Motivation

- Inequality generates need for interpersonal comparisons
  - Decisions about economic policies (R&D, free trade, mergers, safety net, health, education, taxation, etc.)
  - General measurement of societal well-being

- Two common economic methods for resolving interpersonal comparisons
  1. Kaldor Hicks Compensation Principle (Kaldor (1939), Hicks (1939, 1940))
     - Motivates aggregate surplus, or “efficiency”, as normative criteria
     - Ignores issues of “equity”
  2. Social welfare function (Bergson (1938), Samuelson (1947), Diamond and Mirrlees (1971), Saez and Stantcheva (2015))
     - Allows preference for equity
     - Subjective choice of researcher or policy-maker

Revisit Kaldor-Hicks

Modify so that transfers are incentive compatible (Mirrlees (1971))

Kaldor and Hicks envisioned feasible transfers:

“If, as will often happen, the best methods of compensation feasible involve some loss in productive efficiency, this loss will have to be taken into account. (Hicks, 1939)


Provide simple (yet general) empirical method of accounting for these distortions
Key idea: Envelope theorem allows for empirical method to account for distortions

- Corresponds to weighting surplus by the “inequality deflator”
- Turns unequal surplus into equal surplus using modifications to the tax schedule

Inequality deflator is the marginal cost to government of providing $1 of welfare to an income level

- Differs from $1 because of how behavioral response affects government budget (basic PF logic)

Suppose we want to provide transfers to those earning near $y^*$
Modifications to Income Tax Schedule

(c)

Consumption

Earnings (y)

y-T(y)
Modifications to Income Tax Schedule

![Graph showing consumption and earnings relationship]

- Consumption
- Earnings

\[ y^* \]
Modifications to Income Tax Schedule

(c) 

Consumption

Earnings (y)

y - T(y)

y* 

ε
Modifications to Income Tax Schedule

\[ \eta \quad \varepsilon \quad y^* \quad y - T(y) \]

(c)

Consumption

Earnings \( (y) \)
Modifications to Income Tax Schedule

Marginal welfare impact per $\eta$: 
=$1 per mechanical beneficiary
Modifications to Income Tax Schedule

Mechanical cost per $\eta$:

$$= F(y^*+\epsilon/2) - F(y^*-\epsilon/2)$$

$$= $1 per mechanical beneficiary

$y^*$

(c)

Consumption

Earnings $(y)$
Behavioral responses affect tax revenue
But don’t affect utility (Envelope Theorem)
Total cost per η per beneficiary:
=1+FE(y*), FE = “fiscal externality”
Total cost per \( \eta \) per beneficiary:
\[ = 1 + \text{FE}(y^*) \], \text{FE} = “fiscal externality”

Inequality Deflator:
\[ g(y) = \frac{1 + \text{FE}(y)}{\text{E}[1 + \text{FE}(y)]} \]
Formal Model

- Inequality deflator:

\[ g(y) = \frac{1 + FE(y)}{E[1 + FE(y)]} \]

- To first order: $1$ surplus to those earning \( y \) can be turned into \( \frac{g(y)}{n} \) surplus to everyone through modifications to tax schedule.
Inequality deflator:

\[ g(y) = \frac{1 + FE(y)}{E[1 + FE(y)]} \]

To first order: $1$ surplus to those earning \( y \) can be turned into \( g(y) / n \) surplus to everyone through modifications to tax schedule.

Fiscal externality logic does not rely on functional form assumptions.
- Allows for each person to have her own utility function and arbitrary behavioral responses.
- Extends to multiple policy dimensions.
- Nests a lot of optimal tax expressions as special case (e.g. Mirrlees).
Inequality deflator: 

\[ g(y) = \frac{1 + FE(y)}{E[1 + FE(y)]} \]

To first order: $1$ surplus to those earning \( y \) can be turned into $g(y)/n$ surplus to everyone through modifications to tax schedule.

Fiscal externality logic does not rely on functional form assumptions.

- Allows for each person to have her own utility function and arbitrary behavioral responses.
- Extends to multiple policy dimensions.
- Nests a lot of optimal tax expressions as special case (e.g. Mirrlees).

Key assumption: “partial equilibrium” / “local incidence”

- Behavioral response only induces a fiscal externality.
- Other incidence/externalities would need to be accounted for.
Formal Model

- Inequality deflator:

\[ g(y) = \frac{1 + FE(y)}{E[1 + FE(y)]} \]

- To first order: $1$ surplus to those earning $y$ can be turned into $g(y)/n$ surplus to everyone through modifications to tax schedule.

- Fiscal externality logic does not rely on functional form assumptions:
  - Allows for each person to have her own utility function and arbitrary behavioral responses
  - Extends to multiple policy dimensions
  - Nests a lot of optimal tax expressions as special case (e.g. Mirrlees)

- Key assumption: “partial equilibrium” / “local incidence”:
  - Behavioral response only induces a fiscal externality
  - Other incidence/externalities would need to be accounted for

- Inequality Deflator can turn unequal surplus into equal surplus.
Example: Alternative Environment Benefits Poor

Example: Alternative environment benefits the poor and harms the rich
Given $s(y)$, two ways of neutralizing distributional comparisons
Given \( s(y) \), two ways of neutralizing distributional comparisons

“EV”: modify status quo tax schedule
Given $s(y)$, two ways of neutralizing distributional comparisons

“EV”: modify status quo tax schedule

By how much can everyone be made better off in modified status quo world relative alternative environment?
“EV” Example

- Replicate surplus in status quo environment

Graph:
- Consumption vs. Earnings
- \( y - T(y) \)
Replicate surplus in status quo environment

\[ y - T(y) \]

\[ y - T(y) + s(y) \]
Replicate surplus in status quo environment

Is $\hat{T}$ budget feasible?

$y - T(y)$

$y - \hat{T}(y)$

$y - T(y) + s(y)$
Is \( \hat{T} \) budget feasible?
Case 1: YES
"EV" Example

---

\[ y - T(y) \]

\[ s(y) \]

Is \[ T \] budget feasible?

Case 2: NO

---

Graph showing consumption and earnings. The graph illustrates the relationship between consumption and earnings, with curves for different scenarios. The notation \[ s(y) \] and \[ y - T(y) \] are used to denote specific points or calculations in the context of the graph.
Inequality Deflated Surplus: $S^{ID} = E[s(y)g(y)]$

How much better off is everyone in the alternative environment relative to a modified status quo?
Inequality Deflated Surplus: $S^{ID} = E[s(y)g(y)]$

How much better off is everyone in the alternative environment relative to a modified status quo?

$S^{ID} < 0$  Modified status quo delivers Pareto improvement

$S^{ID} > 0$  $s(y)$

$y - T(y)$  $\hat{y}$

Consumption  Earnings $(y)$
Given $s(y)$, two ways of neutralizing distributional comparisons:

- "EV": modify status quo tax schedule
  - By how much can everyone be made better off in modified status quo world relative alternative environment?

- "CV": modify alternative environment tax schedule
  - By how much can everyone be made better off in modified alternative environment relative to status quo?
Compensate Lost Surplus in Alternative environment
“CV” Example

Compensate Lost Surplus in Alternative environment
“CV” Example

Compensate Lost Surplus in Alternative environment

Consumption

Earnings (y)

\[ \gamma - T^a(y) \]

\[ -s(y) \]

\[ S^{ID} > 0 \]
"CV" Example

Compensate Lost Surplus in Alternative environment

\[ y_{-T^a(y)} \]

\[ y_{^\wedge T^a(y)} \]
Inequality Deflated Surplus: $S^D = E[s(y)g(y)]$

How much better off is everyone in the modified alternative environment relative to the status quo?
Inequality Deflated Surplus: $S^{ID} = E[s(y)g(y)]$

How much better off is everyone in the modified alternative environment relative to the status quo?

$S^{ID} > 0$ Modified alternative environment delivers Pareto improvement
If \( g(y) \) is similar in status quo and alternative environment, these two interpretations of inequality deflated surplus are first-order equivalent.

- Similar to first order equivalence of CV and EV

When surplus is homogeneous conditional on income:

- \( S^{ID} \) provides first-order characterization of potential Pareto comparisons
- \( S^{ID} \) quantifies difference between environments without making inter-personal comparisons
  - By how much is everyone better off?
Heterogeneous Surplus

- Redistribution based on income, not individual-specific
  - Two people with same income, \( y(\theta) \), can have different surplus, \( s(\theta) \)
  - Income tax is a “blunt instrument”
  - \( \int s(\theta) g(y(\theta)) = \) how much on average is each income level better off
    - Search for potential Pareto comparisons more difficult
  - But inequality deflator can still be used to characterize Pareto comparisons

**Proposition**

- Define
  
  \[
  S^{ID} = E \left[ \min \{ s(\theta) | y(\theta) = y \} g(y) \right] > 0 \\
  \overline{S}^{ID} = E \left[ \max \{ s(\theta) | y(\theta) = y \} g(y) \right] < 0
  \]

- Modified alternative environment delivers Pareto improvement iff \( S^{ID} > 0 \)
- Modified status quo offers Pareto improvement iff \( \overline{S}^{ID} < 0 \)
Heterogeneous Surplus

- No potential Pareto ranking when $S^{ID} < 0 < \bar{S}^{ID}$
- Potential solution: Add more status quo policies
- Marginal cost $1 + FE(X)$ as opposed to $1 + FE(y)$
  - Augment both tax schedule and Medicaid
- Inequality deflator well-suited for comparisons in which surplus does not vary conditional on income, so that $S^{ID} = S^{ID} = \bar{S}^{ID}$
Existing evidence on behavioral responses to taxation provides guidance on \( 1 + FE(y) \)
Existing evidence on behavioral responses to taxation provides guidance on $1 + FE(y)$

- EITC causes people to:
  - Enter the labor force (summary in Hotz and Scholz (2003))
  - Distort earnings (Chetty et al 2013).
  - $1 + FE(y) \approx 1.14$ for low-earners (calculation in Hendren 2013)
Existing evidence on behavioral responses to taxation provides guidance on $1 + FE(y)$

- EITC causes people to:
  - Enter the labor force (summary in Hotz and Scholz (2003)).
  - Distort earnings (Chetty et al 2013).
  - $1 + FE(y) \approx 1.14$ for low-earners (calculation in Hendren 2013)

- Taxing top incomes causes:
  - Reduction in taxable income (review in Saez et al 2012)
  - Implies $1 + FE(y) \approx 0.50 - 0.75$
  - Disagreement about amount, but general agreement on the sign: $FE(y) < 0$
Existing evidence on behavioral responses to taxation provides guidance on $1 + FE(y)$

- EITC causes people to:
  - Enter the labor force (summary in Hotz and Scholz (2003))
  - Distort earnings (Chetty et al. 2013).
  - $1 + FE(y) \approx 1.14$ for low-earners (calculation in Hendren 2013)

- Taxing top incomes causes:
  - Reduction in taxable income (review in Saez et al. 2012)
  - Implies $1 + FE(y) \approx 0.50 - 0.75$
  - Disagreement about amount, but general agreement on the sign: $FE(y) < 0$

Reduced form empirical evidence suggests deflator values poor more so than the rich
Existing evidence on behavioral responses to taxation provides guidance on $1 + FE(y)$

- **EITC causes people to:**
  - Enter the labor force (summary in Hotz and Scholz (2003))
  - Distort earnings (Chetty et al 2013).
  - $1 + FE(y) \approx 1.14$ for low-earners (calculation in Hendren 2013)

- **Taxing top incomes causes:**
  - Reduction in taxable income (review in Saez et al 2012)
  - Implies $1 + FE(y) \approx 0.50 – 0.75$
  - Disagreement about amount, but general agreement on the sign: $FE(y) < 0$

- **Reduced form empirical evidence suggests deflator values poor more so than the rich**
  - Despite evidence that taxable income elasticities may be quite stable across the income distribution (e.g. Chetty 2012)
Use optimal tax approach to write $FE(y)$ as function of taxable income elasticities.
Use optimal tax approach to write $FE(y)$ as function of taxable income elasticities.

Let

- $\epsilon^c(y) = \text{avg comp. elasticity for those earning } y$
- $\zeta(y) = \text{avg inc. effect for those earning } y$
- $\epsilon^P(y) = \text{avg LFP rate elasticity for those earning } y$
Optimal Tax Expression

For every point, $y^*$, such that $T'(y)$ and $\epsilon^c(y^*)$ are locally constant and the distribution of income is continuous:

$$FE(y^*) = -\epsilon^p(y^*) \frac{T(y) - T(0)}{y - T(y)} - \zeta(y^*) \frac{\tau(y^*)}{1 - \frac{T(y^*)}{y^*}} - \epsilon^c(y^*) \frac{\tau(y^*)}{1 - \tau(y^*)} \alpha(y^*)$$

where $\alpha(y) = -\left(1 + \frac{y f'(y)}{f(y)}\right)$ is the local Pareto parameter of the income distribution.

- Heterogeneity in $FE(y)$ depends on:
  1. Shape of income distribution, $\alpha(y)$
  2. Shape and size of behavioral elasticities
  3. Shape of tax rates

General Formula
Optimal Tax Expression

For every point, $y^*$, such that $T'(y)$ and $\epsilon^c(y^*)$ are locally constant and the distribution of income is continuous:

$$FE(y^*) = -\epsilon^P(y^*) \frac{T(y) - T(0)}{y - T(y)} - \zeta(y^*) \frac{\tau(y^*)}{1 - \frac{T(y^*)}{y^*}} - \epsilon^c(y^*) \frac{\tau(y^*)}{1 - \tau(y^*)} \alpha(y^*)$$

where $\alpha(y) = -\left(1 + \frac{yf'(y)}{f(y)}\right)$ is the local Pareto parameter of the income distribution.

- Heterogeneity in $FE(y)$ depends on:
  1. Shape of income distribution, $\alpha(y)$
  2. Shape and size of behavioral elasticities
  3. Shape of tax rates

- Generalized version of “uni-dimensional” formulas (e.g. Bourguignon and Spadaro (2012), Zoutman (2013a, 2013b))
Estimation Approach

- Calibrate behavioral elasticities from existing literature on taxable income elasticities

Calibration Details

- Estimate shape of income distribution and marginal income tax rate using universe of US income tax returns
- Account for covariance between elasticity of income distribution and marginal tax rate

Estimation Details

Nathaniel Hendren (Harvard)
Cost of Inequality
July, 2016
Estimation Approach

- Calibrate behavioral elasticities from existing literature on taxable income elasticities
  - Assess robustness to range of estimates (e.g. compensated elasticity of 0.1, 0.3, and 0.5)

Calibration Details

Estimate shape of income distribution and marginal income tax rate using universe of US income tax returns

Account for covariance between elasticity of income distribution and marginal tax rate

Calibration Details
Estimation Approach

- Calibrate behavioral elasticities from existing literature on taxable income elasticities
  - Assess robustness to range of estimates (e.g. compensated elasticity of 0.1, 0.3, and 0.5)
- Estimate shape of income distribution and marginal income tax rate using universe of US income tax returns
Estimation Approach

- Calibrate behavioral elasticities from existing literature on taxable income elasticities
  - Assess robustness to range of estimates (e.g. compensated elasticity of 0.1, 0.3, and 0.5)

- Estimate shape of income distribution and marginal income tax rate using universe of US income tax returns
  - Account for covariance between elasticity of income distribution and marginal tax rate
Application #1: Income Inequality

After-Tax Household Income Distribution by Quintile

Source: CBO; Supplemental Tables 43373, Table 7
1. Income Distributions

- Define surplus function

\[ s(\theta) = Q_a(\alpha(\theta)) - Q_0(\alpha(\theta)) \]
Define surplus function

\[ s(\theta) = Q_a(\alpha(\theta)) - Q_0(\alpha(\theta)) \]

How much growth if tax schedule distributed it equally?
1. Income Distributions

- Define surplus function

\[ s(\theta) = Q_a(\alpha(\theta)) - Q_0(\alpha(\theta)) \]

- How much growth if tax schedule distributed it equally?
  - Could do other counterfactual experiments…
1. Income Distributions

- Define surplus function

\[ s(\theta) = Q_a(\alpha(\theta)) - Q_0(\alpha(\theta)) \]

- How much growth if tax schedule distributed it equally?
  - Could do other counterfactual experiments...
    - Search for Pareto comparisons using particular counterfactuals (e.g. SBTC vs no SBTC)
1. Income Distributions

- Define surplus function

\[ s(\theta) = Q_a(\alpha(\theta)) - Q_0(\alpha(\theta)) \]

- How much growth if tax schedule distributed it equally?
  - Could do other counterfactual experiments...
    - Search for Pareto comparisons using particular counterfactuals (e.g. SBTC vs no SBTC)

- Quantile stability implements Kaldor (1939)’s idea of holding distribution constant + Hicks (1939) idea of doing it in cheapest manner possible
1. Income Distributions

- Define surplus function

\[ s(\theta) = Q_\alpha(\alpha(\theta)) - Q_0(\alpha(\theta)) \]

- How much growth if tax schedule distributed it equally?
  - Could do other counterfactual experiments...
    - Search for Pareto comparisons using particular counterfactuals (e.g. SBTC vs no SBTC)

- Quantile stability implements Kaldor (1939)’s idea of holding distribution constant + Hicks (1939) idea of doing it in cheapest manner possible
  - More costly to make the rich poor and the poor rich than to keep everyone rich and poor
Deflated Growth

Raw and Deflated Household Income Change

Difference Relative to 2012

Year

Mean HH Inc (CPI Deflated)  Inequality Deflated HH Inc
Low Elasticity Specification  High Elasticity Specification

Social Cost of Increased Income Inequality

119K * (Mean - Deflated Change in HH Income)

Year

Baseline
Low Elasticity (e = 0.1)
High Elasticity (e = 0.5)
Country Comparison to US

Inequality Deflated Surplus vs. 2012 GNI Per Capita

Inequality Deflated Surplus vs. 2012 GNI Per Capita

Other Country Orderings

Preference for “pro poor” policies?

Two conceptualizations of policy experiments:

Non-Budget Neutral (“should the government spend money on...”)

Budget Neutral

Budget neutral policies: weight surplus by inequality deflator

Cost of Inequality

July, 2016 45 / 49
Example: Producer versus Consumer Surplus

- Suppose budget neutral policy with benefits to producers $S^P$ and consumers $S^C$
  - Extreme assumption: producer surplus falls to top 1%
  - Consumer surplus falls evenly across income distribution
- Optimal weighting:
  $$S^{ID} = 0.77S^P + S^C$$
- “Consumer surplus standard” requires top tax rate near Laffer curve
  - France should have tighter merger regulations?
- Key assumption: policy is budget neutral (inclusive of fiscal externalities)
- What about non-budget neutral policies?
Suppose $G$ affects those with income $y$

Construct

$$MVPF_G = \frac{s(y)}{1 + FE^G}$$

- WTP per unit gov’t revenue (Mayshar 1990; Slemrod and Yitzhaki 2001; Hendren 2013)
- Depends on causal effects ($FE^G$) and WTP for non-market good

Additional spending on $G$ desirable iff

$$\text{Value of } G \geq \frac{1}{1 + FE(y)} \\
\text{Value of } T(y)$$
MVPF of Targeted Policies


Income is average income of policy beneficiaries normalized to 2012 income using CPI-U

- **Section 8**
- **Food Stamps**
- **JTPA**
Inequality isn’t just a transfer!

Policy implications
- Compare policies to the efficiency of the tax schedule
- Weighting individual WTP by inequality deflator provides method to do this

General idea: use marginal costs of feasible redistribution + envelope theorem + Pareto principle instead of a SWF

Key question: what policies are more efficient than the income tax schedule at redistribution?