the project tree structure at a high level.

Project_Folder/

- README.pdf
- Code/
    - ado/
    - analysis/
    - cleaning/
    - pathnames.do
- Data/
    - Raw/
    - Interim/
    - Clean/
- Exhibits/

Before running the code, it is necessary to update the global paths in the pathnames.do file for the relevant file structure on your computer. If you set up your directories as in the above diagram, then it is only necessary to update the global $PROJ with the location of the project folder.

Within each sub-directory in the cleaning directory, there is a single code file that calls all of the other do-files in that directory (if relevant). In order to run all of the cleaning code in order, use the do-file

[1] This is make_z_ujive.do, make_z_ujive_1yrmoreless.do, make_z_ujive_6momoreless.do, make_z_ujive_alt.do and make_z_ujive_intensive.do.

"master_cleaning.do" in the "code/cleaning" directory. In order to run all of the analysis code, use the do-file "replication_analysis.do" in the "code/analysis" directory. Our code also uses a number of custom ado files, which can be found in the "code/ado" directory. See the below tables for details on each of the individual do-files we use for cleaning and analysis.

The analysis output is all placed in the Exhibits folder, where each of the tables and figures are named based on their name in the paper (e.g. TableA1.tex).

Software and Computational Requirements

This section describes the software and computational requirements for running the relevant analysis code, as well as the relevant packages that must be installed.

- Stata version 15
  - ftools (as of 2020-08-17)
  - reghdfe (as of 2020-08-17)
  - outreg (as of 2020-08-17)
  - texsave (as of 2020-08-17)
  - estout (as of 2020-08-17)
  - coefplot (as of 2020-08-17)
  - geodist (as of 2020-08-17)
  - unique (as of 2020-08-17)
- R version 4.0.3
  - data.table (1.13.6)
  - tidyverse (1.3.1)
  - scales (1.1.1)
  - stringi (1.5.3)
  - reshape2 (1.4.4)
  - stargazer (5.2.2)
  - tables (1.13.6)
  - haven (2.4.1)
  - ggplot (2_3.3.3)
  - foreign (0.8-81)

  - readstata13 (0.9.2)
  - stringr (1.4.0)
  - plyr (1.8.6)
  - dplyr (1.0.2)
- Python version 3.5

Memory and Runtime Requirements

The code was last run on an Intel ® Core ™ i5-5200U CPU with 8 GB of RAM on Windows 10. Total run time for the analysis code is approximately 12 hours. Total run time for the cleaning code is approximately 5 days.

Project data

This table describes the different data sets used in the projects as well as how to get access to them.

Name of data	Description of data	Instructions for access
Cleveland Metropolitan School data	This contains educational data from the Cleveland Metropolitan School District for students between 2010 and 2017. This contains name, date of birth, standardized test scores and attendance.	CMSD welcomes research requests, which can be made at https://www.clevelandmetroschools.org/Page/14819 .
Cleveland Municipal Court records	Court dockets containing defendant name, date of birth, accused crime, all recorded information on the case, and case outcome.	Details on this data are available from https://clevelandmunicipalcourt.org/public-access . To request access to the data, contact the Clerk of Courts (ClerkofCourts@cmcoh.org)
Cuyahoga County Common Pleas Court records	Court dockets containing defendant name, date of birth, accused crime, all recorded information on the case, and case outcome.	Details on this data are available from https://cpdocket.cp.cuyahogacounty.us/ and https://coc.cuyahogacounty.us/en-us/public-records-request.aspx . To request access to the data, contact the Case Record Requests (coccfr@cuyahogacounty.us).
Cuyahoga County Juvenile Court records	Juvenile court records containing defendant name, date of birth, charges, and case outcome.	Details on accessing this data are available from http://juvenile.cuyahogacounty.us/en-US/Clerks-Office.aspx. To request access to the data, contact Linda Brooks (LBrooks@cuyahogacounty.us)
Eviction Lab court records	Matched data from Eviction Lab on whether the individual had a case filed against them for eviction	The Eviction Lab will match data by name and date of birth, and return an anonymized version. Research requests can be made at https://evictionlab.org/data-request/ .
Franklin County Municipal and Common Pleas court records	Court dockets containing defendant name, date of birth, accused crime, all recorded information on the case, and case outcome.	To request access to this data, contact the Public Records Request office using the request form here.
Hamilton County Municipal and Common Pleas court records	Court dockets containing defendant name, date of birth, accused crime, all recorded information on the case, and case outcome.	To request access to this data, contact the Public Records Request office using the request form here.

Ohio Department of Health Birth Certificates	Restricted access birth certificate data from the Ohio Department of Health. This contains name and date of birth of the child, name and age of mother and father, and residential address of the mother.	To request access to this data, submit a request to the Ohio Department of Health Institutional Review Board. Details and instructions can be found here.
Ohio ACS data	This is census block group level data from the American County Survey waves 2010-2015. Census block group level data is matched to addresses in the court records and birth records for use in placebo checks.	This data is public access and can be downloaded from the Census website. To download the relevant tables, go to https://data.census.gov/cedsci/ and click "Advanced Search". Then under "Browse Filters", select Geography, Block Group, Ohio, All Block Groups Within Ohio; under "Surveys", select "ACS 5-Year Estimates Detailed Table' and under Years, select 2015. Finally, in the "Search Query" field, individually type in the name of each table to be downloaded. Our analysis uses tables B01003, B22010, B02001, C17002, B03003, B15003, B19013, B19049, B19301, B19052, B19055, B19056, B19057, B25002, B25003, B25018, B27010 (see clean_ohio_census.do).
Ohio Voter Data	This data contains information on all active and inactive voters in the state of Ohio between 2000 and 2016, including their name, address, date of birth, and voting history.	This data can be obtained via a Public Records Request through the office of the Ohio Secretary of State. Contact information for making these requests can be found at https://www.ohiosos.gov/records/ . Direct queries can be made to David Bowling (dbowling@ohiosecretaryofstate.gov)

Cleaning do-files

The below table describes the do-files that are used for the cleaning in the paper. To run all of these, use master_cleaning.do to call all of the cleaning code in order.

	Code	Code description
	master_cleaning.do	Runs all of the cleaning files in the correct order
	bc/master_bc.do	Runs all of the cleaning files to clean the Ohio birth certificates data.
	bc/clean_bc.do	Cleans and processes the birth certificates data for geocoding and matching to the other data sets
	bc/geocode_oh_bc.do	This cleans the addresses on the Ohio birth certificate data, creates output files for geocoding in ArcGIS, and then re-reads the geocoded ArcGIS output to create a dataset of geocoded locations. Note it is necessary to output the addresses, then use ArcGIS or another software to geocode the addresses, and then take the geo-coded outputs to feed into the rest of the code.
	bc/add_addresses.do	Takes the geocoded output and merges it into the birth certificate data
	census_data/ clean_ohio_census.do	Reads in census block-group level tables of variables from the American Community Survey for Ohio. Creates a Stata dataset at the census block group level of characteristics of these block groups
	cuya_juvenile/ clean_cuya_juvenile.do	Cleans the Cuyahoga County juvenile court records and creates a cleaned Stata file for analysis
	cuya_schools/student_info.do	Cleans the raw Cleveland schools data files, merges them, and produces panel dataset
	cuyahoga/ master_clean_cuyahoga.do	Runs all of the files to clean the Cuyahoga Common Pleas and Cleveland municipal courts data
	cuyahoga/arraigned_indicator.do	Generates an indicator for whether the case was ever arraigned
	cuyahoga/case_costs_mun.do	Takes the raw municipal court data and creates outcomes of costs and fines for the case
	cuyahoga/case_dates_mun.do	Takes the raw municipal court data and cleans trial dates and docket assignments
	cuyahoga/case_dates.do	Takes the raw common pleas court data and cleans trial dates
	cuyahoga/clean_detainers.do	Takes the raw common pleas court data and cleans information on whether the defendant was already in jail
	cuyahoga/clean_judge.do	Takes the raw common pleas court data and cleans judge names and assignment order

	cuyahoga/ combine_clean_combined.do	Merges all of the interim data files created by the previous do files to create the analysis data set for Cuyahoga county
	cuyahoga/ combine_docket_entries_mun.do	Merges interim municipal data files on outcomes and parole violations
	cuyahoga/ combine_docket_entries.do	Merges interim common pleas data files on outcomes and parole violations
	cuyahoga/date_returned.do	Formats data on date returned from parsing.
	cuyahoga/ def_characteristics_mun.do	Takes raw municipal court data and cleans defendant-level information
	cuyahoga/def_characteristics.do	Takes raw common pleas court data and cleans defendant-level information
	cuyahoga/def_charges_mun.do	Takes the raw municipal docket data and creates a file with the set of charges against the defendant in the case
	cuyahoga/def_charges.do	Takes the raw common pleas docket data and creates a file with the set of charges against the defendant in the case
	cuyahoga/ def_disposition_mun.do	Takes raw data and cleans case disposition and case transfer information
	cuyahoga/ def_docket_fulloutcomes_mun.R	Takes raw municipal docket data and cleans punishment and parole information
	cuyahoga/ def_docket_fulloutcomes.R	Takes raw common pleas docket information and cleans punishment and parole information for the 2000s
	cuyahoga/ def_docket_limitedoutcomes.R	Takes raw common pleas docket information and cleans punishment and parole information for the 1990s and early 2000s
	cuyahoga/drop_guilty.do	Takes raw common pleas data and identifies cases where the plea and arraignment occur on the same day
	cuyahoga/drop_same_plea.do	Takes raw common pleas data and identifies cases where a plea occurs on the same day as sentencing in another case
	cuyahoga/dropped_cases.do	Takes raw common pleas data and identifies cases where formal charges were dropped
	cuyahoga/extract_docket_mun.R	Reshapes municipal court data to be at the correct level for data cleaning
	cuyahoga/extract_docket.R	Reshapes common pleas court data to be at the correct level for data cleaning
	cuyahoga/find_pre_judge.do	Takes raw data and extracts first judge assigned for early cases in study period
	cuyahoga/ geocode_cuyahoga_mun.do	Merges geocoded courts records with information from the ACS about the neighborhood
	cuyahoga/judge_mun.do	Takes raw municipal court data and cleans judge names and assignment order
	cuyahoga/judge_transfers.do	Takes in judge assignment data and determines whether the case was transferred between judges
	cuyahoga/make_id.do	Generates unique identifiers for individuals

	cuyahoga/ pred_black_cuyahoga.do	Predicts defendant likelihood of being black, using voter roll information
	evictions/master_evictions.do	Runs all of the cleaning files for the merged eviction court data
	evictions/eviction_post_id.do	Updates the returned merged data from Eviction lab with unique identifiers
	evictions/eviction_post_clean.do	Combines the anonymized courts data with the anonymized evictions records
	evictions/ eviction_post_clean_panel.do	Combines the anonymized courts data with the anonymized evictions records and creates a quarterly panel data set
	franklin/master_clean_franklin.do	Runs all cleaning files for Franklin county
	franklin/clean_bdays.do	Clean file of defendant birthdates
	franklin/create_geocoded_addresses.do	Output addresses to be geocoded and matched to census block groups.
	franklin/clean_closed.do	Get date when case was closed.
	franklin/clean_entries.do	Clean docket of events that happened during case.
	franklin/clean_charges.do	Clean and standardize charges.
	franklin/clean_courtrooms.do	Clean courtroom number.
	franklin/clean_codefendants.do	Identify groups of codefendants.
	franklin/clean_txt.do	Clean non-docket information from cases and merge in information about outcomes.
	franklin/clean_bdates.do	Adjusts birthdates to account for transposition mistakes that make the same individual tried multiple times appear to be a different individual.
	franklin/combine_bdates.do	Combine cleaned birthdates and rest of the data.
	franklin/clean_txt_mun.do	Clean municipal court records.
	franklin/make_id.do	Make a combined identifier to identify the same individual across Common Pleas and Municipal Court.
	franklin/combine_clean.do	Combine cleaned municipal and CP data. Construct instruments and define analysis sample.
	hamilton/ master_clean_hamilton.do	Runs all of the files to clean the Hamilton Common Pleas and Municipal courts data
	hamilton/clean_defendants.do	Takes the raw docket data and creates a file with the set of defendant characteristics in the case
	hamilton/clean_charges.do	Takes the raw docket data and creates a file with the set of charges against the defendant in the case
	hamilton/clean_outcomes.do	Takes the raw docket data and creates a file with the set of sentences against the defendant in the case
	hamilton/ create_clean_charges.do	Merges the data on charges and outcomes at the charge level to produce the outcome for each individual charge
	hamilton/ concord_ohio_charges.do	Matches the offense types in Hamilton to the rest of the county data

	hamilton/judge_assignment.do	Takes the raw docket data and creates a panel of the judges ever on the case. This also determines the randomly assigned judge on the case
	hamilton/clean_attorney.do	Takes the raw docket data and creates a dataset with the characteristics of the attorneys on the case.
	hamilton/clean_plea_fines.do	Takes the raw docket data and determines whether the case involved any pleas or fines
	hamilton/ create_geocoded_addresses.do	Takes the geocoded address for the defendant listed on the court filing and merges in ACS information about that neighborhood.
	hamilton/ create_probation_record.do	Uses the raw docket data to determine whether the defendant was incarcerated due to probation or parole violations
	hamilton/create_clean_cases.do	Merges all of the interim data files created by the previous do files to create the analysis data set for Hamilton county
	family/master_family.do	Runs all of the files to merge the interim data sets (e.g. courts, birth certificate) in order to construct the final data sets used in the analysis of family spillovers
	family/parent_instruments.do	Creates instruments that differ for parents and non-parents
	family/child_outcomes.do	Merges interim datasets in order to create the main analysis data for looking at the effect of parental incarceration on child criminal activity, pregnancy, and neighborhood of residence
	family/child_school.do	Merges interim datasets in order to create the main analysis data for looking at the effect of parental incarceration on child school outcomes
	family/make_panel_child.do	Merges interim datasets in order to create a panel data set for looking at the effect of parental incarceration on child outcomes
	family/sibling_outcomes.do	Merges interim datasets in order to create the main analysis data for looking at the effect of sibling incarceration on child criminal activity, pregnancy, and neighborhood of residence
	family/sibling_school.do	Merges interim datasets in order to create the main analysis data for looking at the effect of sibling incarceration on child school outcomes
	family/make_panel_sibling.do	Merges interim datasets in order to create a panel data set for looking at the effect of sibling incarceration on child outcomes
	family/spouse_outcomes.do	Merges interim datasets in order to create the main analysis data for looking at the effect of spousal incarceration on criminal activity, pregnancy, and neighborhood of residence. Note that for spouse we do not necessarily mean that the individuals were in a formal

		domestic partnership, but just that they shared a child (based on birth certificates).
	family/ make_panel_spouse.do	Merges interim datasets in order to create a panel data set for looking at the effect of spousal incarceration. Note that for spouse we do not necessarily mean that the individuals were in a formal domestic partnership, but just that they shared a child (based on birth certificates).
	family/helper/ construct_siblings.do	Creates the set of sibling matches for the analysis on sibling incarceration.
	family/helper/ prep_date_voting.do	Prepares the voter information to be merged in
	family/helper/ prep_voter_rolls.do	Prepares the neighborhood information (based on the voter rolls) to be merged in
	helper/ calc_prev_incar.do	Creates tempfile with previous incarceration history
	helper/ calc_while_incar.do	Creates tempfile with whether during the period of incarceration
	helper/ docket_helper.R	Set of functions to aid extracting docket information.
	helper/ make_pid.do	Creates individual specific unique IDs for defendants
	helper/ make_z.do	Creates standard leave-out judge instruments for incarceration
	helper/ make_z_ujive.do	Creates UJIVE instruments for incarceration
	helper/ make_z_ujive_1yrmoreless.do	Creates UJIVE instruments for incarceration to sentences of more than one year/less than one year
	helper/ make_z_ujive_6momoreless.do	Creates UJIVE instruments for incarceration to sentences of more than 6 months/less than 6 months
	helper/ make_z_ujive_intensive.do	Creates UJIVE instruments for length of incarceration
	matching/master_matching.do	Run all matching files
	matching/make_panel_court.do	Make panel dataset to study dynamic effects of incarceration
	matching/make_parent_indicator.do	Using defendant-birth record matches, make indicator of parents.
	matching/match_children_crime_cuyaJuv.do	Match children of Cuyahoga defendants to juvenile court records.
	matching/match_children_crime.do	Match children of defendants to adult court records to measure crime outcomes.
	matching/match_children_pregnancy.do	Match children of defendants to parents on birth records to measure teen parenthood.
	matching/match_children_school.do	Match children of Cuyahoga defendants to CMSD school records. Make Figure A9.
	matching/match_children_voter.do	Match children of defendants to voter records.
	matching/match_court_bc_parents.do	Match court data to parents on birth records. Create Figure A7.
	matching/match_court_bc.do	Match defendants to birth records.

	matching/match_court_death.do	Match defendants to SSA death records.
	matching/match_court_death_oh.do	Match defendants to Ohio death records.
	matching/match_court_voter.do	Match defendants to voter records.
	matching/match_death_bc.do	Match birth records to death records.
	matching/match_sibling_crime_cuyaJuv.do	Match siblings of defendants to Cuyahoga juvenile crime records.
	matching/match_sibling_school.do	Match siblings of defendants to CMSD school records.
	matching/match_sibling_voter.do	Match siblings of defendants to voter records.
	matching/match_siblings_pregnancy.do	Match siblings of defendants to parents on birth records to measure pregnancy.
	matching/match_siblings.do	Find siblings by finding individuals with same mother or father.
	matching/match_spouse_voter.do	Match co-parents of defendants (identified because on same birth record) to voter records.
	matching/name_rarity_bc_yr.do	Helper file used to estimate likelihood name is unique among Ohio births in a year.
	panel/combine_panels.do	Combines the panel data sets from all three counties
	parsing/parse_cuya_fugitive.py	Parses fugitive status in Cuyahoga common pleas data.
	parsing/parse_cuya_returned.py	Parses dates of return in Cuyahoga common pleas data.
	voterrolls/ master_voter.do	Runs all of the do-files related to the cleaning of the Ohio voter rolls data
	voterrolls/ geocode_ohio.do	Geocodes the voter rolls for Ohio for matching to census data on neighborhood characteristics
	voterrolls/ merge_ohio_census.do	This merges the Ohio voter rolls into the ACS data on characteristics of census block groups
	voterrolls/ vr_ohio.do	Cleans the voter rolls data in Ohio
	voterrolls/ vr_ohio_find_duplicates.do	Creates the voter rolls files with duplicate voters removed
	voterrolls/ vr_ohio_full_voter_history.do	Creates data set with all voting outcomes for all in the Ohio voter rolls
	voterrolls/auxiliary/ rename_noactive_varnames.do	Renames the variables in the voter rolls

Analysis do-files

This table describes the do-files that are used for the analysis in the paper.

	Code	Code description
	analysis/replication_analysis.do	Creates all of the exhibits in the paper. To run all of the exhibits, set RunAll equal to 1. To run an individual exhibit, set the local at the top of the do-file equal to one for the relevant exhibit
	analysis/helper/ load_analysis.do	This loads the set of court cases across all three counties that have been randomly assigned to judges
	analysis/helper/ load_sib_def_incar_yr_mo.do	This loads panel data for the sibling analysis.
	analysis/helper/ load_sibling_outcomes.do	This loads the outcome data for the sibling analysis.
	analysis/helper/ load_parent_indicator.do	Creates a tempfile with an indicator for whether the defendant is a parent that can be merged into the main data
	analysis/helper/ add_depmean.do	Adds dependent variable mean to the regression output
	analysis/helper/ load_child_outcomes.do	Creates the data set for analysis of parental incarceration on child outcomes.
	analysis/helper/ prep_child_school.do	Puts together data for analysis of parental incarceration on child school outcomes.
	analysis/helper/ prepSummary.do	Generates measures of offense severity
	analysis/helper/ stringify_coef_dollars.do	Auxiliary do-file for the cost-benefit analysis for formatting the output

Analysis Programs

This lists where the code producing each table/figure can be found in the analysis code.

Figure/Table #	Program	Line Number
Figure 1a	replication_analysis.do	63-124
Figure 1b	replication_analysis.do	63-124
Figure 2	replication_analysis.do	126-212
Figure 3a	replication_analysis.do	215-285
Figure 3b	replication_analysis.do	215-285
Figure 3c	replication_analysis.do	215-285
Figure 3d	replication_analysis.do	215-285
Figure 3e	replication_analysis.do	215-285
Figure 3f	replication_analysis.do	215-285
Figure 4a	replication_analysis.do	287-308
Figure 4b	replication_analysis.do	312-327
Figure 4c	replication_analysis.do	330-338
Figure 5a	replication_analysis.do	341-369, 374-438
Figure 5b	replication_analysis.do	341-369, 374-438
Figure 5c	replication_analysis.do	341-369, 374-438
Table 1	replication_analysis.do	441-595
Table 2	replication_analysis.do	595-751
Table 3	replication_analysis.do	756-943
Table 4	replication_analysis.do	946-1210
Table 5	replication_analysis.do	1213-1302
Table 6	replication_analysis.do	1305-1373
Table 7	replication_analysis.do	1378-1508
Appendix Figure A1a	replication_analysis.do	1515-1598

Appendix Figure A1b	replication_analysis.do	1515-1598
Appendix Figure A1c	replication_analysis.do	1515-1598
Appendix Figure A1d	replication_analysis.do	1515-1598
Appendix Figure A1e	replication_analysis.do	1515-1598
Appendix Figure A2	replication_analysis.do	1601-1632
Appendix Figure A3a	replication_analysis.do	1635-1709
Appendix Figure A3b	replication_analysis.do	1635-1709
Appendix Figure A3c	replication_analysis.do	1635-1709
Appendix Figure A4	replication_analysis.do	1712-1787
Appendix Figure A5	replication_analysis.do	1789-1996
Appendix Figure A6	not empirical	
Appendix Figure A7a	replication_analysis.do	1999-2160
Appendix Figure A7b	replication_analysis.do	1999-2160
Appendix Figure A8a	replication_analysis.do	2162-2200
Appendix Figure A8b	replication_analysis.do	2162-2200
Appendix Figure A8c	replication_analysis.do	2162-2200
Appendix Figure A9	replication_analysis.do	2203-2436
Appendix Figure A10	not empirical, provides an example of how extensivity might be violated	
Appendix Figure A11a	replication_analysis.do	2441-2636
Appendix Figure A11b	replication_analysis.do	2441-2636
Appendix Table A1	replication_analysis.do	2643-2793
Appendix Table A2	not empirical	
Appendix Table A3	replication_analysis.do	2796-2983
Appendix Table A4	replication_analysis.do	2985-3168
Appendix Table A5	replication_analysis.do	3170-3216

Appendix Table A6	replication_analysis.do	3218-3347
Appendix Table A7	replication_analysis.do	3349-4095
Appendix Table A8	replication_analysis.do	4098-4218
Appendix Table A9	replication_analysis.do	4220-4340
Appendix Table A10	replication_analysis.do	4220-4340
Appendix Table A11	replication_analysis.do	4343-4613
Appendix Table A12	replication_analysis.do	4615-4721
Appendix Table A13	replication_analysis.do	4724-4942
Appendix Table A14	replication_analysis.do	4944-5036
Appendix Table A15	replication_analysis.do	5037-5179
Appendix Table A16	replication_analysis.do	5181-5348
Appendix Table A17	replication_analysis.do	5350-5495
Appendix Table A18	replication_analysis.do	5496-5642
Appendix Table A19	replication_analysis.do	5643-5920
Appendix Table A20	replication_analysis.do	5921-6015
Appendix Table A21	replication_analysis.do	6016-6173
Appendix Table A22	replication_analysis.do	6174-6242
Appendix Table A23	replication_analysis.do	6244-6456
Appendix Table A24	replication_analysis.do	6458-6608
Appendix Table A25	replication_analysis.do	6610-6873
Appendix Table A26	replication_analysis.do	6875-7067
Appendix Table A27	replication_analysis.do	6875-7067
Appendix Table A28	not empirical, provides estimates of victim costs of crime from literature review of listed sources	