Revealed preference and aggregation

Bram De Rock (Brussels)

Joint work with Laurens Cherchye (Leuven, Tilburg),
Ian Crawford (Oxford) and Frederic Vermeulen (Tilburg)

November 25, 2010

Quotes

Samuelson (1947, p.4) proposed *meaningful theorems* as the primary objective of economic research:

> *By a meaningful theorem I mean simply a hypothesis about empirical data which could conceivably be refuted.*

Quotes

Samuelson (1947, p.4) proposed *meaningful theorems* as the primary objective of economic research:

> *By a meaningful theorem I mean simply a hypothesis about empirical data which could conceivably be refuted.*

MasCollel, Whinston and Green (1995, p. 105) on aggregation:

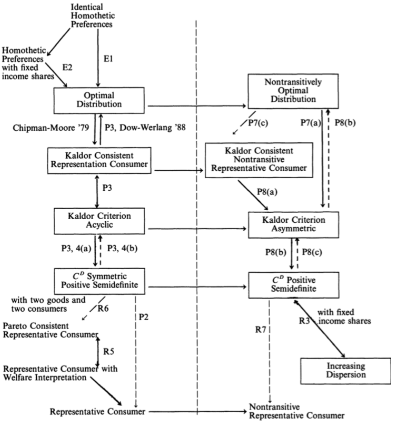
> *For most questions in economics, the aggregate behavior of consumers is more important than the behavior of any single consumer.*

Contribution of this paper

Meaningful theorems on aggregate demand and representative consumers

- Based on Afriat inequalities
  - Easy to apply
  - No functional specifications needed
- Proper investigation of the restrictions imposed by aggregation

Jerison (1994, figure 1)



Outline

1. Representative consumers
2. Gorman Polar Form
3. Empirical illustration
4. Conclusion

Some notation

- The micro data are a balanced panel: $\{\mathbf{p}_t, \mathbf{q}_t^h\}_{t\in\tau}^{h\in\eta}$
  - $\eta = \{1, \ldots, H\}$ is the index set for the households
  - $\tau = \{1, \ldots, T\}$ is the index set for time
  - $\mathbf{p}_t \in \mathbb{R}_{++}^N, \mathbf{q}_t^h \in \mathbb{R}_+^N$
  - Prices are assumed common across households
- The household data: $\{\mathbf{p}_t, \mathbf{q}_t^h\}_{t\in\tau}$
- The macro data: $\{\mathbf{p}_t, \sum_{h=1}^H \mathbf{q}_t^h\}_{t\in\tau}$

The positive representative consumer

Only the macro data are important:

$$\max_{\sum_h \mathbf{q}^h \in \mathbb{R}^N_+} W(\sum_h \mathbf{q}^h) \text{ subject to } \mathbf{p}'(\sum_h \mathbf{q}^h) \leq Y_t = \mathbf{p}_t'(\sum_h \mathbf{q}_t^h)$$

The positive representative consumer

Only the macro data are important:

$$\max_{\sum_h \mathbf{q}^h \in \mathbb{R}^N_+} W(\sum_h \mathbf{q}^h) \text{ subject to } \mathbf{p}'(\sum_h \mathbf{q}^h) \leq Y_t = \mathbf{p}_t'(\sum_h \mathbf{q}_t^h)$$

- As if the aggregated data is obtained from a rational agent
- Households at the micro level can be irrational
- No welfare implications for the micro data

The positive representative consumer

Only the macro data are important:

$$\max_{\sum_h \mathbf{q}^h \in \mathbb{R}^N_+} W(\sum_h \mathbf{q}^h) \text{ subject to } \mathbf{p}'(\sum_h \mathbf{q}^h) \leq Y_t = \mathbf{p}'_t(\sum_h \mathbf{q}^h_t)$$

- As if the aggregated data is obtained from a rational agent
- Households at the micro level can be irrational
- No welfare implications for the micro data
- Gorman, circa 1976, reprinted in Blackorby et al (1995):

  Rather an odd chap ...he is as likely as not to be radiantly happy when those he represents are miserable and vice versa

The positive representative consumer: definition

Definition (Positive representative rationalisation)

A (well-behaved) utility function W provides a positive representative rationalization of the macro data $\{\mathbf{p}_t, \sum_{h=1}^{H} \mathbf{q}_t^h\}_{t\in\tau}$ if for each observation t we have

$$W(\sum_h \mathbf{q}_t^h) \geq W(\sum_h \mathbf{q}^h)$$

for all $\sum_h \mathbf{q}^h$ with $\mathbf{p}_t' \sum_h \mathbf{q}^h \leq \mathbf{p}_t' \sum_h \mathbf{q}_t^h$

The positive representative consumer: theorem

Theorem

The following two statements are equivalent for the macro data $\{\mathbf{p}_t, \sum_{h=1}^{H} \mathbf{q}_t^h\}_{t\in\tau}$:
(i) There exists a ***positive*** *representative rationalization*
(ii) There exists numbers $W_t, \lambda_t \in \mathbb{R}_{++}$ *such that for all* $t, s \in \tau$:

$$W_s \leq W_t + \lambda_t \mathbf{p}_t'(\sum_h \mathbf{q}_s^h - \sum_h \mathbf{q}_t^h)$$

- Standard Afriat theorem (see Afriat 1967)
- See Varian (1982, 1984) for more discussion

The normative representative consumer

Both the micro and macro data are important:

$$\max_{\mathbf{q}^1,\dots,\mathbf{q}^H \in \mathbb{R}^N_+} W(u^1(\mathbf{q}^1), \dots, u^H(\mathbf{q}^H)) \text{ subject to } \sum_{h=1}^{H} \mathbf{p}'\mathbf{q}^h \leq Y_t$$

The normative representative consumer

Both the micro and macro data are important:

$$\max_{\mathbf{q}^1,\ldots,\mathbf{q}^H \in \mathbb{R}^N_+} W(u^1(\mathbf{q}^1), \ldots, u^H(\mathbf{q}^H)) \text{ subject to } \sum_{h=1}^{H} \mathbf{p}'\mathbf{q}^h \leq Y_t$$

- All households act rationally
  - There exists well-behaved utility functions u^h
- The income distribution maximizes the macro-utility function
  - Aggregate income in observation t: $\mathbf{p}_t'(\sum_h \mathbf{q}_t^h) = Y_t$
  - Individual income for household h in observation t : $\mathbf{p}_t'\mathbf{q}_t^h$
- Direct link with the micro data makes welfare judgements possible

The normative representative consumer: definition

Definition (Normative representative rationalisation)

The (well-behaved) utility functions $W, u^1, ..., u^H$ provide a normative aggregate rationalization of the micro data $\{\mathbf{p}_t, \mathbf{q}_t^h\}_{t \in \tau}^{h \in \eta}$ if for each observation t we have

$$W(u^1(\mathbf{q}_t^1), ..., u^H(\mathbf{q}_t^H)) \geq W(u^1(\mathbf{q}^1), ..., u^H(\mathbf{q}^H))$$

for all $\{\mathbf{q}^h\}^{h \in \eta}$ with $\mathbf{p}_t' \mathbf{q}^h \leq \mathbf{p}_t' \mathbf{q}_t^h$

The normative representative consumer: theorem

Theorem

The following two statements are equivalent for the micro data $\{\mathbf{p}_t, \mathbf{q}_t^h\}_{t\in\tau}^{h\in\eta}$*:*
(i) There exists a ***normative*** *representative rationalisation*
(ii) There exists numbers $W_t, \lambda_t, u_t^h, b_t^h \in \mathbb{R}_{++}$ *such that for all* $t, s \in \tau, h \in \eta$:

$$\begin{aligned} W_s &\leq W_t + \lambda_t \mathbf{b}_t'(\mathbf{u}_s - \mathbf{u}_t) \\ u_s^h &\leq u_t^h + \frac{1}{b_t^h}\mathbf{p}_t'(\mathbf{q}_s^h - \mathbf{q}_t^h) \end{aligned}$$

with $\mathbf{u}_t = (u_t^1, \ldots, u_t^H)$ *and* $\mathbf{b}_t = (b_t^1, \ldots, b_t^H)$

The normative representative consumer

New result, however...

- Closely related to weak and latent separability
  - Revealed preference characterizations: Varian (1983) and Crawford (2004)

The normative representative consumer

New result, however...

- Closely related to weak and latent separability
  - Revealed preference characterizations: Varian (1983) and Crawford (2004)
- Nonlinear system of inequalities
  - Empirically less attractive
  - A lot of existing tests for weak separability that are either necessary or sufficient
  - See, e.g., Varian (1983), Swofford and Whitney (1987, 1994), Fleissig and Whitney (2003, 2008)

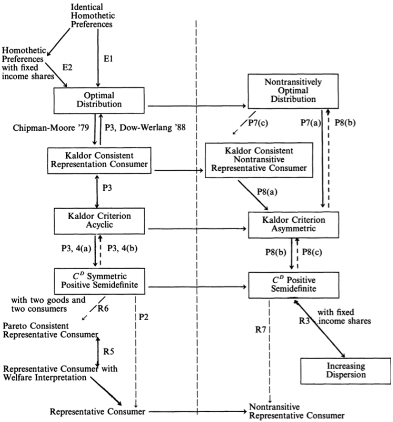
The normative representative consumer

- The above focuses simultaneously on
  - Existence of a representative consumer
  - Existence of an optimal income distribution rule
- Representative consumer does not imply some optimal income distribution rule
  - See, e.g., Samuelson (1956), Chipman and Moore (1979) and Jerison (1984, 1994)

The normative representative consumer

- The above focuses simultaneously on
  - Existence of a representative consumer
  - Existence of an optimal income distribution rule
- Representative consumer does not imply some optimal income distribution rule
  - See, e.g., Samuelson (1956), Chipman and Moore (1979) and Jerison (1984, 1994)
- Can be problematic if, e.g.,
  - Income distribution is assumed to be given
  - Or aggregate demand is assumed to be (locally) independent of income distribution
  - Think of IO models only caring for market demand, equilibrium models focussing on supply side, welfare results concerning consumer surplus,...

Jerison (1994, figure 1)



Commercial break

Next workshop on revealed preferences

- Royal Economic Society conference
- Afriat, Diewert and Varian
- Also Cemmap workshop?
- April 19, 2011
- Oxford? London?
- Hope to see you there!

Outline

1. Representative consumers
2. Gorman Polar Form
3. Empirical illustration
4. Conclusion

Independent of income distribution

- We focus on aggregate demand being independent of all distributions of the given aggregate income
  - Sufficient condition for all other scenarios (e.g. the existence of a normative representative consumer)
- Strong assumption that asks for empirical verification
  - I.e., impact of income changes are the same across households and initial income levels
  - Each household faces a linear expansion path for given prices
  - All expansion paths are parallel across households for given prices

Independent of income distribution

Gorman (1953, 1961) presents the characterization:

- Each household has preferences of **Gorman Polar Form** (or, equivalently, quasi-homothetic preferences)
- The marginal propensity to spend is the **same for all households**
- This holds 'locally'

$$\begin{aligned} c^h(\mathbf{p}, u^h) &= a^h(\mathbf{p}) + b(\mathbf{p})u^h \\ v^h(\mathbf{p}, y^h) &= \frac{y^h - a^h(\mathbf{p})}{b(\mathbf{p})} \\ \mathbf{q}^h(\mathbf{p}, u^h) &= \nabla a^h(\mathbf{p}) + \nabla b(\mathbf{p})u^h \end{aligned}$$

Blackorby, Boyce and Russell (1978, figures 2-5)

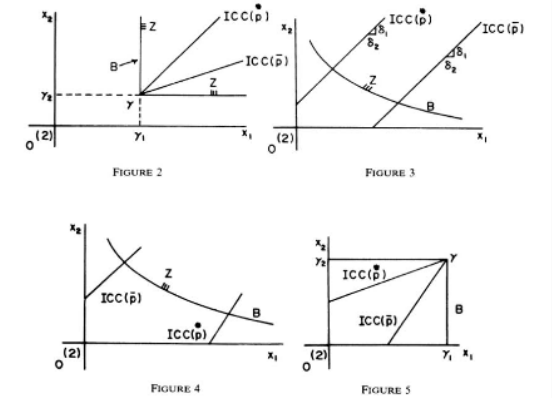
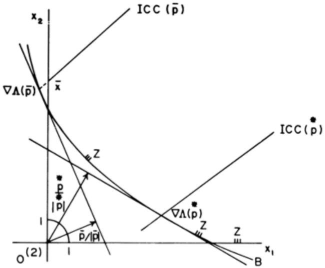


FIGURE 2 FIGURE 3 FIGURE 4 FIGURE 5

Blackorby, Boyce and Russell (1978, figure 1)

Key geometrical ingredients and intuition for our results

- A convex "base indifference curve" $u^h = 0$ where $\mathbf{q}^h(\mathbf{p}, 0) = \nabla a^h(\mathbf{p})$
- Linear expansion paths



Gorman Polar Form: definition

Definition (Gorman Polar Form (GPF) rationalisation)

The household data $\{\mathbf{p}_t, \mathbf{q}_t^h\}_{t\in\tau}$ are GPF rationalizable, if they are rationalisable by a well-behaved utility function u^h and there exists a cost function $c^h(\mathbf{p}, u^h) = a^h(\mathbf{p}) + b^h(\mathbf{p})u^h$, where $a^h(\mathbf{p})$ and $b^h(\mathbf{p})$ are concave, homogeneous of degree one price indices

Gorman Polar Form: theorem

Theorem

*The following two statements are equivalent for the **household** data* $\{\mathbf{p}_t, \mathbf{q}_t^h\}_{t\in\tau}$*:*
*(i) There exists a **GPF** rationalisation*
(ii) There exists $u_t^h \in \mathbb{R}_{++}$ *and* $\alpha_t^h, \beta_t^h \in \mathbb{R}^N$ *such that for all* $t, s \in \tau$ *:*

$$\begin{aligned} \mathbf{q}_t^h &= \alpha_t^h + \beta_t^h u_t^h \\ \mathbf{p}_t' \alpha_t^h &\leq \mathbf{p}_t' \alpha_s^h \\ 0 &< \mathbf{p}_t' \beta_t^h \leq \mathbf{p}_t' \beta_s^h \\ \alpha_t^h &= \alpha_s^h \text{ and } \beta_t^h = \beta_s^h \text{ if } \mathbf{p}_t = \delta \mathbf{p}_s \end{aligned}$$

Independent of income distribution

Theorem

The following two statements are equivalent for the ***micro*** *data* $\{\mathbf{p}_t, \mathbf{q}_t^h\}_{t\in\tau}^{h\in\eta}$*:*
(i) Aggregate demand is independent of the income distribution
(ii) There exists $u_t^h \in \mathbb{R}_{++}$ *and* $\alpha_t^h, \beta_t \in \mathbb{R}^N$ *such that for all* $t, s \in \tau, h \in \eta$:

$$\begin{aligned} \mathbf{q}_t^h &= \alpha_t^h + \beta_t u_t^h \\ \mathbf{p}_t'\alpha_t^h &\leq \mathbf{p}_t'\alpha_s^h \\ 0 &< \mathbf{p}_t'\beta_t \leq \mathbf{p}_t'\beta_s \\ \alpha_t^h &= \alpha_s^h \text{ and } \beta_t = \beta_s \text{ if } \mathbf{p}_t = \delta\mathbf{p}_s \end{aligned}$$

Independent of income distribution

Some remarks:

- Boundary conditions need to be added

Independent of income distribution

Some remarks:

- Boundary conditions need to be added
- This result generalizes several special cases that also generate the independence of the income distribution
- Two notable example are:
  - *Identical homothetic preferences*; see Varian (1983) for the RP characterization of homothetic preferences
  - *Quasilinear preferences* with respect to the same good; see Brown and Calsamiglia (2007) for the RP characterization

Independent of income distribution

Some remarks:

- Boundary conditions need to be added
- This result generalizes several special cases that also generate the independence of the income distribution
- Two notable example are:
  - *Identical homothetic preferences*; see Varian (1983) for the RP characterization of homothetic preferences
  - *Quasilinear preferences* with respect to the same good; see Brown and Calsamiglia (2007) for the RP characterization
  - These special cases are linear characterizations, while ours is nonlinear
- However, we have the following **linear** necessary condition

Linear test

Proposition

Consider the micro data $\{\mathbf{p}_t, \mathbf{q}_t^h\}_{t\in\tau}^{h\in\eta}$. *Then aggregate demand is independent of the income distribution only if there exists* $u_t^h \in \mathbb{R}_{++}$ *and* $\beta_t \in \mathbb{R}^N$ *such that for all* $t, s \in \tau, h \in \eta$:

$$\begin{aligned} u_s^h - u_t^h &\leq \frac{1}{\mathbf{p}_t'\beta_t}\mathbf{p}_t'(\mathbf{q}_s^h - \mathbf{q}_t^h) \\ 0 &< \mathbf{p}_t'\beta_t \\ \mathbf{p}_t'\beta_t &= \delta\mathbf{p}_s'\beta_s \text{ if } \mathbf{p}_t = \delta\mathbf{p}_s \end{aligned}$$

Linear test

Some remarks:

- Same notation as before but is indeed linear
- Very strong condition
  - Rational households with same marginal propensity to spend
  - Ignores base indifference curve
  - Locally sufficient
- Allows two step procedure

Outline

1. Representative consumers
2. Gorman Polar Form
3. Empirical illustration
4. Conclusion

The data

The data used here are the Spanish Continuous Family Expenditure Survey (ECPF)

- Quarterly budget survey of Spanish households which interviews about 3,200 households every quarter
- Subsample of couples with and without children, in which the husband is in full-time employment in a non-agricultural activity and the wife is out of the labour force
- We form a balanced panel of $T = 5$, $H = 342$ and $N = 14$

Results

Are the households rationalizable?

- I.e. do they **seperately** satisfy the necessary condition, but with a household specific β_t^h?
- 326 out of 342 pass
- We drop the 16 ‘irrational’ households

Results

Could it be that the aggregate demand of the households is independent of the income distribution?

- I.e. do they **simultaneously** satisfy the necessary condition, but now with a common β_t?
- No for the 326 households

Results

Could it be that the aggregate demand of the households is independent of the income distribution?

- I.e. do they **simultaneously** satisfy the necessary condition, but now with a common β_t?
- No for the 326 households
- Stratifying on observables (age bands, schooling, household size, children) does not help
  - Even subgroups of 2 households reject

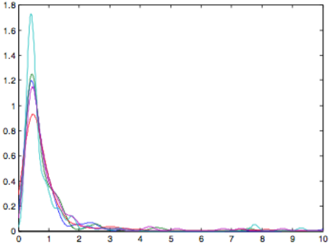
Interpretation

- The rejections can only be caused by β_t being common to all households
- Or, the marginal propensity to spend is not constant

Interpretation

- The rejections can only be caused by β_t being common to all households
- Or, the marginal propensity to spend is not constant
- Next step: do we need a lot of heterogeneity in the marginal propensity to spend to pass the necessary condition?
  - I.e. minimize the $\sum_{t,h}(\epsilon_t^h)^2$ in order to find $\frac{1}{\mathbf{p}_t'\beta_t^h} = \frac{1}{\mathbf{p}_t'\beta_t} + \epsilon_t^h$ that allow for a solution of the necessary condition
  - We consider three age classes but patterns are roughly the same

Heterogeneity in marginal propensity to spend



- Age ≤ 40 (134 households)
- Each color represents an observed time period
- $(\mathbf{p}_1'\beta_1, \ldots, \mathbf{p}_5'\beta_5) = (0.549,\ 0.556,\ 0.583,\ 0.539,\ 0.606)$
- Not much heterogeneity is needed to pass necessary condition

Conclusion

- I presented revealed preference characterizations for
  - Normative representative consumer
  - Gorman Polar Form
  - Aggregate demand being independent of income distribution
- The empirical illustration
  - Focuses on a (linear) necessary condition
  - Illustrates that already this condition is very stringent
  - Suggests that not much heterogeneity is needed for the data at hand
  - Needs to be elaborated