GR 6307
Public Economics and Development

1.1 Detour:
Applied Welfare Analysis

Michael Best

Fall 2019

Outline

Outline

Chetty (2009): 2 Competing Paradigms

- We can characterize 2 competing paradigms for policy evaluation & welfare analysis

1. **Structural:** specify a *complete* model, and estimate or calibrate the model's primitives.
2. **Reduced form:** prioritize clean *identification* of causal effects. Accept narrower scope of analysis.

- PRO structural / CON reduced form:
    1. Estimate statistics that are policy-invariant parameters of models.
    2. Can simulate effects of changes in policies on behavior and welfare.
- PRO reduced form / CON of structural approach:
    1. (quasi-)experimental research designs achieve compelling estimates of treatment effects
    2. Need to estimate all primitive parameters. Impossible to be compelling (selection, simultaneity, omitted variables etc)

Chetty (2009): A Bridge Between the 2

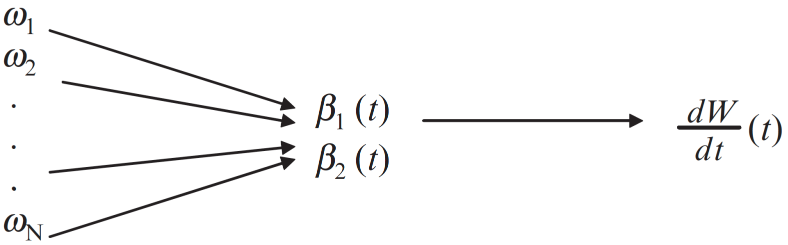
- Public Economics has pioneered an approach to compromise between the two: **Sufficient Statistics.**
- Setup:
  - Policy instrument t
  - Social welfare $W(t)$ (e.g. $\sum_h \gamma_h V_h(t)$)

$$\text{What is } \frac{dW(t)}{dt}??$$

- Structural approach:
  1. Write model with primitives $\boldsymbol{\omega} = (\omega_1, \ldots, \omega_N)$
  2. Derive
  $$\frac{dW(t)}{dt} = f(\boldsymbol{\omega})$$
  3. Estimate $\boldsymbol{\omega}$
  4. Calculate $dW(t)/dt$

Chetty (2009): A Bridge Between the 2

Primitives | **Sufficient statistics** | **Welfare change**



ω = preferences, constraints

ω not uniquely identified

$\beta = f(\omega, t)$

$y = \beta_1 X_1 + \beta_2 X_2 + \varepsilon$

β identified using program evaluation

dW/dt used for policy analysis

Chetty (2009): Benefits

1. Simpler to estimate.
   1.1 Less data and variation needed to identify marginal treatment effects than full structural model
   1.2 Especially beneficial with heterogeneity and discrete choice (lots of primitives, still few MTEs)
2. Weaker assumptions and design-based empirical methods.
   2.1 more transparent and empirically credible estimates.
3. Can be implemented even when we're uncertain about what the right model is.

Chetty (2009): Costs

1. Each question requires its own sufficient-statistics formula
   1.1 e.g. unemployment benefit level vs duration of unemployment benefits; tax rate vs tax base etc.
   1.2 In some settings it might be hard to characterize the sufficient statistics formula.
2. More potential to be misapplied: A little bit of knowledge is a dangerous thing!
   2.1 One can draw policy conclusions from a sufficient-statistics formula without evaluating the validity of the model it is based on. Structural requires full estimation of the model so can only draw conclusions from models that fit the data.

Precedent: Harberger (1964)

- Remember Harberger's deadweight loss triangle?
- That's the first sufficient statistics formula!
- The sufficient statistic is the elasticity of equilibrium quantity of the taxed good wrt its after-tax price
- The structural primitives are the demand- and the supply-elasticities of all the goods in the economy

Precedent: Harberger (1964)

- Consider a static, general equilibrium model.
- An individual is endowed with Z units of numeraire good y (think of it as labor)
- Firms use the numeraire as input to production of J consumption goods $\boldsymbol{x} = (x_1, \ldots, x_J)$ with convex cost functions $c_j(x_j)$
- Total cost of production is $c(\boldsymbol{x}) = \sum_{j=1}^{J} c_j(x_j)$. Production is perfectly competitive.
- Government taxes good 1 at rate t. $\boldsymbol{p} = (p_1, \ldots, p_J)$ is the vector of (endogenous) pretax prices

Precedent: Harberger (1964)

- Consumer takes prices as given and maximizes quasi-linear utility:

$$\max_{\boldsymbol{x},y} u\left(x_1,\ldots,x_J\right)+y$$
$$\text{s.t. } \boldsymbol{p}\cdot\boldsymbol{x}+tx_1+y=Z$$

- Firms take prices as given and solve

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

- These two problems give us demand and supply of the J goods: $\boldsymbol{x}^D\left(\boldsymbol{p}\right)$ and $\boldsymbol{x}^S\left(\boldsymbol{p}\right)$
- Markets clear to close the model: $\boldsymbol{x}^D\left(\boldsymbol{p}\right)=\boldsymbol{x}^S\left(\boldsymbol{p}\right)$

Precedent: Harberger (1964)

- What is the welfare cost of the tax t? It's the loss of social surplus from transactions that fail to take place because of the tax.
- Conceptual experiment: what is the loss in welfare if we raise the tax rate and rebate the revenues lump sum to consumers?

$$W(t) = \underbrace{\left\{\max_{\boldsymbol{x}} u(\boldsymbol{x}) + Z - tx_1 - \boldsymbol{p}(t) \cdot \boldsymbol{x}\right\}}_{\text{consumer surplus } CS(\boldsymbol{x};t)} + \underbrace{\left\{\max_{\boldsymbol{x}} \boldsymbol{p}(t) \cdot \boldsymbol{x} - c(\boldsymbol{x})\right\}}_{\text{producer surplus } PS(\boldsymbol{x};t)} + \underbrace{tx_1}_{\text{tax revenue}}$$

- Note: consumers don't take account of change in size of rebate when choosing x_1: It is a "*fiscal externality*"

Precedent: Harberger (1964)

- So how can we estimate $dW(t)/dt$?

1. Estimate J good demand and supply system to get $u(\boldsymbol{x})$ and $c(\boldsymbol{x})$. The problem is simultaneity: To get the slope of the demand and supply curves, we need $2J$ instruments!
2. Harberger's simpler approach: Exploit the power of the envelope theorem. Consumers and producers are choosing $\boldsymbol{x}$ optimally so we can ignore behavioral responses $d\boldsymbol{x}/dt$ in the curly brackets:

$$\frac{dCS(\boldsymbol{x};t)}{dt} = \frac{\partial CS(\boldsymbol{x};t)}{\partial \boldsymbol{x}}\frac{d\boldsymbol{x}}{dt} + \frac{\partial CS(\boldsymbol{x};t)}{\partial t} = \frac{\partial CS(\boldsymbol{x};t)}{\partial t}$$

(and similarly for producer surplus)

Precedent: Harberger (1964)

- Let's demonstrate this for consumer surplus
- Consumer's FOCs are

$$\frac{\partial u\left(\boldsymbol{x}\right)}{\partial x_1} = p_1 + t \quad \frac{\partial u\left(\boldsymbol{x}\right)}{\partial x_j} = p_j,\ j = 2, \ldots, J$$

- Totally differentiating $CS\left(\boldsymbol{x}; t\right)$

$$\frac{dCS\left(\boldsymbol{x}; t\right)}{dt} = \underbrace{\sum_{j=1}^{J} \left(\frac{\partial u\left(\boldsymbol{x}\right)}{\partial x_j} - p_j \right) \frac{\partial x_j}{\partial t} - t \frac{\partial x_1}{\partial t}}_{\partial CS(\boldsymbol{x};t)/\partial \boldsymbol{x}\ \partial \boldsymbol{x}/\partial t} \underbrace{- \frac{\partial \boldsymbol{p}\left(t\right)}{\partial t} \cdot \boldsymbol{x} - x_1}_{\partial CS(\boldsymbol{x};t)/\partial t}$$

$$= - \frac{\partial \boldsymbol{p}\left(t\right)}{\partial t} \cdot \boldsymbol{x} - x_1 = \frac{\partial CS\left(\boldsymbol{x}; t\right)}{\partial t}$$

Precedent: Harberger (1964)

- Using the power of the envelope theorem we have

$$\begin{aligned}\frac{dW(t)}{dt} &= \frac{dCS(\boldsymbol{x};t)}{dt} + \frac{dPS(\boldsymbol{x};t)}{dt} + \frac{dtx_1}{dt} \\ &= \frac{\partial CS(\boldsymbol{x};t)}{\partial t} + \frac{\partial PS(\boldsymbol{x};t)}{\partial t} + \frac{dtx_1}{dt} \\ &= \left(-\frac{\partial \boldsymbol{p}(t)}{\partial t}\cdot \boldsymbol{x} - x_1\right) + \left(\frac{\partial \boldsymbol{p}(t)}{\partial t}\cdot \boldsymbol{x}\right) + \left(x_1 + t\frac{dx_1}{dt}\right) \\ &= t\frac{dx_1(t)}{dt}\end{aligned}$$

- $dx_1(t)/dt$ is a **sufficient statistic** for the welfare loss

1. Taxes induce behavioral responses $d\boldsymbol{x}/dt$ but these have no first-order effects on welfare because households and firms are optimizing (envelope theorem)
2. Taxes induce changes in prices $d\boldsymbol{p}/dt$ but these have no first-order effects on welfare, they only redistribute surplus between producers and consumers

A General Cookbook

- Here's a general cookbook (we'll focus on a single agent in a static model, but easy to generalize)
- Step 1: Specify the structure of the model. What are the agent's choices and constraints?

$$\max_{x} U(\boldsymbol{x}) \quad s.t.\ G_1(\boldsymbol{x}, t, T), \ldots, G_M(\boldsymbol{x}, t, T)$$

  where $\boldsymbol{x} = (x_1, \ldots, x_J)$ are choices, t is "tax" on x_1, $T(t)$ is transfer in units of x_J

- Solution to this problem defines welfare as a function of the policy instrument

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

A General Cookbook

- Step 2: Express $dW\left(t\right)/dt$ in terms of multipliers:

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

- We know $\partial T/\partial t$ from the government budget constraint, and can calculate $\partial G_m/\partial T$ and $\partial G_m/\partial t$. The key unknowns are the λ_ms.

A General Cookbook

- Step 3: Substitute Multipliers by marginal utilities. The agent's FOCs imply
$$\frac{\partial u\left(\boldsymbol{x}\right)}{\partial x_j} = -\sum_{m=1}^{M} \lambda_m \frac{\partial G_m}{\partial x_j}$$
- This maps the λs to the marginal utilities. Let's make an assumption on the structure of the constraints:

$$\frac{\partial G_m}{\partial t} = k_t\left(\boldsymbol{x}, t, T\right) \frac{\partial G_m}{\partial x_1} \; \forall m = 1, \ldots, M$$
$$\frac{\partial G_m}{\partial T} = -k_T\left(\boldsymbol{x}, t, T\right) \frac{\partial G_M}{\partial x_J} \; \forall m = 1, \ldots, M$$

A General Cookbook

- Now we have

$$\begin{aligned}
\frac{dW(t)}{dt} &= \sum_{m=1}^{M} \lambda_m \left\{ \frac{\partial G_m}{\partial T}\frac{\partial T}{\partial t} + \frac{\partial G_m}{\partial t} \right\} \\
&= \sum_{m=1}^{M} \lambda_m \left\{ -k_T(\boldsymbol{x}, t, T)\frac{\partial G_M}{\partial x_J}\frac{\partial T}{\partial t} + k_t(\boldsymbol{x}, t, T)\frac{\partial G_m}{\partial x_1} \right\} \\
&= -k_T \frac{\partial T}{\partial t}\sum_{m=1}^{M} \lambda_m \frac{\partial G_M}{\partial x_J} + k_t \sum_{m=1}^{M} \lambda_m \frac{\partial G_m}{\partial x_1} \\
&= k_T \frac{\partial T}{\partial t} u'(x_J(t)) - k_t u'(x_1(t))
\end{aligned}$$

A General Cookbook

- Step 4: Recover the marginal utilities from observed choices.
- Sometimes we make assumptions about the marginal utilities (e.g. quasilinear utility in the Harberger example means that $u'(x_J) = 1$)
- In general, try and use the fact that the marginal utilites are inputs into observed choices and then recover them from how choices change in response to price/policy changes. e.g. in Harberger example, $u'(x_1) = p_1 + t$

A General Cookbook

- Step 5: Empirical implementation
- The work so far tells us which empirical objects we need to try and estimate:

$$\text{e.g. } \frac{dW(t)}{dt} = f\left(\frac{\partial x_1}{\partial t}, \frac{\partial x_1}{\partial Z}, t\right)$$

- Estimate these objects using policy/price changes
- Step 6: Model evaluation

1. Test predictions of the model that was needed to get the sufficient statistics model
2. Identify at least one set of structural parameters ω that is consistent with the model

Outline

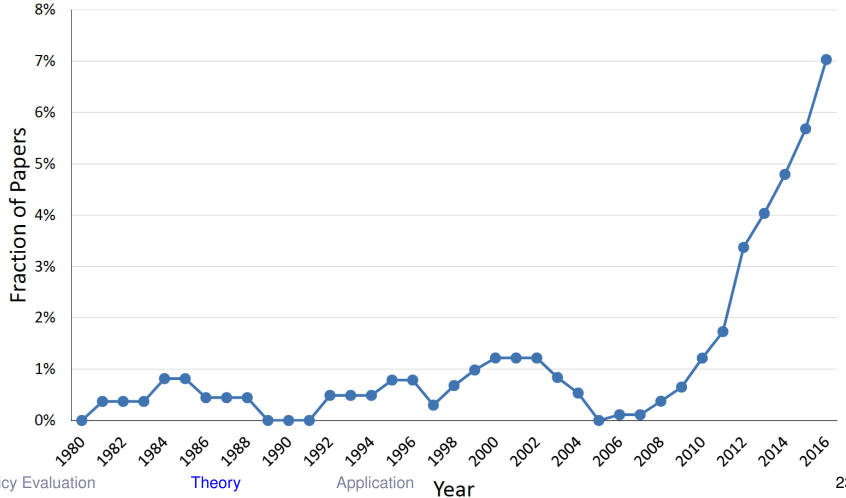
Outline

Theory: Welfare Concepts and Sufficient Statistics

- Kleven (2019) *Sufficient Statistics Revisited*
- Finkelstein & Hendren (2020) *Welfare Analysis Meets Causal Inference*

Kleven (2019): Sufficient Statistics Approaches Have Become Popular

FIGURE 1: FRACTION OF NBER WORKING PAPERS IN PUBLIC ECONOMICS REFERRING TO THE SUFFICIENT STATISTICS APPROACH



Kleven (2019): Overview

- The sufficient statistics approach we just saw relies on 3 things

1. The reform being analyzed is small
2. There are no other distortions in the economy
3. Decisions about the structure of the environment that pin down which elasticities are sufficient in a particular setting

- This paper:

1. Provide a general framework to showcase how general sufficient statistics approach is
2. Show how it generalizes when we relax 1. and 2.
3. Show how quickly the estimation requirements go up!

Kleven (2019): General Model

- Continuum of individuals indexed by i. Discrete set of goods $j = 0, \ldots, J$
$$u^i\left(x_0^i, \ldots, x_J^i\right) = u^i\left(\boldsymbol{x}^i\right)$$
- Budget constraint (normalize pre-tax prices to 1)
$$\sum_{j=1}^{J} x_j^i + T\left(x_0^i, \ldots, x_J^i\right) = y^i$$
where $T(\boldsymbol{x})$ is a tax fundtion that need not be separable embodying all taxes and transfers.
- Assume that $T\left(x_0^i, \ldots, x_J^i\right)$ is piecewise linear and denote marginal tax rates $\partial T / \partial x_j^i \equiv \tau_j^i$. BC becomes
$$\sum_{j=0}^{J} \left(1 + \tau_j^i\right) x_j^i = Y^i$$
where $Y^i = y^i + \sum_{j=0}^{J} \tau_j^i x_j^i - T\left(x_0^i, \ldots, x_J^i\right)$ is virtual income.

Kleven (2019): General Model

- Solution given by FOCs

$$\frac{\partial u^i}{\partial x_j^i} - \lambda^i \left(1 + \tau_j^i\right) = 0 \; \forall j$$

- Yielding indirect utility

$$v^i \left(1 + \tau_0^i, \ldots, 1 + \tau_J^i, Y^i\right) = u^i \left(x_0^i \left(1 + \tau_0^i, \ldots, 1 + \tau_J^i, Y^i\right), \ldots \right.$$
$$\left. \ldots, x_J^i \left(1 + \tau_0^i, \ldots, 1 + \tau_J^i, Y^i\right)\right)$$

- Remember 1st year micro results:

$$\frac{\partial v^i}{\partial Y^i} = \lambda^i, \quad \underbrace{\frac{\partial v^i}{\partial \left(1 + \tau_k^i\right)} = -\lambda^i x_k^i}_{\text{Roy's identity}} \quad \underbrace{\frac{\partial x_j^i}{\partial \left(1 + \tau_k^i\right)} = \frac{\partial \tilde{x}_j^i}{\partial \left(1 + \tau_k^i\right)} - x_k^i \frac{\partial x_j^i}{\partial Y^i}}_{\text{Slutsky decomposition}}$$

The Welfare Effect of Small Reforms

- Specify the tax policy as a function of treatment parameters θ: $T\left(x_0^i, \ldots, x_J^i, \theta\right)$, and $\tau\left(x_0^i, \ldots, x_J^i, \theta\right)$
- Consider a small reform $d\theta \approx 0$
- Money-metric measure of effect on individuals' utility:

$$\begin{aligned}
\frac{dv^i}{d\theta} &= \sum_{j=0}^{J} \frac{\partial v^i}{\partial\left(1+\tau_j^i\right)} \frac{d\tau_j^i}{d\theta} + \frac{\partial v^i}{\partial Y^i} \frac{dY^i}{d\theta} \\
&= -\lambda^i \left(\sum_{j=0}^{J} x_j^i \frac{d\tau_j^i}{d\theta} + \frac{dY^i}{d\theta} \right) \\
&= -\lambda^i \frac{\partial T^i}{\partial \theta}
\end{aligned}$$

  where the last equality uses $\frac{dY^i}{d\theta} = \sum_{j=0}^{J} x_j^i \frac{d\tau_j^i}{d\theta} - \frac{\partial T^i}{\partial \theta}$:
- The utility effect is equal to the mechanical revenue effect

The Welfare Effect of Small Reforms

- Now let's look at the social welfare effect. Define

$$W(\theta) = \int_i \omega^i v^i(\theta)\, di + \mu \int_i T^i(\theta)\, di$$

where ω^i is a Pareto weight on individual i and μ is the marginal value of government revenue.

- Differentiating

$$\frac{dW(\theta)/d\theta}{\mu} = \int_i \left[\frac{\omega^i}{\mu}\frac{dv^i}{d\theta} + \frac{dT^i}{d\theta}\right] di = \int_i \left[\frac{dT^i}{d\theta} - g^i \frac{\partial T^i}{\partial \theta}\right] di$$

$$= \int_i \left(\underbrace{\left[\frac{dT^i}{d\theta} - \frac{\partial T^i}{\partial \theta}\right]}_{\text{efficiency}} + \underbrace{\left(1 - g^i\right)\frac{\partial T^i}{\partial \theta}}_{\text{equity}} \right) di$$

where $g^i = \frac{\omega^i \lambda^i}{\mu}$ is each individual's social marginal welfare weight

The Welfare Effects of Small Reforms

- The efficiency term is the *fiscal externality*: behavioral changes reduce government revenue and reduce the potential transfer others can receive
- So what are the sufficient statistics? (and are they going to be externally valid?) Using $T\left(x_0^i, \ldots, x_J^i, \theta\right)$ the fiscal externality can be rewritten as

$$\left.\frac{dW/d\theta}{\mu}\right|_{g^i=1} = \int_i \sum_{j=0}^{J} \tau_j^i \left[\sum_{k=0}^{J} \frac{\partial x_j^i}{\partial\left(1+\tau_k^i\right)} \frac{d\tau_k^i}{d\theta} + \frac{\partial x_j^i}{\partial Y^i} \frac{dY^i}{d\theta}\right] di$$

$$= \int_i \sum_{j=0}^{J} \tau_j^i \left[\sum_{k=0}^{J} \frac{\partial x_j^i}{\partial\left(1+\tau_k^i\right)} \frac{d\tau_k^i}{d\theta} + \frac{\partial x_j^i}{\partial Y^i} \left(\sum_{k=0}^{J} \frac{d\tau_k^i}{d\theta} x_k^i - \frac{\partial T^i}{\partial \theta}\right)\right] di$$

The Welfare Effect of Small Reforms

$$\left.\frac{dW/d\theta}{\mu}\right|_{g^i=1} = \int_i \sum_{j=0}^{J} \left[\sum_{k=0}^{J} \tau_j^i x_j^i \varepsilon_{jk}^i \frac{d\tau_k^i/d\theta}{1+\tau_k^i} - \tau_j^i x_j^i \eta_j^i \frac{\partial T^i/\partial\theta}{Y^i} \right]$$

where $\varepsilon_{jk}^i \equiv \frac{\partial \tilde{x}_j^i}{\partial\left(1+\tau_k^i\right)} \frac{1+\tau_k^i}{x_j^i}$ is the Hicksian (compensated) price elasticity and $\eta_j^i \equiv \frac{\partial x_j^i}{\partial Y^i} \frac{Y^i}{x_j^i}$ is the income elasticity.

- Therefore, the sufficient statistics for evaluating the reform are $\left\{\varepsilon_{jk}^i, \eta_j^i\right\}_{\forall j,k,i}$
- Completely general given 1) small reform; and 2) no non-policy imperfections
- It is also a *general equilibrium* result, allowing for cross market effects etc.

But You Said That Sufficient Statistics was Simple!

- Of course, the problem is that J is potentially very large. So we require restrictions on either a) the tax policy space (what the $d\tau/d\theta$ and $\partial T/\partial\theta$ terms look like) or on behavioral responses (the εs and ηs)
- How can we get back to that nice simple Harberger equation?
- Assume

1. Utility is quasi-linear: $\Rightarrow \eta_j^i = 0 \; \forall j, i$
2. Only one good is taxed: $\tau_0 \neq 0; \; \tau_j = 0 \; \forall j = 1, \ldots, J$
3. The tax is linear: $T = \tau_0 x_0$

- Now that big equation collapses down to

$$\left.\frac{dW/d\theta}{\mu}\right|_{g^i=1} = \bar{\varepsilon}_0 \frac{\tau_0}{1+\tau_0}\frac{d\tau_0}{d\theta}$$

where $\bar{\varepsilon}_0 = \int_i x_0^i \varepsilon_{00}^i di$, the demand-weighted average elasticity is the sufficient statistic

But You Said That Sufficient Statistics was Simple!

- A different way to do this is to assume quasi-linearity but that goods $0, \ldots, J_0$ are taxed at rate τ_0 while goods $J_0 + 1, \ldots, J$ are taxed at rate τ_1 $(= 0$ wlog$)$
- Now the sufficient statistic is $\bar{\varepsilon}_0 = \int_i \left[\sum_{j=0}^{J_0} \sum_{k=0}^{J_0} x_j^i \varepsilon_{jk}^i \right] di$ is the demand-weighted elasticity across goods $0, \ldots, J_0$ with respect to their tax rate τ_0. Still a single elasticity, but a different one.
- e.g. 1: Goods $0, \ldots, J_0$ are labor supply in different periods and the others are consumption in those periods. Now we can interpret $\bar{\varepsilon}_0$ as the elasticity of *lifetime* rather than contemporaneous earnings.
- e.g. 2: Goods $0, \ldots, J_0$ are multiple dimensions of labor supply (hours, effort, occupation, training etc). Now the elasticity is the elasticity of total labor income: the *elasticity of taxable income* (Feldstein, 1999)

Large Reforms

- When reforms are large, the small-reform sufficient statistics are the first-order Taylor approximation to the welfare effect.
- Can we do better? Consider a large reform to the tax policy $T\left(x_0^i, \ldots x_J^i, \theta\right)$ and its associated marginal tax rates $\tau_j^i\left(\theta\right)$. Define marginal tax rates as $\tau_j^i + \theta\Delta\tau_j^i$ where Δ_j^i are the reform-induced changes to the MTRs. Then we consider going from $\theta_0 = 0$ to $\theta_1 = 1$.
- The change in welfare is

$$\Delta W = W\left(1\right) - W\left(0\right) = \int_0^1 \frac{dW}{d\theta} d\theta$$

Large Reforms

- The efficiency effect of a large reform is

$$\left.\frac{\Delta W}{\mu_0}\right|_{g^i=1} = \int_0^1 \left.\frac{dW/d\theta}{\mu_0}\right|_{g^i=1} d\theta$$

where

$$\frac{dW/d\theta}{\mu_0} \approx \int_i \sum_{j=0}^{J} \left[\sum_{k=0}^{J} \left(\tau_j^i + \theta \Delta t_j^i\right) x_j^i(\theta)\, \varepsilon_{jk}^i(\theta) \frac{\Delta \tau_k^i}{1 + \tau_k^i + \theta \Delta \tau_k^i} \right.$$
$$\left. - \left(\tau_j^i + \theta \Delta \tau_j^i\right) x_j^i(\theta)\, \eta_j^i(\theta) \frac{\partial T^i / \partial \theta}{Y^i(\theta)} \right] di$$

- Small-reform expression gets it wrong because
  - wedges change over the reform path
  - elasticities change over the reform path

Large Reforms

- Let's simplify things a little

1. Instead of the full integral, consider the trapezoid approximation

$$\Delta W \approx \frac{1}{2}\left[\frac{dW\left(0\right)}{d\theta} + \frac{dW\left(1\right)}{d\theta}\right]$$

2. Assume quasi-linear utility and a single tax rate on taxed goods τ_0

- Now we get

$$\left.\frac{\Delta W}{\mu_0}\right|_{g^i=1} \approx \frac{1}{2}\left\{\bar{\varepsilon}_0\left(0\right)\frac{\tau_0}{1+\tau_0}\Delta\tau_0 + \bar{\varepsilon}_0\left(1\right)\frac{\tau_0+\Delta\tau_0}{1+\tau_0+\Delta\tau_0}\Delta\tau_0\right\}$$

Large Reforms

- Now we can see what the correction term is we need to apply to the small-reform sufficient statistics formula:

$$\left.\frac{\Delta W}{\mu_0}\right|_{g^i=1} \approx \underbrace{\bar{\varepsilon}_0 \frac{\tau_0}{1+\tau_0}\Delta\tau_0}_{\text{Small reform formula}} + \frac{1}{2}\left\{\bar{\varepsilon}_0 \Delta\left[\frac{\tau_0}{1+\tau_0}\right] + \Delta\bar{\varepsilon}_0 \frac{\tau_0}{1+\tau_0} + \Delta\bar{\varepsilon}_0 \Delta\left[\frac{\tau_0}{1+\tau_0}\right]\right\}\Delta\tau_0$$

- Note that now the sufficient statistics are the elasticity $\bar{\varepsilon}_0$ and the elasticity change $\Delta\bar{\varepsilon}_0 = \bar{\varepsilon}_0(1) - \bar{\varepsilon}_0(0)$
- With iso-elastic preferences $(\Delta\bar{\varepsilon}_0 = 0)$

$$\left.\frac{\Delta W}{\mu_0}\right|_{g^i=1} \approx \bar{\varepsilon}_0 \left(\frac{\tau_0}{1+\tau_0} + \frac{1}{2}\Delta\left[\frac{\tau_0}{1+\tau_0}\right]\right)\Delta\tau_0$$

Non-Government Distortions

- What if there are other sources of wedges between private and social incentives?

$$u^i\left(x_0^i, \ldots, x_J^i; E_0^i, \ldots, E_J^i\right)$$

where $E_j^i \equiv \int_{\hat{i}} \phi_j^{i\hat{i}} x_j^{\hat{i}} d\hat{i}$ are externalities on individual i of consumption of good j

1. $\phi^{i\hat{i}} = 1$: Atmospheric externality, depends only on sum of consumption
2. $\phi^{i\hat{i}} = 0 \ \forall i \neq \hat{i}$ and $\phi^{ii} = 1$: Internality. Gap between decision (takes E_j^i as given) and experienced utility
3. $\phi^{i\hat{i}} = -1 \ \forall i \neq \hat{i}$ and $\phi^{ii} = 1$: Relative consumption concerns

Non-Government Distortions

- Effect on money-metric utility of a small reform:

$$\frac{dv^i/d\theta}{\lambda^i} = -\frac{\partial T^i}{\partial \theta} + \sum_{j=0}^{J} \frac{\partial v^i/\partial E^i_j}{\lambda^i} \frac{dE^i_j}{d\theta}$$

- The fiscal externality is still there, but there's also an externality effect
- The effect of small reforms on efficiency is

$$\left.\frac{dW/d\theta}{\mu}\right|_{g^i=1} = \int_i \left[\frac{dT^i}{d\theta} - \frac{\partial T^i}{\partial \theta} + \sum_{j=0}^{J} \frac{\partial v^i/\partial E^i_j}{\lambda^i} \frac{dE_j}{d\theta}\right] di$$

Non-Government Distortions

- To go from this to a sufficient statistics formula, assume

1. $x_j^i = x_j^i\left(1+\tau_0^i, \ldots, 1+\tau_J^i, Y^i\right)$: Demand doesn't depend on the externalities
2. The externalities take the form $\phi^i = \phi_{I_j}^i \mathbf{1}\left\{\hat{i}=i\right\} + \phi_{E_j}^i$

- Now we get

$$\left.\frac{dW/d\theta}{\mu}\right|_{g^i=1} = \int_i \sum_{j=0}^{J}\left[\left(\tau_j^i + \tau_{I_j}^i + \tau_{E_j}\right)\frac{dx_j^i}{d\theta}\right] di$$

$$= \int_i \left[\sum_{j=0}^{J}\sum_{k=0}^{J}\hat{\tau}^i x_j^i \varepsilon_{jk}^i \frac{d\tau_k^i/d\theta}{1+\tau_k^i} - \sum_{j=0}^{J}\hat{\tau}_j x_j^i \eta_j^i \frac{\partial T^i/\partial\theta}{Y^i}\right] di$$

where $\tau_{I_j}^i \equiv \frac{\partial v^i/\partial E_j^i}{\lambda^i}\phi_{I_j}^i$, $\tau_{E_j} \equiv \int_i \frac{\partial v^i/\partial E_j^i}{\lambda^i}\phi_{E_j}^i$, and $\hat{\tau}_j^i = \tau_j^i + \tau_{I_j}^i + \tau_{E_j}$

Non-Government Distortions

- If we specialize to the Harberger case, this becomes

$$\left.\frac{dW/d\theta}{\mu}\right|_{g^i=1} = \bar{\varepsilon}_0 \frac{\hat{\tau}_0}{1+\tau_0} \frac{d\tau_0}{d\theta}$$

- The formula is almost the same.
- But we need to estimate $\hat{\tau}_0$ so the sufficient statistics are $\{\bar{\varepsilon}_0, \hat{\tau}_0\}$

Outline

Theory: Welfare Concepts and Sufficient Statistics

- Kleven (2019) *Sufficient Statistics Revisited*
- Finkelstein & Hendren (2020) *Welfare Analysis Meets Causal Inference*

Finkelstein & Hendren (2020): Overview

- In Public Economics the question we ultimately care about is what the *welfare* effects of a policy are (rather than "just" the effects on earnings, employment, education, etc.)
- How can we measure this *empirically*?
- Problem:
  - Marginal deadweight loss depends on *compensated* elasticities, but we usually estimate uncompensated elasticities
  - The Marginal Cost of Public Funds (MCPF) approach needs to monetize the welfare impacts of a bunch of changes in non-monetary outcomes, and relate them to 1 + MCPF (0.3??)
- A metric popularized by Hendren (2016) is the **Marginal Value of Public Funds (MVPF)**

$$MVPF = \frac{\text{Marginal Benefit of Policy}}{\text{Marginal Cost of Policy}} = \frac{\text{Beneficiaries' Willingness to Pay}}{\text{Net Cost"to Government}}$$

Application to Targeted Transfer

- Consider a small increase of $1 in the transfer to a specific income group (e.g. through EITC, or anti-poverty program)
- Assume recipients are optimizing, and that the policy does not have externalities
- Start with the numerator, the benefit. There are two types of individuals

1. Infra-marginal beneficiaries I. They value the extra $1 at $1.
2. Marginal beneficiaries M. They changed their behavior to become just eligible. Envelope theorem tells us they are, to first order, indifferent to this change.

- $\Rightarrow$ benefit is $\$1 \times$number of infra-marginal beneficiaries

Application to Targeted Transfer

- Now consider the denominator, the cost. This also comes from two sources:

1. The **mechanical cost**. \$1 per inframarginal recipient: \$I
2. The **fiscal externality** FE. The impact of behavioral changes on the government budget.

$$\Rightarrow MVPF = \frac{1}{1 + FE}$$

- This is great because the Fiscal Externality is the causal effect of the policy change on the government's budget. This object is ideally suited to the use of modern empirical methods to estimate it (in a way that, say, compensated elasticities are not)

Policies that Affect Different Income Groups

- We can compare policies that affect the same income group by comparing their MVPF.
- But how can we compare policies that affect different income groups?
- A natural benchmark is to compare the MVPF of a policy affecting income group y_i to the MVPF of a $1 tax cut at y_i, what Hendren (2017) terms the "Efficient Welfare Weight" v_i
- Now we can compare a policy that transfers resources to income group y_i $(MVPF_i)$ to a policy that transfers resources to income group y_j $(MVPF_j)$: We prefer policy i to policy j if

$$MVPF_i \frac{v_j}{v_i} > MVPF_j$$

Policies that Affect Different Income Groups

- We can also provide a way to benchmark a single policy: Is it's MVPF bigger than a tax cut for that group?
- If $MVPF_i > v_i$ it suggests we raise taxes on that group by \$1 and use the money to do policy i.

Relaxing Assumptions

1. Large policy changes. Now we need the area under the demand curve for WTP. See Kleven (2019) above.
2. Non-privately optimal decisions: Now there's a wedge between choices (demand) and welfare (preferences). Need behavioral welfare analysis methods
3. In-kind transfers. Now need to estimate willingness to pay for the in-kind transfer in the numerator.
4. Externalities. Now need to estimate net benefits of all people who are affected in the numerator.

Outline

Outline

Applications

- Hendren & Sprung-Keyser (2020) *A Unified Welfare Analysis of Government Policies*

Hendren & Sprung-Keyser (2020): Overview

- . . .Over to you guys