An approach to measuring informal activity in California

Allegra Saggese

ECON221 Economic slack
Final results presentation

December 17, 2025

Question and positioning

How large is the informal economy in CA?

- Measuring the magnitude of **informal economy throughout business cycles** (Bracha & Burke 2017, Restrepo-Echavarria 2018, Leyva & Urrutia 2020, Valleta et al. 2020, Schneider 2023, Pappada & Rogoff 2023).
  $\rightarrow$ *do people take on informal work in response to macroeconomic conditions?*
  $\rightarrow$ *is this behavior permanent or temporary?*

- **Tax evasion behavior** has various motivations (Slemrod 2007), where an improved measurement and theory for informal work can sharpen estimates of income rate based evasion (Zucman et al 2019)

Data: what are we measuring?

	Illicit activity	**Legal activity**
Taxable	Tax fraud including **informal work activities** i.e. under the table pay or cash-only businesses	All recorded transactions and taxable receipts from businesses and income generated by US individuals
Untaxable	Criminal activity (drug trade, other black market activities)	Tax avoidance behavior, tax exempt purchases and sales specific to jurisdictions

Data: sources

Core data: building a synthetic tax base: I assemble a set of data that is a synthetic taxable base, using county-level recorded taxable transactions to estimate revenue of taxation.
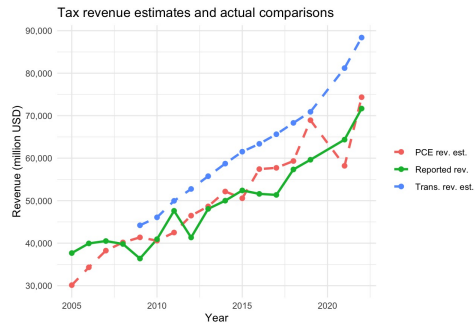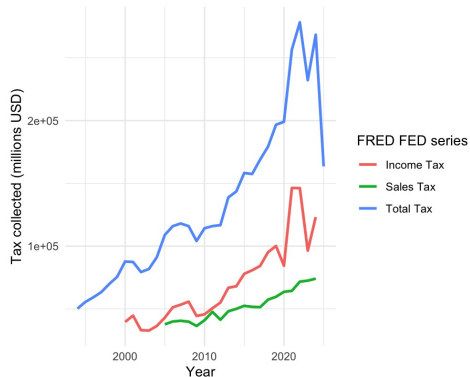
1. Get quintile-level disposable income estimates for California (five income categories)
2. Classify counties in CA by median income level in quintiles
3. Estimate a tax base from disposable income spending for each county
4. Apply state and local tax rate to each counties tax base
5. Sum across counties, compare to total collected taxes in the state

44 %

Data and model: motivation for model

Recall the model I was using was a VAT-correction approach. I have simplified to a simple ratio of actually collected sales and transaction revenue vs. the estimated amount of the tax base:
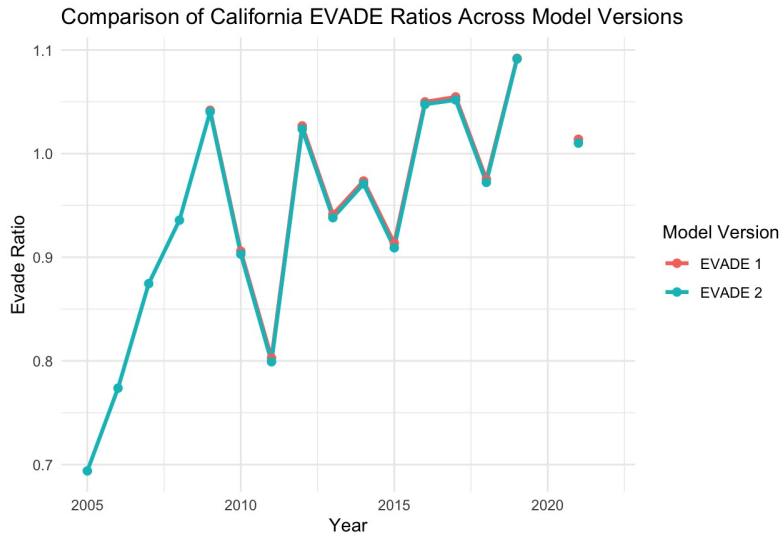
$$\text{EVADE}_t = \frac{tax_t^{\text{full}} - tax_t}{tax_t}$$

where $tax_t^{\text{full}} = \sum c_{it} \tau_{i,t,j}$ (all possible transactions at their taxable rate)

Tax revenue estimates and actual comparisons

Comparison of California EVADE Ratios Across Model Versions
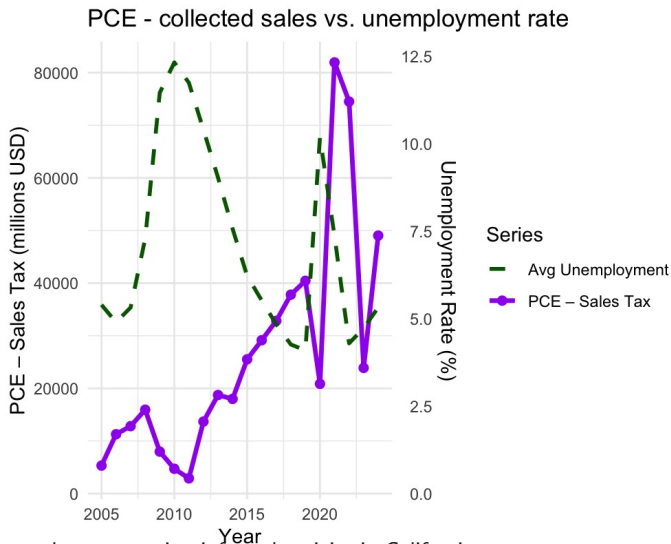
Results: Tax collection over time and unemployment

- How does unemployment vary with the gap between spending and collected sales tax?

- Unemployment is high, spending drops

- Transfers can be seen post-2020 (COVID trend)

PCE - collected sales vs. unemployment rate

Implication of results

- Note: haven't been able to adjust the EVADE measure significantly, so we cannot say definitively what the size of the informal economy is

  $\rightarrow$ further consideration should be made to more high frequency data, as data estimates short run responses by workers during recessions

  $\rightarrow$ large issue with being able to estimate measurement error vs. informal activity

- Next steps: what is the best way to visually demonstrate business cycles fluctuations in the rate of unemployment, and the potential metric for informal work?