SUFFICIENT STATISTICS REVISITED

Henrik Jacobsen Kleven

Princeton University

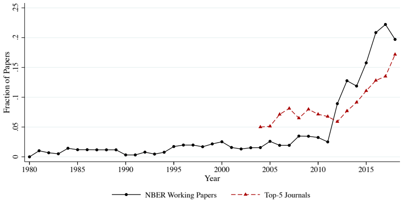
September 2020

Sufficient Statistics Approach

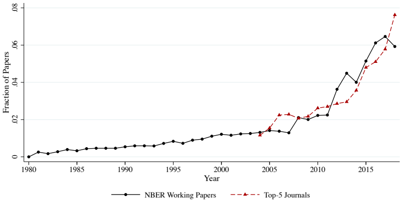
- The idea is that the welfare effect of policy changes can be written in terms of estimable reduced-form elasticities
  - Allows for policy evaluation without estimating the structural primitives of fully specified models
- A bridge between reduced-form approaches (credible identification) and structural approaches (welfare predictions)
- Chetty (2009) coined the phrase, but the intellectuel origins are very old
  - DWL: Harberger (1964), Feldstein (1999), Kleven-Kreiner (2005)
  - MCF: Browning (1976), Kleven-Kreiner (2006)
  - OT: Ramsey (1927), Diamond-Mirrlees (1971), Saez (2001)

Papers Using Sufficient Statistics Terminology

Public Economics



All Economics

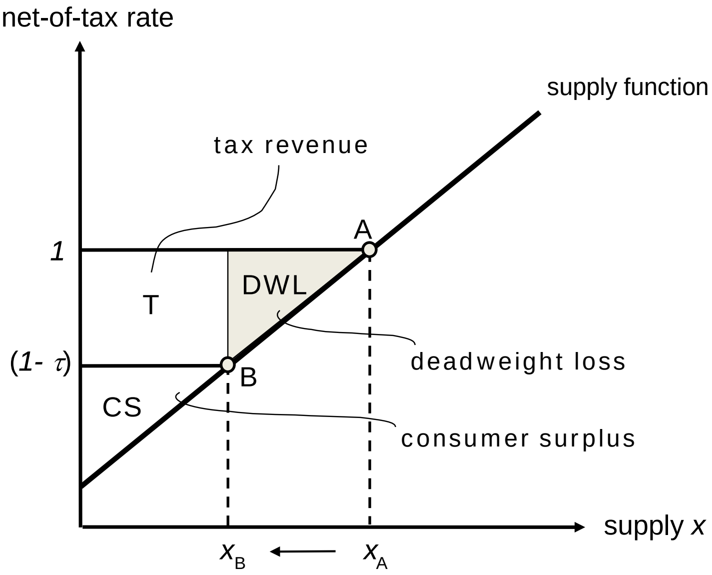


Source: Kleven (2020)

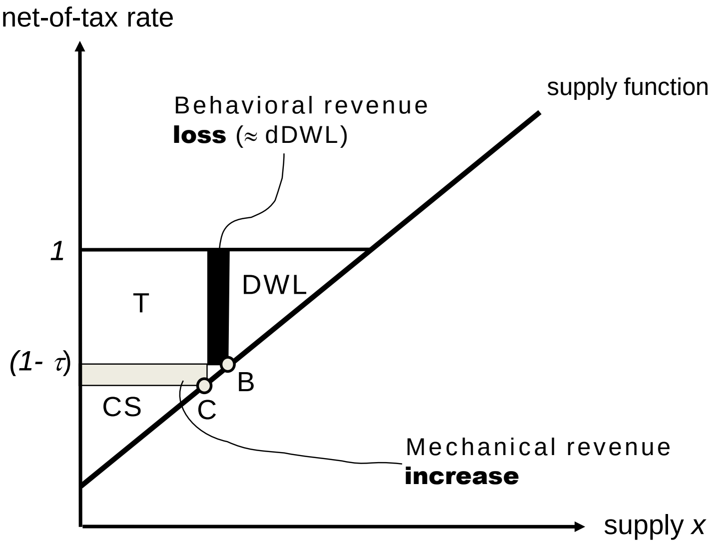
Outline

- Harberger Triangles
- A standard sufficient statistics approach
  - Assume small reforms and no non-government distortions
  - Use the language of taxation, but the analysis is more general
  - Provide general formula and Harberger-style formulas
- Link between sufficient statistics approach and optimal taxation
- Generalizations of the approach
  - Large reforms
  - Non-government externalities or internalities

Harberger Triangle: Total DWL



Harberger Triangle: Marginal DWL



Marginal DWL Formula

- From the graphical analysis, we have

$$dDWL \approx -\tau \cdot dx = \frac{\tau}{1-\tau} \cdot x \cdot \varepsilon \cdot d\tau,$$

  where $\varepsilon \equiv \frac{dx/x}{d(1-\tau)/(1-\tau)}$ is the elasticity of supply x with respect to the net-of-tax rate $(1-\tau)$

- Conditional on the observables $(x, \tau, d\tau)$, the reduced-form elasticity ε is a "sufficient statistic" for welfare evaluation

- How general is this argument?

The Basic Principle

- Define $DWL = L - T$ where L is a money-metric utility loss from taxation and T is tax revenue
- The marginal DWL from tax reform is $dDWL = dL - dT$
  - We have $dT = dM + dB$, where dM is the **mechanical effect** and dB is the **behavioral effect**
  - Assuming small reform and no non-government externalities, we have $dL = dM$ (**envelope theorem**)
  - This implies $dDWL = dL - dT = dM - (dM + dB) = -dB$
- Hence, the marginal efficiency cost equals the behavioral revenue loss (**fiscal externality**)

A Formal Model

- Utility of individual i:

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

where $j = 0, ..., J$ index different goods (demands and supplies) as well as time

- Budget constraint

$$\sum_{j=0}^{J} x_j^i + T\left(x_0^i, \ldots, x_J^i\right) = y^i \quad \Leftrightarrow \quad \sum_{j=0}^{J} \left(1 + \tau_j^i\right) x_j^i = Y^i,$$

where the piecewise linear function $T(\cdot)$ embodies all taxes and transfers, $\tau_j^i \equiv \partial T / \partial x_j^i$ are marginal tax rates, and $Y^i \equiv y^i + \sum_{j=0}^{J} \tau_j^i x_j^i - T\left(x_0^i, ..., x_J^i\right)$ is virtual income

Individual Optimization

- FOC for good j:

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

- Indirect utility: $v^i \left(1 + \tau_0^i, \ldots, 1 + \tau_J^i, Y^i\right)$

- Derivatives of indirect utility (**envelope conditions**):

$$\frac{\partial v^i}{\partial Y^i} = \lambda^i, \qquad \frac{\partial v^i}{\partial \left(1 + \tau_k^i\right)} = -\lambda^i x_k^i \tag{1}$$

Utility Effect of Small Reforms

- Specify tax policy in terms of a treatment parameter θ, i.e. $T\left(x_0^i, \ldots, x_J^i, \theta\right)$ and $\tau_j^i(\theta) \quad \forall j$
- Changes in θ may capture any changes in $\tau_0^i, \ldots, \tau_J^i$ and $T(\cdot)$. For now we focus on small reforms, $d\theta \approx 0$
- Using the envelope conditions (1) and the definition of Y^i:

$$\frac{dv^i/d\theta}{\lambda^i} = -\frac{\partial T^i}{\partial \theta}$$

  $\rightarrow$ The utility effect of any arbitrary, small reform equals the mechanical revenue effect

From Individual Welfare to Social Welfare

- Social welfare function:

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

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

- Social welfare effect of reform:

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

where $g^i \equiv \frac{\omega^i \lambda^i}{\mu}$ is the social welfare weight on individual i

Efficiency Effect = Fiscal Externality

The effect of any small tax reform on economic efficiency equals

$$\left.\frac{dW/d\theta}{\mu}\right|_{g^i=1} = \int_i \left[\frac{dT^i}{d\theta} - \frac{\partial T^i}{\partial \theta}\right] di,$$

i.e., the difference between the total and mechanical revenue effects, corresponding to the behavioral revenue effect = fiscal externality

Why Estimate Elasticities?

- Elasticities are irrelevant for the welfare evaluation of actually implemented reforms
  - We can directly estimate revenue effects net of mechanical effects
- But evaluating actually implemented reforms is a limited objective
- Two reasons for estimating elasticities:
  1. Assessing counterfactual reforms and policy design → requires **externality valid** elasticities
  2. Providing a normalized, **reform-invariant** measure of behavioral response that can be compared across settings

Sufficient Statistics Formula

The effect of any small reform on economic efficiency can be written in terms of "reduced-form" elasticities:

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

where ε_{jk}^i is the Hicksian price elasticity of good j wrt. the price on good k, and η_j^i is the income elasticity of good j

- The sufficient statistics for evaluating reform are $\left\{\varepsilon_{jk}^i, \eta_j^i\right\}_{\forall j,k,i}$

Notes on Sufficient Statistics Formula

- The formula accounts for **general equilibrium effects**
  - It accounts for distortions in all markets and allows for any possible cross-effects between markets
  - Explicitly accounting for pre-tax price changes (tax incidence) will not change the formula
- The formula accounts for **dynamics**
  - Indices j, k capture any time dimension
- But the parameter space is too large for empirical implementation $\rightarrow$ impose more structure by (i) restricting policy space and/or (ii) restricting behavioral responses

Harberger-Style Formula

- Assumptions:
    1. Utility is quasi-linear ($\eta_j^i = 0 \;\; \forall j, i$)
    2. Only one good is taxed (e.g. good 0)
    3. Linear tax (marginal rate τ_0)

- We then obtain:

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

where $\bar{\varepsilon}_0 \equiv \int_i \left[x_0^i \varepsilon_{00}^i\right] di$ is the demand-weighted average Hicksian elasticity in the population

Harberger-Style Formula Generalized

- Relaxed assumptions:
  1. Utility is quasi-linear ($\eta_j^i = 0 \;\; \forall j, i$)
  2. Goods $0, ..., J_0$ are taxed at rate τ_0 and goods $J_0 + 1, ..., J$ are taxed at rate τ_1 $\rightarrow$ normalize $\tau_1 = 0$

- We then obtain the same formula:

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

  but where $\bar{\varepsilon}_0 \equiv \int_i \left[\sum_{j=0}^{J_0} \sum_{k=0}^{J_0} x_j^i \varepsilon_{jk}^i \right] di$ is a demand-weighted average Hicksian elasticity across goods $0, ..., J_0$ with respect to the tax price $1 + \tau_0$ on all those goods

Labor Supply Interpretation

- Goods $0, ..., J_0$ are labor supplies in different periods; goods $J_0 + 1, ..., J$ are consumption in different periods
- The sufficient statistic $\bar{\varepsilon}_0$ is a **lifetime earnings-weighted elasticity of labor supply wrt a permanent labor tax**
- If $0, ..., J_0$ also include multi-dimensional earnings components, $\bar{\varepsilon}_0$ is a **lifetime earnings-weighted elasticity of taxable income wrt permanent labor tax** (generalizes Feldstein 1999)
- Uniform tax on consumption in different periods rules out capital taxes. To allow for capital taxes and still have just one sufficient statistic $\bar{\varepsilon}_0$, we have to add separability assumptions between labor and consumption

Sufficient Statistics vs Structural Approaches

- A general approach depends on too many parameters to be empirically tractable
- Sufficient statistics and structural approaches can be viewed as different ways of reducing the dimensionality of the parameter space:
  - Sufficient Stats: Reduce dimensionality by restricting (i) tax policy space and/or (ii) decision environment and preferences
  - Structural: Reduce dimensionality by assuming parametric functional forms
- Preceding example provides an illustration:
  - Sufficient Stats: Small reform, permanent labor tax, zero capital tax, quasi-linearity $\rightarrow$ estimate lifetime labor supply elasticity
  - Structural: Parametric utility function $\rightarrow$ estimate structural "primitives"

Trade-offs Between Approaches

- Data requirements
  - Often stronger under the sufficient statistics approach
- Assumptions
  - Both approaches make strong (but different) assumptions
- Transparency
  - Sufficient statistics approach is more transparent
- Empirical design: Quasi-experimental vs observational variation
  - This divide is due to research culture rather than any deep difference between the approaches

Link to Optimal Taxation

- For optimal taxation, we must have $dW/d\theta = 0$ for any θ, i.e.

$$\left.\frac{dW/d\theta}{\mu}\right|_{g^i=1} = \int_i \left(g^i - 1\right) \frac{\partial T^i}{\partial \theta} di$$

  where LHS is the efficiency effect and RHS is the equity effect

- In the Harberger-style special case, this implies

$$\frac{\tau_0}{1+\tau_0} = \frac{\int_i \left(g^i - 1\right) x_0^i di}{\bar{\varepsilon}_0}$$

- Tempting to call $\bar{\varepsilon}_0$ a sufficient statistic for optimal taxation, but this is now the elasticity at the optimal point $\rightarrow$ requires global knowledge of demand functions $\rightarrow$ structural approach

Generalizations of the Sufficient Statistics Approach

We will generalize the approach by relaxing its two key assumptions:

1. Allow for large reforms
2. Allow for non-government externalities and internalities
   - Some sufficient statistics papers (e.g. Piketty et al. 2014) model specific market imperfections, but we develop a general approach

Welfare Effect of Large Reforms

- Specify marginal tax rates as $\tau_j^i + \theta\Delta\tau_j^i$ where $\Delta\tau_j^i$ is the reform-induced tax rate change $\rightarrow$ pre-reform policy corresponds to $\theta_0 = 0$ and the post-reform policy corresponds to $\theta_1 = 1$

- An exact welfare formula:

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

- Trapezoid approximation of welfare formula:

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

Sufficient Statistics for Large Reforms: Trapezoid

Assuming quasi-linear utility and a single tax rate τ_0 on taxed goods, the efficiency effect of large reforms can be approximated as

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

where $\bar{\varepsilon}_0$ is the elasticity at the initial policy and $\Delta\bar{\varepsilon}_0$ is the elasticity change due to the policy

- The sufficient statistics are the elasticity level $\bar{\varepsilon}_0$ and the elasticity change $\Delta\bar{\varepsilon}_0$

Sufficient Statistics for Large Reforms: Iso-Elastic

Assuming quasi-linear, iso-elastic utility and a single tax rate τ_0 on taxed goods, the efficiency effect of large reforms can be approximated as

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

a standard Harberger-style formula using a modified tax wedge, $\frac{\tau_0}{1+\tau_0} + \frac{1}{2}\Delta\left[\frac{\tau_0}{1+\tau_0}\right]$

Sufficient Statistics for Large Reforms: Conclusions

- In general, we have to account for the entire path of wedges and elasticities between the pre-reform and post-reform equilibria
- We can provide approximations that depend on the changes in wedges and elasticities from before to after the reform
  - Sufficient statistics include elasticity levels and elasticity changes
- If elasticity changes cannot be estimated, we have to impose more structure
  - The simplest solution is to assume iso-elastic utility $\rightarrow$ a fully structural approach

Welfare Effect with Non-Government Distortions

- Utility of individual i:

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

where E_j^i is the externality on individual i due to the consumption of good j:

$$E_j^i = \int_{\hat{\imath}} \phi_j^{i\hat{\imath}} x_j^{\hat{\imath}} d\hat{\imath},$$

where $\phi_j^{i\hat{\imath}}$ captures the externality that individual $\hat{\imath}$ imposes on individual i when consuming good j

- Different assumptions on $\phi_j^{i\hat{\imath}}$ capture a wide range of cases, e.g. atmospheric externalities, relative consumption or relative earnings externalities, and internalities

Fiscal Externality and Non-Government Externalities

In the presence of non-government externalities, the effect of any small tax reform on economic efficiency is given by

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

where the first term is the fiscal externality and the second term includes any non-government externalities

- The additivity between tax distortions and non-tax externalities echoes Sandmo (1975) for the case of atmospheric externalities, but here we show that this additivity property is very general

Simplifying Assumptions

To obtain simple sufficient statistics results, we make two assumptions:

1. Demands and supplies are independent of the external effects, i.e. $x_j^i = x_j^i\left(1+\tau_0^i, ..., 1+\tau_J^i, Y^i\right)$
2. The externality parameter equals $\phi_j^{i\hat{\imath}} = \phi_{I_j}^i \cdot \mathbf{1}\left(\hat{\imath} = i\right) + \phi_{E_j}^i$, where $\phi_{I_j}^i$ is an internality and $\phi_{E_j}^i$ is an externality
   - Different combinations of $\phi_{I_j}^i$ and $\phi_{E_j}^i$ encompass all of the examples provided earlier

Sufficient Statistics with Non-Government Externalities

In the presence of non-government externalities, the effect of any small tax reform on economic efficiency can be written as

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

where $\hat{\tau}_j^i \equiv \tau_j^i + \tau_{I_j}^i + \tau_{E_j}$ is an externality-adjusted tax wedge on good j for individual i

- The sufficient statistics are $\left\{\varepsilon_{jk}^i, \eta_j^i, \hat{\tau}_j^i\right\}_{\forall j,k,i}$

Harberger-style special case

- Assumptions:
  1. Utility is quasi-linear ($\eta_j^i = 0 \; \forall j, i$)
  2. Only good 0 is distorted ($\hat{\tau}_j^i = 0$ for $j \geq 1$)
  3. Both the tax and non-tax distortions are homogeneous across individuals ($\tau_0^i = \tau_0$ and $\hat{\tau}_0^i = \hat{\tau}_0$)

- We then obtain:

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

  where $\bar{\varepsilon}_0 \equiv \int_i \left[x_0^i \varepsilon_{00}^i \right] di$

- The sufficient statistics are $\bar{\varepsilon}_0$ and $\hat{\tau}_0$

Sufficient Statistics with Non-Government Externalities

- Under a general formulation, we have developed sufficient statistics formulas that retain the standard form:
  - Interactions between elasticities and modified "tax wedges"
  - The modified wedges include any externality/internality effects
  - Welfare evaluation does not require a fully specified model of each market imperfection, but can be written in terms of reduced-form gaps between private and social prices
- The empirical challenge is that we have to estimate both the elasticities and the externality-adjusted wedges

Bridge Between Structural and Reduced-Form Estimation?

- The sufficient statistics approach builds on a powerful logic, provides clear intuition, and establishes a transparent link between theory and data
- The approach does not (always) require specific functional forms, but it requires other assumptions on policy parameters, decision environment, and preferences
- The claim that simple sufficient statistics formulas (depending on one or two reduced-form parameters) are robust across a "broad class" of models is tenuous
- The transparency and intuition of the approach is important, but in feasible implementations it *is* a structural approach