Financial Econometrics –
Cross Section and Panel Data

Discussion of "Opioid Crisis and Real Estate Prices" by Custodio et al. (WP 2024)

Andreas Fuster
Swiss Finance Institute @ EPFL

Reminder – questions for discussion

- What is/are the economic question(s) the paper is trying to answer? What is the paper's "unique selling point" (USP), i.e. how does it move the literature forward?
- What is the empirical approach? Potential endogeneity issues & how does the paper address them?
- Data used & main results? Economic interpretation?
- What do you like about the paper?
- What could be improved / wasn't clear to you?

Try to link in particular to things we discussed in the lectures. Also think about the way results are communicated (tables/figures/writing).

Economic question and the paper's USP

- How does more widespread usage of opioids (fentanyl etc.) – the "opioid epidemic" in the US – affect real estate values?
  — and through which channels?

- USP: use of various identification strategies based on state laws that restricted opioid prescriptions (and therefore usage)
  — staggered DiD; RDD
  — robustness: IV analysis using different variation in prescriptions
- But note: other papers have done similar things (see p. 6)

- All public data (I think)

Setting and data

- County-level panel over 2006-2018 of
  - Opioid prescription rates (prescriptions per 100 people)
  - Zillow home value index, 2019 version
    - entire history of these indices gets updated over time!
  - Various demographics from Census (on population composition in terms of age, race/ethnicity, poverty, unemployment, …)
  - Data on doctors per capita and payments from pharma firms to doctors and hospitals (in particular opioid-related ones)

- State–level passage of laws that limit prescriptions
  - 32 states in total; 9 in 2016, 18 in 2017, 5 in 2018

Analysis, part 1: correlation between home values and prescription rates

or state X year FE

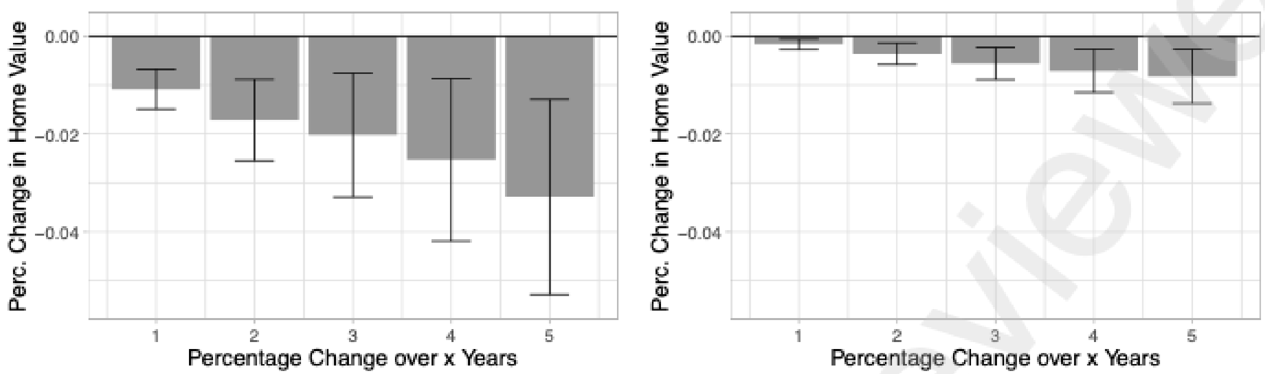
$$PCHomeValue_{c,t-x\ to\ t} = \alpha + \beta PrescriptionRate_{c,t-x} + \gamma Controls_{c,t-x} + \boxed{\theta_c + \tau_t} + \epsilon_{ct} \quad (1)$$

The dependent variable $PCHomeValue_{c,t-x\ to\ t}$ in equation 1 is the log percentage change of average county c home values, $(log(HV_t/HV_{t-x}) * 100)$ over $X = \{1,2,3,4,5\}$ years. $PrescriptionRate_{c,t-x}$ captures county c prescription rate at $t-x$. We also include a vector of time-varying county-level controls $Controls_{c,t-x}$, measured with a lag at time $t-x$.

FIGURE 2: HOME VALUE AND OPIOID PRESCRIPTION RATE: CORRELATIONS

(A) COUNTY & YEAR FIXED EFFECTS

(B) STATE-YEAR FIXED EFFECTS



	over 1 year	over 2 years	over 3 years	over 4 years	over 5 years
x-year lagged Prescription Rate	-0.011*** (0.002)	-0.017*** (0.004)	-0.020*** (0.006)	-0.025*** (0.008)	-0.033*** (0.010)
Adj. R2	0.599	0.648	0.699	0.735	0.767
No. of Obs	25,026	22,510	19,993	17,487	15,075
SD Pres. Rate	41.203	41.670	42.136	42.459	42.620

	over 1 year	over 2 years	over 3 years	over 4 years	over 5 years
x-year lagged Prescription Rate	-0.002*** (0.001)	-0.004*** (0.001)	-0.006*** (0.002)	-0.007*** (0.002)	-0.008*** (0.003)
Adj. R2	0.759	0.806	0.838	0.854	0.854
No. of Obs	25081	22572	20077	17590	15157
SD Pres. Rate	41.247	41.714	42.216	42.563	42.640

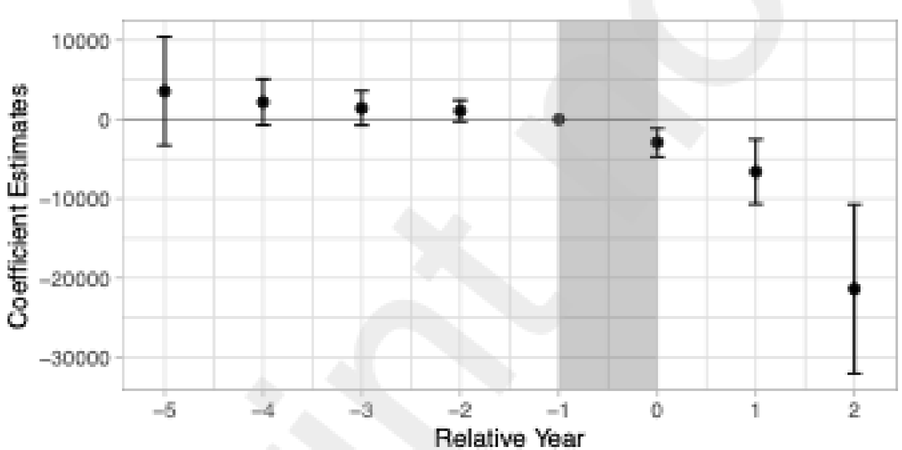
- Interpretation of (A) vs. (B)? Magnitudes?
- Why doesn't the paper end here?

Analysis, part 2: staggered DiD based on adoption of opioid-limiting state laws

- First outcome: prescription rates (to see if laws work)
- Use Sun and Abraham (2021) approach – described in detail on pp. 12/13
  - — why not just use TWFE?

$$PrescriptionRate_{c,t} = \alpha + \sum_{e \in \{16,17,18\}} \sum_{l=-5,\neq -1}^{2} \delta_{e,l} \mathbf{1}\{E_i = e\} D_{ct}^{\ell} + \gamma Controls_{c,t-1} + \theta_c + \tau_t + \epsilon_{c,t}$$

(A) COUNTY PRESCRIPTIONS

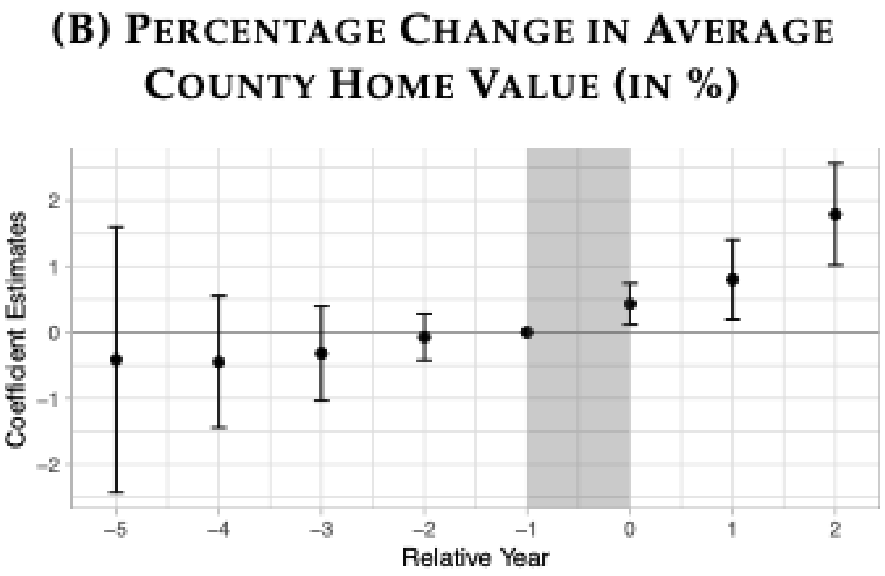


- How is the +2 effect identified (which data)?
- Interpretation of magnitude?
- Clustering?

Very unclear to me why they end in 2018; should add at least 2019 (while 2020 may be Covid-affected)

Analysis, part 2: staggered DiD based on adoption of opioid-limiting state laws

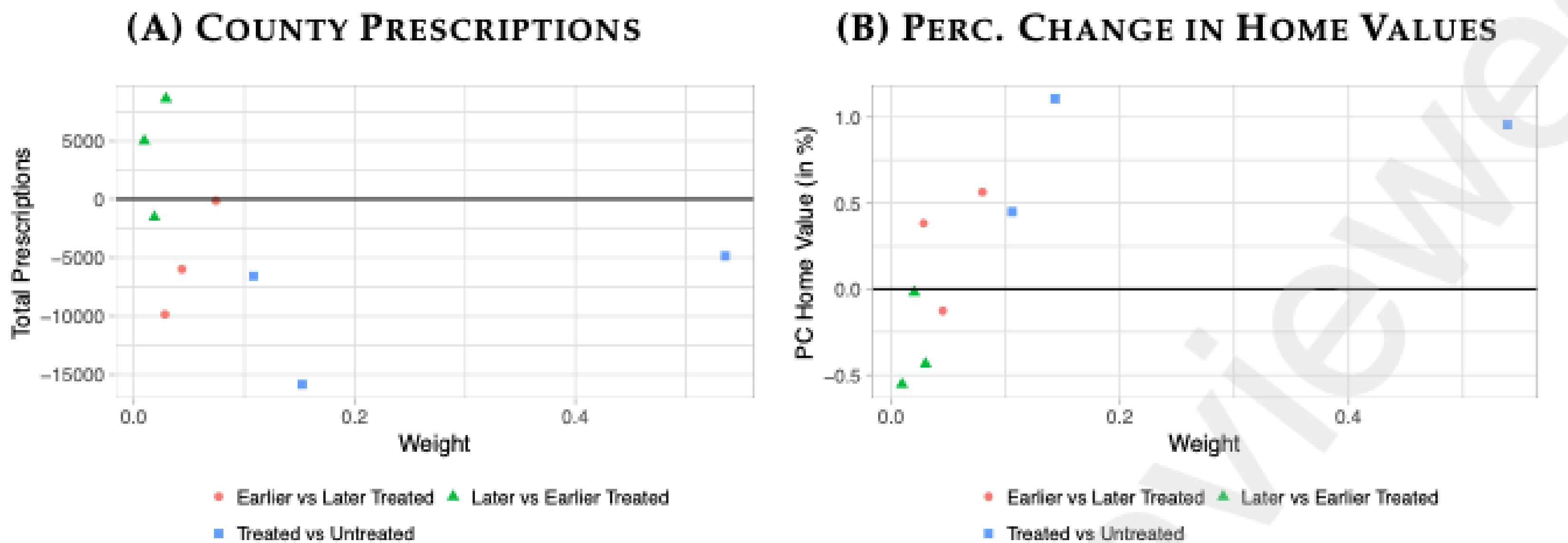
- Second outcome: one-year home-value log change
  — same Sun-Abraham method as before



- Interpretation?
- Identifying assumption? What are concerns about this?
- Table A.I: adoption non-random, but “only” related to overdose death rate, not econ conditions
- Figure A.III: using Roth (2022) to construct hypothesized trend such that prob. of testing pre-trend test is 50% (see p.14)

Analysis, part 2: staggered DiD based on adoption of opioid-limiting state laws

FIGURE 4: GOODMAN-BACON DECOMPOSITION



Notes: The sample period is 2013 to 2018. The dependent variable is *Prescriptions* in Panel A and *PCHomeValue* in Panel B. We show the Goodman-Bacon (2021) decompositions for the TWFE regression $Dep.\ variable_{ct} = \alpha + \beta Post_{ct} + \theta_c + \tau_t + \epsilon_{ct}$. We do not include any controls in the regression.

- Why 9 dots in each chart? Interpretation?
- Why are they showing this?

Analysis, part 2: staggered DiD based on adoption of opioid-limiting state laws

- County-level evidence: now use TWFE set-up but interacting the Post dummy with indicators for high-opioid-supply
  - — top tercile of #Drs per capita or pharma company payments in 2011-2015

TABLE II: OPIOID SUPPLY PROPENSITY INTERACTION

	(1)	(2)	(3)	(4)	(5)	(6)
	Prescription Rate			Percentage Change Home Prices		
Post	−2.533*	−1.658	0.261	0.731**	0.673**	0.569*
	(1.310)	(1.423)	(1.376)	(0.319)	(0.317)	(0.319)
Post X Physicians per capita Tercile 3		−2.266*			0.148	
		(1.204)			(0.185)	
Post X Phys. Opioid Payment Rate Tercile 3			−5.998***			0.346**
			(1.415)			(0.160)
R2	0.950	0.950	0.950	0.589	0.590	0.590
N	15199	15199	15199	14695	14695	14695

- What's the point of these regressions? Are results convincing?

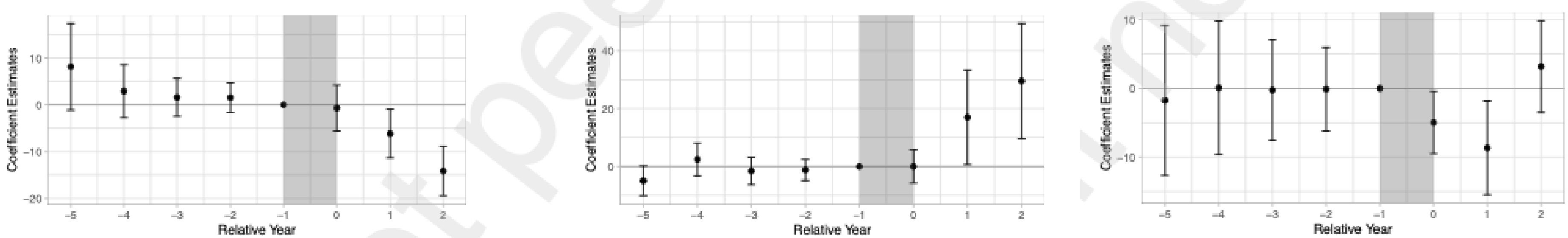
Economic mechanisms

- Use additional data on mortgage delinquencies, number of home improvement loans, and property vacancy rates
- Table 3: regress 5-year relative (percentage) changes on 5-year-lagged prescription rates – takeaways/limitations?

	(1)	(2)	(3)	(4)	(5)	(6)
	5-year percentage change in					
	Mortgage delinquency rate		No. of home improvement loans		Residential vacancy rate	
Lag Prescription Rate	0.837***	0.192***	−0.175***	−0.024**	0.267***	0.062***
	(0.188)	(0.047)	(0.045)	(0.010)	(0.047)	(0.022)
R2	0.904	0.901	0.672	0.661	0.758	0.337
N	2350	2320	14721	14794	9488	9589
Std. dev. prescription rate	27.11	27.08	43.29	43.34	43.55	43.63
County F.E.	Yes	No	Yes	No	Yes	No
Year F.E.	Yes	No	Yes	No	Yes	No
State-Year F.E.	No	Yes	No	Yes	No	Yes

Economic mechanisms

- Use additional data on mortgage delinquencies, number of home improvement loans, and property vacancy rates
- Figure 5: use same Sun-Abraham estimator as before based on introduction of prescription-limiting laws



(A) Mortgages 90 plus days past due (B) Home improvement loans (C) Residential vacancy rate

In all panels, *y* is the log percentage *change* in a variable – not clear to me that shouldn't just be the level (A and C) or log-level (B).

E.g. as is, in panel A, a change in delinquency rate from 2% to 1% is treated the same as a change from 0.4% to 0.2% -> reasonable?

Economic mechanisms

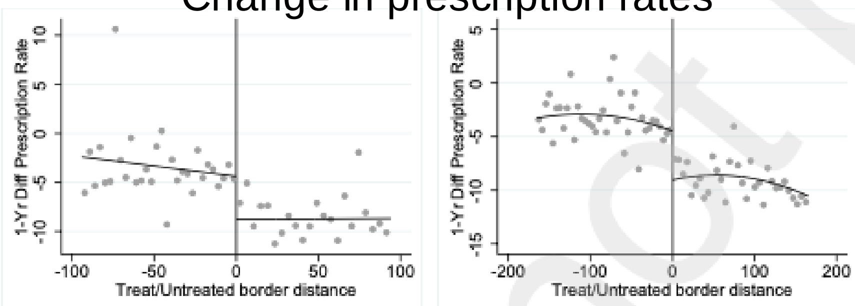
- They then do something similar for county level “migration” (inflows/outflows in terms of #households or individuals, and total income) – see Table 4 & Figure 6
  — here, they report results for different transformations of dependent variable & using opioid overdose death rates as additional independent variable… Will skip the details.

- Section 5.1 discusses interpretations of the results
- 5.2 tries to translate estimates into aggregate effects
  — acknowledging that they can’t assess general eqm effects
- 5.3: limits to internal and external validity

Robustness: state-border RDD

- Idea: use counties close to state borders (where one state limits prescriptions while the other state doesn't)

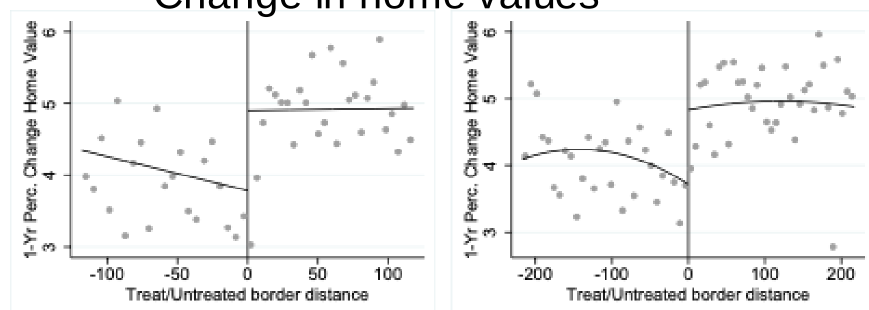
Change in prescription rates



(A.1) 1-YR DIFFERENCE & LINEAR POLYNOMIAL

(A.2) 1-YR DIFFERENCE & QUADRATIC POLYNOMIAL

Change in home values



(A.1) 1-YR PERC. CHANGE & LINEAR POLYNOMIAL

(A.2) 1-YR PERC. CHANGE & QUADRATIC POLYNOMIAL

- Table 5: effect on prescription rates sign. at $p<0.01$; house price growth "only" $p<0.1$ (a bit surprising?)
  - — using Calonico et al (2014) optimal bandwidth
- What are the assumptions here? Do the authors provide support for these assumptions? Are you convinced?

Robustness: instrumental variables

- Finally, in addition to DiD and RDD, they also use IV!

- Instruments from Cornaggia et al. (2022)
  - — Purdue marketing of OxyContin in 1997-2003
  - — "Leaky" supply chains and addictiveness of products

- Results in Tables 6 and 7 – first stage appears strong in both cases (although F-stats in Table 6 too high?), but 2SLS estimate at most marginally significant

- Other comments on execution of this analysis? Which assumptions have to hold for the IV estimates to be "valid"?

What I liked / what could be improved

- Important policy question, solid (public) data effort, results nicely presented
- Good setting to see recent methods for staggered DiD in action (although not that much staggering, and post-period short)
  — also the RDD robustness check (although could do more there)

- Room for improvement: I think paper is doing "too much" – Should focus more on comparing different estimates and making sure they are sensible
- Some choices questionable/ad-hoc (e.g. ending sample in 2018; using # prescriptions in Fig 3(A); the 4y- and 5y-changes used in IV tables)

- No discussion at all of treating all counties as equally important in analysis – see next slides

US counties are EXTREMELY heterogenous in size – part 1

	Total Population	# counties (or equivalent)	Average Pop
Texas	28,995,881	254	114,157
Georgia	10,617,423	159	66,776
Virginia	8,535,519	133	64,177
Kentucky	4,467,673	120	37,231
Missouri	6,137,428	115	53,369
Kansas	2,913,314	105	27,746
Illinois	12,671,821	102	124,234
North Carolina	10,488,084	100	104,881
Iowa	3,155,070	99	31,869
Tennessee	6,833,174	95	71,928
(…)			
New York	19,453,561	62	313,767
California	39,512,223	58	681,245
(…)			
Massachusetts	6,949,503	14	496,393
Vermont	623,989	14	44,571
New Hampshire	1,359,711	10	135,971
Connecticut	3,565,287	9	445,661
Hawaii	1,415,872	5	283,174
Rhode Island	1,059,361	5	211,872
Delaware	973,764	3	324,588
District of Columl	705,749	1	705,749

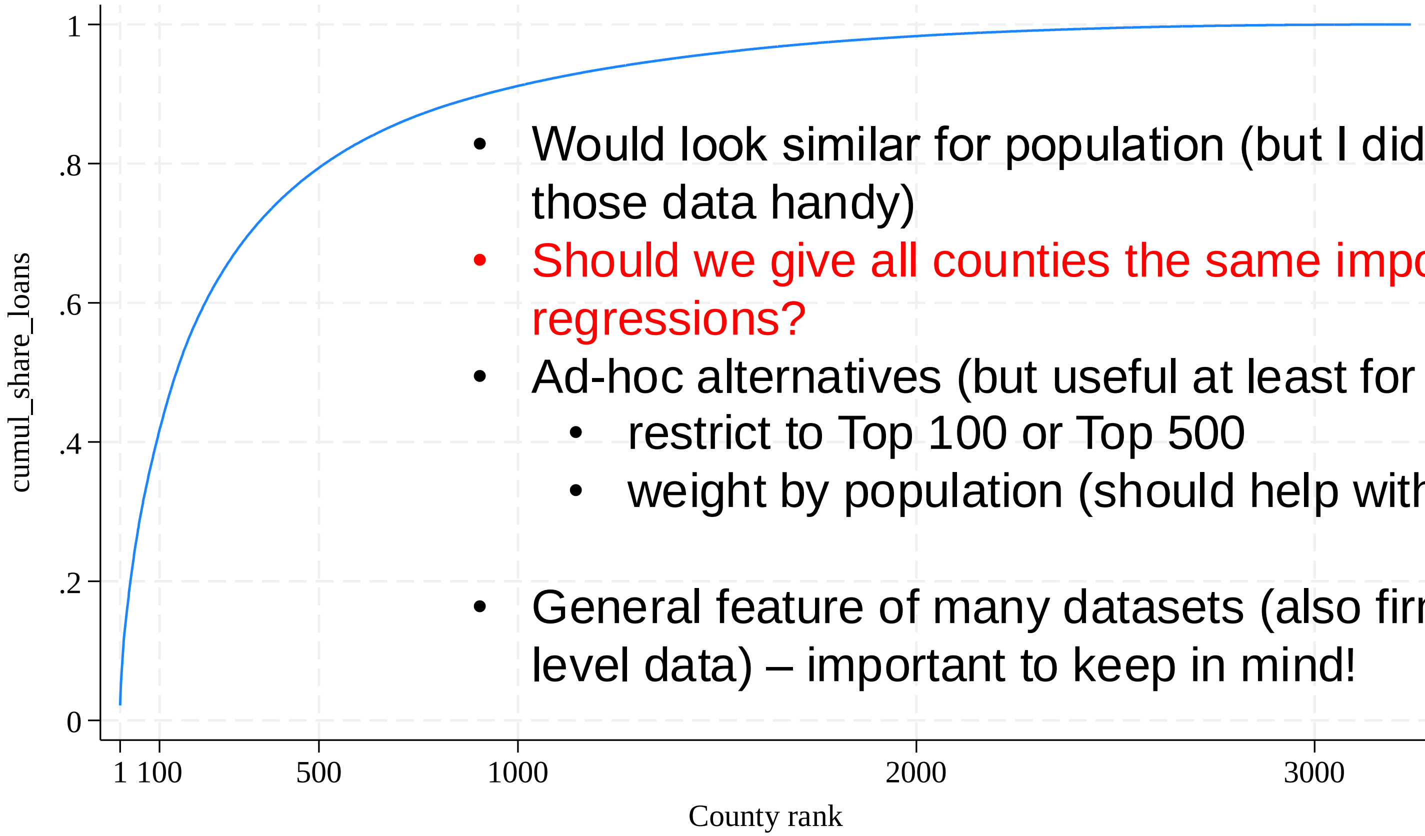
Different US states have very different approaches to subdividing into counties.

This leads to very different average population numbers, and to some states accounting for many more observations than others despite similar total population (e.g. KS vs. CT)

(Source: https://en.wikipedia.org/wiki/County_(United_States))

US counties are EXTREMELY heterogenous in size – part 2

- Example: number of mortgage originations by county in 2018 (from HMDA data) – top 100 counties account for about 40% of all loans; top 500 for about 80%



- Would look similar for population (but I didn’t have those data handy)
- Should we give all counties the same importance in our regressions?
- Ad-hoc alternatives (but useful at least for robustness):
  - restrict to Top 100 or Top 500
  - weight by population (should help with precision)
- General feature of many datasets (also firm- or bank-level data) – important to keep in mind!

Ending with some positive news…

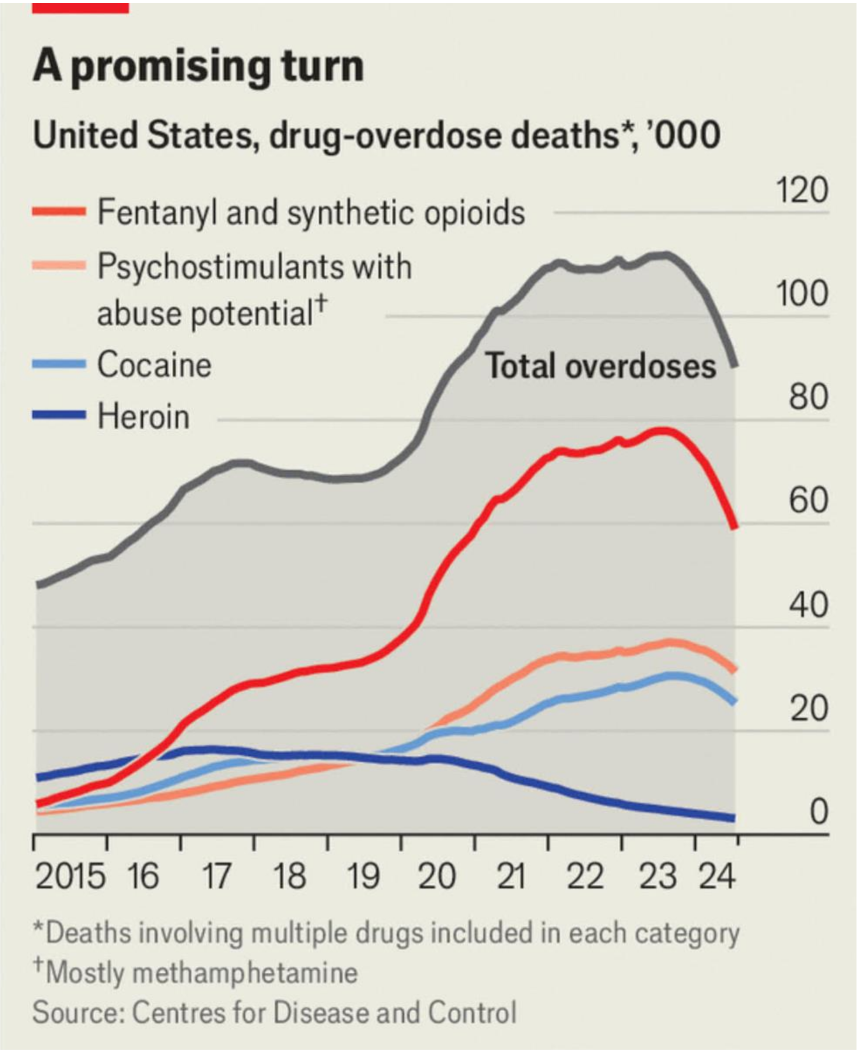


IMAGE: THE ECONOMIST