Empirical Research on Economic Inequality
Equivalent variation and welfare

Maximilian Kasy

Harvard University, fall 2015

Welfare versus observables

- Previous classes: distribution of observable variables
  - (top) incomes, as found in tax records
  - invitations to job interviews, as collected in audit experiments
  - earnings, for union and non-union members
  - incomes, again, and their relation to parental income
- Recall notions of individual welfare which *do not* correspond to these variables:
  - Utility
  - Capabilities
  - Opportunities

Price changes

- Suppose trade liberalization changes price of goods such as rice.
- Should affect different people differently:
  - Poor city dwellers consuming rice.
  - Small farmers producing for own consumption.
  - Larger farmers producing for the market.
  - ...
- Can we quantify the welfare impact of such a change in prices?

Deaton, A. (1989). Rice prices and income distribution in Thailand: a non-parametric analysis. The Economic Journal, *pages 1–37.*

Reminder: Utility

- General setup:
  - Choice set C_i
  - Utility function $u_i(x)$, for $x \in C_i$
  - Realized welfare

$$v_i = \max_{x \in C_i} u_i(x).$$

- Double role of utility
  - Determines choices (individuals choose utility-maximizing x)
  - Normative yardstick (welfare is realized utility)

Can we measure utility?

- Utility can not be observed.
- But we *do* observe choice sets and choices!
- Trick: change the question in two ways
  1. **Changes** in utility, rather than **levels** of utility.
  2. Transfers of **money** that would induce similar changes of utility, rather than changes in **utility** itself.
- $\Rightarrow$ Equivalent variation

The consumer problem

- Assume: Individuals choose **consumption** to **maximize utility.**
- Constraint: Expenses may not exceed income.
- Two consumption goods, good 1 and good 2, with prices p_1 and p_2
- Individuals i
- Income y_i
- Choose consumption $x_i = (x_{1,i}, x_{2,i})$ to maximize their utility u_i.
- Special case of general setup discussed 2 slides ago!

Consumer problem, continued

- Utility maximization:

$$x_i(p, y_i) = \underset{x}{\operatorname{argmax}}\ u_i(x),$$

- subject to the budget constraint

$$x_{1,i} \cdot p_1 + x_{2,i} \cdot p_2 \leq y_i.$$

- Utility v_i that a household can achieve for given prices and income is equal to the utility of the chosen consumption bundle,

$$v_i(p, y_i) = u_i(x_i(p, y_i)).$$

Questions for you

- Write x_1 as function of x_2 under the budget constraint.
- Substitute this into the utility maximization problem.
- Calculate the first order condition for the utility maximization problem.

Solution

- Budget constraint:

$$x_{1,i} = \frac{1}{p_1}(y_i - x_{2,i} \cdot p_2)$$

- Substitute into utility maximization problem:

$$x_{2,i} = \underset{x_2}{\operatorname{argmax}}\ u_i\left(\frac{1}{p_1}(y_i - x_2 \cdot p_2), x_2\right)$$

- First order condition:

$$\frac{\partial}{\partial x_2}\left[u_i\left(\frac{1}{p_1}(y_i - x_2 \cdot p_2), x_2\right)\right] = 0$$

- Rewriting the first order condition:

$$\frac{\partial_{x_1} u_i(x_i)}{p_1} = \frac{\partial_{x_2} u_i(x_i)}{p_2}$$

- "The ratio of marginal benefits to marginal costs has to be the same for both goods."

The welfare effect of changing prices

What is $\partial_{p_2} v_i(p, y_i) = \partial_{p_2} u_i(x_i(p, y_i))$?

Questions for you

Try to calculate this

1. using the chain rule,
2. substituting for $\partial_{p_2} x_1$ using the rewritten budget constraint,
3. rearranging, and
4. using the first order condition of utility maximization.

Solution

$$\begin{aligned}
\partial_{p_2} v(p,y) &= \partial_{x_1} u(x) \cdot \partial_{p_2} x_1 + \partial_{x_2} u(x) \cdot \partial_{p_2} x_2 \\
&= \partial_{x_1} u(x) \cdot \left(-\frac{x_2}{p_1} - \partial_{p_2} x_2 \frac{p_2}{p_1} \right) + \partial_{x_2} u(x) \cdot \partial_{p_2} x_2 \\
&= -\partial_{x_1} u(x) \cdot \frac{x_2}{p_1} + \left(-\frac{\partial_{x_1} u(x)}{p_1} + \frac{\partial_{x_2} u(x)}{p_2} \right) \cdot p_2 \cdot \partial_{p_2} x_2 \\
&= -x_2 \cdot \frac{\partial_{x_1} u(x)}{p_1}.
\end{aligned}$$

- Last step is key!
- Uses
  1. household utility maximization, and
  2. welfare = utility.
- $\Rightarrow$ behavior changes “drop out.”

The welfare effect of changing income

By a very similar calculation:

$$
\begin{aligned}
\partial_y v(p,y) &= \partial_{x_1} u(x) \cdot \partial_y x_1 + \partial_{x_2} u(x) \cdot \partial_y x_2 \\
&= \partial_{x_1} u(x) \cdot \left(-\frac{1}{p_1} - \partial_y x_2 \frac{p_2}{p_1} \right) + \partial_{x_2} u(x) \cdot \partial_y x_2 \\
&= -\frac{\partial_{x_1} u(x)}{p_1} + \left(-\frac{\partial_{x_1} u(x)}{p_1} + \frac{\partial_{x_2} u(x)}{p_2} \right) \cdot p_2 \cdot \partial_{p_2} x_2 \\
&= -\frac{\partial_{x_1} u(x)}{p_1}.
\end{aligned}
$$

- Again: Behavior changes “drop out.”

How does a change in prices compare to a change in income?

- Equivalent variation:

$$EV = \frac{\partial_{p_2} v(p,y) \cdot dp_2}{\partial_y v(p,y)} = \frac{-x_2 \cdot \frac{\partial_{x_1} u(x)}{p_1} \cdot dp_2}{-\frac{\partial_{x_1} u(x)}{p_1}} = -x_2 \cdot dp_2.$$

Questions for you

Interpret this equation.

Takeaway

- Suppose the prices p_j of various goods change.
- The effect of this change on utility of a given individual i is the same as the effect of a change in her income of

$$dy_i = EV_i = -\sum_j x_{ij} dp_j.$$

- The right hand side is a price index, using the individual's "consumption basket" x_i to weight price changes.

Aggregation and disaggregated reporting

- Equivalent variation measures utility changes expressed in monetary units.
- Can aggregate to social welfare, if we have welfare weights:

$$dSWF = \sum_i \omega_i \cdot EV_i$$

- ω_i measures value of an additional \$ for person i
- Could also report welfare changes in a disaggregated way:
  1. Average for various demographic groups, or
  2. average conditional on income.

Average conditional on income – nonparametric regression

- We would like to estimate

$$E[EV|y].$$

- But (almost) nobody has exactly income y,
  since income is continuously distributed.
- This is a **nonparametric regression** problem.
- Various methods exist to estimate this.
- We discuss k-nearest neighbor regression.
- Deaton uses a very similar method, kernel regression.

k-nearest neighbor regression

- Suppose that for a given value y, you want to estimate $E[EV|y]$.
- Find the k observations $i_1, \ldots, i_k$
  with the smallest distance in income, $|y_i - y|$.
- Calculate the average.
  This gives the k-nearest neighbor estimator,

$$\widehat{E}[EV|y] = \frac{1}{k} \sum_{j=1}^{k} EV_{i_j}.$$

- Can do this for any value y.
- Can also do this if we want to estimate conditional averages given additional variables, $E[EV|y,z]$, say.

Variance-bias tradeoff

- We need to choose k.
- That choice involves a tradeoff between variance and bias.
- Larger k means
  1. smaller variance,
  2. but larger bias.
- Formally:

$$\text{Var}(\widehat{E}[EV|y]) = \frac{1}{k}\text{Var}(EV|y)$$

$$\text{Bias} = \frac{1}{k}\sum_{j=1}^{k}(E[EV|y_i] - E[EV|y])$$

- The difference in the last term is generally larger, the further y_i is from y.