

Integrating Third-Party Data on Establishment Births and Deaths to Improve Monthly State Retail Sales Estimates

Drew Keller

Supervised by Rebecca Hutchinson

Economic Indicators Division

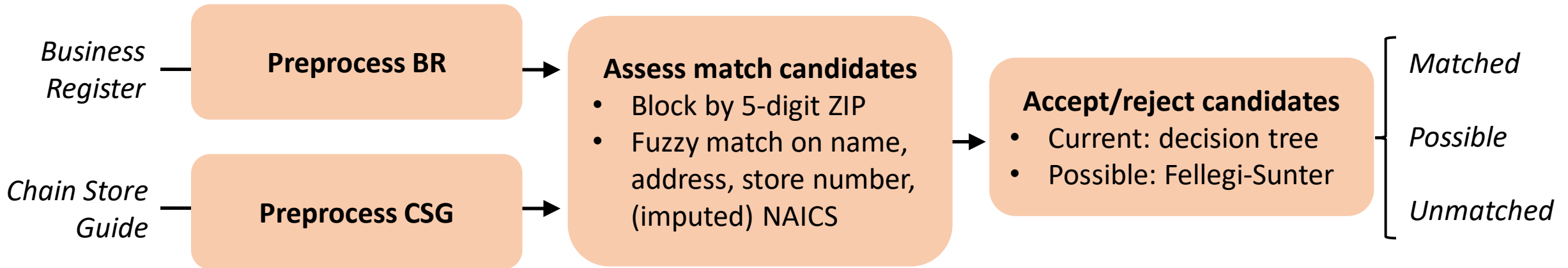
Introduction

- Monthly State Retail Sales (MSRS): experimental data product
 - Monthly Retail Trade Survey + administrative payroll data + third-party data
- Challenge: Accounting for establishment openings/closures in real time
 - 1-2 year lag in MSRS frame due to Business Register (BR) update timeline
- Third-party data to the rescue? Chain Store Guide dataset:

Example (Fake) Chain Store Guide Data

INDUSTRY	NAME	ADDR_1	ZIP	STORE	PARENT INDUSTRY	PARENT NAME	OPENED	CLOSED
Category Killer	Store A	1212 Example St.	12345	5	Card/Gift/Novelty	Store A Inc.	Jan 2018	
Apparel	Store B	10 Example Ave.	02468		Apparel Stores	Holding Company D	Jan 2018	Aug 2020
Convenience Stores	Store C	Hwy 1 & Rte 2	13579	Y-27	Convenience Stores	Store C Partners		Sept 2021
etc.								

1. Link Chain Store Guide to Business Register



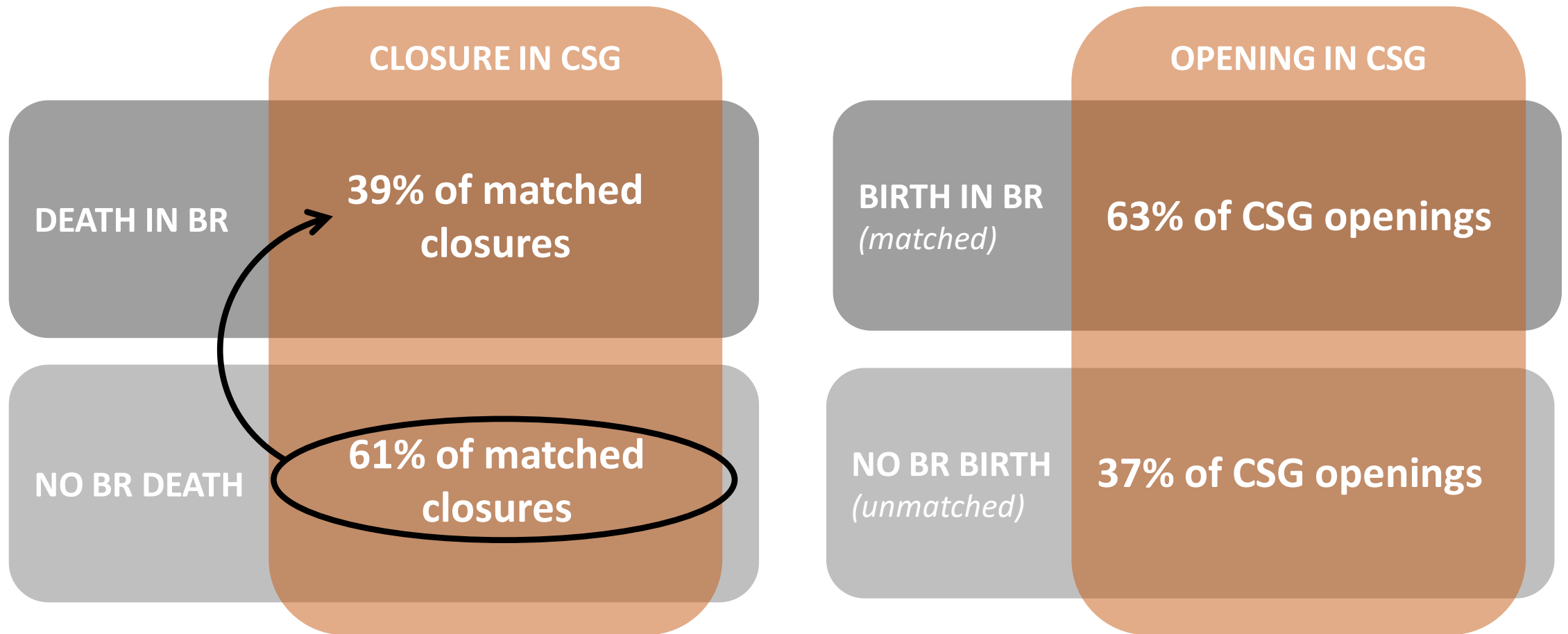
Matching Results

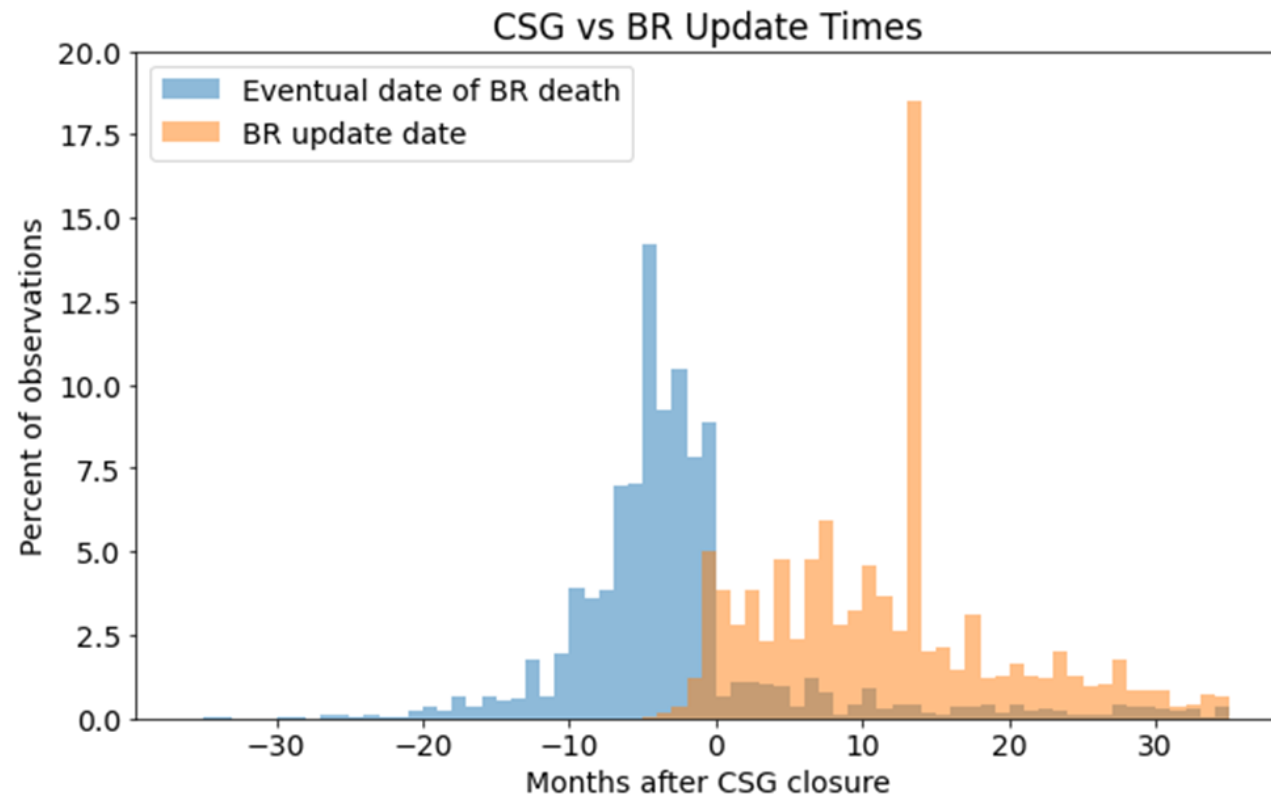
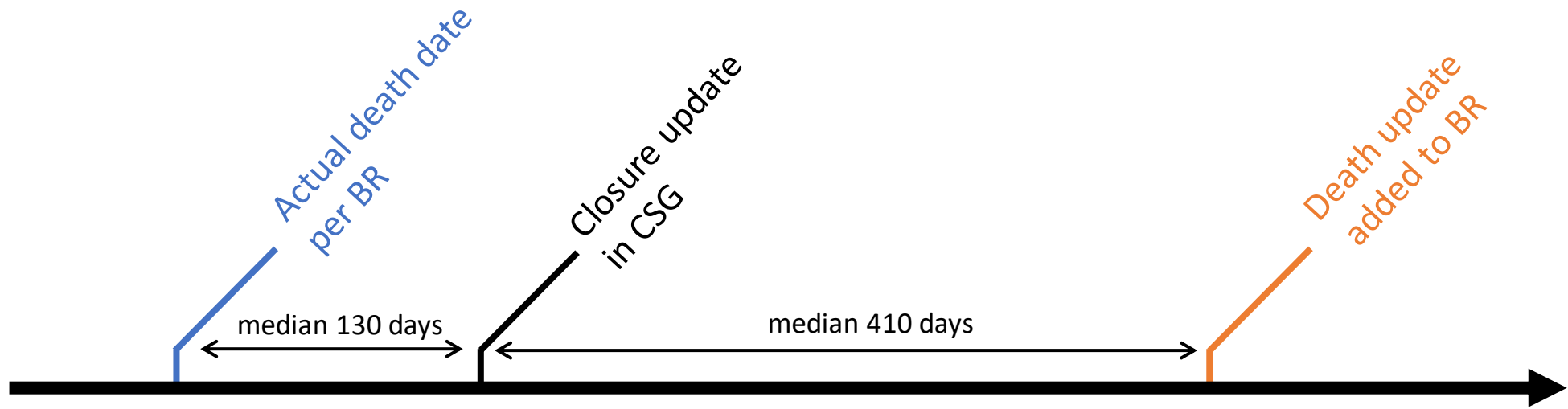
Estimated Run Time	Match Rate		Positive Predictive Value*	
	Closures	Openings	Closures	Openings
30 CPU-hours	62%	63%	91%	79%

* 95% confidence lower bound

differing objectives

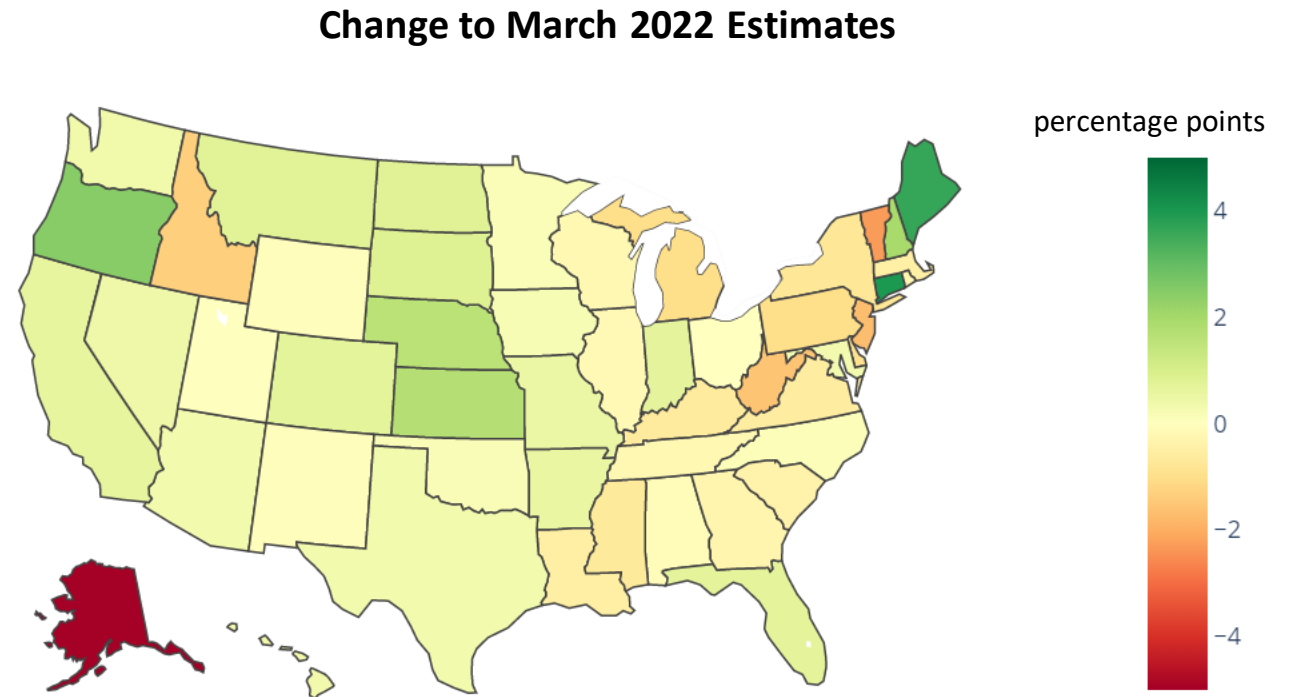
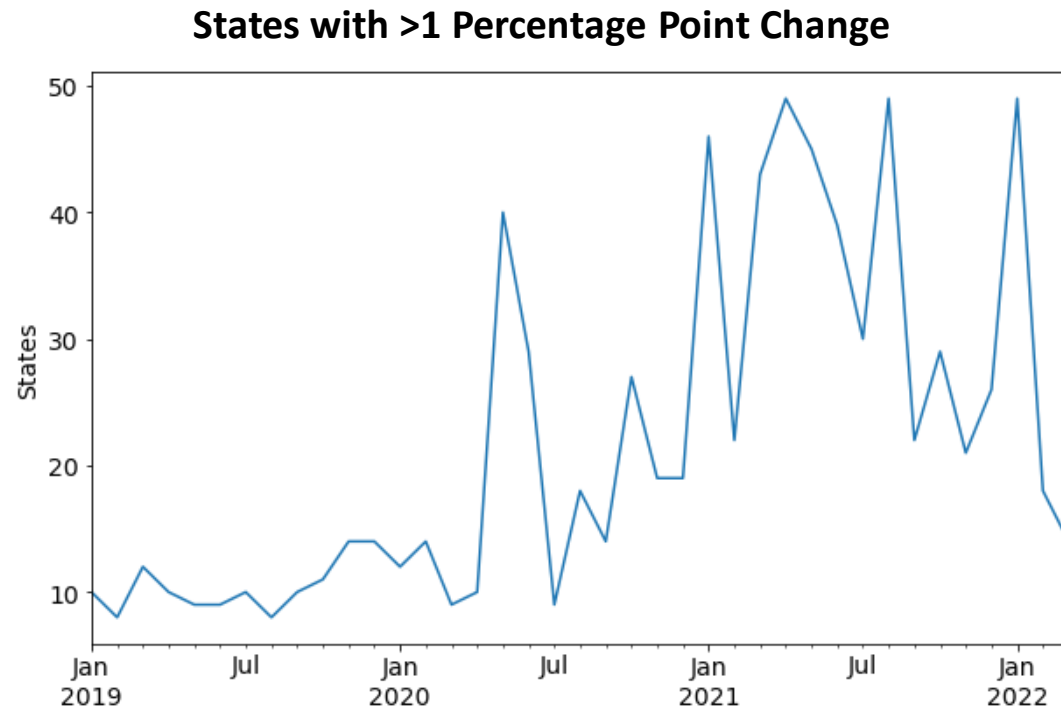
2. Compare CSG and BR Openings/Closings





3. Incorporate Births/Deaths into MSRS

Changes to Apparel (NAICS 448) Sales Estimates When Incorporating CSG Deaths



Conclusion

Third-party data can improve timeliness of establishment birth/death updates in the Business Register by >1 year, changing BR-derived Monthly State Retail Sales estimates.

Limitations

- Matching challenges (franchises!)
- CSG has data quality issues too
- Harder to account for births than deaths

Next Steps

- Evolve birth imputation process
- Model and account for CSG update lag
- Explore additional third-party data sources and applications