GR 6307
Public Economics and Development

1.1 Detour:
Applied Welfare Analysis

Michael Best

Fall 2019

Outline

2 Approaches to Policy Evaluation

Theory: Welfare Concepts and Sufficient Statistics

An Application

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

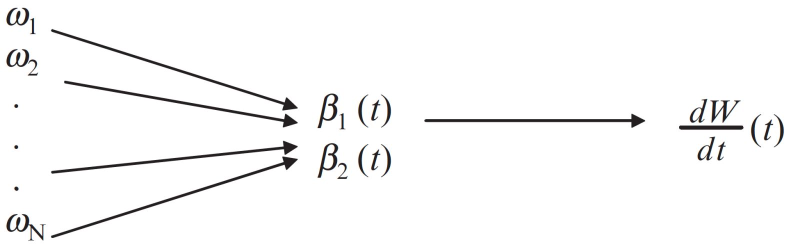
- Public Economics has pioneered an approach to compromise between the two: **Sufficient Statistics.**
- Setup:
  - Policy instrument t
  - Social welfare $W(t)$ (e.g. $\sum_h \gamma_h V_h(t)$)

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

- 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	$\beta = f(\omega, t)$ $y = \beta_1 X_1 + \beta_2 X_2 + \varepsilon$	dW/dt used for policy analysis
ω not uniquely identified	β identified using program evaluation	

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

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

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

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

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

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 $dW(t)/dt$ in terms of multipliers:

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

- 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

2 Approaches to Policy Evaluation

Theory: Welfare Concepts and Sufficient Statistics

An Application

Outline

Theory: Welfare Concepts and Sufficient Statistics

- Kleven (2019) *Sufficient Statistics Revisited*
- Finkelstein (2019) *Welfare Analysis Meets Causal Inference: A Suggested Interpretation of Hendren*

Kleven (2019)

- ...

Outline

Theory: Welfare Concepts and Sufficient Statistics

- Kleven (2019) *Sufficient Statistics Revisited*
- Finkelstein (2019) *Welfare Analysis Meets Causal Inference: A Suggested Interpretation of Hendren*

Finkelstein (2019): Overview

- ...

Outline

2 Approaches to Policy Evaluation

Theory: Welfare Concepts and Sufficient Statistics

An Application

Outline

Hendren & Sprung-Keyser (2019): Overview

- . . .