

Smoothing 2020 Census Age Distributions with P-TOPALS

Sigurd Dyrting, Charles Darwin University
Abraham Flaxman, University of Washington
Ethan Sharygin, Portland State University

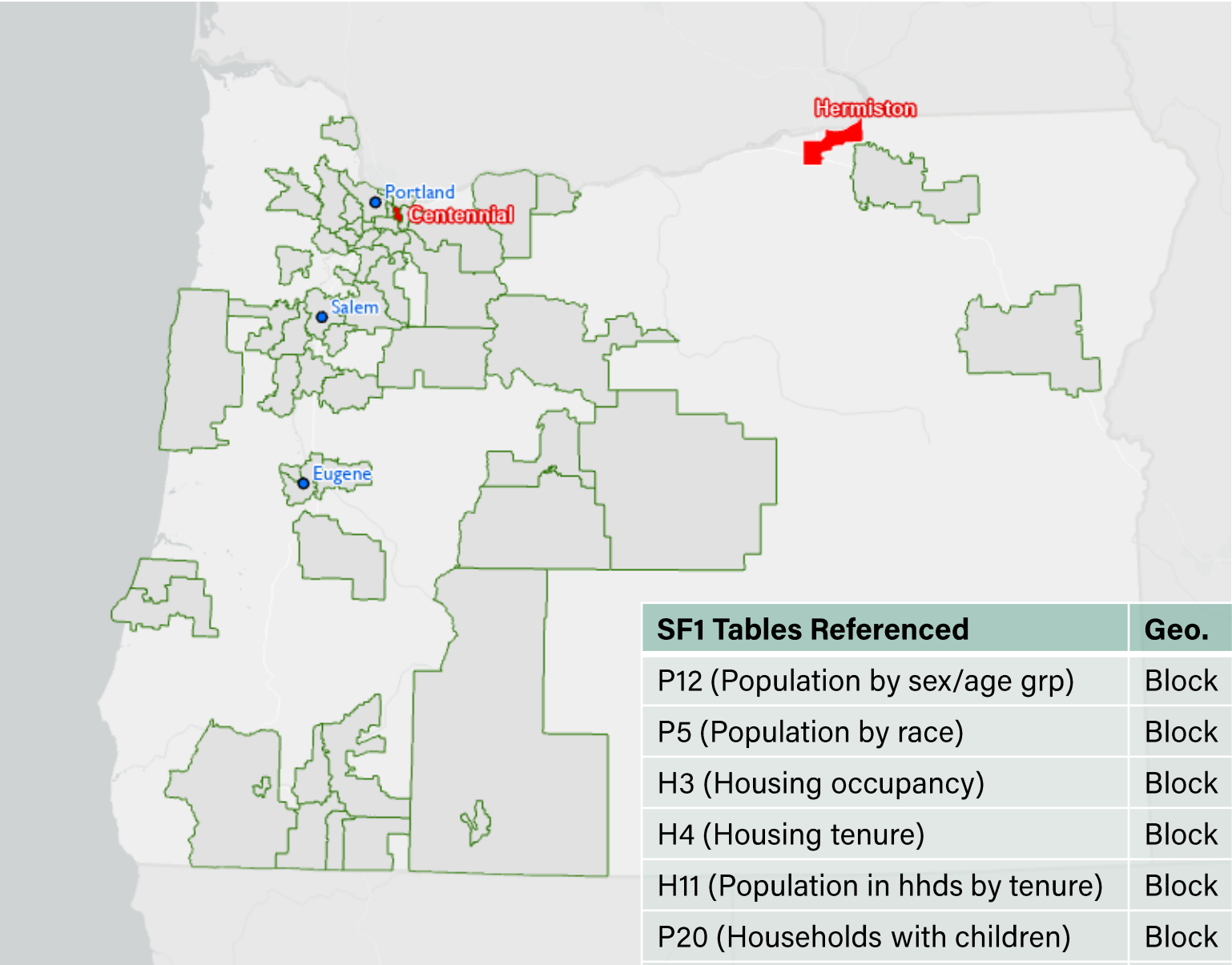


CASE STUDY 1. School enrollment forecasts (ORS 195.11)	School enrollment forecasts are less accurate when made from the 2020-05 or 2022-03 forecasts compared to baseline. 2022-03 forecasts are more accurate.
CASE STUDY 2. Urban Growth Boundary population (ORS 195.033)	UGB total population estimates were not affected, although housing characteristics (average household size, occupancy/vacancy rates) were less accurate from demonstration data than from published data.
CASE STUDY 3. County and UGB population forecasts (ORS 195.033)	P-TOPALS, a nonparametric smoothing technique, can be adapted to smoothing age distributions, and recovered significantly less noisy/more accurate age distributions in both the 2020-05 and 2022-03 demonstration data products.

Case Study 1.

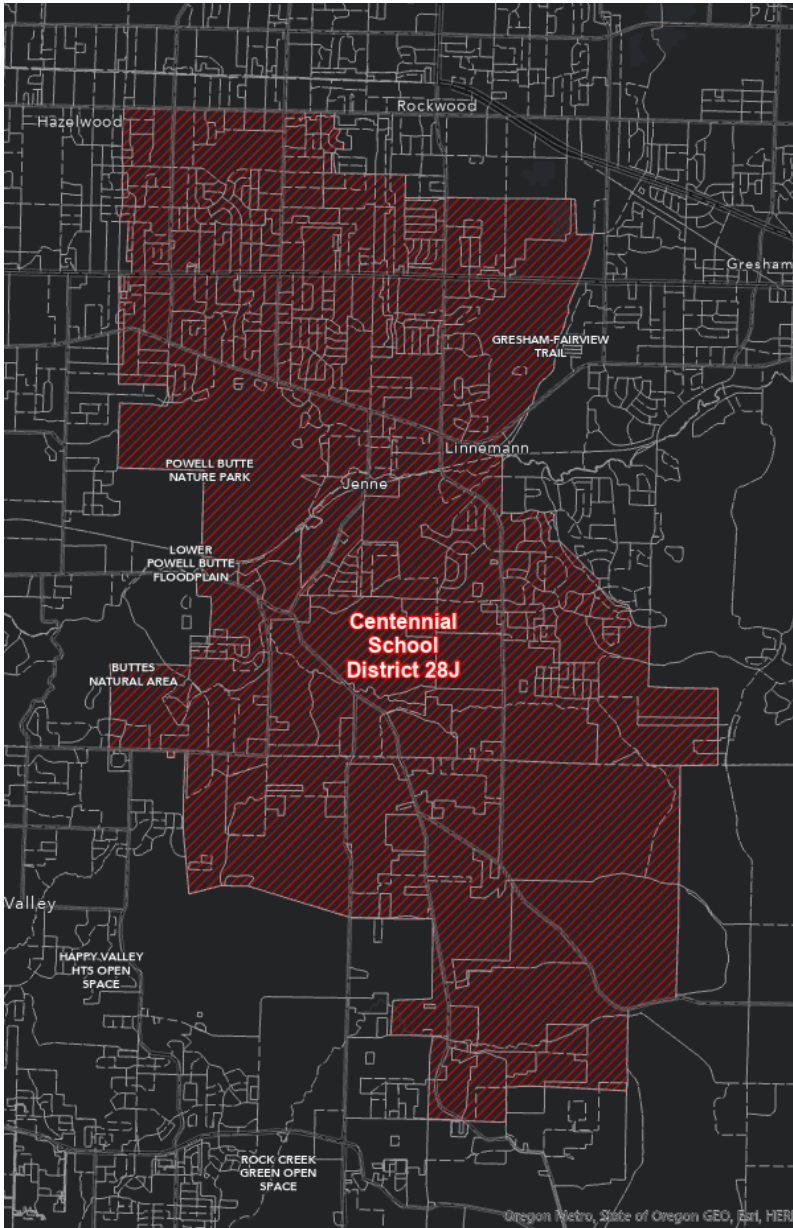
ORS 195.110
School facility plan for
large school districts:

*“A city or county containing
a large school district shall:
[produce] a school facility
plan [...] The school facility
plan must cover a period of
at least 10 years and must
include, but need not be
limited to, the following
elements: (A) Population
projections by school age
group. [...]”*



SF1 Tables Referenced	Geo.	Status
P12 (Population by sex/age grp)	Block	DHC
P5 (Population by race)	Block	DHC
H3 (Housing occupancy)	Block	DHC
H4 (Housing tenure)	Block	DHC
H11 (Population in hhds by tenure)	Block	D-DHC*
P20 (Households with children)	Block	DHC (P17)
P14 (Age/Sex single years 0-19)	Block	DHC

“large school district” means a school district that has an enrollment of over 2,500 students (ORS 195.110)
* D-DHC table equivalent to 2010 SF1 H11 is planned for block group and larger areas only



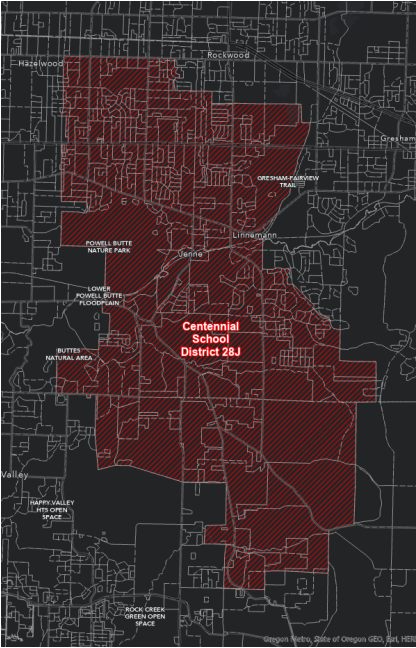
Centennial School District

45,000 residents (+4,000 since 2010).

8,000 school age children (no change)

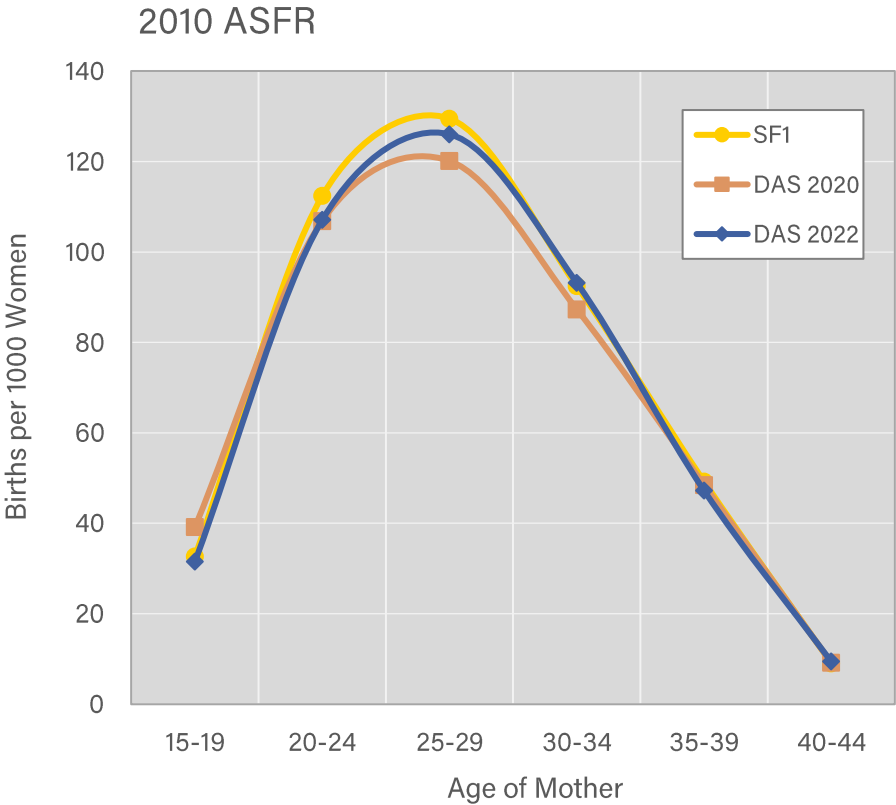
6,000 enrolled in CSD (-400 since 2010).

Budget: \$78 million (2020-21).

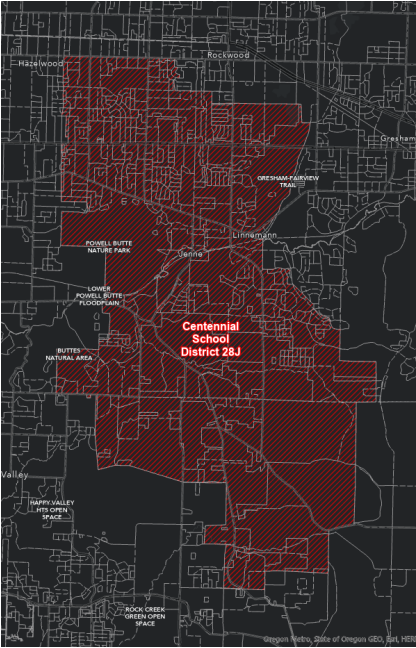


Centennial School District

Implied birth rates differ from SF1.
2022 improved.

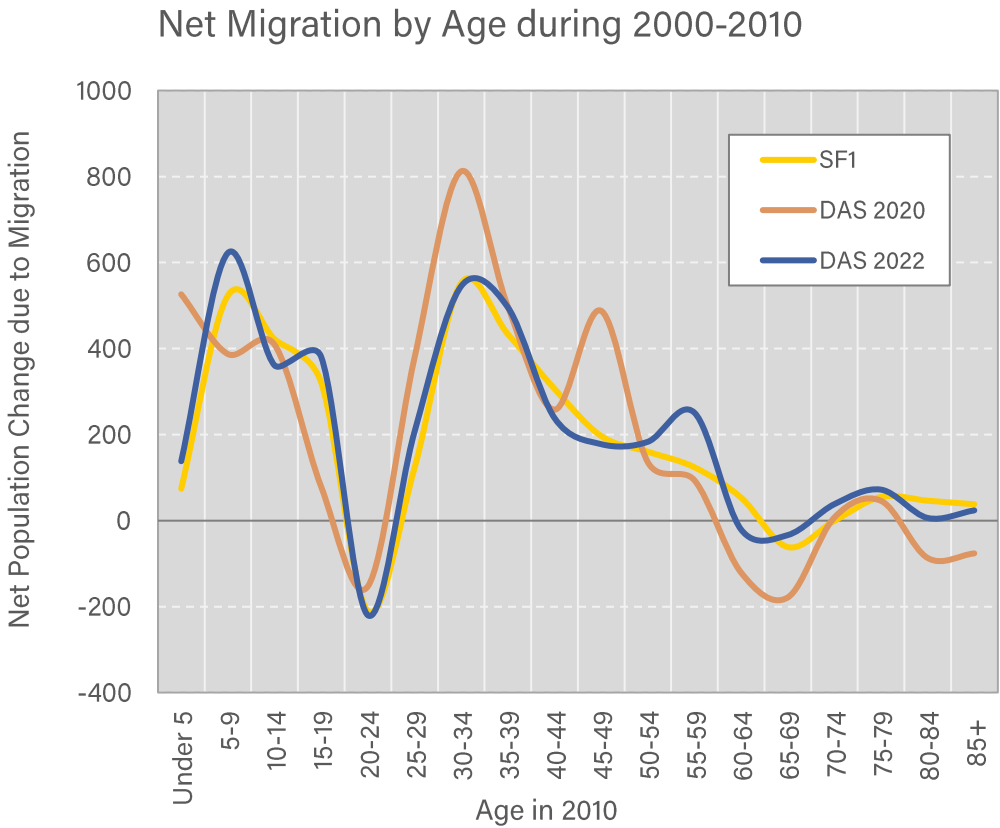


		2010 births/1000 women based on:		
	2000	SF1	DAS 2020	DAS 2022
15-19	45.0	32.7	39.2	31.6
20-24	130.8	112.4	106.9	107.1
25-29	124.9	129.5	120.1	126.0
30-34	85.6	92.4	87.3	93.2
35-39	33.1	49.3	48.5	47.3
40-44	5.7	9.0	9.1	9.5
TFR	2.13	2.13	2.06	2.07

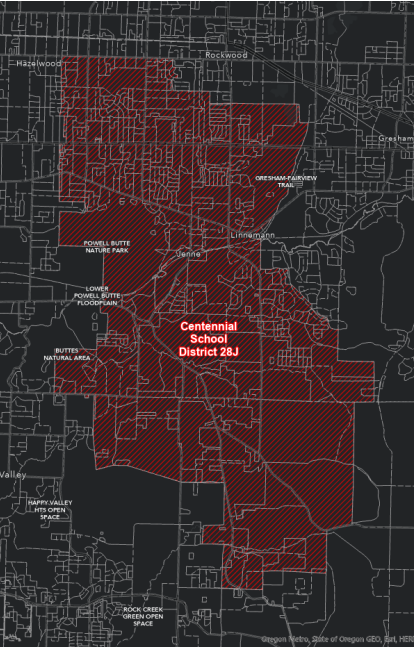


Centennial School District

Similar total net migrants. 2020 highly erratic;
2022 improved.

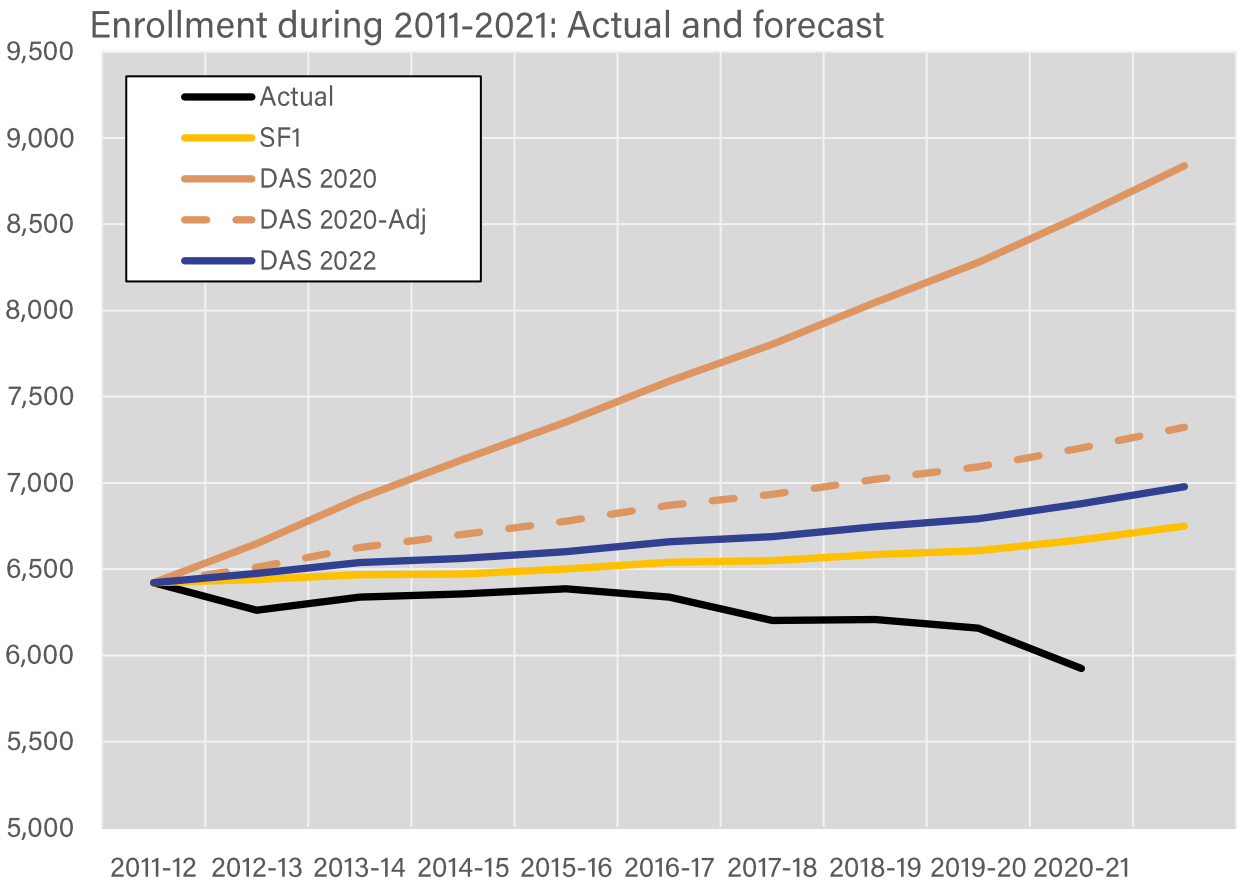


	2000-10 net migrants based on:		
Age in 2010	SF1	DAS 2020	DAS 2022
Under 5	74	526	138
5-9	523	387	622
10-14	422	409	361
15-19	322	79	381
20-24	-211	-153	-220
25-29	122	378	206
30-34	552	812	545
35-39	434	498	496
40-44	307	258	239
45-49	197	488	178
50-54	160	136	183
55-59	124	94	251
60-64	54	-120	-20
65-69	-62	-179	-34
70-74	-1	7	38
75-79	55	46	72
80-84	46	-87	6
85+	38	-76	24
Total	3156	3503	3466



Centennial School District

2020 Forecast exceptionally high, even after adjustment.
2022 improved (+1.5 pp higher error than baseline).



		2020-21 enrollment forecast based on:		
Grade	Actual	SF1	DAS 2020	DAS 2022
K	368	459	627	473
1	431	493	676	511
2	422	495	682	515
3	416	497	688	519
4	461	500	696	523
5	422	499	683	507
6	454	520	754	542
7	512	524	793	558
8	440	534	800	575
9	462	516	520	519
10	446	509	509	510
11	462	519	518	519
12	417	401	399	403
UN	212	205	205	205
K-12	5,925	6,671	8,550	6,879

Case Studies 2-3.

ORS 195.033

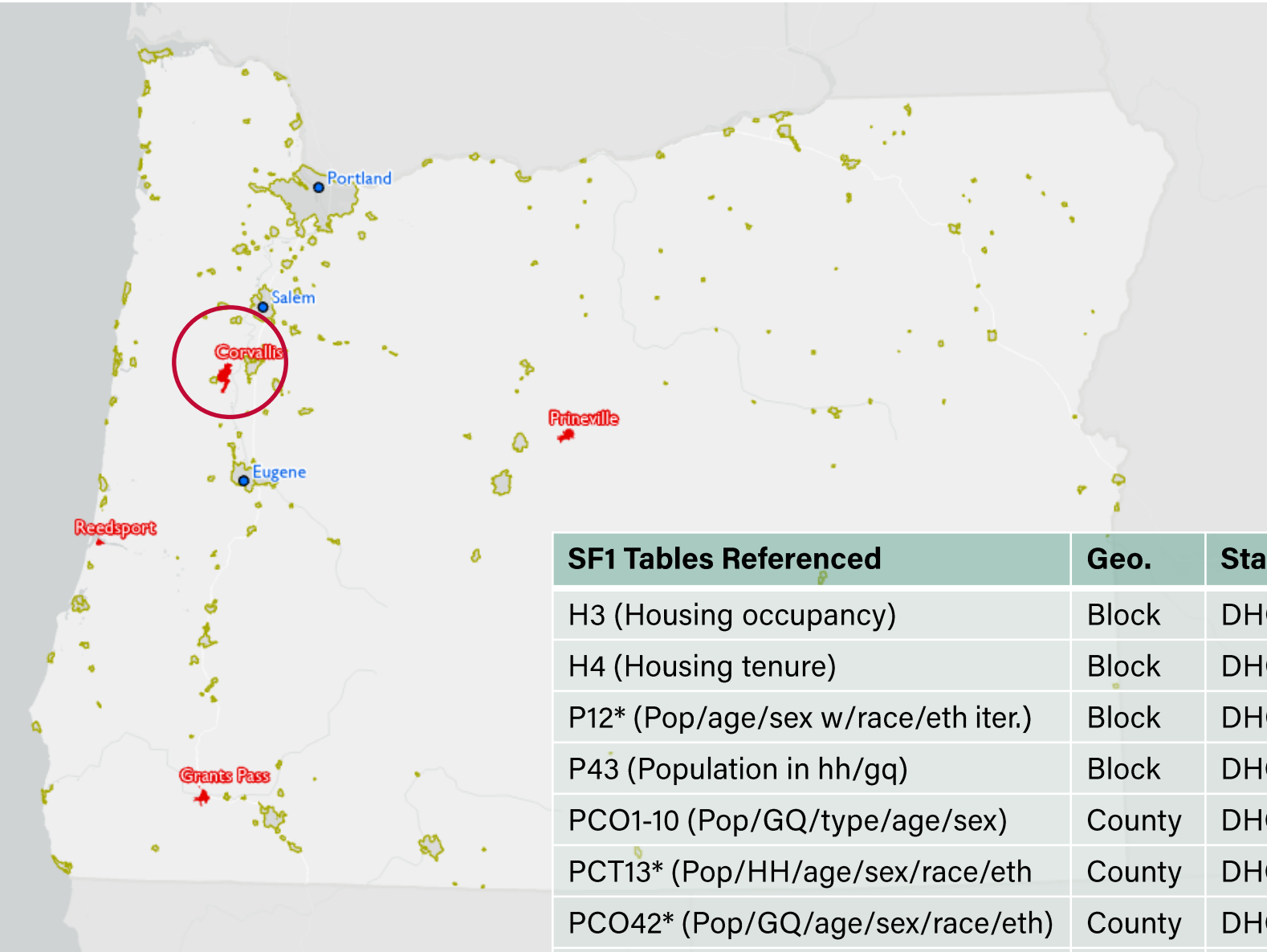
Area population forecasts

"For the purpose of land use planning, the center shall issue a population forecast for: (a) Each county [except Metro] and (c) The area within each urban growth boundary..."*

ORS 197.296

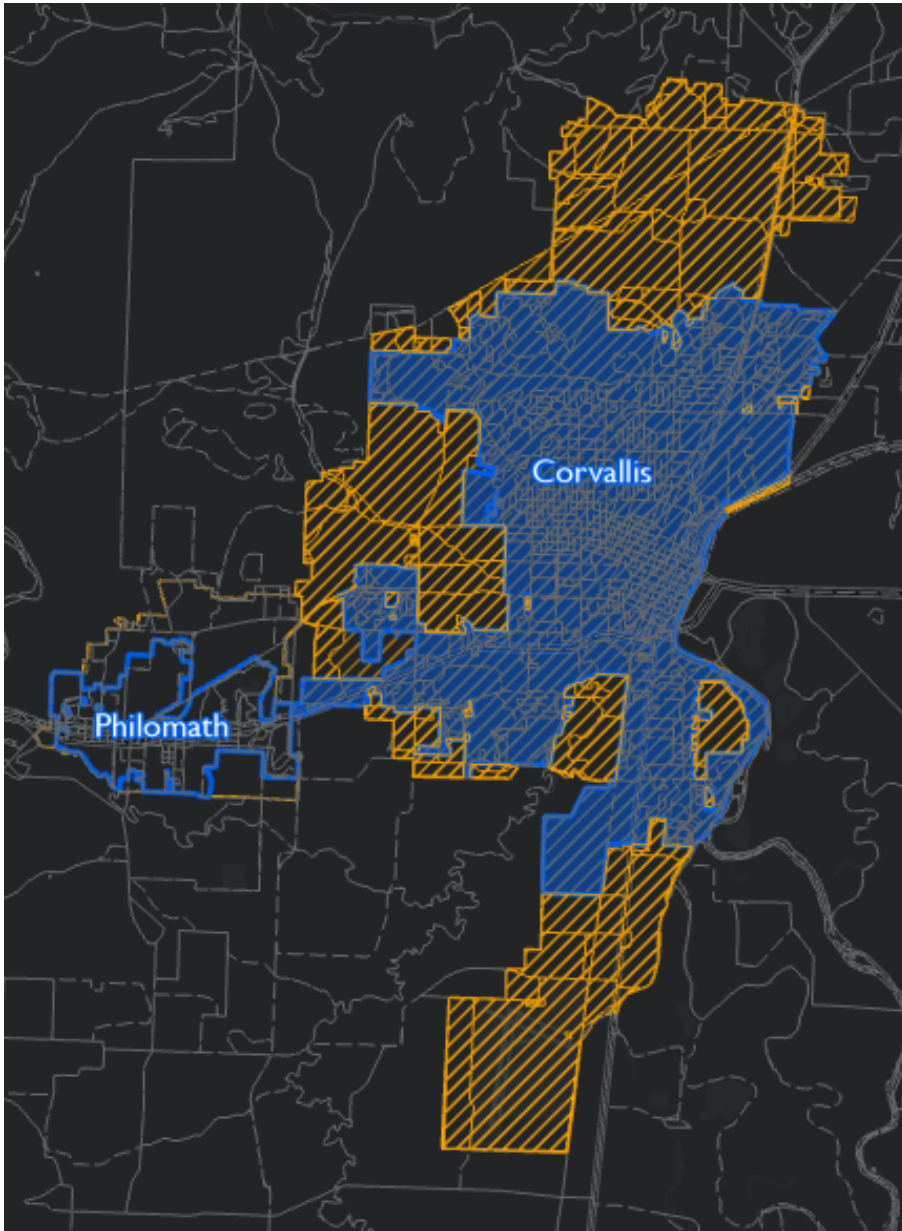
Factors to establish sufficiency of buildable lands within UGB

"A local government shall demonstrate [...] sufficient buildable lands within the urban growth boundary [...] to accommodate estimated housing needs for 20 years"



SF1 Tables Referenced	Geo.	Status
H3 (Housing occupancy)	Block	DHC
H4 (Housing tenure)	Block	DHC
P12* (Pop/age/sex w/race/eth iter.)	Block	DHC
P43 (Population in hh/gq)	Block	DHC (P20)
PCO1-10 (Pop/GQ/type/age/sex)	County	DHC (PCO12-23)
PCT13* (Pop/HH/age/sex/race/eth)	County	DHC (PCO24*)
PCO42* (Pop/GQ/age/sex/race/eth)	County	DHC (PCO20*)
PCT12 (Sex by single year of age)	County	DHC (PCT1)

* "Metro" means area in the jurisdiction of the Metro regional government/metropolitan planning organization.

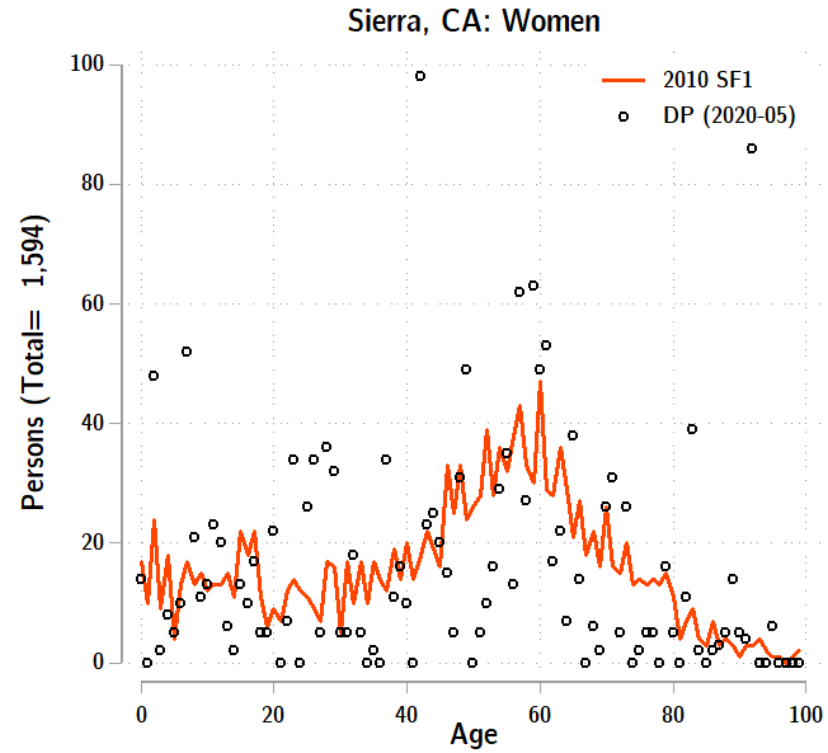
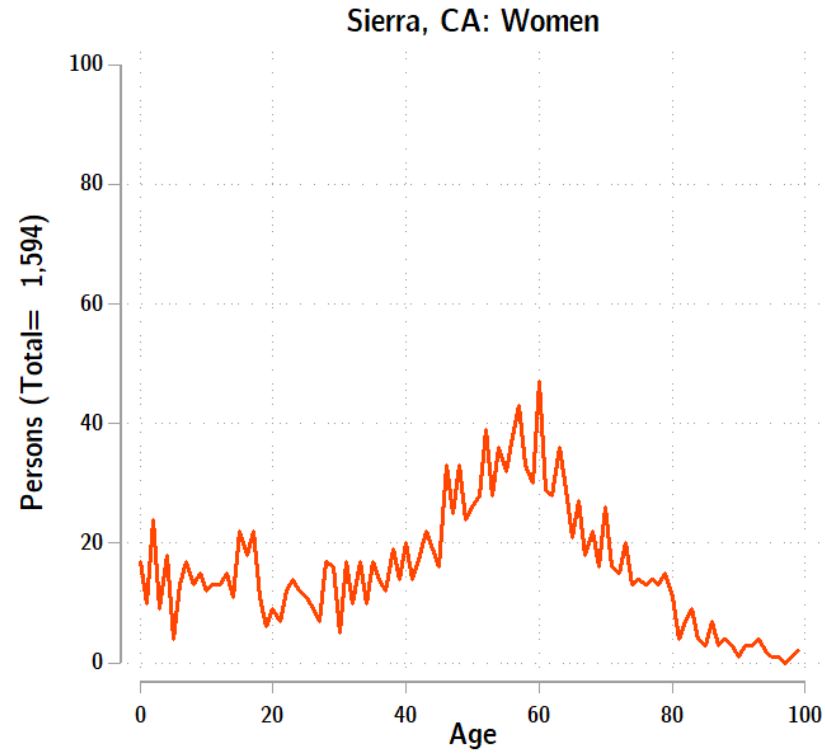


Corvallis

2022 shows very small differences; minimal effects on UGB est.

	2000	2010		
<u>City (total)</u>	SF1	SF1	DAS 2020	DAS 2022
Population	49322	54462	54318	54455
Households	44435	49563	49444	49604
Group Quarters	4887	4899	4874	4851
Housing Units	20909	23423		23423
Occupied	19630	22283		22273
Vacant	1279	1140		1150
Share vacant	6.1%	4.9%		4.9%
Avg household size	2.26	2.22		2.23
	2000	2010		
<u>UGB (margins)</u>	SF1	SF1	DAS 2020	DAS 2022
Population	2947	3012	3188	3015
Households	2943	2975	3170	2983
Group Quarters	4	37	18	32
Housing Units	1295	1330		1330
Occupied	1210	1250		1251
Vacant	85	80		79
Share vacant	6.6%	6.0%		5.9%
Avg household size	2.43	2.38		2.38

2020-05 DAS

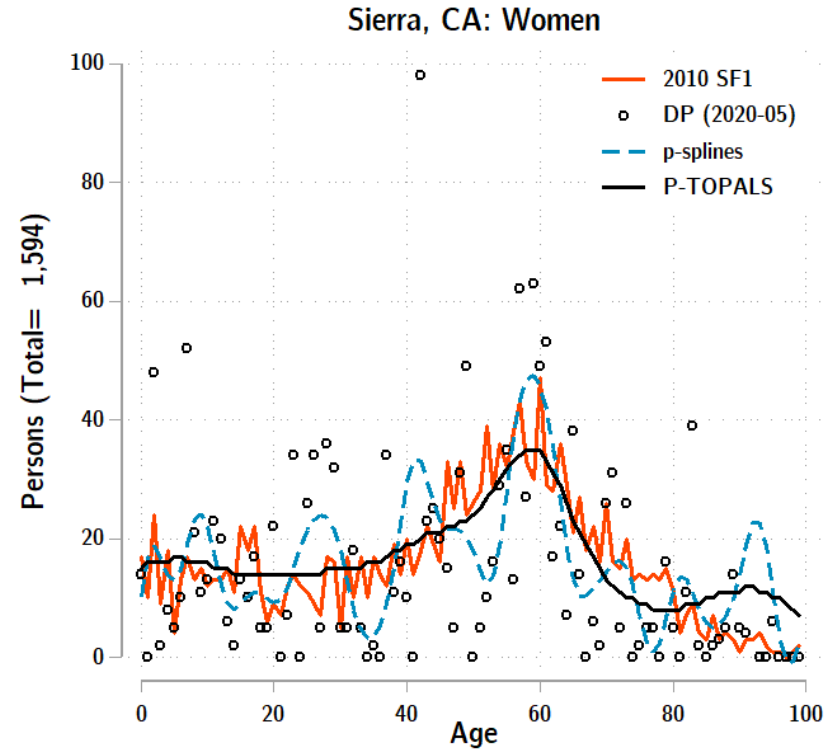
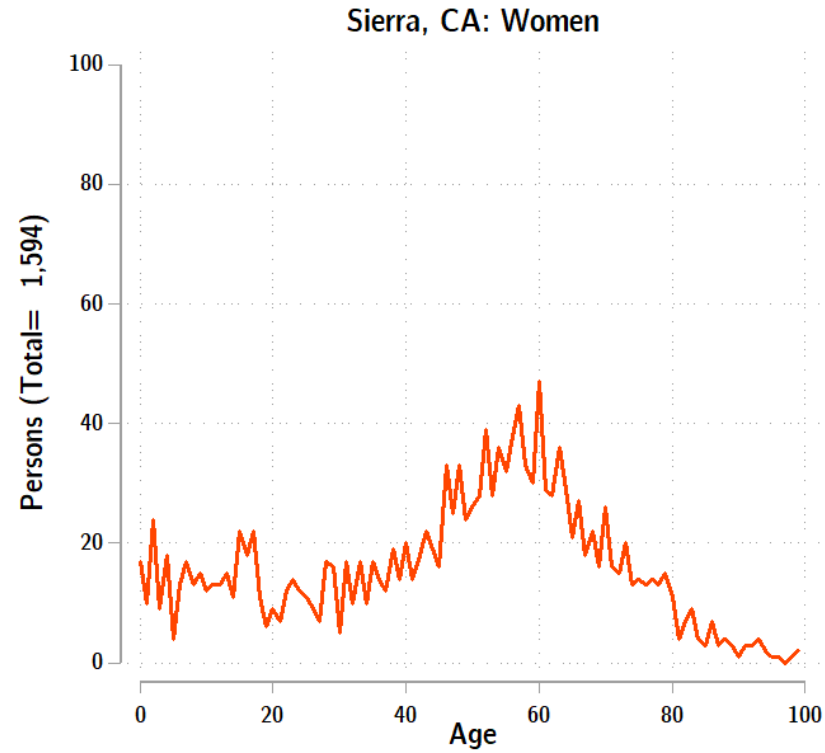


P-TOPALS

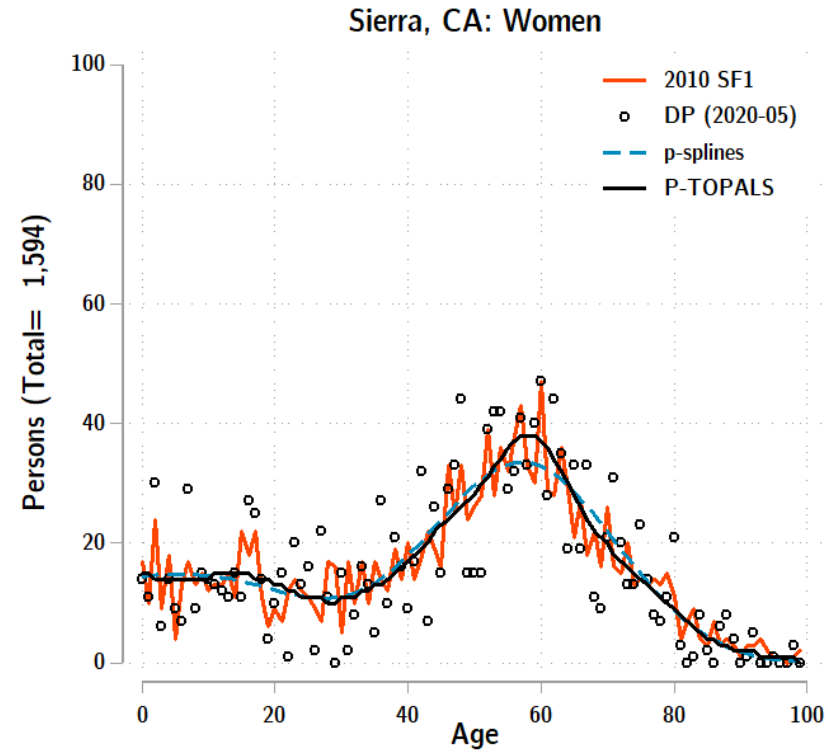
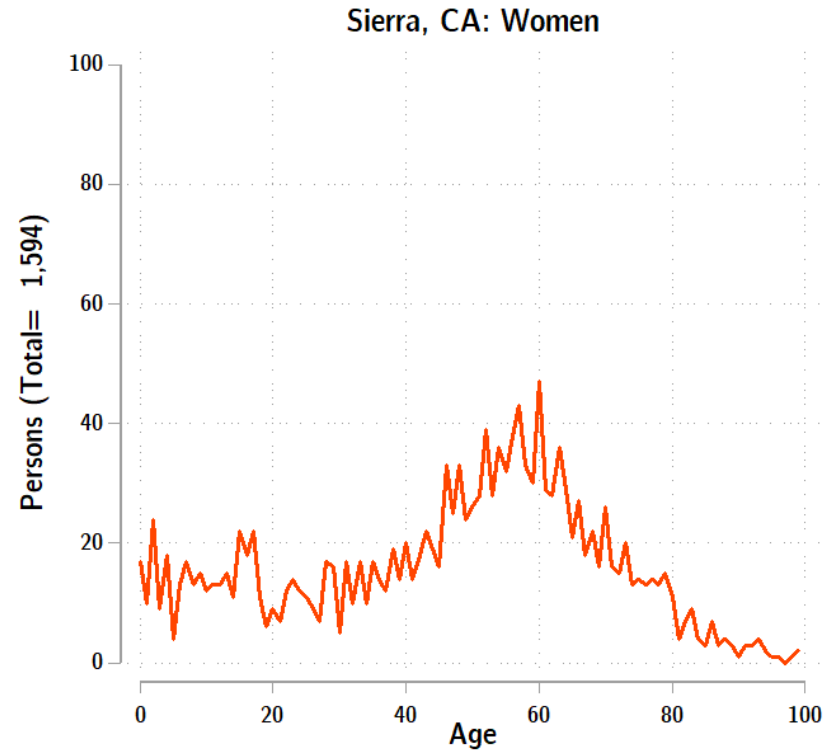
Demographic methods for smoothing age distributions:

1. Age heaping in censuses UN [1956](#); Arriaga [1968](#)
2. Nuptiality, fertility, mortality; p-splines Eilers and Marx [1996](#); Currie et al. [2004](#)
3. Migration: TOPALS de Beer [2011](#), [2012](#) and P-TOPALS Dyrting [2020](#)

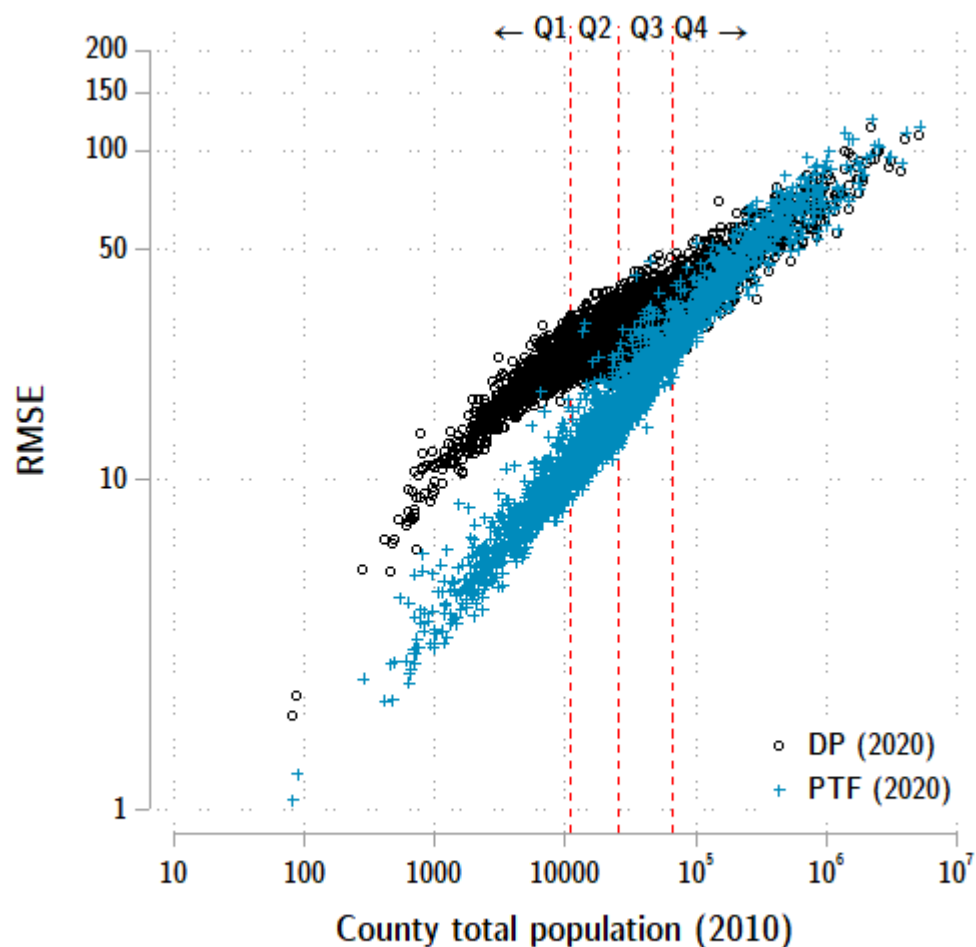
2020-05 DAS



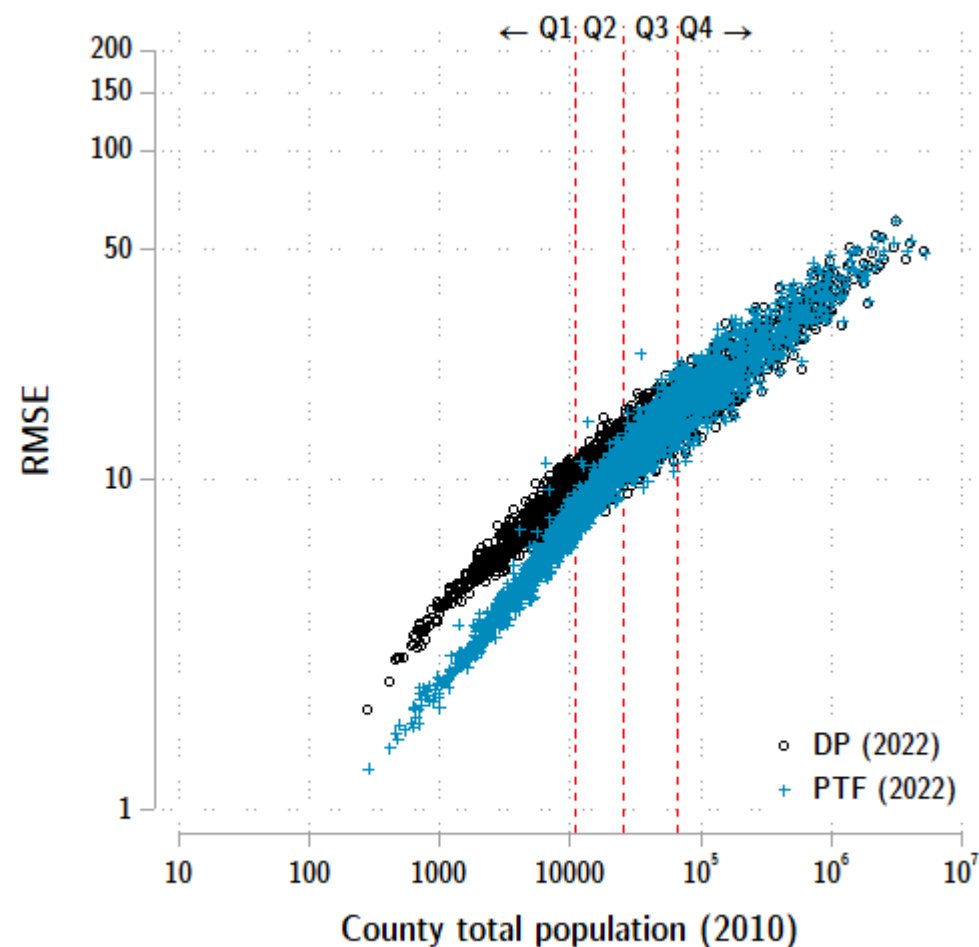
2022-03 DAS



Results: PCT1



Note: Q1-Q4 denote quartiles of counties by population size



Note: Q1-Q4 denote quartiles of counties by population size

Conclusion

2020-05 PPMF shows problematic levels of noise and leads to significant errors in calculating rates, population estimates, and forecasts. 2022-03 DHC shows significant improvement and is fit for many use cases.

P-TOPALS improved accuracy of data and real-world usability of age distributions, even in new releases.

Future Work:

More informative priors

Apply P-TOPALS as part of TDA or using noisy measurement file?

Contact: sharygin@pdx.edu

DYRTING Sigurd, Charles Darwin University
FLAXMAN Abraham, University of Washington
SHARYGIN Ethan, Portland State University