Inattention to and Misperception of Taxes

Dmitry Taubinsky

May 2, 2022

Tax misperceptions

Core assumption in public finance:

- All that matters are the financial incentives that are in place, and not how they are implemented and framed
  - “Implementation invariance”
  - Underlies mechanism design approaches to taxation

But recent empirical work challenges this assumption:

- People may be inattentive, confused, employ heuristics, etc.

Proof-of-concept illustration

Abeler and Jager (2015): Real effort experiment where labor is “taxed”

Holding incentives constant, consider

1. tax schedules described by 22 tax rules
2. tax schedules instead described by of one transparent tax rule

Finding: People less responsive to incentives in complex condition

Tax salience

Visibility of the tax matters:

1. Road tolls (Finkelstein 2009)
2. Sales taxes (Chetty et al. 2009; Goldin and Homonoff 2013; Feldman and Ruffle 2015; Taubinsky and Rees-Jones 2018; Bradley and Feldman 2020)
    - People more responsive to sales taxes included in the posted price
3. Income tax rules (Miller and Mumford 2015)
    - Taxpayers ignore how CDCC interacts with CTC, in a way that leads to suboptimal claiming behavior

Confusion (or limited knowledge)

- Many tax-payers don't know their marginal tax (Fuji and Hawley 1988; Blaufus et al. 2013; Gideon 2015; Rees-Jones Taubinsky 2020)
- And differ vastly in their understanding of EITC rules (Chetty, Friedman, Saez 2013)

Confusion (or limited knowledge)

- Many tax-payers don't know their marginal tax (Fuji and Hawley 1988; Blaufus et al. 2013; Gideon 2015; Rees-Jones Taubinsky 2020)
- And differ vastly in their understanding of EITC rules (Chetty, Friedman, Saez 2013)

Specific example:

- Feldman et al (2013) use a regression discontinuity design with CTC to show that loss of CTC reduces reported wage income
- However, CTC does not actually affect marginal tax rates, and thus the returns to labor supply
- $\Rightarrow$ People might confuse average and marginal tax rates

Outline for remainder of talk

Zoom in on measurement and implications of

1. Tax salience
2. Income tax misperceptions

Tax salience

Chetty et al.: Empirical Framework

- Economy with two goods, x and y
- Prices: Normalize the price of y to 1 and let p denote the (fixed) pretax price of x.
- Taxes: y untaxed, x subject to an ad valorem sales tax τ (not included in posted price)
  - Tax-inclusive price of x is $q = (1 + \tau)p$
- Let demand for good x be denoted by $x(p, \tau)$

Chetty et al.: Empirical Framework

- If agents optimize fully, demand should only depend on the total tax-inclusive price: $x(p, \tau) = x((1+\tau)p, 0)$
- Full optimization implies price elasticity equals gross-of-tax elasticity:

$$\varepsilon_{x,p} \equiv -\frac{\partial \log x}{\partial \log p} = \varepsilon_{x,1+\tau^s} \equiv -\frac{\partial \log x}{\partial \log(1+\tau)}$$

- To test this hypothesis, log-linearize demand fn. $x(p, \tau)$ to obtain estimating equation:

$$\log x(p, \tau) = \alpha + \beta \log p + \sigma\beta \log(1+\tau)$$

- σ measures degree to which agents under-react to the tax:σ

Chetty et al.: Two Empirical Strategies

Two strategies to estimate σ:

1. **Manipulate tax salience:** make sales tax as visible as pre-tax price

   - Effect of intervention on demand:

   $$v = \log x((1+\tau)p, 0) - \log x(p, \tau)$$

   - Compare to effect of equivalent price increase to estimate σ:

   $$(1-\sigma) = -\frac{v}{\varepsilon_{x,p} \log(1+\tau)}$$

2. **Manipulate tax rate:** compare $\varepsilon_{x,p}$ and $\varepsilon_{x,1+\tau}$

   $$\sigma = \varepsilon_{x,1+\tau} / \varepsilon_{x,p}$$

Chetty et al.: Strategy 1

- Experiment manipulating salience of sales tax implemented at a supermarket that belongs to a major grocery chain
  - 30% of products sold in store are subject to sales tax
  - Posted tax-inclusive prices on shelf for subset of products subject to sales tax (7.375% in this city)
- Data: Scanner data on price and weekly quantity sold by product

Chetty et al.: Strategy 1

Chetty et al.: Research Design

- Quasi-experimental difference-in-differences
- Treatment group:
  - *Products*: Cosmetics, Deodorants, and Hair Care Accessories
  - *Store*: One large store in Northern California
  - *Time period*: 3 weeks (February 22, 2006 – March 15, 2006)
- Control groups:
  - *Products*: Other prods. in same aisle (toothpaste, skin care, shave)
  - *Stores*: Two nearby stores similar in demographic characteristics
  - *Time period*: Calendar year 2005 and first 6 weeks of 2006

Experimental results

Effect of Posting Tax-Inclusive Prices: Mean Quantity Sold

TREATMENT STORE

Period	Control Categories	Treated Categories	Difference
Baseline	26.48 (0.22)	25.17 (0.37)	-1.31 (0.43)
Experiment	27.32 (0.87)	23.87 (1.02)	-3.45 (0.64)
Difference over time	0.84 (0.75)	-1.30 (0.92)	DD_{TS} = **-2.14** (0.64)

CONTROL STORES

Period	Control Categories	Treated Categories	Difference
Baseline	**30.57** **(0.24)**	**27.94** **(0.30)**	**-2.63** **(0.32)**
Experiment	**30.76** **(0.72)**	**28.19** **(1.06)**	**-2.57** **(1.09)**
Difference over time	**0.19** **(0.64)**	**0.25** **(0.92)**	DD_{CS} **= 0.06** **(0.90)**
		DDD Estimate	***-2.20*** (0.58)

Implies $\sigma = 0.35$

Documenting Salience Mechanism

- Concern with posting tax inclusive prices: may have influenced behavior through various channels besides salience
- Common problem in field experiments termed “Hawthorne effects”
- How does left-digit bias interact with this?
- Difficult to rule out all mechanisms, but helpful to present evidence that mechanism of interest is very powerful

Chetty et al.: Strategy 2

- Compare effects of price changes and tax changes
- Alcohol subject to two state-level taxes in the U.S.:
  - Excise tax: included in price
  - Sales tax: added at register, not shown in posted price
- Exploiting state-level changes in these two taxes, estimate σ
  - Addresses concern that experiment may have induced a Hawthorne effect

Results

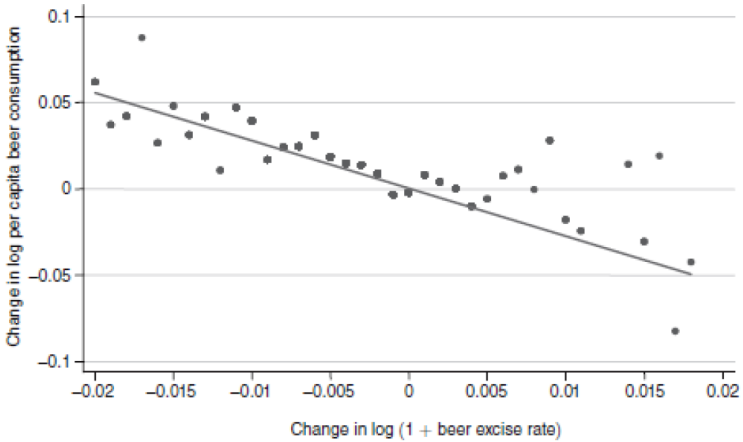


FIGURE 2A. PER CAPITA BEER CONSUMPTION AND STATE BEER EXCISE TAXES

Results

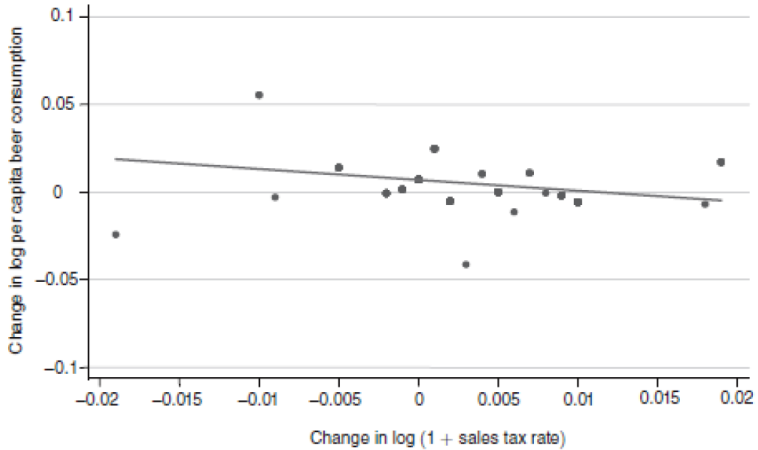


FIGURE 2B. PER CAPITA BEER CONSUMPTION AND STATE SALES TAXES

Results

Effect of Excise and Sales Taxes on Beer Consumption

Dependent Variable: **ΔLog(per capita beer consumption)**

	Baseline	Bus Cyc, Alc Regs.	3-Year Diffs	Food Exempt
	(1)	(2)	(3)	(4)
ΔLog(1+Excise Tax Rate)	**-0.87**	**-0.89**	**-1.11**	**-0.91**
	(0.17)***	(0.17)***	(0.46)**	(0.22)***
ΔLog(1+Sales Tax Rate)	**-0.20**	**-0.02**	**-0.00**	**-0.14**
	(0.30)	(0.30)	(0.32)	(0.30)
Business Cycle Controls		x	x	x
Alcohol Regulation Controls		x	x	x
Year Fixed Effects	x	x	x	x
F-Test for Equality of Coeffs.	0.05	0.01	0.05	0.04
Sample Size	1,607	1,487	1,389	937

Source: Chetty, Looney, and Kroft (2009)

Digging deeper

Taubinsky and Rees-Jones (2018)

Morrison and Taubinsky (forthcoming)

Two key facts:

- People pay more attention to bigger taxes
- There is a lot of heterogeneity: Some consumers actually over-react!

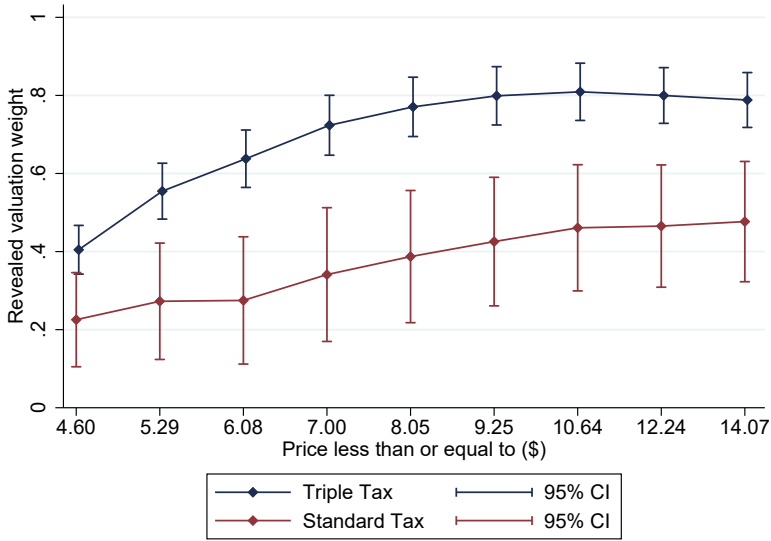
Morrison and Taubinsky: Sample and decisions overview

- Online experiment with demographically diverse sample ($N = 1545$) from the 45 states with positive sales taxes
  - Approximates US population on basic demographics
  - Panel provided by *ClearVoice Market Research*
- Series of *real* purchase decisions about common household products

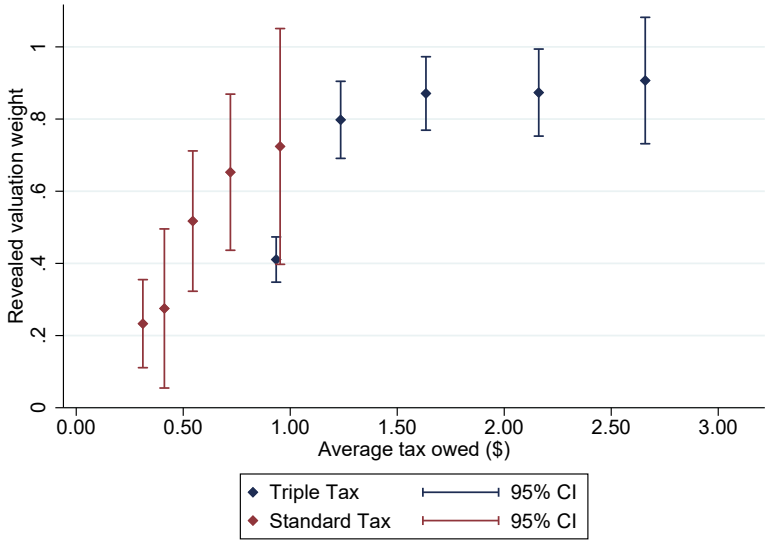
3 “stores” and 3 different items for each individual

- Store A: No sales tax
- Store B: “The standard sales tax that you pay in your city of residence on standard, non-tax-exempt items.”
- Store C: “Triple the standard sales tax that you pay in your city of residence on standard, non-tax-exempt items.”

$\mathbb{E}[\sigma]$ (weight on tax) for prices at or below a cutoff



$\mathbb{E}[\sigma]$ (weight on tax) as a function of absolute tax size



Bounds on dispersion

	Standard	Triple	Standard-Triple
Variance (Lower Bound)	0.84	0.72	0.81
	[0.43]	[0.55]	[0.39]
Supremum (Lower Bound)	2.24	1.71	0.86
	[1.51]	[1.48]	[0.26]

Notes: 5% confidence bounds in brackets

Implications for welfare

Assumptions for welfare analysis (application of Bernheim and Rangel 2009)

1. Utility depends only on the final consumption bundle
2. Consumers optimize perfectly in the absence of taxes
3. ~~Consumers optimize perfectly in the *presence* of taxes~~

$\sigma \neq 1$ corresponds to "mistake," and may arise from, e.g., inattention, imperfect perception, incorrect beliefs, rounding heuristics, etc.

Implications for welfare

Recall again “sin tax” formulas:

$$dW_\theta = \underbrace{-\gamma_\theta \frac{dx_\theta}{dt} dt}_{\text{Bias correction}} \quad \underbrace{+t \frac{dx_\theta}{dt} dt}_{\text{Fiscal externality}}$$

Implications for welfare

Recall again “sin tax” formulas:

$$dW_\theta = \underbrace{-\gamma_\theta \frac{dx_\theta}{dt} dt}_{\text{Bias correction}} \quad \underbrace{+t \frac{dx_\theta}{dt} dt}_{\text{Fiscal externality}}$$

- Misperception of purchase value is $\gamma = (1 - \sigma)t$, where σ is the “attention weight on tax”

Implications for welfare

Recall again "sin tax" formulas:

$$dW_\theta = \underbrace{-\gamma_\theta \frac{dx_\theta}{dt} dt}_{\text{Bias correction}} \quad \underbrace{+t \frac{dx_\theta}{dt} dt}_{\text{Fiscal externality}}$$

- Misperception of purchase value is $\gamma = (1 - \sigma)t$, where σ is the "attention weight on tax"

- Welfare effect of increasing tax by dt is

$$\begin{aligned} dW_\theta &= -(1 - \sigma)t \frac{dx_\theta}{dt} dt + t \frac{dx_\theta}{dt} \\ &= \sigma t \frac{dx_\theta}{dt} dt \quad \text{(Chetty et al. 2009)} \end{aligned}$$

Implications for welfare

Recall again "sin tax" formulas:

$$dW_\theta = \underbrace{-\gamma_\theta \frac{dx_\theta}{dt} dt}_{\text{Bias correction}} \quad \underbrace{+t \frac{dx_\theta}{dt} dt}_{\text{Fiscal externality}}$$

- Misperception of purchase value is $\gamma = (1-\sigma)t$, where σ is the "attention weight on tax"

- Welfare effect of increasing tax by dt is

$$\begin{aligned} dW_\theta &= -(1-\sigma)t \frac{dx_\theta}{dt} dt + t \frac{dx_\theta}{dt} \\ &= \sigma t \frac{dx_\theta}{dt} dt \quad \text{(Chetty et al. 2009)} \end{aligned}$$

- But since $\frac{dx_\theta}{dt} \propto \sigma$, deadweight loss $\propto$ second moment of σ (Taubinsky&Rees-Jones, 2018)

Efficiency costs from small taxes

Proposition (TRJ 2018): For a small tax t (terms $O(t^3)$ negligible):

$$EB(t) \approx -\frac{1}{2}t^2 \left[\overbrace{E[\sigma|p,t]\frac{d}{dt}D(p(t),t)}^{\text{Chetty et al. (2009)}} + \overbrace{Var[\sigma|p,t]\frac{\partial}{\partial p}D(p,t)|_{p=p(t)}}^{\text{New}} \right]$$

Intuition: Heterogeneous mistakes $\Rightarrow$ consumers purchasing the good are *not* the ones who value it the most $\Rightarrow$ inefficient allocation

Corollary: Even when equilibrium quantity doesn't change because supply is inelastic ($\varepsilon_{S,p} = 0$), still have efficiency costs $EB(t) \approx Var(\sigma|p,t)\frac{\partial}{\partial p}D$.

More generally, salient taxes better whenever $Var(\sigma) > 0$ and supply inelastic relative to demand

Income tax misperceptions

Rees-Jones and Taubinsky (2020): Measuring "Schmeduling"

Liebman and Zeckhauser (2004) proposed two heuristic ways in which people might think about complicated price schedules

1. Ironing:
   1.1 People have a sense of the average price that they pay, but incorrectly think that the marginal price equals the average price
2. Spotlighting:
   2.1 People know marginal price, but assume that it doesn't change

RJT study these heuristics in the context of income taxes

Elicitation of tax schedules

*This next group of questions is about Fred, a hypothetical taxpayer who is very similar to you. Fred is your age, and has a lifestyle similar to yours. Fred filed his 2014 Federal Tax Return claiming **X** exemption(s) and **Y** filing status, like you did. Fred also claimed the standard deduction, like you did. However, Fred's tax computation is particularly simple, since all of his taxable income comes from his annual salary. He has no other sources of taxable income, and is not claiming additional credits or deductions.*

Elicitation of tax schedules

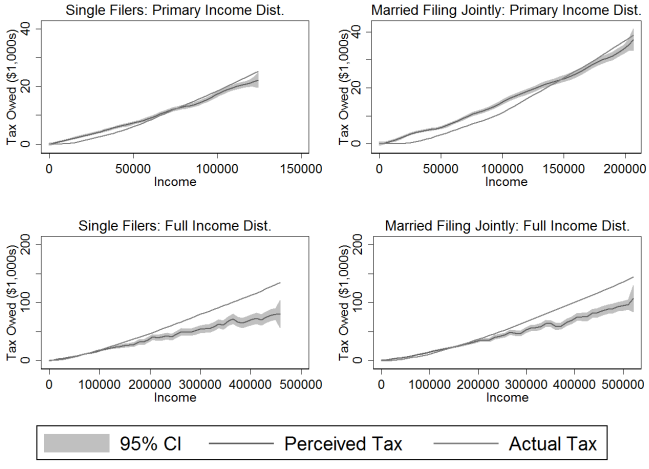
Question: If Fred's total income for the year were **X**, the total federal income tax that he has to pay would be...

Task: Respondents asked to repeatedly make this forecast for different, randomly drawn, amounts of income.

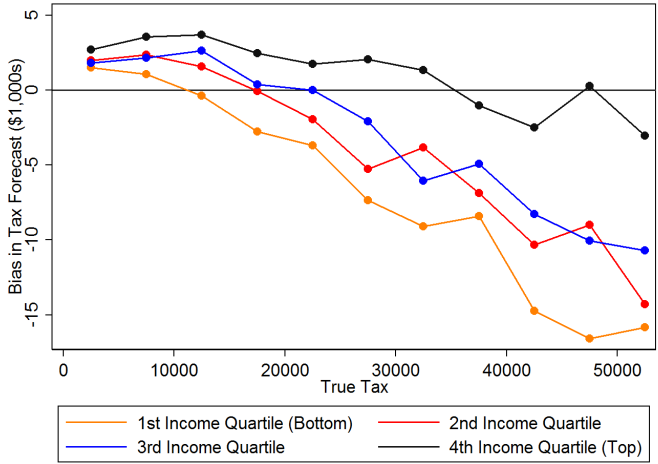
- 14 random draws
- 10 draws from mid-sampling distribution.
- Drawing from first four income brackets.
- 2 income amounts near their own income, which can be combined with random draws that fall within their own tax bracket to form "local" sample.

4,197 respondents, 58,758 forecasts

Local-polynomial estimates of $\tilde{T}(z)$

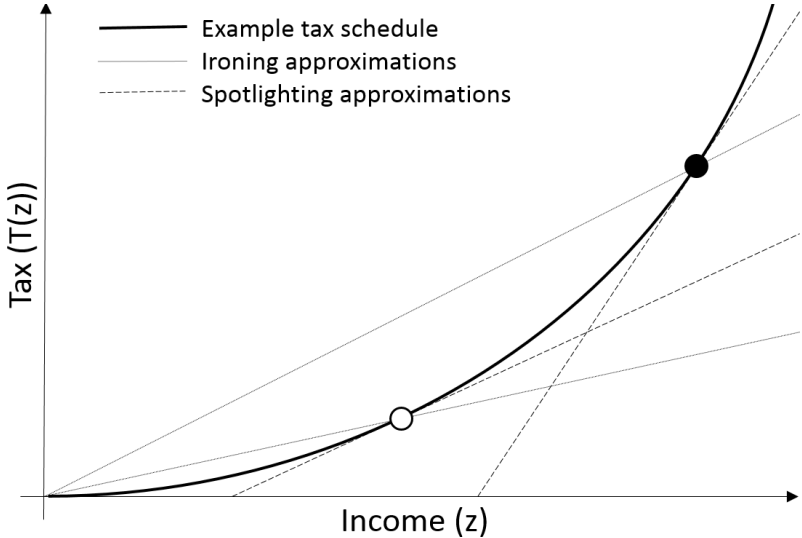


Bias by respondents' income



Five patterns in the data

1. Overestimation of taxes on poor, underestimation of taxes on rich.
2. Higher-income individuals exhibit less underestimation of taxes on rich and more overestimation on poor.
3. Individuals underestimate MTRs.
4. Perceived average slope of the *whole* tax schedule is too shallow.
5. Perceived average slope of the *whole* tax schedule increases with income.

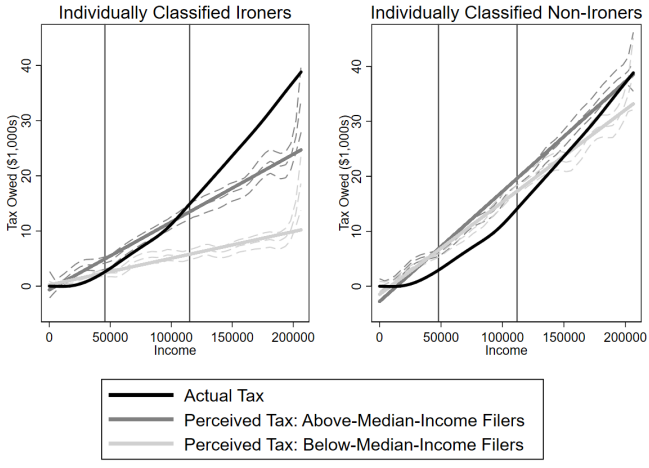


All consistent with Ironing, but only 5 (possibly 4) is consistent with Spotlighting.

Mixture model estimates

	(1)	(2)	(3)	**(4)**
γ_I: weight on ironing forecast	0.21***	0.29***	0.47***	**0.43*** **
	(0.037)	(0.052)	(0.048)	**(0.095)**
γ_S: weight on spotlighting forecast	-0.09*	-0.02	-0.03	**-0.02**
	(0.050)	(0.057)	(0.062)	**(0.076)**
Residual misperception function included	No	Yes	No	**Yes**
Income sampling distribution	Mid	Mid	Full	**Full**
Respondents	4197	4197	4197	**4197**
Forecasts	41970	41970	58758	**58758**

Mixture model in action



Simulated welfare gains from bias in standard Mirleesian model

Structural Elasticity ($\frac{1}{k}$)	Increase in Tax Rev. (%)	*Net Welfare Increase (%)* Low λ $\lambda = U'_{50}$	— $\lambda = \bar{U}'$	High λ $\lambda = U'_{90}$
1/2	3.7	3.2	3.4	3.6
1/3	2.5	2.2	2.3	2.5
1/4	1.9	1.7	1.8	1.9
1/5	1.6	1.4	1.4	1.5

$U = log\left(z - T(z) - \frac{(z/w)^{1+k}}{1+k}\right)$. All tax rev. spent on public good. Value of public good (λ) benchmarked to 50th, average (w/ income floor), or 90th percentile values of U'.

Reducing bias through simplification

Relevant application: quantifying gains/losses from simplification.

Example: Going to a “flat tax.”

- Consider a flat tax that would result in the same revenue, assuming no behavioral response.
- Approximately 11% tax rate.

This policy would:

- Mechanically change the redistribution achieved in the tax system.
- Eliminate effects of bias (since “ironers” understand linear taxes).

Simulated welfare consequences of flat tax

Structural Elasticity ($\frac{1}{k}$)	*All correct forecasters* Δ Tax Rev. (%)	Δ Welfare (%)	*43% ironers* Δ Tax Rev. (%)	Δ Welfare (%)
1/2	5.3	-9.9	3.0	-12.3
1/3	3.5	-11.6	2.0	-13.2
1/4	2.6	-12.4	1.5	-13.6
1/5	2.1	-12.9	1.2	-13.8

$U = log\left(z - T(z) - \frac{(z/w)^{1+k}}{1+k}\right)$. All tax revenue spent on public good. Value of public good (λ) benchmarked to 50th, average (w/ income floor), or 90th percentile values of U'.

Summary and conclusion

- When incentives are complex and/or opaque, people mis-react to them
- Welfare implications of mistakes are subtle
  - In representative agent framework, under-reaction increases efficiency
  - but heterogeneity lowers efficiency
  - and covariation with income matters as well
- Some questions for future work:
  - People seem to dislike and distrust complexity. Why? Implications?
  - Should we modify our standard welfare frameworks to accommodate distaste for confusing people?