Disposable Income and Redistribution over the Lifecycle

Christian Hilscher   **Holger Stichnoth**

Discussion by

Cameron LaPoint
Yale SOM

9th Annual Mannheim Taxation Conference

ZEW Mannheim

September 9th, 2022

# THIS PROJECT: SIMULATION APPROACH TO MEASURE INCOME

- Early-stage project proposing a series of ambitious goals:
  - ▶ Use longitudinal panel to study inequality in lifetime <u>disposable</u> income, and compare to earnings inequality in other studies (e.g. Kopczuk, Saez, Song 2010 using SSA data)
  - ▶ Simulate counterfactuals for different family structures, tax/transfer regimes
  - ▶ Evaluate whether ML model offers potential improvement over structural approaches

- What are the economic hypotheses?
  - ▶ To extent that rich are more adept at tax avoidance/evasion and financial planning, **disposable income inequality** $>$ **(after-tax) earnings inequality**
  - ▶ Workhorse computational models of life-cycle income dynamics like Kaplan & Violante (2014) limited by # of state variables $\longrightarrow$ "hump" shapes in wealth/consumption

- No paper yet, so I will focus my remarks on possible directions and **differentiating** the work w.r.t. existing literatures in public finance & macro

		Wealth Growth	Total Income Reported		Total Taxes Paid		True Tax Rate
	Warren Buffett *Berkshire Hathaway Inc.*	$24.3B	$125M		$23.7M		**0.10%**
	Jeff Bezos *Amazon.com Inc.*	$99.0B	$4.22B		$973M		**0.98%**
	Michael Bloomberg *Bloomberg LP*	$22.5B	$10.0B		$292M		**1.30%**
	Elon Musk *Tesla Inc.*	$13.9B	$1.52B		$455M		**3.27%**

Source: ProPublica, "The Secret IRS Files: Trove of Never Before Seen Records Reveal How the Wealthiest Avoid Income Tax," June 8, 2021.

- Common mechanisms: preferential treatment of LT gains, reporting losses, trusts, etc.
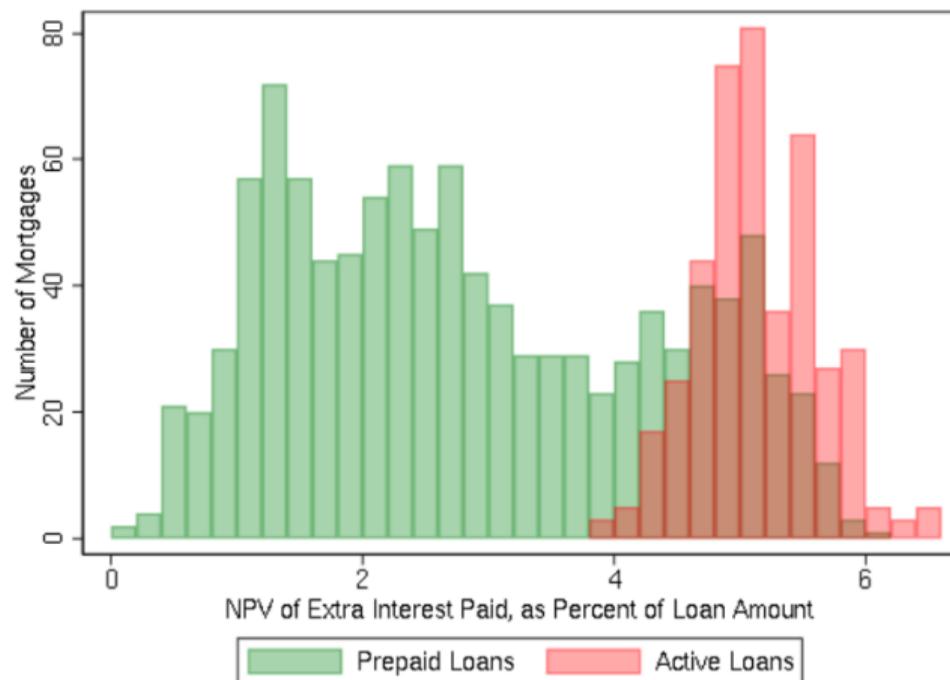
UNREPORTED INCOME SPIKES FOR PROFESSIONAL EARNERS
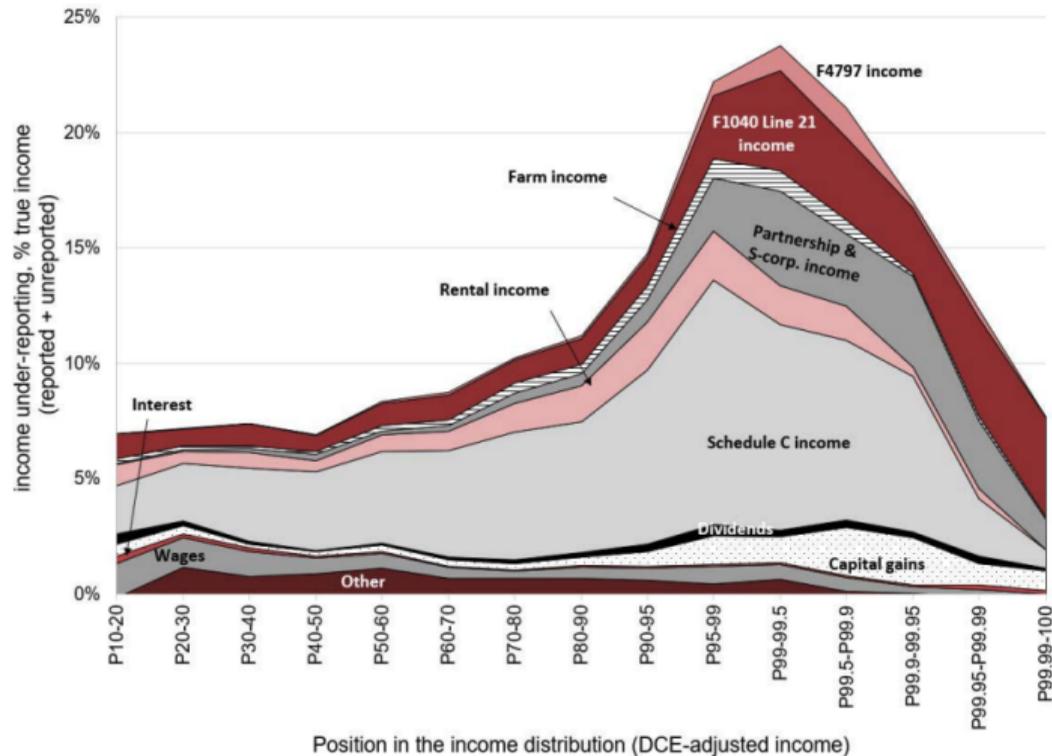
- In U.S., upper-middle class professionals more likely to hide income/wealth through C-corps and partnerships (Smith, Zidar, Zwick 2021)

- ML methods vs. tax authority's audit-based approaches (Feinstein 1991, maximum likelihood) to uncover tax evasion?

Source: Guyton et al. (2021): "Tax Evasion at the Top of the Income Distribution: Theory and Evidence," NBER Working Paper, No. 28542.

FINANCING EXAMPLE: MORTGAGE CONTRACT OPTIMIZATION

Source: Zhang (2022): "Closing Costs, Refinancing, and Inefficiencies in the Mortgage Market," *mimeo*, Rice University.

- Figure: capitalizing upfront costs into the loan amount $\implies$ disposable income $\downarrow$ b/c effective interest rate $\uparrow$

- Actively refinancing borrowers gain 3% of the loan amount relative to inactive borrowers

- Liquidity constraints $\implies$ poor get poorer

- Other examples in mortgage context: ARMs vs. FRMs, early repayment, etc.

COMMENT #1: IDEAS FOR DEALING WITH MISSING DATA

- SOEP data has many features in common with PSID in the U.S.
  - Longitudinal panel dimension: SOEP starts at age 39, PSID at 12 for most variables
  - Covers tax/transfer system, household structure, labor market outcomes
  - Panel gaps also makes this look like repeated cross-sectional data like SCF

- **Be more precise on what SOEP might have that PSID does not**

- Authors propose using standard forward simulation, backfilling, interpolation methods
  - More details needed here – can validity be checked through out-of-sample merge?

- **Bayesian imputation methods as an alternative approach**
  - Gibbs sampler/EM approach: Tanner & Wong (1987); Kong, Liu, Wong (1994 *JASA*)
  - Identifying restrictions: Linero & Daniels (2018 *Stat Sci*) e.g. we know some missing data is due to attrition $\implies$ monotonicity conditions

COMMENT #2: POSITION WITHIN LIFE-CYCLE INCOME RESEARCH

- **Suppose all the proposed study goals are met, what would be main contribution relative to the life-cycle income literature?**

- Guvenen et al. (2021 *ECMA*); Guvenen et al. (2022 *AEJ Applied*) use SSA earnings files

  ▶ Comprehensive coverage, but no way to credibly map earnings to disposable income

  ▶ Many assumptions required...but isn't the same true here? e.g. you need mortgage payment calculator, pension contribution schedule, etc.

  ▶ Tradeoffs between survey and administrative data

- Kaplan & Violante (2014) provide set of building blocks to model how life-cycle profiles differentially respond to income shocks depending on **disposable income**

  ▶ "Wealthy hand-to-mouth" is another way of saying little cash on hand

  ▶ Collapse tax/transfer schedules into single pension term and proportional τ

  ▶ But still, can do complex counterfactuals like estimate MPC out of tax rebates

COMMENT #3: USING ML TO EVALUATE OPTIMAL TAX MODELS

- ML potentially powerful tool to predict redistributive capacity of various policy initiatives
  - Example: IHT reforms in Germany aimed at closing exemptions for corporate assets
  - Given attrition and relatively short panel, can application to this dataset produce any **long-run or intergenerational forecasts**? (wide confidence intervals)

- **Perhaps this paper could be a vehicle for bringing ML to the tax authority**
  - Main application: automate auditing process and cutdown on administrative costs
  - Merge SOEP panel to tax returns to gauge extent of unreported income or misoptimization

- Standard Mirrlees model:
  - Govt. has (income) redistribution motives + faces revenue constraint
  - But cannot observe ability to pay ("types"), just noisy measure of earnings
  - ML brings govt. closer to perfectly discriminating monopolist $\longrightarrow$ approach first-best allocation, relax incentive compatibility constraints

ROADMAP FOR FUTURE WORK

- Solve missing data problems due to gaps + attrition
  - ⟶ standard imputation methods vs. Bayesian approaches

- Clarify what is new here on the data side (SOEP vs. PSID/SCF)

- Quantify advantages to moving away from structural macro models
  - Apply workhorse models of life-cycle dynamics and compare to ML

- Counterfactual analysis in simulation model can be used to inform...
  1. Moving from second-best tax regime closer to first-best outcomes predicted by optimal tax models ⟶ help govt. figure out spanning set of variables needed to identify ability to pay
  2. Degree of under-reporting of earnings ⟶ subject to constraints that panel survey may be censored for high-income respondents

THANKS!