Lecture I: Sufficient Statistics and Labor Supply

Alice Kügler and Attila Lindner

November 4 2019

Lecture outline

Sufficient statistics:

- Motivation and definitions
- Harberger's deadweight loss
- Framework for deriving formulae for welfare analysis
- Illustration of the approach:
  - Feldstein (1999) on total earned income elasticity
  - Saez (2001) on optimal income tax rates

This lecture follows the review by Chetty (2009a)

Classifications used

- **Reduced form**: estimate high-level behavioral elasticities that are qualitatively relevant for policy analysis, but do not provide quantitative welfare results
- **Structural**: estimate or calibrate primitives to make predictions about welfare
- **Sufficient statistic**: make predictions about welfare without estimating or specifying primitives

Sufficient statistics

- Reduced-form strategies: transparent and credible identification
- Structural models: complete models of economic behavior that allow for making predictions about counterfactual outcomes and welfare
- Combining these two approaches: develop formulae for the welfare consequences of policies that are functions of reduced-form elasticities rather than structural primitives

Sufficient statistics: conditional on the statistics in the formula, other statistics that can be calculated from the same sample provide no additional information about the welfare consequences of the policy

Harberger's deadweight loss

Harberger (1964), "The Measurement of Waste":

- Precedent to the modern literature on sufficient statistics
- Measures the excess burden of a commodity tax with a simple elasticity-based formula

Setup of the static general equilibrium model:

- Individual endowed with Z units of the numeraire y, with price normalized to 1
- J other consumption goods $x = (x_1, ..., x_J)$, with cost $c(x)$ of production
- The government levies tax t on good 1
- Vector of pre-tax prices for the produced goods $p = (p_1, ..., p_J)$
- For simplification, ignore income effects by assuming quasi-linear utility in y

The consumer takes prices as given and solves:

$$\max_{x,y} u(x_1, , ..., x_J) + y \quad s.t. \quad p \cdot x + tx_1 + y = Z$$

The representative firm takes prices as given and solves

$$\max_{x} p \cdot x - c(x)$$

With market clearing, $x^D(p) = x^S(p)$.

How to measure the efficiency cost of the tax t?

- Net loss in welfare from raising the tax rate and returning the tax revenue to the taxpayer through a lump-sum

Social welfare is the sum of the consumer's utility, producer profits and tax revenue:

$$W(t) = \{\max_x u(x) + Z - tx_1 - p(t) \cdot x\} + \{\max_x p(t) \cdot x - c(x)\} + tx_1$$

$$= \{\max_x u(x) + Z - tx_1 - c(x)\} + tx_1$$

- Harberger's simple solution for calculating the efficiency costs of tax changes:

$$\frac{dW(t)}{dt} = -x_1 + x_1 + t\frac{dx_1}{dt} = t\frac{dx_1(t)}{dt}$$

- The effect of the tax on equilibrium quantity in the taxed market, $\frac{dx_1(t)}{dt}$, is a sufficient statistic

- When calculating $\frac{dW}{dt}$, behavioral responses $\frac{dx}{dt}$ can be ignored because of envelope conditions from consumer and firm optimization
  - Social welfare has already been optimized by individuals and firms
- Prices can also be ignored
  - Price changes redistribute income from producers to consumers without changing aggregate surplus

Alternative: estimate full structural demand system

- Estimate a J good demand and supply system to recover the utility function and cost function

- Almost Ideal Demand System (Deaton and Muellbauer, 1980):

$$x_i^D(p) = (\alpha_i^D - \beta_i^D \alpha_0^D) + \sum_{j=1}^{J} \gamma_{ij}^D \log p_j +$$
$$+ \beta_i^D \left[log x + \sum_{k=1}^{J} \alpha_k^D \log p_k - \frac{1}{2} \sum_{k=1}^{J} \sum_{j=1}^{J} \gamma_{kj}^D \log p_k \log p_j \right] +$$

- Estimate the model with maximum likelihood
- But estimating such a simultanous equation model is quite challenging
- At the very least $2J$ instruments needed to identify the slopes of supply and demand curves

- When calculating $\frac{dW}{dt}$, behavioral responses $\frac{dx}{dt}$ can be ignored because of envelope conditions from consumer and firm optimization
    - Social welfare has already been optimized by individuals and firms
- Prices can also be ignored
    - Price changes redistribute income from producers to consumers without changing aggregate surplus

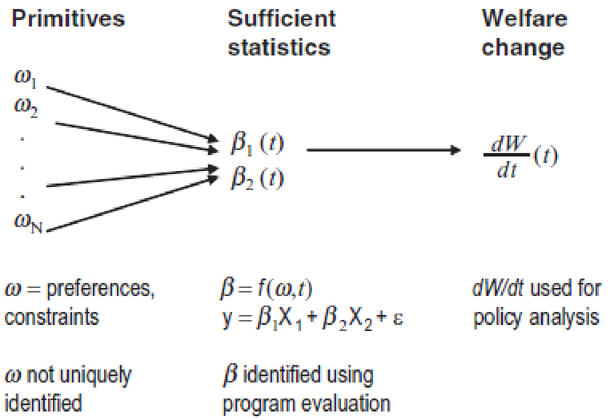
Limitations when calculating (only) $\frac{dx_1(t)}{dt}$:

- For example, $\frac{dx_1(t)}{dt}$ as a sufficient statistic does not allow for pre-existing distortions in the other markets
- Cannot be used to evaluate counter-factual policy changes

Benefits of the approach, for example, in the case of:

- Heterogeneous preferences across individuals
- Discrete choice of only one of the J products

The sufficient statistics approach



Advantages of sufficient statistics

1. Simpler to implement empirically

   - less data and variation are needed to identify marginal treatment effects (MTEs) than to fully identify a structural model
   - in structural models primitives are often calibrated rather than formally estimated using microdata, because of estimation challenges
   - no need to fully calibrate a structural model: particularly useful with large set of primitives but small set of MTEs needed for welfare evaluation, such as for models with heterogeneity and discrete choice

Advantages of sufficient statistics (continued)

1. Weaker assumptions required
   - the identification of structural models often requires strong assumptions given available data and variation
   - for sufficient-statistics formulae, unnecessary to identify all primitives
   - results are more transparent and empirically credible
2. Implementation possible even with uncertainty about the positive model that generates observed behavior
   - welfare analysis based on a structural model may be impossible
   - e.g., Chetty et al. (2009) derive formulae for the deadweight cost of taxation in a model in which agents make arbitrary optimization errors with respect to taxes

Limitations of sufficient statistics

1. A new sufficient-statistic formula must be derived for each question
   - by contrast, if one had estimated the structural primitives of the model used to derive the formula, different policy simulations could be conducted
   - for some questions, it may be difficult to derive a sufficient-statistic formula, and a structural approach may be the only feasible option
2. Sufficient-statistic formulae are more easily misapplied than structural methods
3. One can draw policy conclusions from a sufficient-statistic formula without assessing the validity of the model upon which it is basedA new sufficient-statistic formula must be derived for each question

Comparing structural modeling and sufficient statistics

- Precision of the extrapolation in out-of-sample predictions
    - One can use the sufficient statistic approach for out-of-sample predictions by estimating MTEs as a function of the policy instrument and making a statistical extrapolation
    - structural methods do not require such ad hoc extrapolations, as the primitive structure is by definition policy invariant
    - however, in practice, structural models often rely on extrapolations based on functional form assumptions (such as constant-elasticity utilities)
    - statistical extrapolations from sufficient statistics may be less reliable than extrapolations guided by an economic model that imposes restrictions on how behavior changes with policies
- Structural and sufficient-statistic methods can be combined to address the shortcomings of each strategy

Recent examples of structural, reduced-form, and sufficient-statistic studies

	Structural	Reduced form	Sufficient statistic
Taxation	Hoynes 1996	Eissa & Liebman 1996	Diamond 1998
	Keane & Moffitt 1998	Blundell et al. 1998	Feldstein 1999
	Blundell et al. 2000	Goolsbee 2000	Saez 2001
	Golosov & Tsyvinksi 2007	Meyer & Rosenbaum 2001	Goulder & Williams 2003
	Weinzierl 2008	Blau & Khan 2007	Chetty 2009
Social insurance	Rust & Phelan 1997	Anderson & Meyer 1997	Gruber 1997
	Golosov & Tsyvinski 2006	Gruber & Wise 1999	Chetty 2006a
	Blundell et al. 2008	Autor & Duggan 2003	Shimer & Werning 2007
	Einav et al. 2008b	Lalive et al. 2006	Chetty 2008
	Lentz 2009	Finkelstein 2007	Einav et al. 2008a
Behavioral models	Angeletos el at. 2001	Genesove & Mayer 2001	Bernheim & Rangel 2008
	İmrohoroğlu et al. 2003	Madrian & Shea 2001	Chetty et al. 2009
	Liebman & Zeckhauser 2004	Shapiro 2005	
	DellaVigna & Paserman 2005	Ashraf et al. 2006	
	Amador et al. 2006	Chetty & Saez 2009	

A general sufficient statistics framework in six steps

Think about government policies as levying a tax t to finance a transfer $T(t)$, with applications, for example, in:

- redistributive taxation - a transfer to another agent
- social insurance - a transfer to another state of the world
- excess-burden calculations (as above) - to finance a public good

Step 1: specify the structure of the model

- For simplicity, a static model with a single agent
- Vector of choices for the representative agent $x = (x_1, ..., x_J)$
- A unit tax t is levied on choice x_1
- The agent faces $M \leq J$ constraints $G_1(x, t, T), ..., G_M(x, t, T)$, such as budget constraints, restrictions on insurance or borrowing, hours constraints, etc.
- The agent maximizes utility taking t and T as given:

$$\max U(x) \quad s.t. \quad G_1(x, t, T) = 0, ..., G_M(x, t, T) = 0$$

- Solved so that social welfare is a function of the policy instrument:

$$W(t) = \max_x U(x) + \sum_{m=1}^{M} \lambda_m G_m(x, t, T)$$

Step 2: express $\frac{dW}{dt}$ in terms of multipliers

- Differentiate W, and use the envelope conditions from optimization in the private sector

$$\frac{dW}{dt} = \sum_{m=1}^{M} \lambda_m \left\{ \frac{\partial G_m}{\partial T} \frac{dT}{dt} + \frac{\partial G_m}{\partial t} \right\}$$

- λ_m are the Lagrange multipliers, for example, the marginal value of relaxing the budget constraint

- the government's budget constraint gives $\frac{dT}{dt}$, and $\frac{\partial G_m}{\partial T}$ and $\frac{\partial G_m}{\partial t}$ can be calculated mechanically
  - in the Harberger example, $T(t) = tx_1$, so that $\frac{dT}{dt} = x_1 + t\frac{dx_1}{dt}$, and $\frac{dG_1}{dT} = 1$ and $\frac{dG_1}{dt} = -x_1$, and $\frac{dW}{dt} = \lambda_1 t \frac{dx_1}{dt}$

Step 3: substitute multipliers by marginal utilities

- Exploit restrictions from the agent's first order conditions to recover λ_m multipliers
- Optimization leads to:

$$u'(x_j) = -\sum_{m=1}^{M} \lambda_m \frac{\partial G_m}{\partial x_j}$$

- Assumption 1: x_1 and t, and x_J and T enter every constraint interchangeably
- Then we can write:

$$\frac{dW}{dt} = k_T \frac{dT}{dt} u'(x_J(t)) - k_t u'(x_1(t))$$

- This distills local welfare analysis to recovering a pair of marginal utilities

Step 4: recover marginal utilities from observed choices

- Recover the marginal utilities from choice data using the model structure specified in step 1
- Different ways to do this, but usually marginal utilities are elements in the first-order conditions for choices

Step 5: empirical implementation

- Derivatives may require holding different variables fixed:
  - for instance, the Harberger formula measures the total derivative $\frac{dx_1}{dt}$, which incorporates general equilibrium effects and price changes in all markets
  - If the elasticity called for by the formula cannot be credibly identified, it may be possible to use approximations
- Assessing the efficiency cost of a discrete policy change:
  - by estimating inputs as non-parametric functions of the policy instrument, or instead (if insufficient power for estimation), reduced-form studies often estimate local average treatment effects
  - then integrate the welfare change function between any two tax rates t_1 and t_2 within the support of observed policies
  - extrapolate the estimates out of sample to make predictions about welfare changes outside the observed support

Step 5: empirical implementation (continued)

Two options:

1. Bound the average welfare gain over the observed range:

$$W(t_2) - W(t_1) \;=\; \int_{t_1}^{t_2} \frac{dW}{dt} dt \;=\; \int_{t_1}^{t_2} t \frac{dx_1}{dt} dt$$

$$\Rightarrow t_1 \frac{\Delta x_1}{\Delta t} \geq \overline{dW/dt} \geq t_2 \frac{\Delta x_1}{\Delta t}$$

2. Approximate $x_1(t)$ to calculate $\overline{dW/dt}$

   - For example, by approximating that $\frac{dx_1}{dt}$ is constant over the observed range

Step 6: model evaluation

- There are no theory-free statements about welfare
- Important to assess the validity of these assumptions:

1. Try to falsify the central assumptions underlying the sufficient statistics:
   - For example, the Harberger model assumes that individuals treat prices and taxes identically, and choose based on the total price of a good: Chetty et al. (2009) test this assumption by comparing price and tax elasticities of demand
2. Identify at least one vector of structural parameters ω that is consistent with the sufficient statistics estimated in step 5

Application: Feldstein (1999)

Efficiency costs of taxation in a model with multidimensional labor-supply choices:

- In addition to hours of work, affect choice of training, effort, occupation, but also tax avoidance and evasion
- Sufficient statistic for calculating welfare loss: elasticity of taxable income (ETI) with respect to a tax rate
- Total taxable income is $TI = \sum_{j=1}^{J} w_j x_j - e$, with e of earnings sheltered from the tax authority
- Excess burden of a tax calculated by assuming that the tax revenue is returned to the individual as a lump-sum transfer T(t) (as above)

Feldstein (1999): model

- Taxpayers problem:

$$\max u(x, e) = (1 - t) \left[\sum_{j=1}^{J} w_j x_j - e \right] + e - g(e) - \sum_{j=1}^{J} \psi_j(x_j)$$

- FOCs:

$$(1 - t) w_j = \psi_j'(x_j)$$

$$t = g'(e)$$

Feldstein (1999): model

- The Social planner problem:

$$\max u(c, x, e) = c - g(e) - \sum_{j=1}^{J} \psi_j(x_j)$$

$$s.t. \;\; T(t) = t \cdot TI$$

$$s.t. \;\; G_1(c, x, t) = T + (1 - t)TI + e - c$$

- Social welfare:

$$W(t) = (1 - t)TI + e + g(e) - \sum_{j=1}^{J} \psi_j(x_j) + tTI$$

Feldstein (1999): model

- Totally differentiate social welfare $W(t)$, and recover marginal utilities by exploiting first-order conditions (as in step 4):

$$\frac{dW(t)}{dt} = \frac{dTI}{dt} + \frac{de}{dt}(1 - g'(e)) + \sum_{j=1}^{J} \psi_j'(x_j)\frac{dx_j}{dt}$$

since $\psi_j'(x_j) = (1-t)w_j$ implies
$\sum_j \psi_j'(x_j)\frac{dx_j}{dt} = \sum_j (1-t)w_j \frac{dx_j}{dt} = (1-t)\frac{d(TI+e)}{dt}$

- As a result:

$$\frac{dW}{dt} = t\frac{dTI}{dt}$$

- Alterantively, differentiate $W(t)$ and use the envelope conditions, that behavioral responses have no first-order effect on private surplus

Feldstein (1999): implications

- Only need to measure the response of ETI to the change in the tax rate, rather than changes in hours, occupation or avoidance behaviors
- Need, however, to investigate the structural parameters (step 6)

  For example, for ETI effects Chetty (2009b) points out:

  - Marginal social cost of tax avoidance may not be equal to the tax rate at the optimum, violating the first order condition
  - Then, if no large resource cost of avoidance and sheltering, changes in e have little efficiency cost: only the real labor-supply response matters for deadweight loss
- Sufficient statistics approaches are not model-free

Application: Saez (2001)

The literature on optimal taxation focused on the optimal progressivity of non-linear income tax system

- Seminal work by Mirrlees (1971) that formalized the optimal tax problem as functions of primitive parameters, giving little insight into the forces that determine optimal taxes
  - For an overview of the evolution of literature on optimal taxation, see the Handbook chapter by Piketty and Saez (2013)
- Saez (2001) expressed these optimality conditions in the Mirrlees model in terms of empirically estimable sufficient statistics, building on the work by Diamond (1998)

Saez (2001) framework

- Start by analyzing the optimal tax rate on top incomes
- Linear tax τ (note change in notation) on earnings above threshold $\bar{z}$
- For a given $\bar{z}$, individuals maximize utility:

$$u(c, l) = c - \phi(l) \;\; s.t. \;\; G_1(c, l) = (1 - \tau) max(wl - \bar{z}, 0) + \bar{z} - c = 0$$

- Pre-tax earnings are $z = wl$, and $c(w, \tau)$ and $l(w, \tau)$ are optimal choices, and $z(w, \tau) = wl(w, \tau)$ is the optimized earnings function
- Tax revenue generated by the top bracket is $R = \tau(z_m(\bar{z}) - \bar{z})$

Saez (2001): social planner optimum

- The social planner maximizes a weighted average of individuals' utilities, with social welfare weights $\tilde{G}(u)$ reflecting the redistributive preferences of the planner:

$$W = \left\{ \int_0^\infty \tilde{G}(u(c(w,\tau), wl(w,\tau)))dF(W) \right\} + \tau(z_m(\bar{z}) - \bar{z})$$

- Individuals with income below $\bar{z}$ are unaffected by the tax increase
- Envelop condition: behavioral responses $\frac{\partial l}{\partial \tau}$ have no first-order effect on private surplus, as $wu_c(w,\tau) = \psi'(l(w,\tau))$

Saez (2001): welfare effects

$$\frac{dW}{d\tau}(\tau) = \underbrace{-(z_m(\bar{z}) - \bar{z})\bar{g}}_{\text{loss to top tax bracket}} + \left[\underbrace{(z_m(\bar{z}) - \bar{z})}_{\text{mechanical gain}} + \underbrace{\tau \frac{dz_m}{d\tau}}_{\text{behavioral response}} \right]$$

- $\bar{g}$ is the social value of giving \$1 more income to individuals in the top bracket relative to the value of public expenditure.:

$$\bar{g} = \frac{\int_{\bar{w}}^{\infty} \tilde{G}_u(u)(z - \bar{z})dF(W)}{\int_{\bar{w}}^{\infty}(z - \bar{z})dF(w)}$$

- If $\bar{g} = 1$: equal weight of the consumption of individuals taxed and public expenditure
- If $\bar{g} < 1$: loss to individuals in the top bracket from having to pay more taxes (captured by the first term above)

Saez (2001): sufficient statistics

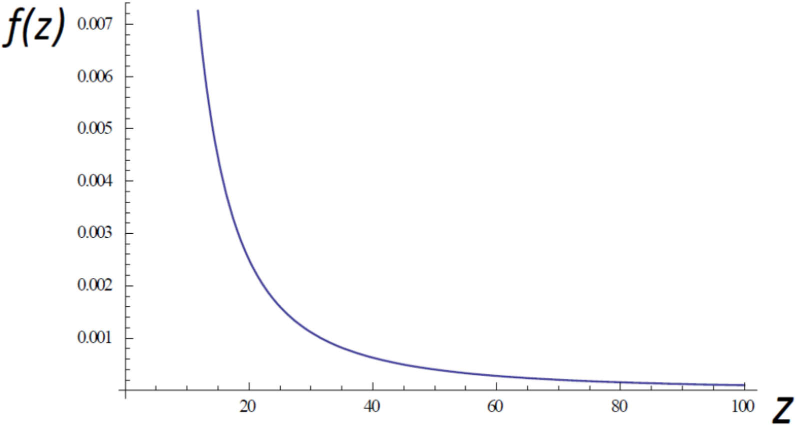
Three parameters together are sufficient statistics for the welfare gain of increasing top income-tax rates:

- The **effect of taxes on earnings** $\frac{dz_m}{d\tau}$, that quantifies the tax distortions
- The **shape of the earnings distribution** $z_m(\bar{z})$, capturing the mass of individuals whose behavior changes because of the tax
- The **marginal social-welfare weight** $\bar{g}$, which measures the planner's redistributive preferences
  - $\bar{g}$ would be the relevant marginal utility for step 4, but taken as external to choices here
  - $\bar{g}$ is determined by the shape of the earnings distribution and the exogenously specified social-welfare function

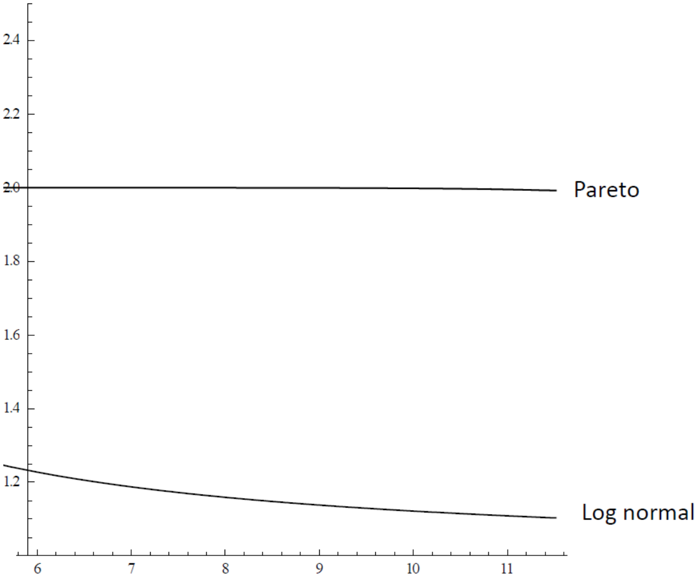
Saez (2001): sufficient statistics

- No need to identify preferences ψ or the shape of the skill distribution $F(w)$ to calculate $\frac{dW}{d\tau}$
- Arbitrary heterogeneity across skill types in preferences possible without changing the formula
- Disadvantage, however, that the three sufficient parameters are endogenous to τ
  - Level of earnings and the social planner weight on top earners likely to decrease with τ
  - Depending on the shape of the $\psi(l)$ function $\frac{dz_m}{d\tau}$ may vary with τ
- Thus, $\frac{dW}{d\tau}(\tau)$ measures only the marginal welfare gain at a given τ, and needs to be estimated at all values of τ to find the optimal tax rate τ^* that maximizes W

Pareto distribution, with Prob$[z > \bar{z}] = (constant) \cdot \bar{z}^{-a}$, for $a > 1$



The value of $z_m/\bar{z}$, comparing Pareto and log normal distributions



Saez (2001): U.S. earnings data

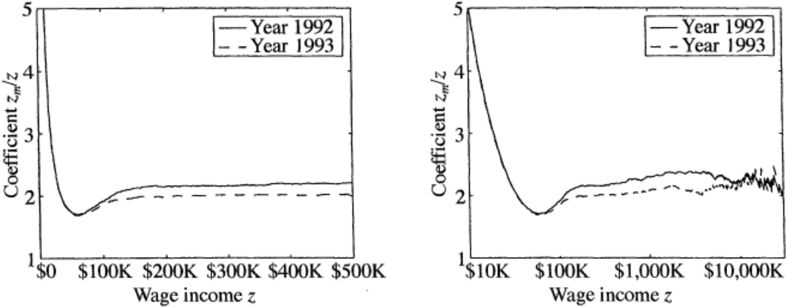


FIGURE 2

Ratio mean income above z divided by z, z_m/z, years 1992 and 1993

Saez (2001): deriving an optimal top-income tax formula

- Observe that the ratio $\frac{z_m(\bar{z})}{\bar{z}}$ is approximately constant in the upper tail of the empirical distribution of earnings in the U.S.:
  - Well-described by a Pareto distribution with parameter a, so that $\frac{z_m(\bar{z})}{\bar{z}} = \frac{a}{a-1}$ for all $\bar{z}$
- Rearranging the expression for $\frac{dW}{d\tau}(\tau)$ above and observing that the optimal tax rate τ satisfies $\frac{dW}{d\tau}(\tau) = 0$ gives

$$\frac{\tau^*}{1-\tau^*} = \frac{1-\bar{g}}{a\varepsilon}$$

with the taxable income elasticity in the top bracket

$$\varepsilon = \frac{dz_m}{d(1-\tau)}\frac{1-\tau}{z_m}$$

- This is an explicit formula for the optimal asymptotic top income-tax rate for an exogenous social-welfare weight $\bar{g}$

Optimal tax rates for high income earners

	Uncompensated elasticity = 0·0			Uncompensated elasticity = 0·2			Uncompensated elasticity = 0·5	
	Compensated elasticity			Compensated elasticity			Compensated elasticity	
	0·2	0·5	0·8	0·2	0·5	0·8	0·5	0·8
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Social marginal utility with infinite income g = 0								
Pareto parameter								
1·5	91	80	71	77	69	63	57	53
2	83	67	56	71	59	50	50	43
2·5	77	57	45	67	51	42	44	37
Panel B: Social marginal utility with infinite income g = 0·25								
Pareto parameter								
1·5	88	75	65	71	63	56	50	45
2	80	60	48	65	52	43	43	37
2·5	71	50	38	60	44	32	38	31

g is the ratio of social marginal utility with infinite income over marginal value of public funds. The Pareto parameter of the income distribution takes values 1·5, 2, 2·5. Optimal rates are computed according to formula (9).

Saez (2001): optimal tax at *any* income level

- The government chooses a schedule of total tax paid $T(z)$ on income z, to maximize social welfare subject to resource and incentive-compatibility constraints

- Applying envelope and perturbation arguments, gives the optimal tax schedule at all z:

$$\frac{T(z)}{1 - T(z)} = \frac{1}{\varepsilon(z) z h(z)} \int_z^{\infty} (1 - g(z'))h(z')dz'$$

- Determined by the same three parameters as before, that are endogenous to the tax regime

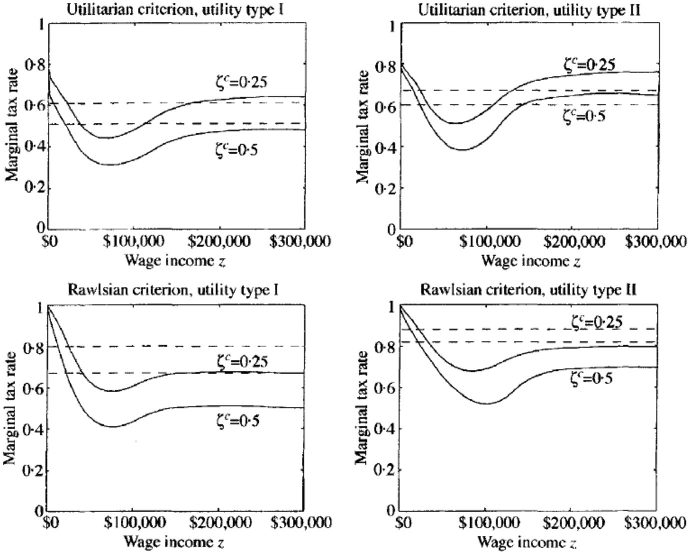
- By contrast to the top-income tax rate, no explicit formula can be derived for an arbitrary income level z by appealing to limit convergence results

Saez (2001): optimal tax at any income level

- Instead of an explicit formula, use the optimal tax schedule to evaluate the effect on welfare of perturbing the existing tax system $T(z)$
  - Saez (2001) matches the skill distribution $F(w)$ from empirically observed distributions in the current tax system, and simulates the optimal tax schedule in a calibrated model
  - Result: an inverse-U shaped optimal income-tax schedule, with a large lump-sum grant to non-workers and marginal rates from 50-80%
- More structure comes at the cost of strong assumptions for identification, and less confidence in the calculations for the optimal tax-schedule at any income z

Optimal tax simulations

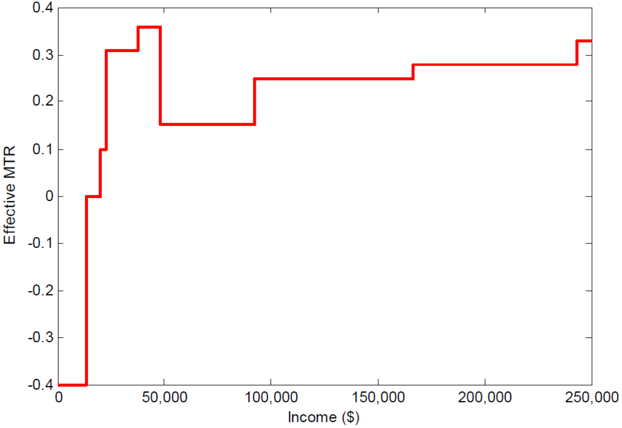
No income effect: Income effect:



Marginal rates in practice

- Most countries have a similar U-shaped pattern
- Low earners tend to face near-zero or negative marginal rates

E.g., effective rates for U.S. married couple with two children in 2013:



Concluding thoughts

- We covered the advantages and shortfalls of different approaches
- While some problems lead to easily implementable formulas (e.g. top linear tax rate), the derived sufficent statistics often very complicated (optimal non-linear taxes)
- Still the Saez (2001) approach revoultionzied public finances:
  - Tried quantify some theoretical discussions on the optimal taxes
  - Drew the attention to some key reduced form elasitcities (behavioral responses to taxes, moral hazar effects in UI)
- The sufficent statistic approach became a standard tool in public economics though still disliked in large fraction of the profession
  - People dislike relying too much on the "envelop theorem" (especially, that behavioral responses has only second order effects)
  - Often the elasticity estimated in the data is not the elasticity that is derived by theory (people are quite sloppy about this)

Bibliography I

Chetty, R., Looney, A., and Kroft, K. (2009). Salience and Taxation: Theory and Evidence. *American Economic Review*, 99(4): 1145–77.

Diamond, P. A. (1998). Optimal Income Taxation: An Example with a U-Shaped Pattern of Optimal Marginal Tax Rates. *American Economic Review*, 88(1): 83–95.

Feldstein, M. (1999). Tax Avoidance and the Deadweight Loss of the Income Tax. *Review of Economics and Statistics*, 81(4): 674–680.

Harberger, A. C. (1964). The Measurement of Waste. *American Economic Review*, 54(3): 58–76.

Mirrlees, J. A. (1971). An Exploration in the Theory of Optimum Income Taxation. *Review of Economic Studies*, 38(2): 175–208.

Piketty, T. and Saez, E. (2013). Optimal Labor Income Taxation. volume 5 of *Handbook of Public Economics*, pages 391 – 474. Elsevier.

Saez, E. (2001). Using Elasticities to Derive Optimal Income Tax Rates. *Review of Economic Studies*, 68(1): 205–229.