Discussion of
“Who wins, who loses? Tools for distributional policy evaluation”

James Heckman

September 26, 2014
Kasy Paper: Many Moving Parts

Claims

(1) Point or set identification of “expected” individual welfare gains conditional on choice variables
(2) Distribution of welfare effects
(3) No restriction on dimension of the heterogeneity vis-à-vis dimension of endogenous variables
(4) Aggregation over the population using a SWF
(5) GE effects
(6) Study of the EITC impact on welfare

Some Key Assumptions

(1) One dimensional parameterization of policies
(2) Functions differentiable in policies
(3) Strong support conditions
Table 1: Comparison of Steady States Under Alternative Tax Regimes

<table>
<thead>
<tr>
<th>Variable</th>
<th>Percentage difference from benchmark progressive case</th>
<th>Flat tax</th>
<th>Flat consumption tax</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>PE</td>
<td>GE</td>
</tr>
<tr>
<td>Fraction attending college</td>
<td></td>
<td>18.79</td>
<td>0.26</td>
</tr>
<tr>
<td>Aggregate output</td>
<td></td>
<td>-0.09</td>
<td>1.15</td>
</tr>
<tr>
<td>Aggregate consumption</td>
<td></td>
<td>-0.08</td>
<td>0.16</td>
</tr>
<tr>
<td>Mean wage, college graduate</td>
<td></td>
<td>3.39</td>
<td>2.60</td>
</tr>
<tr>
<td>Mean wage, high school graduate</td>
<td></td>
<td>2.44</td>
<td>2.44</td>
</tr>
</tbody>
</table>

Placing Paper in Literature on Distributional Treatment Effects

- Two outcome model: \((Y_0, Y_1)\)
- Observe only one coordinate and that subject to selection bias
- \(D = 1\) if person gets treatment; \(D = 0\) otherwise
- \(Y = DY_1 + (1 - D)Y_0\)
Two Problems

I. From data on outcomes $F_1 (y_1 \mid D = 1, X), F_0 (y_0 \mid D = 0, X)$, under what conditions can one recover $F_1 (y_1 \mid X)$ and $F_0 (y_0 \mid X)$, respectively?

II. Construct the joint distribution $F (y_0, y_1 \mid X)$ from the marginal distributions.
Why Bother Identifying Joint Distributions?
Depends on the criterion

\[ Pr(Y_1 \geq Y_0) : \text{Percentage of people voting} \]
\[ Pr(Y_1 > Y_0 | Y_0 < y_0) : \text{Gains to poor people in base state} \]
Solutions

- Two basic approaches in the literature to solving the problem of identifying $F(y_0, y_1 | X)$

(A) Bounds
(B) Solutions that postulate assumptions about dependence between $Y_0$ and $Y_1$.
(C) Solutions based on information from agent participation rules and choice data.
Bounds

(1) Frechet Bounds
(2) Makarov Inequality Bounds (For $Y_1 - Y_0$)
Solutions Based on Conditional Independence or Matching

- **Q**: Conditioning Variable
- \( F_0 (y_0 \mid D = 0, X, Q) = F_0 (y_0 \mid X, Q) \)
- \( F_1 (y_1 \mid D = 1, X, Q) = F_1 (y_1 \mid X, Q) \).
- All of the dependence between \((Y_0, Y_1)\) given \(X\) comes through \(Q\)
- \( F (y_1, y_0 \mid X, Q) = F_1 (y_1 \mid X, Q) F_0 (y_0 \mid X, Q) \).
Common Coefficient:

\[ Y_1 - Y_0 = \Delta \]  \hspace{1cm} (1)

\( \Delta \) is a constant given \( X \).
Quantile Treatment Effects

- $Y_1 = F_1^{-1}(F_0(Y_0))$.
- This is the tight upper bound of the Fréchet bounds.
- Alternative assumption: $Y_1 = F_1^{-1}(1 - F_0(Y_0))$.
- Tight Fréchet lower bound.
Constructing Distributions from Assuming Independence of the Gain from the Base

C-1

\[ Y_1 = Y_0 + \Delta \]
\[ Y_0 \perp \Delta | X. \]

M-1

\( (Y_0, Y_1) \perp D | X, \)

- Identify \( F(y_0, y_1 | X) \) from the cross section outcome distributions of participants and non-participants and estimate the joint distribution by using deconvolution.
Information From Revealed Preference

- e.g. Roy Model, Generalized Roy Model
Additional Information

- Dependence through factor structure or other assumptions on copulas.
Kasy: $Y_d$ on a continuum

$$d = \phi(\alpha)$$

Same unobservables across all $d$
Identification (Non-parametric)

1. Main challenge: \( \gamma(w, l) = E[\dot{w} \cdot l | w \cdot l, \alpha] \)
2. Causal effect of policy
3. Conditional on endogenous outcomes,

Aggregation

Social welfare & distributional decompositions

1. Welfare weights \( \approx \) derivative of influence function
2. Welfare impact = impact on income - behavioral correction

Inference

1. Local linear quantile regressions
2. Combined with control functions
3. Suitable weighted averages
Setup

\[ u(c, l) : \]

- \( u \) differentiable, increasing in \( c \), decreasing in \( l \), quasiconcave, does not depend on \( \alpha \)
- \( \gamma(\alpha) \) continuous in \( \alpha \).
Objects of Interest: \( W \): conditioning variables not affected by \( \alpha \)

Sets of winners and losers:

\[
W := \{(y, W) : \gamma(y, W) \geq 0\}
\]
\[
L := \{(y, W) : \gamma(y, W) \leq 0\}
\]
Question:

Treatment of heterogeneity among winners and losers?
Want to compute

\[ Pr(\gamma(y, W) \geq 0) \]

Need more than \( E(\gamma(y, W)) \)

- Given \( y, W \) what is the distribution of \( \gamma \)?
Identification of disaggregated welfare effects

- **Goal:** identify \( \gamma(y, W) = E[\hat{\varepsilon}|y, W, \alpha] \)

- Simplified case:
  - no change in prices, or unearned income
  - no covariates, *just tax change*

- Then
  \[
  \gamma(y) = E[l \cdot (1 - \partial z_t) \cdot \hat{w}|l \cdot w, \alpha]
  \]
  \[ z = lw \]

- Denote \( x = (l, w) \).
  Seek to identify
  \[
  g(x, \alpha) = E[\dot{x}|x, \alpha]
  \]  \hspace{1cm} (2)
  from
  \[
  f(x|\alpha).
  \]

- Made necessary by combination of
  1. utility-based social welfare
  2. heterogeneous wage response.
Question: Do we really know $f(x|\alpha)$? (Random assignment with a continuum of treatments)
If so, can do table look up for policy effects.
How dense is the set of policies?
Assumptions:

1. \( x = x(\alpha, \epsilon), x \in \mathbb{R}^k \)
2. \( \alpha \perp \epsilon \) (really “\( \perp \perp \) ”)
3. \( x(., \epsilon) \) differentiable (\( \epsilon \) explicitly introduced for the first time)
Analogy from fluid dynamics:

- $x(\alpha, \epsilon)$: position of particle $\epsilon$ at time $\alpha$
- $f(x|\alpha)$: density of gas / fluid at time $\alpha$, position $x$
- $\dot{f}$: change of density
- $h(x, \alpha) = E[\dot{x}|x, \alpha] \cdot f(x|\alpha)$: “flow density”
• Knowledge of $f(x|\alpha)$
  • identifies $\nabla \cdot h = \sum_{j=1}^{k} \partial_{x_j} h^j$
    ($k =$ number of endogenous variables)
  • where $h = E[\dot{x}|x, \alpha] \cdot f(x|\alpha)$
  • identifies nothing else.

• Add to $h$
  • $\tilde{h}$ such that $\nabla \cdot \tilde{h} \equiv 0$

• cannot identify true $h(= h^0)$ from $\tilde{h}$ perturbations
Density and flow

\[ \dot{f} = -\nabla \cdot h \]
• This relationship is a property of differentiability of functions.
Question: Additional restrictions from properties of expenditure function? Properties of demand functions?
Theorem

The identified set for $h$ is given by

$$h^0 + \mathcal{H}$$

where

$$\mathcal{H} = \{ \tilde{h} : \nabla \cdot \tilde{h} \equiv 0 \}$$

$$h^{0j}(x, \alpha) = f(x|\alpha) \cdot \partial_\alpha Q(v^j|v^1, \ldots, v^{j-1}, \alpha)$$

$$v^j = F(x^j|x^1, \ldots, x^{j-1}, \alpha)$$
Theorem

1. Suppose $k = 1$. Then

$$\mathcal{H} = \{ \tilde{h} \equiv 0 \}. \quad (5)$$

2. Suppose $k = 2$. Then

$$\mathcal{H} = \{ \tilde{h} : \tilde{h} = A \cdot \nabla H \text{ for some } H : \mathcal{X} \to \mathbb{R} \}. \quad (6)$$

where

$$A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$  

3. Suppose $k = 3$. Then

$$\mathcal{H} = \{ \tilde{h} : \tilde{h} = \nabla \times G \}. \quad (7)$$

where $G : \mathcal{X} \to \mathbb{R}^3$. □
Question: Initial Conditions? Boundary values? How identified?
Theorem

Assume

\[ \frac{\partial}{\partial x^j} E[\dot{x}^i|x, \alpha] = 0 \text{ for } j > i. \] (8)

Then \( h \) is point identified, and equal to \( h^0 \) as defined before.

In particular

\[ g^j(x, \alpha) = E[\dot{x}^j|x, \alpha] = \partial_\alpha Q(v^j|v^1, \ldots, v^{j-1}, \alpha). \] \( \square \)

Triangularity (e.g. as in Blundell and Matzkin, 2014).
Multiple component policies?

(a) Does result generalize to PDEs?
(b) Convert to ODE? Method of characteristic curves?
(c) What restrictions give linear and hyperbolic PDEs?
Aggregation

- Relationship
  social welfare ⇔ distributional decompositions?

- public finance welfare weights
  $\approx$ derivative of dist decomp influence functions

- Alternative representations of $\dot{SWF}$
  $\Rightarrow$ alternative ways to estimate $\dot{SWF}$:
  1. weighted average of individual welfare effects $\dot{e}, \gamma$
  2. distributional decomposition for counterfactual income $\tilde{y}$
     (holding labor supply constant)
  3. distributional decomposition of realized income minus behavioral correction
Estimation

1. First estimate the disaggregated welfare impact

\[
\gamma(y, W) = E[\hat{\gamma}|y, W, \alpha] = E[l \cdot \hat{w} \cdot (1 - \partial_z t) - \dot{t} + y_0 - c \cdot \hat{p}|y, W, \alpha] \tag{9}
\]

2. Then estimate other objects by plugging in \(\hat{\gamma}\):

\[
\hat{W} = \{(y, W) : \hat{\gamma}(y, W) \geq 0\}
\]

\[
\hat{L} = \{(y, W) : \hat{\gamma}(y, W) \leq 0\}
\]

\[
\hat{\text{SWF}} = E_{N}[\omega_i \cdot \hat{\gamma}(y_i, W_i)]. \tag{10}
\]
Question:

Can SWF depend on $x$?
• Use restrictions from economic theory in the estimation of nonparametric models of consumer behaviour
• Develop nonparametric methods that can be applied to general systems of demands
• Can be used to construct $\gamma$. 

Consumer Demand with Multidimensional Nonseparable Unobserved Heterogeneity
Blundell, Kristensen, and Matzkin 2014
• Particular attention is given to discrete prices, multiple goods and nonseparable unobserved heterogeneity.
• Demand systems are the reduced form of models of simultaneous equations.
• Develop methods for simultaneous equations.
• Use them to identify the effect on any particular individuals of a change in his/her budget set.
• Identifying individuals across different Engel curves allows to impose Revealed Preference inequalities on the demand of any particular individual.
• **Z**: set of conditioning variables

• Identification when:
  1. A unimodal restriction with respect to \( Z \) on the conditional density of the vector of unobserved heterogeneity.
  2. Rank condition on the conditional density of the unobserved heterogeneity given \( Z \).

• Methods to estimate the value of the vector of unobserved tastes of each consumer and the demand function of each consumer.

• Methods to estimate the effect of finite and infinitesimal changes in prices and income on the demand of each individual consumer.

• Estimators are consistent and asymptotically normal.
System of demand functions

\[
Y_1 = d^1(p, l, \varepsilon_1, ..., \varepsilon_G)
\]
\[
Y_2 = d^2(p, l, \varepsilon_1, ..., \varepsilon_G)
\]
\[...
\]
\[
Y_G = d^G(p, l, \varepsilon_1, ..., \varepsilon_G)
\]

$(\varepsilon_1, ..., \varepsilon_G)$ is independent of $(p, l)$ conditional on $Z$

$(\varepsilon_1, ..., \varepsilon_G)$ vector of unobserved heterogeneity (tastes)
They show

- When $Z$ is discrete the rank condition can still be used for point identification results.
- When $Z$ is discrete the unimodal condition can be used for partial (set) identification of unobserved tastes $\varepsilon$.
- Methods can be used for continuously distributed and for discrete prices.
- When prices are discrete, partial identification results for the demand of any particular budget that had not been observed, ‘predicted demands’, can be obtained by extending the results in Blundell, Kristensen, and Matzkin (2011).
• Identify the effect of a discrete change in \((p, l)\) when \(\varepsilon\) stays fixed

\[ d(p', l', \varepsilon) - d(p, l, \varepsilon) \]

• Identify the effect of an infinitesimal change in \((p, l)\) when \(\varepsilon\) stays fixed

\[ \frac{\partial d(p, l, \varepsilon)}{\partial p} \quad \text{and} \quad \frac{\partial d(p, l, \varepsilon)}{\partial l} \]
• System of demand functions

\[ Y_1 = d^1(p, l, \varepsilon_1, \ldots, \varepsilon_G) \]
\[ Y_2 = d^2(p, l, \varepsilon_1, \ldots, \varepsilon_G) \]
\[ \ldots \]
\[ Y_G = d^G(p, l, \varepsilon_1, \ldots, \varepsilon_G) \]

where \((\varepsilon_1, \ldots, \varepsilon_G)\) is independent of \((p, l)\) conditional on \(Z\)
• **Key assumption** (from Matzkin, 2008, and other co-authored papers)

• **Invertibility**

\[
\varepsilon_1 = r^1(Y_1, \ldots, Y_G, p, l) \\
\varepsilon_2 = r^2(Y_1, \ldots, Y_G, p, l) \\
\vdots \\
\varepsilon_G = r^G(Y_1, \ldots, Y_G, p, l)
\]
EITC: Questions

- Many reforms not captured by scalar $\alpha$
- Multiple margins and policy change over time not uniform
- EITC affects skill accumulation and wages at micro level
- Model of skill accumulation estimated affects estimates
- Are estimated wage effects GE effects or the effects on skill?
Does scalar $\alpha$ describe this policy?
Figure 1: Simulated Effects of EITC on Hours Worked (OJT model)

**Figure 2**: Simulated Effects of EITC on Hours Worked (LBD model)

**Source**: Heckman, Lochner and Cossa (2003)
Figure 3: Simulated Effects of EITC on Human Capital (OJT model)

Non–White Females – Less than 10th grade

White Females – HS graduates

**Figure 4**: Simulated Effects of EITC on Human Capital (LBD model)

Non–White Females – Less than 10th grade

White Females – HS graduates

Figure 5: Simulated Effects of EITC on Wage Income (OJT model)

Non–White Females – Less than 10th grade

White Females – HS graduates

Figure 6: Simulated Effects of EITC on Wage Income (LBD model)