Analysis of U.S. immigration detention and enforcement dynamics

Phil Neff, Trung-Anh Nguyen, Maria Vignau-Loria, June Yang

University of Washington Center for Human Rights

February 17, 2023

CENTER FOR HUMAN RIGHTS

HENRY M. JACKSON SCHOOL OF INTERNATIONAL STUDIES

UNIVERSITY of WASHINGTON

Photo: La Resistencia

Context:

- The U.S. immigrant detention system has grown fivefold in the past two decades, and exponentially since 2017 (Ryo and Peacock 2018).
- Detention threatens the physical and mental health, economic welfare, and educational outcomes of migrants, their families, and their communities (Brabeck and Qingwen Xu 2010, Capps et al. 2015, Cohen 2020).
- Research suggests that due to perverse financial and political incentives, detention may drive a need for enforcement, rather than immigration enforcement driving a need for detention (Rubenstein and Gulasekaram 2019, Sanudo-Kretzmann 2018, Gilman and Romero 2018, Sinha 2016).

Substantive question:

How do changes in U.S. immigration detention capacity influence interior immigration enforcement in the surrounding region? What can we expect to happen to enforcement rates if a detention center closes, or expands?

Hypothesis:

An increase in available detention beds will be associated with increased arrests in the surrounding region. Conversely, decrease in available detention beds will be associated with decreased arrests.

Data:

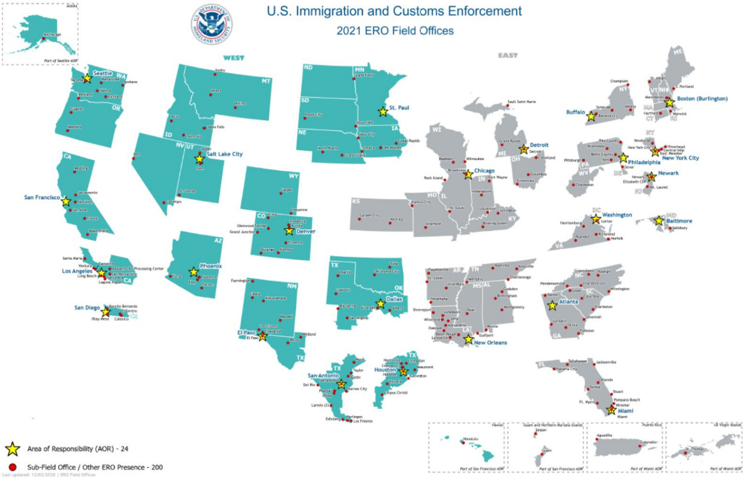
Nation-wide datasets of individual ICE enforcement actions, detention facility characteristics, individual detention placements per facility, and demographics for the time period from 2015-10-01 to 2019-06-30, transformed into monthly panel data.

Datasets:

1. ICE enforcement actions
2. Detention population ("headcount")
3. Detention facility characteristics
4. Border Patrol encounters
5. Demographic information

Geographic unit:

Data grouped by 24 U.S. Immigration and Customs Enforcement Areas of Responsibility (AORs)[1]



[1] ICE field offices exercise some degree of autonomy over detention and enforcement decisions in respective AORs.

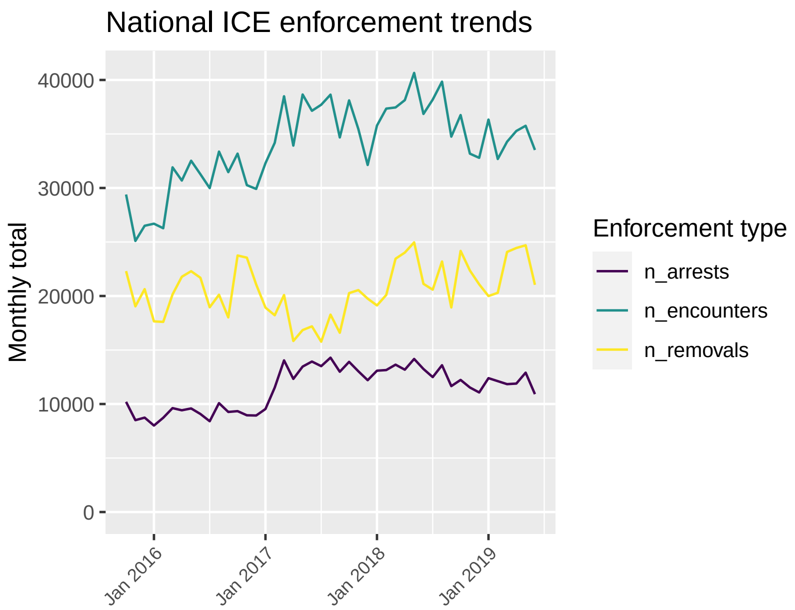
ICE enforcement:

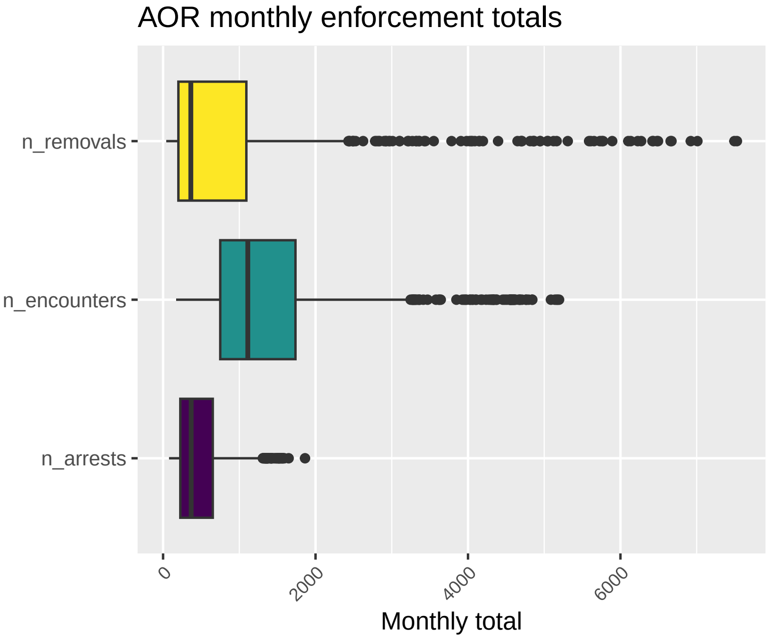
Individual ICE encounters, arrests, and removals from 2015-10-01 to 2019-06-30.

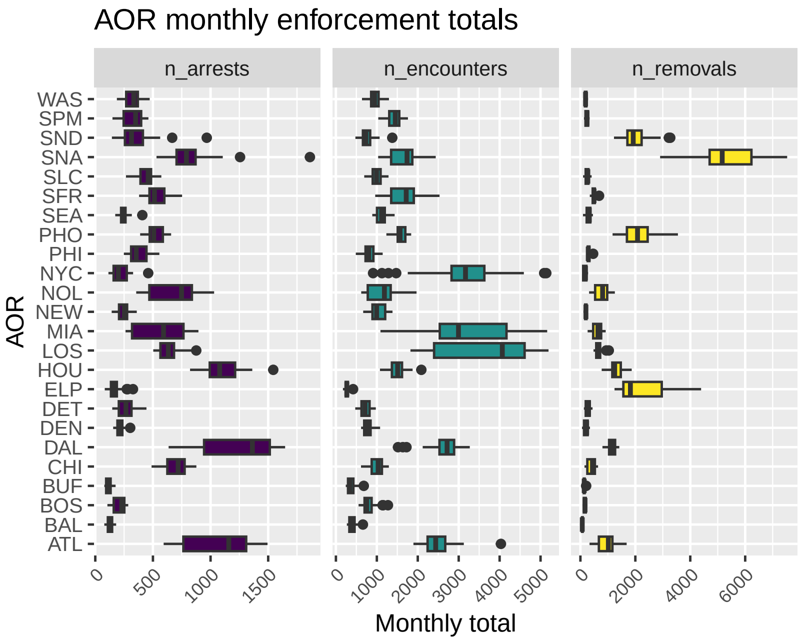
Problem! Lack precise arrest location—"Apprehension landmark" values often ambiguous, differ significantly between AORs, e.g.:

- "SNA GENERAL AREA, NON-SPECIFIC"
- "SEA CAP"

Source: UWCHR FOIA 2019


National ICE enforcement trends
Monthly total
40000
30000
20000
10000
0
Jan 2016
Jan 2017
Jan 2018
Jan 2019
Enforcement type
n_arrests
n_encounters
n_removals


AOR monthly enforcement totals
n_removals
n_encounters
n_arrests
0
2000
4000
6000
Monthly total

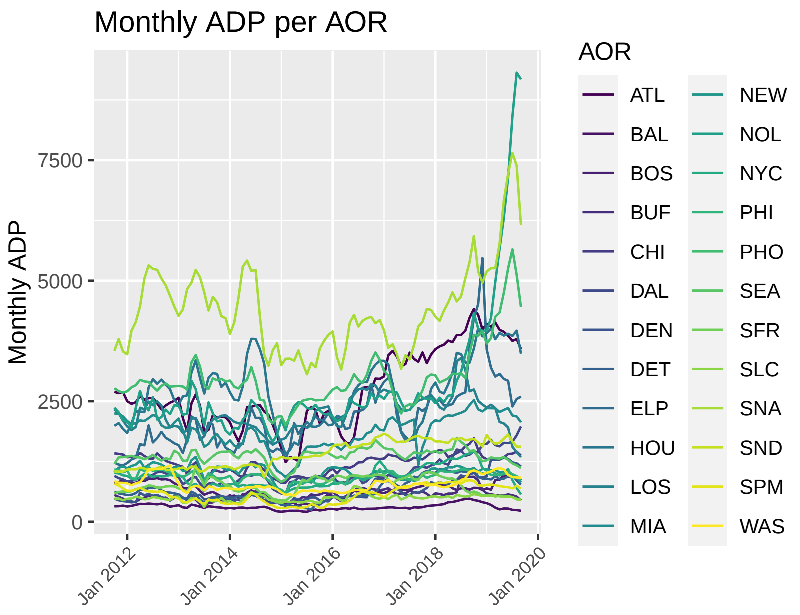

AOR monthly enforcement totals
n_arrests
n_encounters
n_removals
WAS
SPM
SND
SNA
SLC
SFR
SEA
PHO
PHI
NYC
NOL
NEW
MIA
LOS
HOU
ELP
DET
DEN
DAL
CHI
BUF
BOS
BAL
ATL
AOR
0
500
1000
1500
0
1000
2000
3000
4000
5000
0
2000
4000
6000
Monthly total

Detention placements:

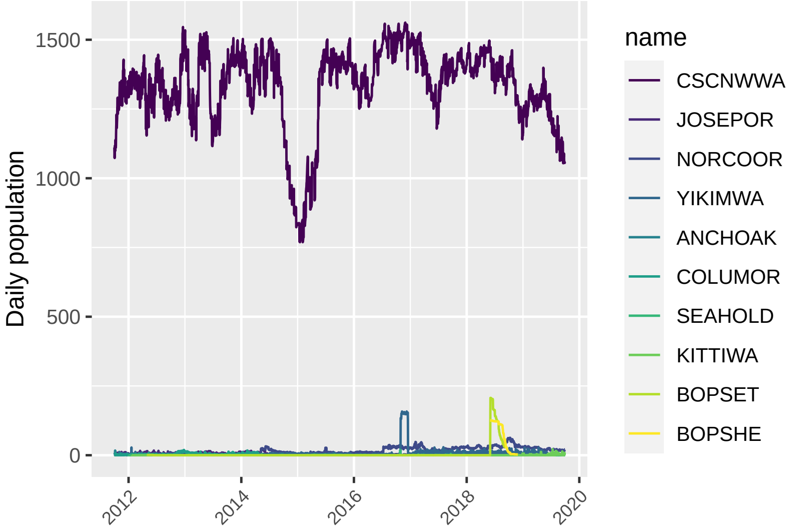
Individual detention placement records per person per facility from 2011-10-01 to 2019-09-30; used to generate midnight daily headcounts per facility, monthly average detained population (ADP).

Problem! Lack individual identifiers, duplication of records year to year and between installments. Lack identification of arresting agency (i.e. ICE, CBP).

Source: Hausmann, ACLU of CA FOIA 2020


Monthly ADP per AOR
Monthly ADP
7500
5000
2500
0
Jan 2012
Jan 2014
Jan 2016
Jan 2018
Jan 2020
AOR
ATL
BAL
BOS
BUF
CHI
DAL
DEN
DET
ELP
HOU
LOS
MIA
NEW
NOL
NYC
PHI
PHO
SEA
SFR
SLC
SNA
SND
SPM
WAS

SEA daily detained population*



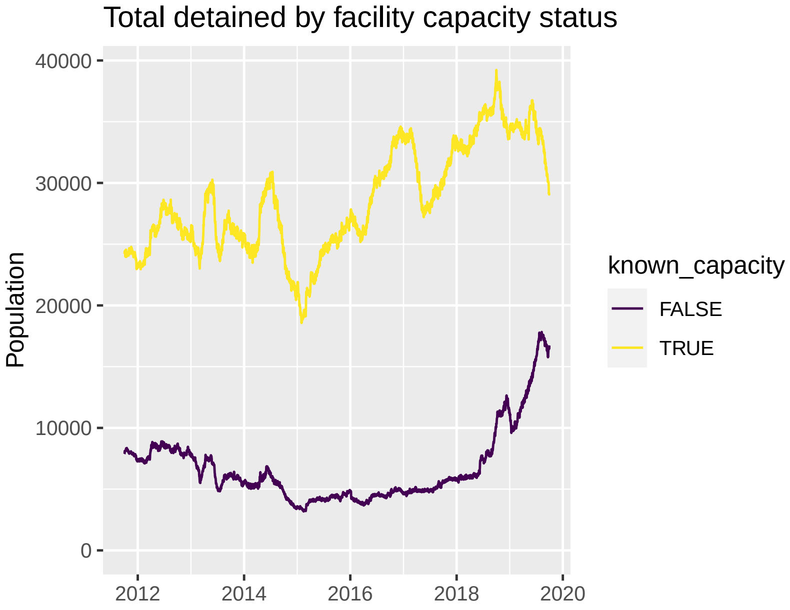
Facilities with historic max population >= 10

Detention facility characteristics:

Characteristics of interest including capacity, guaranteed-minimum beds, over/under-72 hour stays, contract dates for 782 ICE detention facilities.

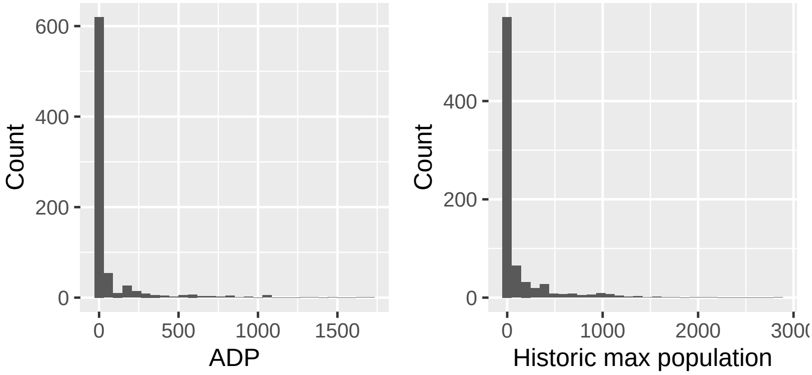
Problem! Many facilities report capacity "AS NEEDED"; missing data for new facilities since late 2017.

Source: NIJC FOIA 2017, UWCHR



Total detained by facility capacity status
Population
40000
30000
20000
10000
0
2012
2014
2016
2018
2020
known_capacity
FALSE
TRUE

Most ICE detention facilities have low ADP and historic maximum population:



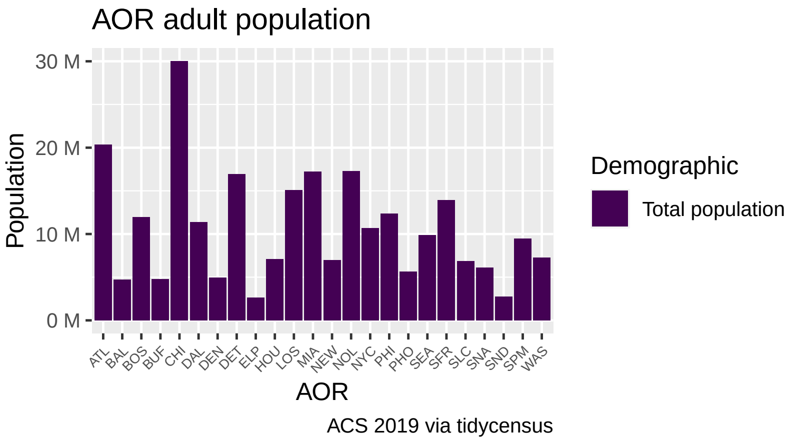
Demographic information:

Census estimates of total adult population, foreign-born population, non-citizen population, and potentially undocumented population; used to calculate immigration enforcement rates per capita.

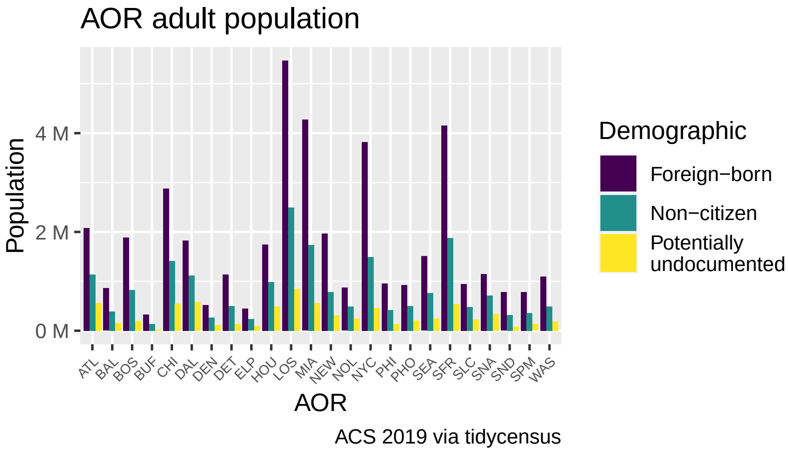
For estimation of adult potentially undocumented population, we use the logical imputation method (see Hall, Greenman, and Farkas 2010; Borjas 2017; Amuedo-Dorantes, Arenas-Arroyo, and Sevilla 2018).

- Adult U.S. potentially undocumented population (2019): $7.41\text{M} (95\% C.I. : 7.2\text{M}, 7.62\text{M})$

Source: ACS via `tidycensus`



AOR adult population
30 M
20 M
10 M
0 M
Population
ATL
BAL
BOS
BUF
CHI
DAL
DEN
DET
ELP
HOU
LOS
MIA
NEW
NOL
NYC
PHI
PHO
SEA
SFR
SLC
SNA
SND
SPM
WAS
AOR
Demographic
Total population
ACS 2019 via tidycensus

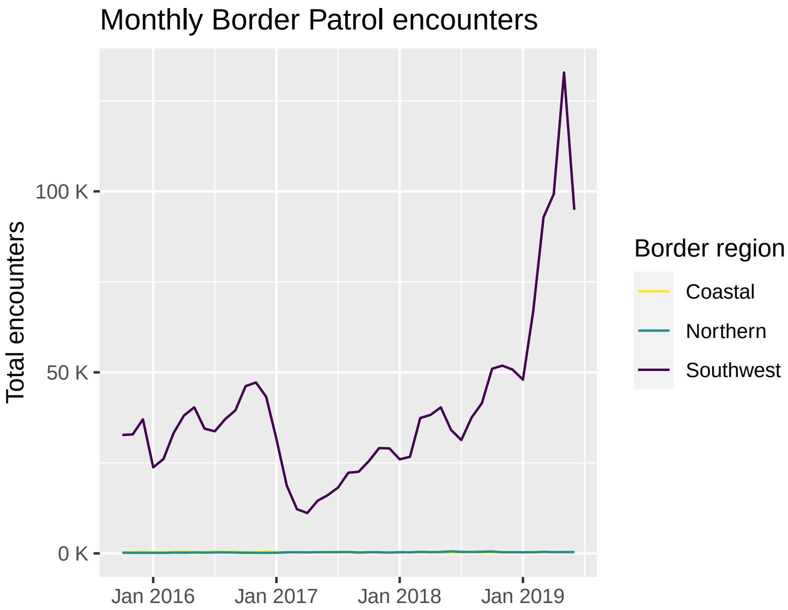

AOR adult population
Population
4 M
2 M
0 M
ATL
BAL
BOS
BUF
CHI
DAL
DEN
DET
ELP
HOU
LOS
MIA
NEW
NOL
NYC
PHI
PHO
SEA
SFR
SLC
SNA
SND
SPM
WAS
AOR
Demographic
Foreign−born
Non−citizen
Potentially undocumented
ACS 2019 via tidycensus

Border Patrol encounters:

Monthly encounters per Border Patrol Sector, FY2000-FY2020. Important variable because major driver of detained population.

Problem! No direct mapping of Border Patrol Sectors to ICE AORs; aggregate statistics rather than individual encounters data; encounters $\neq$ apprehensions.

Source: U.S. Border Patrol


Monthly Border Patrol encounters
Total encounters
100 K
50 K
0 K
Jan 2016
Jan 2017
Jan 2018
Jan 2019
Border region
Coastal
Northern
Southwest

How to calculate "capacity"?

Only a minority of detention facilities used by ICE have a specific reported capacity; however, these facilities hold the majority of the detained population. Capacity changes over time are not recorded.

Most facilities are used by ICE on an "AS NEEDED" basis. Many facilities stay "on the books" but are rarely or never used. How to measure capacity of these facilities?

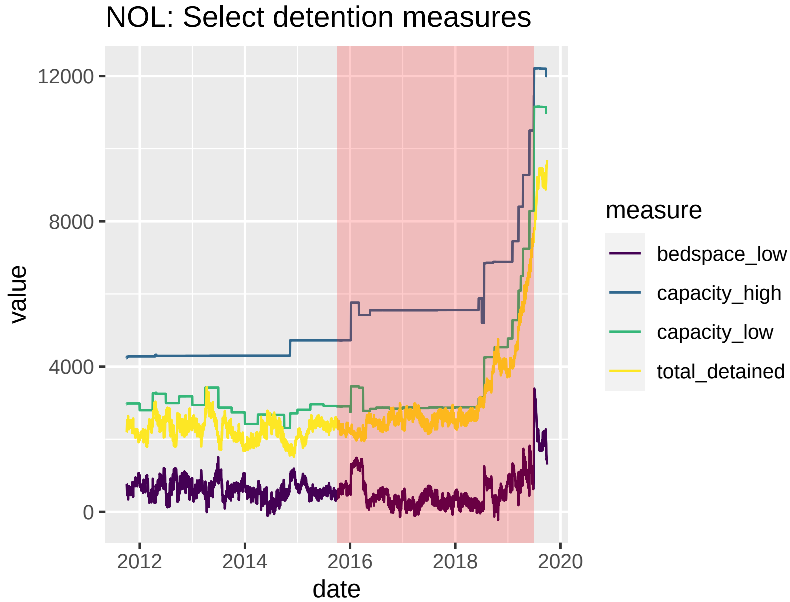
How to calculate “capacity”?

Our approach—Use observed detained population to estimate available capacity of active facilities:

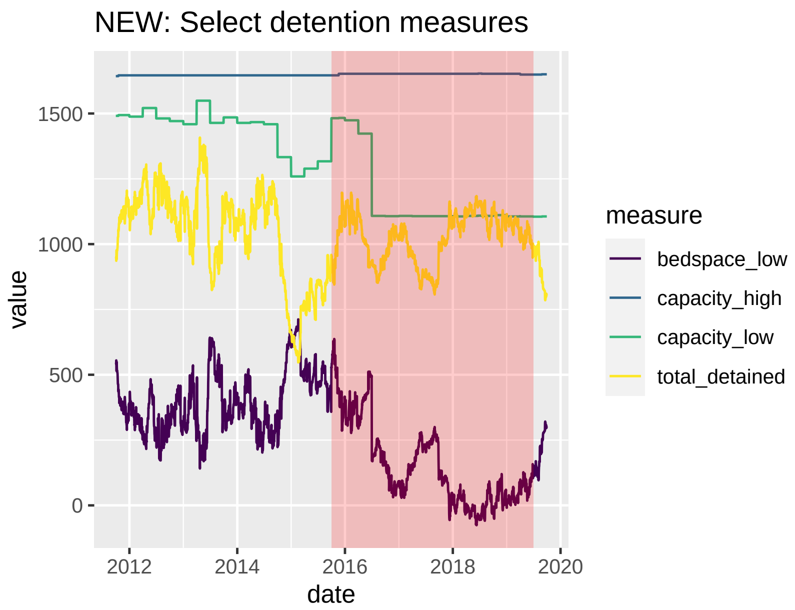
1. `capacity_low`: Lower-bound estimate based on reported ICE capacity (where known) plus quarterly average detained population of “AS NEEDED”/unknown capacity facilities.
2. `capacity_high`: Higher-bound estimate based on historic maximum detained population of all facilities; primarily used as indicator of facility change.

How to calculate "available detention beds"?

"Bedspace": Estimated capacity minus detained population, either expressed as estimate of number of available beds (`bedspace_low`) or percent bed utilization (`bedspace_low_util_pct`).



Date range of enforcement data highlighted.



Date range of enforcement data highlighted.

Preliminary modeling:

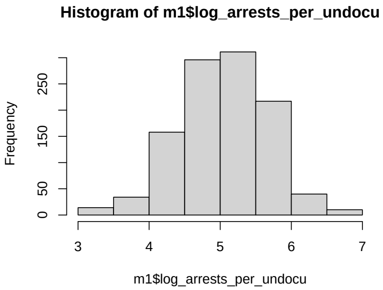
Goals:

- Identify appropriate statistical approaches
- Assess preliminary TWFE and DiD results
- Brainstorm next steps

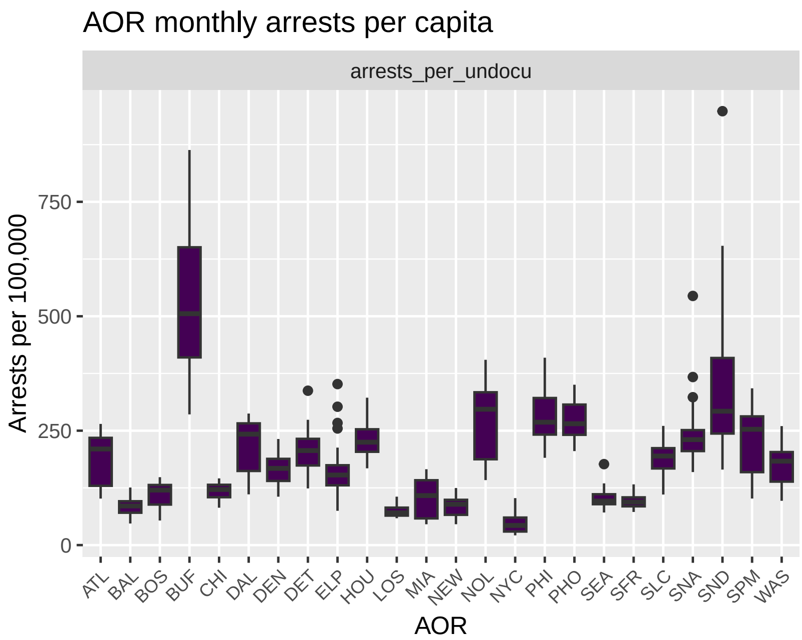
Dependent variable:

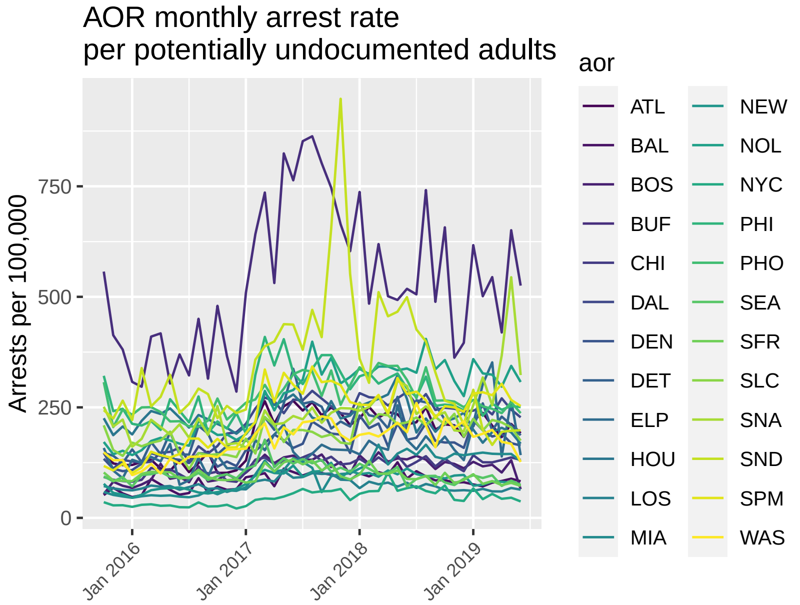
- arrests_per_undocu: Arrests per 100,000 estimated adult undocumented population[2]

Log transformed; we include a lagged term in models to account for auto-correlation (AR1) where appropriate.



[2] Other dependent variables considered: arrests_per_capita, arrests_per_forborn, arrests_per_noncit


AOR monthly arrests per capita
arrests_per_undocu
Arrests per 100,000
0
250
500
750
AOR
ATL
BAL
BOS
BUF
CHI
DAL
DEN
DET
ELP
HOU
LOS
MIA
NEW
NOL
NYC
PHI
PHO
SEA
SFR
SLC
SNA
SND
SPM
WAS



AOR monthly arrest rate
per potentially undocumented adults
Arrests per 100,000
0
250
500
750
Jan 2016
Jan 2017
Jan 2018
Jan 2019
aor
ATL
BAL
BOS
BUF
CHI
DAL
DEN
DET
ELP
HOU
LOS
MIA
NEW
NOL
NYC
PHI
PHO
SEA
SFR
SLC
SNA
SND
SPM
WAS

Independent variables:

One of four measures of detention population and capacity:[3]

- `avg_detained`: Monthly average daily population (ADP)
- `capacity_low`: Estimated detention capacity
- `bedspace_low`: Bedspace—capacity minus monthly ADP
- `bedspace_low_util_pct`: Bedspace utilization (%)

[3] Variables with suffix `_low` based on lower-bound capacity estimate.

Table: Data structure (selected fields)

year_mth	aor	log_arrests_per_undocu	log_arrests_per_undocu.lag1	avg_detained
Oct 2015	ATL	4.992		2040
Nov 2015	ATL	4.814	4.992	2238.767
Dec 2015	ATL	4.783	4.814	2302.71
Jan 2016	ATL	4.786	4.783	2092.548
...	...	...	...	...
Jun 2019	ATL	5.256	5.358	3856.7
Oct 2015	BAL	3.942		233.613
Nov 2015	BAL	4.221	3.942	246.9
Dec 2015	BAL	4.014	4.221	265.935
Jan 2016	BAL	3.853	4.014	266.323
...	...	...	...	...
Jun 2019	BAL	4.417	4.484	272.833
Oct 2015	BOS	3.983		360.774
Nov 2015	BOS	4.404	3.983	502.767
Dec 2015	BOS	4.285	4.404	569.323
Jan 2016	BOS	4.214	4.285	536.29
...	...	...	...	...

Unit: `aor` ($N = 24$); Period: `year_mth` ($T = 45$)[4]

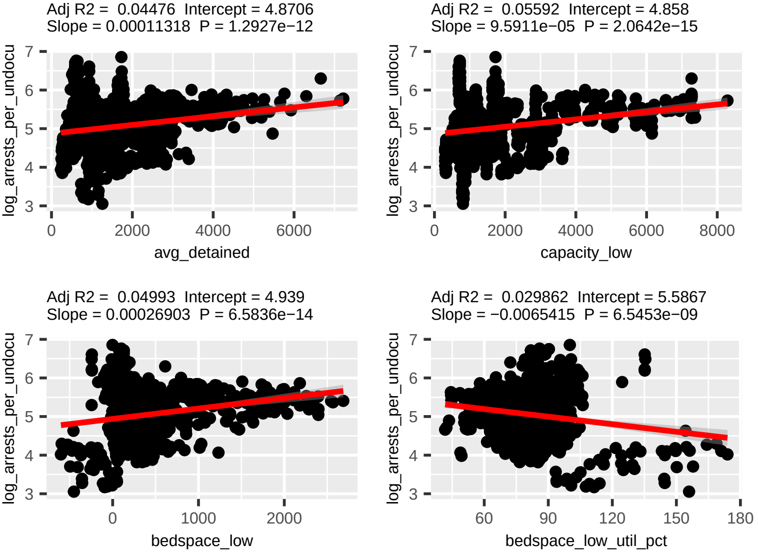
[4] When using lagged terms, $T = 44$

Table: Select variables of interest

Statistic	N	Mean	St. Dev.	Min	Max
n_arrests	1,080	478.440	341.627	78	1,862
arrests_per_undocu	1,080	187.627	120.656	21.223	948.049
log_arrests_per_undocu	1,080	5.052	0.617	3.055	6.854
avg_detained	1,080	1,600.460	1,164.230	233.613	7,217.833
capacity_low	1,080	2,019.467	1,532.510	302.000	8,289.133
bedspace_low	1,080	419.007	516.776	−602.700	2,687.419
bedspace_low_util_pct	1,080	81.782	16.536	41.687	173.933

Simple regressions:

Problem! These OLS regressions do not account for auto-correlation!



TWFE:

We first attempt TWFE regressions to estimate causal effects adjusting for AOR-specific and time-specific confounders.

$$y_{it} = x_{it}\beta_1 + y_{i(t-1)}\beta_2 + c_i + f_t + u_{it}$$

$i = \texttt{aor}$ $(1...N)$, $t = \texttt{year\_mth}$ $(1...T\text{-}1)$

Example specification:

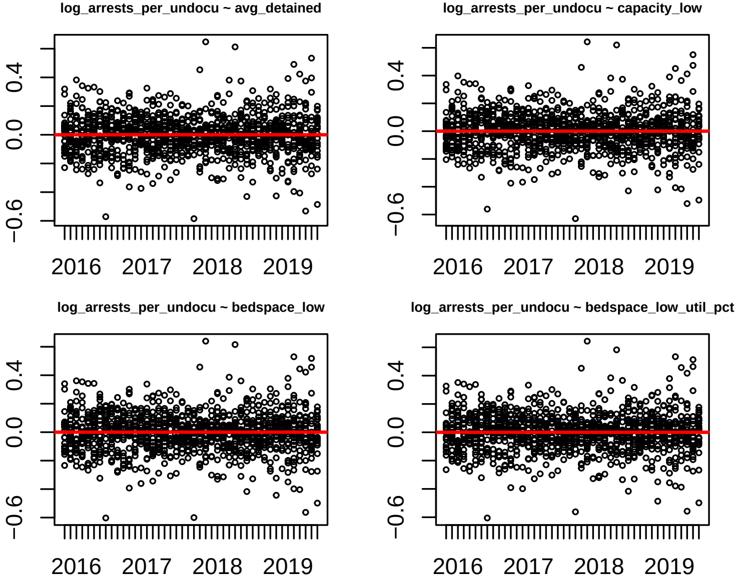
```
plm1 <- plm(log_arrests_per_undocu ~
            avg_detained +
            log_arrests_per_undocu.lag1,
            data = m1, model="within", effect = "twoways")
```

TWFE results:

Table: TWFE results

	Dependent variable:			
	log_arrests_per_undocu			
	(1)	(2)	(3)	(4)
avg_detained	0.0001***			
	(0.00001)			
capacity_low		0.00004***		
		(0.00001)		
bedspace_low			−0.00003	
			(0.00002)	
bedspace_low_util_pct				0.002***
				(0.001)
log_arrests_per_undocu.lag1	0.560***	0.578***	0.585***	0.567***
	(0.026)	(0.026)	(0.026)	(0.026)
Observations	1,056	1,056	1,056	1,056
R^2	0.367	0.360	0.354	0.361
Adjusted R^2	0.324	0.316	0.310	0.317
F Statistic (df = 2; 987)	286.478***	278.183***	270.601***	279.011***
Note:			*p<0.1; **p<0.05; ***p<0.01	

TWFE residuals over time:



TWFE discussion:

- `bedspace_low` not significant.
- Direction of `bedspace_low_util_pct` does not agree with our hypothesis, and big SE.
- Border Patrol encounter controls that do not vary across AOR dropped in `plm` call

Recent criticism of TWFE assumptions: (Kropko and Kubinec 2018)

Difference-in-Differences:

Treatment definition: Earliest significant facility change affecting relative estimated AOR capacity—facility opening or closure; compared to AOR(s) without significant relative permanent facility change.

We expect a positive relationship between treatment and outcome: i.e. capacity increase with increased rate of enforcement; capacity decrease with decreased rate of enforcement.

Need to do more pre-testing for parallel trends; qualitatively verify treatments by researching specific facility case studies.

DiD treatment categories:

Reference:

- MIA
- PHI
- SEA
- SLC
- SNA
- SND
- SPM

Increase:

- ATL
- BAL
- BOS
- CHI
- DAL
- DEN
- DET
- ELP
- HOU
- LOS
- NOL
- PHO
- WAS

Decrease:

- NEW
- SFR

Exclude:

- BUF
- NYC

2×2 DiD results:

- Pairing subjective: We attempt to pair treated AORs with roughly comparable control AORs; e.g. by regions.
- Results inconsistent; approx 50% significant.
- 58% of increase treatment effects in hypothesized direction.
- 67% of decrease treatment effects in hypothesized direction.

Table: Pairwise DiD Model Coefficients for Treatments of Increased Capacity

model_names	coefficients	pvalue	ispositive	issignificant
ATL_PHI	31.845	0.038	1	1
ATL_MIA	30.873	0.001	1	1
BAL_MIA	-51.404	0.000	0	1
CHI_SPM	-98.275	0.001	0	1
DAL_PHO	53.951	0.000	1	1
DAL_SNA	54.533	0.008	1	1
DET_SPM	-59.574	0.000	0	1
LOS_SND	-98.601	0.019	0	1
HOU_SNA	-110.079	0.000	0	1
PHO_SNA	-80.039	0.002	0	1
BAL_PHI	6.216	0.727	1	0
BOS_PHI	21.151	0.512	1	0
DEN_SLC	1.999	0.930	1	0
DEN_SPM	-33.372	0.363	0	0
ELP_SNA	3.766	0.875	1	0
LOS_SEA	6.302	0.379	1	0
NOL_MIA	26.323	0.339	1	0
WAS_MIA	-24.454	0.062	0	0
WAS_PHI	23.998	0.226	1	0

Table: Pairwise DiD Model Coefficients for Treatments of Decreased Capacity

model_names	coefficients	pvalue	ispositive	issignificant
SFR_SND	-86.944	0.042	0	1
NEW_PHI	-17.259	0.434	0	0
SFR_SEA	2.606	0.735	1	0

DiD with multiple time periods:

Problems! Staggered treatments; AORs can have multiple treatments, multiple directions of treatment, pre-treatment.

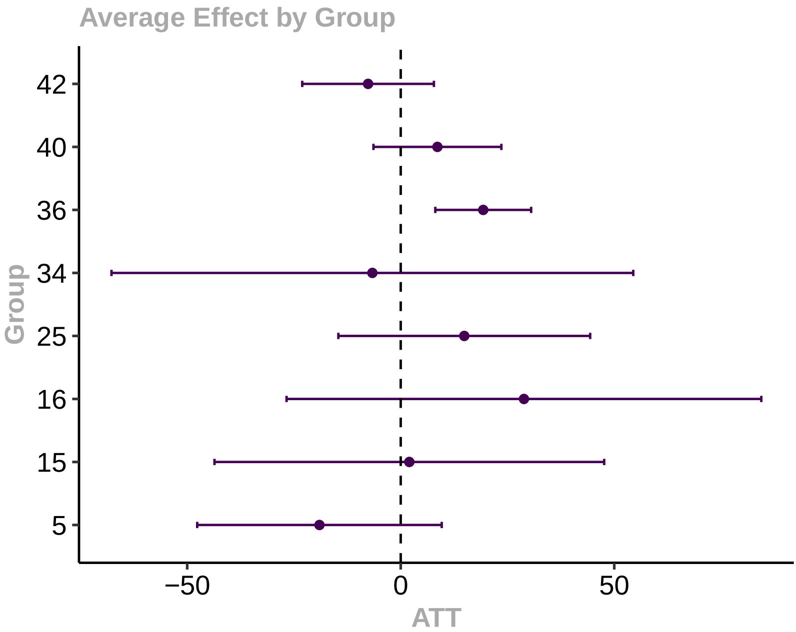
`did`: Callaway and Sant'Anna 2021

Causal parameter: $ATT(g, t)$ group-time average treatment effect

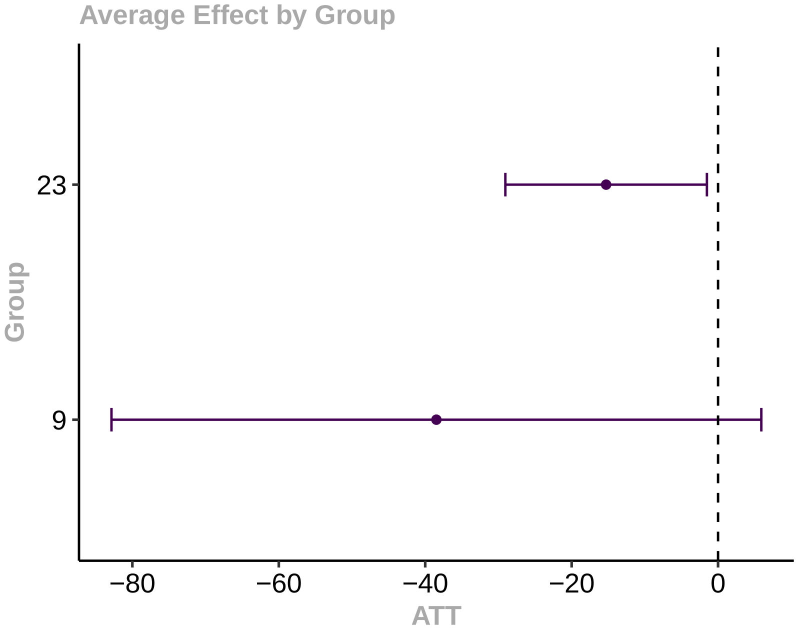
Model specification:

- Exclude "pre-treated" AORs: LOS, NYC, SLC, SNA, SND, WAS
- Compare treated groups by time period against "not-yet-treated" groups
- Allow 1 month of anticipation effect

Increase treatment groups:



Decrease treatment groups:



Next steps:

- Working towards open source publication of data and code via GitHub.
- Data subsets and covariates: GM beds, arrest categories, etc.
- Alternative treatments: increase/decrease GM beds.
- Other methods to consider: TWFE/DiD methods for continuous treatments; cross-lagged panel model with FE (see Leszczensky and Wolbring 2022).

Thank you!

University of Washington Center for Human Rights

Contact: philneff@uw.edu