Empirical Revealed Preference

Ian Crawford

University of Oxford & Nuffield College

Overview

"There is nothing more practical then a good theory" - L. Boltzmann

- An elementary, "structural" approach to the analysis and interpretation of data by means of economic theory.

- It is somewhat distinct from structural econometrics because
  - it avoids having to resort to error terms
  - minimises the use of untestable assumptions
  - is expressed in terms of empirical inequalities.

Overview

- Typically we use economic theory to develop formal statements concerning causes and effects.
  - causes (explanatory variables) which may be observed ($\mathbf{x}$) or unobserved ($\boldsymbol{\eta}$)
  - effects (endogenous variables, $\mathbf{y}$)

- These are linked by structural equations which are theory-driven

$$\mathbf{y} = f(\mathbf{x}, \boldsymbol{\eta}, \theta)$$

  where θ represents a set of unknown parameters or functions.

Overview

Econometricians **always** then append a statistical structure to the economic model in order to account for the fact that the model does not perfectly explain the data.

> "The ... interpretation is that the true utility used by consumers to make choices is deterministic, but due to the researcher's inability to formulate individual behavior precisely, an additional stochastic term is added, thus making utility stochastic from the researcher's point of view (see Manski 1977; McFadden 1981, 1984). This is the interpretation followed in the economics literature" - Nevo (*Annual Reviews of Economics*, 2011, p. 59)

Overview

This relatively recent, authoritative survey echoes identical views expressed nearly 70 years earlier by Haavelmo:

> "Observable economic variables do not satisfy exact relationships (except, perhaps, some trivial identities). Therefore, if we start out with such a theoretical scheme, we have - for the purpose of application - to add some stochastical elements, to bridge the gap between the theory and the facts." - Haavelmo, (*Econometrica*, 1944.)

Overview

- This extra structure entails
  - the introduction of unobservable econometric error terms (ε)
  - statistical assumptions regarding the joint distribution of $(\mathbf{x}, \boldsymbol{\eta}, \varepsilon)$

- When combined these economic and statistical assumptions deliver an empirical model that is capable of rationalising **any** set of observables.

Overview

- The art of structural econometric modelling thus mainly lies in getting this statistical aspect right, because the source and the properties of these econometric errors ε can have a critical impact on the estimation results.

- This can be a challenge because:
  - economic theories, which are by and large completely deterministic, generally have little to say about the statistical model,
  - the data have generally little to say about the unobserved ($\boldsymbol{\eta}$) and unobservable (ε).

Overview

- Empirical revealed preference also begins from economic theory, is entirely different to the "$\mathbf{y} = f(\mathbf{x}, \boldsymbol{\eta}, \theta, \varepsilon)$" type of framework.

- Uses *systems of inequalities* which depend neither on the form of structural functions nor on unobservables.

- Statistical error terms and specific assumptions about the functional structure of the economic model may be added but it is not an essential requirement.

- In a sense empirical revealed preference is concerned with what we can learn simply by combining economic theory with the features of the world that we can observe.

Consumer Theory

$$\max_{\mathbf{q}} u(\mathbf{q}) \text{ subject to } \mathbf{p}_t'\mathbf{q} \leq x_t$$

- Suppose we observe some data on prices and choices $\{\mathbf{p}_t, \mathbf{q}_t\}_{t=1,...,T}$ for an individual consumer.
  - If the data were generated by the model, what properties must they *necessarily* have?
  - If the observed data have these properties, is that *sufficient* to know that they could have been generated by the model?
- What are the necessary and sufficient conditions for this model?

Necessity and Sufficiency

- Necessity and sufficiency are implicational relations between statements.

- Necessity: If $A \Rightarrow B$ then B is a *necessary condition* for A.
  - If A is true B then is necessarily true.
  - In particular $A \Rightarrow B$ is equivalent to $\neg B \Rightarrow \neg A$

- Sufficiency: If $A \Leftarrow B$ then B is a *sufficient condition* for A.
  - If B is true then that is sufficient to know that A is true too.
  - In particular $A \Leftarrow B$ is equivalent to $\neg A \Rightarrow \neg B$

Necessity and Sufficiency

- The assertion that one statement is necessary and sufficient for another means that the former statement is true *if and only if* the latter is true.

$$A \Leftrightarrow B$$

- That is, the two statements are
  - equivalent
  - simultaneously true or simultaneously false.

$$A \Rightarrow B,\ \neg B \Rightarrow \neg A,\ B \Rightarrow A,\ \neg A \Rightarrow \neg B$$

Necessity and Sufficiency

$$\max_{\mathbf{q}} u(\mathbf{q}) \text{ subject to } \mathbf{p}_t'\mathbf{q} \leq x_t$$

- Suppose we observe some data on prices and choices $\{\mathbf{p}_t, \mathbf{q}_t\}_{t=1,...,T}$ for an individual consumer.
  - If the data were generated by the model what properties must the data *necessarily* have?
  - If we observe these properties in some data, is that *sufficient* to know that the data could have been generated by the model?

Afriat's Theorem

We are interested in whether there is agreement between theory and data. We first need to define what that means.

Definition: A utility function $u(\mathbf{q})$ *rationalises* the data $\{\mathbf{p}_t, \mathbf{q}_t\}_{t=1,\ldots,T}$ if $u(\mathbf{q}_t) \geq u(\mathbf{q})$ for all $\mathbf{q}$ such that $\mathbf{p}_t'\mathbf{q}_t \geq \mathbf{p}_t'\mathbf{q}$.

Afriat's Theorem*

The following statements are equivalent:

A. *there exists a utility function* $u(\mathbf{q})$ *which is continuous, non-satiated and concave which rationalises the data* $\{\mathbf{p}_t, \mathbf{q}_t\}_{t=1,...,T}$.

B1. *there exist numbers* $\{U_t, \lambda_t > 0\}_{t=1,...,T}$ *such that*

$$U_s \leq U_t + \lambda_t \mathbf{p}_t' (\mathbf{q}_s - \mathbf{q}_t) \; \forall \; s, t \in \{1, ..., T\}$$

B2. *the data* $\{\mathbf{p}_t, \mathbf{q}_t\}_{t=1,...,T}$ *satisfy the Generalised Axiom of Revealed Preference (GARP).*

C. there *exists a non-satiated utility function* $u(\mathbf{q})$ *which rationalises the data* $\{\mathbf{p}_t, \mathbf{q}_t\}_{t=1,...,T}$.

*Afriat (1967), Diewert (1973), Varian (1982).

Afriat's Theorem

- We are going to take A to be true and work out why it implies the condition $B1$.

- By concavity of $u(\mathbf{q})$ we have

$$u(\mathbf{q}_s) \leq u(\mathbf{q}_t) + \nabla u(\mathbf{q}_t)'(\mathbf{q}_s - \mathbf{q}_t)$$

- Optimising behaviour implies the first-order condition $\nabla u(\mathbf{q}_t) \leq \lambda_t \mathbf{p}_t$ where $\lambda_t > 0$ and with equality when $q_t^k > 0$.

Afriat's Theorem

- Putting the foc into the concavity condition preserves the inequality and gives

$$u(\mathbf{q}_s) \leq u(\mathbf{q}_t) + \lambda_t \mathbf{p}_t' (\mathbf{q}_s - \mathbf{q}_t)$$

- Utility functions are real-valued so there must therefore exist real numbers $\{U_t, \lambda_t > 0\}_{t=1,...,T}$ corresponding to the values of $u(\mathbf{q}_t)$ and λ_t

$$U_s \leq U_t + \lambda_t \mathbf{p}_t' (\mathbf{q}_s - \mathbf{q}_t) \ \forall \ s, t \in \{1, ..., T\}$$

- This is condition $B1$.

Afriat's Theorem

- We are going to take $B1$ to be true (i.e. that we observe some data for which this condition holds) and work out why it implies the condition A.

- The path we are going to take is *constructive*: we are going to build a utility function out of the available raw materials and show that it does indeed rationalise the data.

- The raw materials are a set of $\{U_t, \lambda_t > 0\}_{t=1,\ldots,T}$ which satisfy the inequalities

$$U_s \leq U_t + \lambda_t \mathbf{p}_t' (\mathbf{q}_s - \mathbf{q}_t) \;\; \forall \; s, t \in \{1, ..., T\}$$

Afriat's Theorem

- Let

$$u(\mathbf{q}) = \min_{s\in\{1,..,T\}} \left\{ U_s + \lambda_s \mathbf{p}_s' (\mathbf{q} - \mathbf{q}_s) \right\}_{s=1,\ldots,T}$$

  be our utility function (it's piecewise linear, continuous, non-satiated, and concave).

- We now need to show that this rationalises the data.

- That means that $u(\mathbf{q}_t) \geq u(\mathbf{q})$ for all $\mathbf{q}$ such that $\mathbf{p}_t'\mathbf{q}_t \geq \mathbf{p}_t'\mathbf{q}$.

Afriat's Theorem

- Suppose we have some $\mathbf{q}$ with $\mathbf{p}_t'\mathbf{q}_t \geq \mathbf{p}_t'\mathbf{q}$. We need to show that $\exists$ some $u(\mathbf{q})$ such that $u(\mathbf{q}_t) \geq u(\mathbf{q})$

- Firstly consider the observation $\mathbf{q}_t$.

- What utility number does our function associate with it?

$$u(\mathbf{q}_t) = \min_{s \in \{1,..,T\}} \left\{ U_s + \lambda_s \mathbf{p}_s' (\mathbf{q}_t - \mathbf{q}_s) \right\}$$

Afriat's Theorem

- One element of the set concerns the case where $t = s$ in which case the corresponding element is:

$$U_t + \lambda_t \mathbf{p}_t' (\mathbf{q}_t - \mathbf{q}_t) = U_t$$

- We know that $U_t \le U_s + \lambda_s \mathbf{p}_s' (\mathbf{q}_t - \mathbf{q}_s)\ \forall\ s, t$

- Therefore $U_t \le \min_{s \in \{1,\ldots,T\}} \{U_s + \lambda_s \mathbf{p}_s' (\mathbf{q}_t - \mathbf{q}_s)\}$

- So our utility function assigns

$$u(\mathbf{q}_t) = U_t$$

Afriat's Theorem

- Now consider the affordable alternative $\mathbf{q}$. What utility number does our function associate with it?

$$u(\mathbf{q}) = \min_{s \in \{1,..,T\}} \left\{ U_s + \lambda_s \mathbf{p}_s' (\mathbf{q} - \mathbf{q}_s) \right\}$$

- We know that

$$u(\mathbf{q}) = \min_{s \in \{1,..,T\}} \left\{ U_s + \lambda_s \mathbf{p}_s' (\mathbf{q} - \mathbf{q}_s) \right\} \leq U_t + \lambda_t \mathbf{p}_t' (\mathbf{q} - \mathbf{q}_t)$$

Afriat's Theorem

- Finally $\lambda_t > 0$ and $\mathbf{p}_t'\mathbf{q}_t \geq \mathbf{p}_t'\mathbf{q}$ means $\lambda_t\mathbf{p}_t'(\mathbf{q} - \mathbf{q}_t) \leq 0$. So $u(\mathbf{q}) \leq U_t$

- Therefore we have shown that for any $\mathbf{q}$ with $\mathbf{p}_t'\mathbf{q}_t \geq \mathbf{p}_t'\mathbf{q}$

$$u(\mathbf{q}) \leq u(\mathbf{q}_t)$$

- Hence that our utility function rationalises the data in the required sense.

Afriat's Theorem

- Afriat's Theorem presents necessary and sufficient conditions for the standard utility-maximisation model.

- It shows that if a dataset can be rationalised by any utility function then it can in fact be rationalised by a well-behaved one - with competitive pricing non-convexities are "shrouded in enternal darkness"

- It is therefore *exhaustive*: it summarises ALL of the empirical implications which come from the basic model without making any special assumptions on functional forms (other than those needed for well-behavedness).

Afriat's Theorem

Definition: Given an observation $\mathbf{q}_t$ and a bundle $\mathbf{q}$:
(i) $\mathbf{q}_t$ is *directly revealed preferred* to $\mathbf{q}$, written $\mathbf{q}_t R_0 \mathbf{q}$ if $\mathbf{p}_t' \mathbf{q}_t \geq \mathbf{p}_t' \mathbf{q}$
(ii) $\mathbf{q}_t$ is *strictly directly revealed preferred* to $\mathbf{q}$, written $\mathbf{q}_t P_0 \mathbf{q}$ if $\mathbf{p}_t' \mathbf{q}_t \geq \mathbf{p}_t' \mathbf{q}$;
(iii) $\mathbf{q}_t$ is *revealed preferred* to $\mathbf{q}$, written $\mathbf{q}_t R \mathbf{q}$ if $\mathbf{p}_t' \mathbf{q}_t \geq \mathbf{p}_t' \mathbf{q}_u$, $\mathbf{p}_u' \mathbf{q}_u \geq \mathbf{p}_u' \mathbf{q}_v$, ..., $\mathbf{p}_v' \mathbf{q}_v \geq \mathbf{p}_v' \mathbf{q}$ for some sequence of observations $\mathbf{q}_t, \mathbf{q}_u, \mathbf{q}_v, ..., \mathbf{q},$. In this case we say that the relation R is the transitive closure of the relation R_0.
(iv) $\mathbf{q}_t$ is *strictly revealed preferred* to $\mathbf{q}$, written $\mathbf{q}_t P \mathbf{q}$, if there exist observations $\mathbf{q}_i$ and $\mathbf{q}_j$ such that $\mathbf{q}_t R \mathbf{q}_i$, $\mathbf{q}_i P_0 \mathbf{q}_j$, $\mathbf{q}_j R \mathbf{q}$.

Afriat's Theorem

GARP: $\mathbf{q}_t R \mathbf{q}_s$ implies NOT $\mathbf{q}_s P_0 \mathbf{q}_t$

If a consumption bundle $\mathbf{q}_t$ is revealed preferred to a consumption bundle $\mathbf{q}_s$, then $\mathbf{q}_s$ cannot be strictly directly revealed preferred to $\mathbf{q}_t$.

Afriat's Theorem

Suppose that we have four observations $\{\mathbf{p}_t, \mathbf{q}_t\}_{t=1,\ldots,4}$ and that

$\mathbf{p}_1'\mathbf{q}_1 \geq \mathbf{p}_1'\mathbf{q}_3$: $\mathbf{q}_1 R_0 \mathbf{q}_3$

$\mathbf{p}_2'\mathbf{q}_2 > \mathbf{p}_2'\mathbf{q}_1$: $\mathbf{q}_2 P_0 \mathbf{q}_1$ (also $\mathbf{q}_2 R_0 \mathbf{q}_1$)

$\mathbf{p}_2'\mathbf{q}_2 \geq \mathbf{p}_2'\mathbf{q}_4$: $\mathbf{q}_2 R_0 \mathbf{q}_4$

$\mathbf{p}_3'\mathbf{q}_3 \geq \mathbf{p}_3'\mathbf{q}_2$: $\mathbf{q}_3 R_0 \mathbf{q}_2$

Afriat's Theorem

We can write this into a matrix m where $m_{st} = 1$ if $\mathbf{q}_s R_0 \mathbf{q}_t$ and zero otherwise:

$$\begin{bmatrix} 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

In graph theory and computer science a square matrix made of zeros and ones is used to represent a simple finite directed graph. It's called an "adjacency matrix". The rows and columns label the graph vertices, with a 1 or 0 in (row, col) (s, t) according to whether s and t are adjacent.

Afriat's Theorem

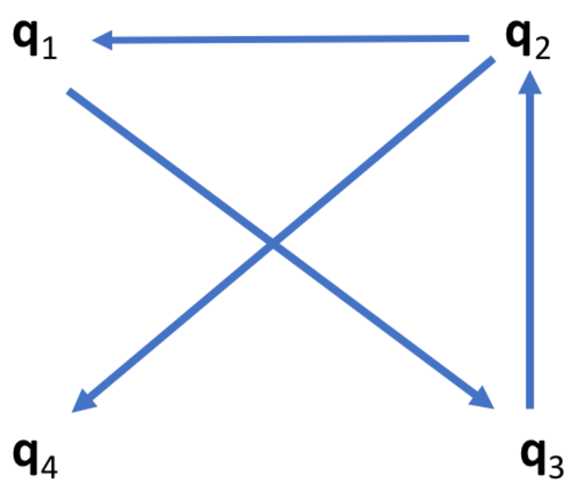
In a "directed graph" the adjacency is directional, there is an edge from s to t connecting them.

A directed graph has a cycle if it is possible to walk from any vertex and follow a consistently-directed sequence of edges that eventually loops back to that same vertex again.

GARP is a kind of "no-cyclic" condition on the directed graph generated by the data.

Afriat's Theorem

If we look at our data we can plot the directed graph for the R_0 relations:



There is clearly a cycle starting/ending at q_1:

$$\mathrm{q}_1 R_0 \mathrm{q}_3, \;\; \mathrm{q}_3 R_0 \mathrm{q}_2, \;\; \mathrm{q}_2 R_0 \mathrm{q}_1$$

The potential problem (for economics) with this is that the first two steps imply that this consumer prefers $\mathbf{q}_1$ to $\mathbf{q}_2$ (albeit indirectly via transitivity) so

$$\mathbf{q}_1 R \mathbf{q}_2$$

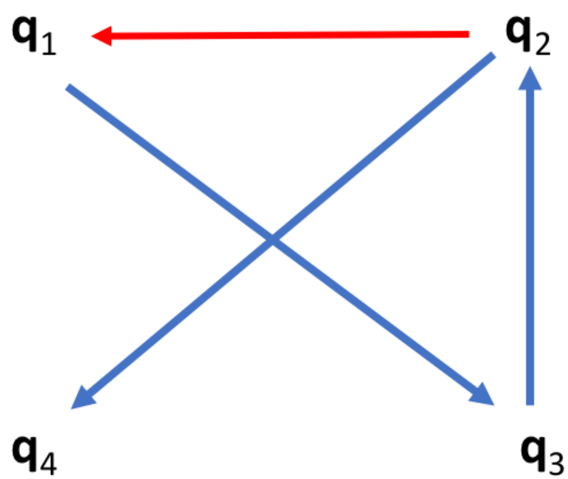
whilst we have the direct relation

$$\mathbf{q}_2 R_0 \mathbf{q}_1$$

This is only a "potential problem": it is OK if and only if none of the revealed preferences are strict - because they may be indifferent between $\mathbf{q}_1$ and $\mathbf{q}_2$. But if there is a strict preference for $\mathbf{q}_2$ over $\mathbf{q}_1$ then GARP is violated and the individual's preferences are not representable by a utility function.

Afriat's Theorem

If we look at our data we can indicate the edges for the P_0 relation:



Now we can see that $\mathbf{q}_1$ is revealed preferred to $\mathbf{q}_2$ and that $\mathbf{q}_2$ is directly strictly revealed preferred to $\mathbf{q}_1$. Another way of putting it is that $\mathbf{q}_1$ is revealed strictly preferred to itself.

Afriat's Theorem

- The Strong Axiom of Revealed Preference,

$$\mathbf{q}_t R \mathbf{q}_s \text{ and } \mathbf{q}_t \neq \mathbf{q}_s \text{ implies NOT } \mathbf{q}_s R^0 \mathbf{q}_t$$

- SARP implies GARP, but not vice versa.

- SARP requires single valued demand functions while GARP is compatible with multivalued demand functions (correspondences).

Afriat's Theorem

- The Weak Axiom of Revealed Preference

$$\mathbf{q}_t R^0 \mathbf{q}_s \text{ and } \mathbf{q}_t \neq \mathbf{q}_s \text{ implies NOT } \mathbf{q}_s R^0 \mathbf{q}_t$$

- Does not involve transitivity.

- Is necessary and sufficient for utility maximisation when there are only two goods - why?

Afriat's Theorem - Some history

- The theory of revealed preference was first introduced by Paul Samuelson in his 1938 *Economica* article.

- Samuelson's view was that economics was really about derivation of "meaningful theorems"

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

P. Samuelson, *Foundations of Economic Analysis,* p.4, (1947).

Afriat's Theorem - Some history

- His aim was to derive testable implications of theory without first postulating a utility function that represents the consumer's preferences.

- He argued that the testable implications of the theory should be based on axioms about observable demands rather than on axioms about unobservable preferences.

- The focus on the *observable* rather than the *unobservable* remains at the heart of the topic.

Afriat's Theorem - Some history

- Houthakker (1950) extended Samuelson's work by introducing the Strong Axiom of Revealed Preference (SARP).

- SARP works for any number of budget sets and works by exploiting transitivity.

- He also demonstrated that demand functions satisfy SARP if and only if they are the result of the maximisation of well-behaved preferences subject to the consumer's budget constraint.

- Clearly, this establishes a close link, also recognised by Samuelson, between the axioms about demand and the axioms about preferences.

Afriat's Theorem - Some history

- But both Samuelson and Houthakker assumed that the researcher could observe the *entire demand system*.

- If you *could* observe the entire demand system then the question of testable implications could as easily be addressed using the "standard" differential approach which goes back to Slutsky (1915) and Antonelli (1886).

- But we do not observe the entire demand system.

- We only ever observed a finite number of observations.

Afriat's Theorem - Some history

- The structural econometric approach makes up for this “data deficit” by fitting functions to the (finite) data.

- These functions are like having an infinite amount of data - once estimated we can evaluate them anywhere/everywhere and check the integrability conditions.

- Of course this is easier said than done.

Afriat's Theorem - Some history

- In particular, in order to estimate them consistently, requires us to make untestable auxiliary statistical assumptions.

- Any test of the hypothesis of maximising behaviour therefore is really a test of a joint hypothesis: the behaviour of interest plus the auxiliary statistical hypotheses required to deliver the estimate.

- This is the essence of the Duhem-Quine problem in the philosophy of science.

Afriat's Theorem - Some history

- Afriat in his 1967 *International Economic Review* article focussed on, and solved, the same problem but with only a finite number of observations.

- This might seems a small thing but it was the key to liberating applied work from the need to rely on assumed properties of unobserved and unobservable quantities.

- In that sense it represents the fruition of Samuelson's quest for a truly *Meaningful Theorem*.

Afriat's Theorem - Some history

- There was a very important further contribution by Erwin Diewert in the *Review of Economic Studies* in 1973.

- He analysed which assumptions on the utility function must be satisfied so that a solution to the utility maximisation problem exists in the first place. This was ignored in Afriat (1967).

- It turned out that the assumption of local non-satiation is crucial in this respect. Without local non-satiation, it may be the case that there is no solution to the utility maximisation problem.

Afriat's Theorem - Some history

- Moreover, without local non-satiation, any set of observed choices can be rationalised by a utility function in a trivial way - by resorting to "thick" indifference curves.

- Diewert (1973) also demonstrated that a linear programme can be constructed to solve the testability and recoverability questions.

- This was the first step on the way to translating Afriat's rather impenetrable work into something which could actually be applied.

Afriat's Theorem - Some history

- Varian's contributions begins with his 1982 *Econometrica* article.

- He solved or simplified many of the most important computational aspects of revealed preference

- He also extended Afriat's and Diewert's work by considering the recoverability and extrapolation questions.

Applying Afriat's Theorem

- Applied consumer theory typically addresses three sorts of issues:

1. *Consistency*. When is observed behaviour consistent with the model?

2. *Recoverability/The inverse problem*. How can we recover preferences given observations on consumer behavior?

3. *Extrapolation/The forward problem.* Given consumer behaviour for some budgets how can we forecast behaviour for other budgets?

Applying Afriat's Theorem

- Doing applied work using RP restrictions requires a completely different set of techniques than does the "$\mathbf{y} = f(\mathbf{x}, \boldsymbol{\eta}, \theta, \varepsilon)$" type of framework.

- The methods are mainly algorithmic or combinatorial - more typically learned when studying operations research than economics.

Applying Afriat's Theorem

- Given some data $\{\mathbf{p}_t, \mathbf{q}_t\}_{t=1,...,T}$ to check for consistency with the theory we can either

1. Determine whether there exist numbers $\{U_t, \lambda_t > 0\}_{t=1,...,T}$ such that

$$U_s \leq U_t + \lambda_t \mathbf{p}_t' (\mathbf{q}_s - \mathbf{q}_t) \ \forall \ s, t \in \{1, ..., T\}$$

2. Determine whether the data satisfy GARP.

Applying Afriat's Theorem

- Suppose we have just two observations $\{\mathbf{p}_1, \mathbf{p}_2; \mathbf{q}_1, \mathbf{q}_2\}$.

- Then the Afriat Inequalities

$$U_s \leq U_t + \lambda_t \mathbf{p}_t' (\mathbf{q}_s - \mathbf{q}_t) \quad \text{and } \lambda_t > 0, \forall\ s, t \in \{1, 2\}$$

are (explicitly)

$$\begin{aligned}
U_1 - U_1 - \lambda_1 \mathbf{p}_1' (\mathbf{q}_1 - \mathbf{q}_1) &\leq 0 \\
U_1 - U_2 - \lambda_2 \mathbf{p}_2' (\mathbf{q}_1 - \mathbf{q}_2) &\leq 0 \\
U_2 - U_1 - \lambda_1 \mathbf{p}_1' (\mathbf{q}_2 - \mathbf{q}_1) &\leq 0 \\
U_2 - U_2 - \lambda_2 \mathbf{p}_2' (\mathbf{q}_2 - \mathbf{q}_2) &\leq 0 \\
-\lambda_1 &< 0 \\
-\lambda_2 &< 0
\end{aligned}$$

Applying Afriat's Theorem

- This can be written, more compactly, as

$$\begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & -1 & 0 & -\mathbf{p}_2'(\mathbf{q}_1-\mathbf{q}_2) \\ -1 & 1 & -\mathbf{p}_1'(\mathbf{q}_2-\mathbf{q}_1) & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} U_1 \\ U_2 \\ \lambda_1 \\ \lambda_2 \end{bmatrix} \leq \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ -\varepsilon \\ -\varepsilon \end{bmatrix}$$

where the ε represent arbitrarily small constants.

Applying Afriat's Theorem

- The check for consistency is therefore whether (or not) there exists a vector $\mathbf{x}$ such that

$$\mathbf{Ax} \leq \mathbf{b}$$

- In essence we are asking whether there exist a solution to a set of linear inequalities

- This is a linear programming problem. Linear programs are problems that can be expressed in the general form:

$$\max_{\mathbf{x}} \mathbf{c}'\mathbf{x} \text{ subject to } \mathbf{Ax} \leq \mathbf{b}$$

Applying Afriat's Theorem

- The "simplex algorithm", developed by George Dantzig in 1947, solves LP problems by

1. constructing a feasible solution iff a solution exists, and then

2. optimising it.

- We are only concerned with Phase 1.

- Dantzig's algorithm can determine whether or not there is a feasible solution in a finite number of steps (a trial and error approach would, conversely, never be guaranteed to stop).

Applying Afriat's Theorem

- In general checking for consistency requires a linear program with $2T$ variables and T^2 constraints.

- The fact that the number of constraints rises as the square of the number of observations can makes this condition computationally demanding in practice for very large datasets.

- Condition B2 (GARP) is sometimes more efficient.

- This requires us to compute the transitive closure of a finite relation but that is certainly a finite problem and Warshall (1962) gives a solution in T^3 steps. It is very easy to implement.

Applying Afriat's Theorem

GARP provides another way to test for utility maximisation.

1. Loop through the data and create the directed adjacency matrix m for the R_0 relation

2. Compute the transitive closure of the graph (Floyd-Warshall algorithm) to fill in the implied R relations.

3. Loop through the data and create adjacency matrix n for the P_0 relations

4. Add the two matrices and check there are no "2's".

Applying Afriat's Theorem

In the example we had when showing the graph interpretation of GARP we had the R_0 graph:

$$m = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

The transitive closure is

$$m = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Applying Afriat's Theorem

In our example the P_0 graph was:

$$n = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Adding data the P_0 graph gives:

$$m + n = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 0 \\ \mathbf{2} & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Applying Afriat's Theorem

The data used here is the Spanish Continuous Family Expenditure Survey (the *Encuesta Continua de Presupuestos Familiares* - ECPF).

The ECPF is a quarterly budget survey of Spanish households which interviews about 3,200 households every quarter.

These households are randomly rotated at a rate of 12.5% each quarter.

It is possible to follow a participating household for up to eight consecutive quarters.

It's nationally representative and its coverage of expenditure is wide.

Applying Afriat's Theorem

The data from the period 1985 to 1997 and are the selected sub-sample 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.

The dataset consists of 21,866 observations on 3,134 households.

It records household non-durable expenditures aggregated into 5 broad commodity groups.

The price data are calculated from published prices aggregated to correspond to the expenditure categories.

We check GARP individually for each member of the panel - no pooling.

Applying Afriat's Theorem

$$Pass\ rate = 0.957$$

Applying Afriat's Theorem

- Suppose we have some data $\{\mathbf{p}_t, \mathbf{q}_t\}_{t=1,\ldots,T}$ which satisfies GARP.

- Since this individual has thus far been observed to behave in a way which is *perfectly* consistent with utility maximisation, we can try to recover their implied preferences.

- It is essential to understand that there may be more than one preference relation which is consistent with the data (*Afriat's Theorem* give us one, there may be others).

- So the recoverability question focuses on identifying the *set* of preferences that are consistent with a given data set.

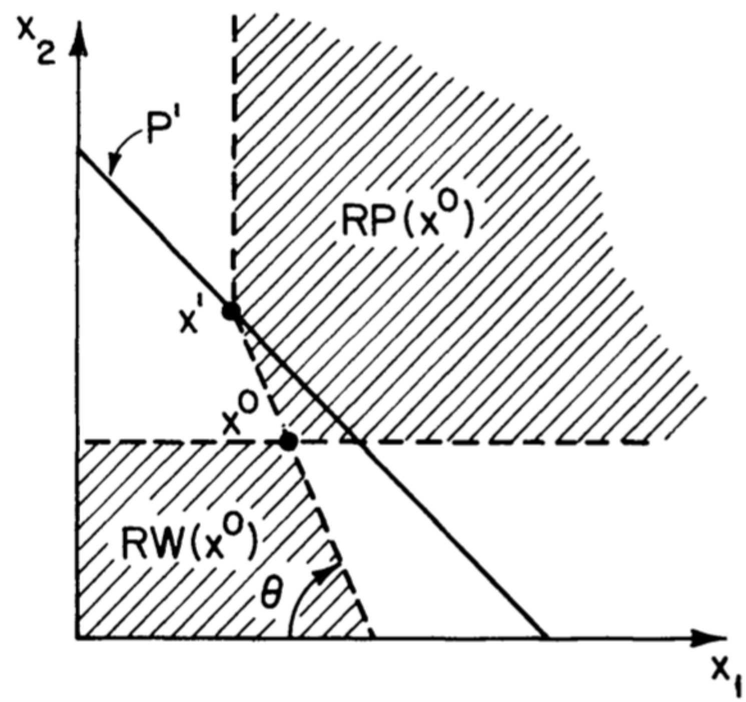
Applying Afriat's Theorem

- Recoverability is based entirely on the restrictions upon behaviour imposed by GARP.

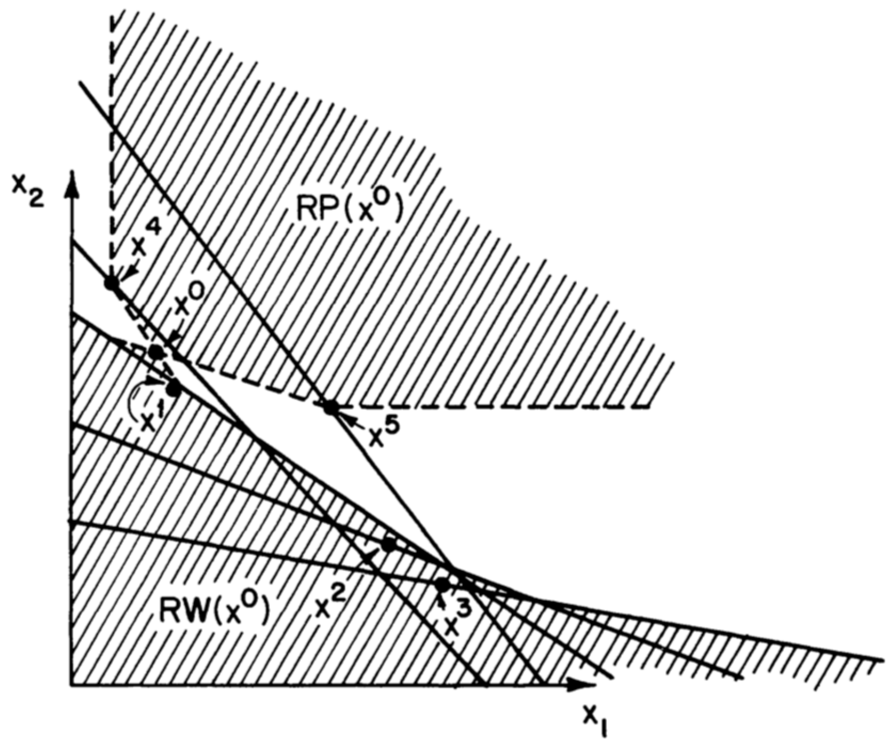
- The recoverability question aims at constructing inner and outer bounds for the indifference curves passing through an arbitrary, not necessarily observed, quantity bundle.

- The essential idea is to squeeze the indifference curve of interest between a set of bundles which are revealed preferred and a set which are revealed worse.

Applying Afriat's Theorem



Applying Afriat's Theorem



Applying Afriat's Theorem

- Suppose we have some data $\{\mathbf{p}_t, \mathbf{q}_t\}_{t=1,\ldots,T}$ which satisfies GARP and then present the consumer with a new budget $\{\mathbf{p}_0, x_0\}$. What will the consumer do?

- We will use the assumption that, since the individual has thus far been observed to behave in a way which is *perfectly* consistent with utility maximisation, she will continue to do so.

- Whatever she does, her new chosen bundle must therefore satisfy GARP in combination with her previously-observed choices..

- Again there will typically be more than one bundle which satisfies this restriction.

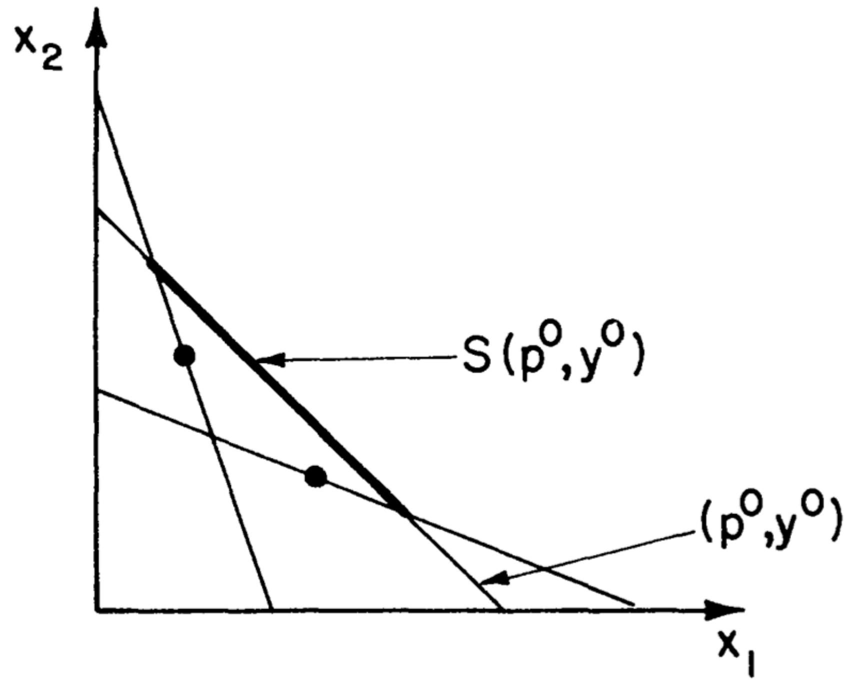
Applying Afriat's Theorem

- The description of the forecast is that the new demand ($\mathbf{q}_0$) must lie in the set defined by

$$S(\mathbf{p}_0, x_0) = \left\{ \mathbf{q}_0 : \begin{array}{c} \mathbf{q}_0 \geq 0, \quad \mathbf{p}_0'\mathbf{q}_0 = x_0 \\ \{\mathbf{p}_0, \mathbf{p}_t; \mathbf{q}_0, \mathbf{q}_t\}_{t=1,\ldots,T} \text{ satisfies GARP} \end{array} \right\}$$

- That is it must be
  - non-negative (although corners are fine)
  - satisfy the budget constraint
  - satisfy GARP when pooled with the observed choices.

Applying Afriat's Theorem



Applying Afriat's Theorem

- As the data becomes dense
  - the RP test for consistency becomes more demanding
  - the bounds on indifference curves become tighter
  - the bounds on demand responses become tighter.
- If the data become perfectly dense (effectively an infinite dataset) we have the indifference curve map and demand curves themselves.

Firms

- The most elementary way in which to describe a firm's technology is by means of its input requirement set $V(y)$.

- This consists of all input vectors $\mathbf{x}$ that can produce at least the output y

$$V(y) = \{\mathbf{x} : \mathbf{x} \text{ can produce at least } y\}$$

- A key property of input requirement sets is that they must be nested:

Definition: Nestedness: *If* $\mathbf{x}$ *is in* $V(y)$ *and* $y \geq y'$, *then* $\mathbf{x}$ *is in* $V(y')$.

Firms

- The following are also typically assumed:

Definition: Monotonicity: *If* $\mathbf{x}$ *is in* $V(y)$ *and* $\mathbf{x}' \geq \mathbf{x}$, *then* $\mathbf{x}'$ *is in* $V(y)$

Definition: Convexity: *If* $\mathbf{x} \in V(y)$ *and* $\mathbf{x}' \in V(y)$, *then*

$$\lambda\mathbf{x} + (1 - \lambda)\,\mathbf{x}' \in V(y)$$

where $\lambda \in [0, 1]$.

Definition: Regularity: $V(y)$ *is a closed non-empty set for all* y.

Firms

- Suppose that we observe a sequence of input vectors, input price vectors and output production for a firm: $\{\mathbf{w}_t, \mathbf{x}_t, y_t\}_{t=1,\dots,T}$

- The most elementary behavioural hypothesis about the firm we can entertain is that it is cost-minimising.

$$\min_{\mathbf{x}} \mathbf{w}_t'\mathbf{x} \text{ such that } \mathbf{x} \in V(y_t)$$

  - If the data were generated by the model what properties must the data *necessarily* have?

  - If the observed data with these properties in some data, is that *sufficient* to know that the data could have been generated by the model?

Firms

We are interested in the agreement between theory and data.

Definition: A family of input requirement sets $V(y)$ *rationalises* the data $\{\mathbf{w}_t, \mathbf{x}_t, y_t\}_{t=1,\ldots,T}$ if $\mathbf{w}_t'\mathbf{x}_t \leq \mathbf{w}_t'\mathbf{x}$ for all $\mathbf{x} \in V(y_t)$.

Firms

Theorem[†]. *The following statements are equivalent:*

A. *there exists a family of nested input requirement sets* $V(y)$ *which rationalises the data* $\{\mathbf{w}_t, \mathbf{x}_t, y_t\}_{t=1,\ldots,T}$.

B. *the data satisfies the* Weak Axiom of Cost-Minimisation: $y_t \leq y_s$ *implies* $\mathbf{w}_t'\mathbf{x}_t \leq \mathbf{w}_t'\mathbf{x}_s$

C. *there exists a family of nested, convex, monotonic input requirement sets* $V(y)$ *which rationalises the data* $\{\mathbf{w}_t, \mathbf{x}_t, y_t\}_{t=1,\ldots,T}$.

[†]Hanoch and Rothschild (1972), Diewert and Parkan (1979), Varian (1984).

Firms

- The empirical condition in statement (B) is known as the Weak Axiom of Cost Minimisation (WACM).

- Compared to Afriat's Theorem it is *very* straightforward to verify - simply inspect the dataset and check the condition directly.

- The equivalence between (A) and (C) means that if the data can be rationalised by *any* family of input requirement sets, then in fact it can be rationalised by a "nice" one and there is no harm is having these extra properties.

- Another way to say the same thing is that, in a finite data setting, these additional properties have no empirical content.

Firms

- We will take A to be true and show that it implies B (WACM).

- Let $V(y)$ be a family of nested input requirement sets that rationalise the data.

- If $y_t \leq y_s$ then nestedness means that $\mathbf{x}_s \in V(y_t)$.

- Since $V(y)$ rationalises the data $\mathbf{w}_t'\mathbf{x}_t \leq \mathbf{w}_s'\mathbf{x}_s$. Thus

$$y_t \leq y_s \textit{ implies } \mathbf{w}_t'\mathbf{x}_t \leq \mathbf{w}_t'\mathbf{x}_s$$

Firms

- Now we will show that if B holds then C holds.

- Again (like Afriat's Theorem) the proof is constructive - we will build a family of input requirement sets out of the raw materials: a dataset $\{\mathbf{w}_t, \mathbf{x}_t, y_t\}_{t=1,\ldots,T}$ which satisfies WACM.

- Let $V(y)$ be the positive convex monotonic hull of the $\{\mathbf{x}_t\}_{t=1,\ldots,T}$ such that $y_t \geq y$

$$V(y) = com^{+}\{\mathbf{x}_t : y_t \geq y\}$$

Firms

- This construction is closed, convex and monotonic.

- We need to show that this construction rationalises the data. That is that for any $\mathbf{x}_t$ it is the case that $\mathbf{w}_t'\mathbf{x}_t \leq \mathbf{w}_t'\mathbf{x}$ for all $\mathbf{x} \in V(y_t)$.

- The trick is to note that since the inequality of interest $\mathbf{w}_t'\mathbf{x}_t \leq \mathbf{w}_t'\mathbf{x}$ is linear we only need to worry about the vertices.

Firms

- The next point is to realise that the construction means that all of the vertices are observed input bundles. And we know that at the observed bundles WACM (condition B) holds.

- Therefore $\mathbf{w}_t'\mathbf{x}_t \leq \mathbf{w}_t'\mathbf{x}_s$ for all $\mathbf{x}_s \in V(y_t)$.

- So $\mathbf{x}_t$ is itself a vertex and lies on a supporting hyperplane with slope given by $\mathbf{w}_t$. Therefore for any $\mathbf{x}_t$ it is the case that $\mathbf{w}_t'\mathbf{x}_t \leq \mathbf{w}_t'\mathbf{x}$.

- Lastly we know that C implies A since it is stronger. Thus we have

$$A \Rightarrow B \Rightarrow C \Rightarrow A \Rightarrow ...$$

Firms

- Instead of considering a set-theoretic object we can use a production function.

Definition: A production function $f(\mathbf{x})$ *rationalises* the data $\{\mathbf{w}_t, \mathbf{x}_t, y_t\}_{t=1,...,T}$ if $f(\mathbf{x}_t) = y_t$ and $f(\mathbf{x}) \geq y_t$ implies $\mathbf{w}_t'\mathbf{x}_t \leq \mathbf{w}_t'\mathbf{x}$ for $t = 1, ..., T$.

- It is useful, in the context of production functions, to add the requirement of continuity.

Firms

Theorem. *The following statements are equivalent:*

A. *there exists continuous production function which rationalises the data* $\{\mathbf{w}_t, \mathbf{x}_t, y_t\}_{t=1,...,T}$.

B1. *the data satisfy the Strong Axiom of Cost Minimisation:* $y_s \leq y_t$ *implies* $\mathbf{w}_s'\mathbf{x}_s \leq \mathbf{w}_s'\mathbf{x}_t$

B2. there exist numbers $\{U_t, \lambda_t > 0\}_{t=1,...,T}$ *such that*

$$U_s \leq U_t + \lambda_t \mathbf{w}_t' (\mathbf{w}_s - \mathbf{w}_t) \;\; \forall \; s,t \in \{1,...,T\}$$

C. *there exists a continuous, monotonic and quasi-concave production function* $f(\mathbf{x})$ *which rationalises the data* $\{\mathbf{w}_t, \mathbf{x}_t, y_t\}_{t=1,...,T}$.

Firms

- Competitive profit maximisation is probably the canonial model in the theory of the firm.

- We let $\mathbf{y}_t$ denote an observed *netput* vector where $\mathbf{y}_t = [y_t^1, ...y_t^K,]$ is a signed vector of net outputs of K goods.

- So if the kth element is positive it's an output, and if it's negative it's an input.

- The prices (of inputs and outputs) are $\mathbf{p}_t = [y_t^1, ...y_t^K,]$.

Firms

- Given the sign convention on netputs the firm's profit is just

$$\mathbf{p}_t'\mathbf{y}_t$$

- You can think of the netput vector being arranged as

$$\mathbf{y}_t = \begin{bmatrix} y_t \\ -\mathbf{x}_t \end{bmatrix}, \qquad \mathbf{p}_t = \begin{bmatrix} p_t \\ \mathbf{w}_t \end{bmatrix}$$

so

$$\mathbf{p}_t'\mathbf{y}_t = p_t y_t - \mathbf{w}_t'\mathbf{x}_t$$

Firms

- The key theoretical object is the *production set*: the set of technically feasible input-output combinations: $\mathbf{Y}$.

- Generally we assume that Y is negative monotonic (free disposal). This says

$$\text{if } \mathbf{y} \in \mathbf{Y} \text{ and } \mathbf{y}' \leq \mathbf{y} \text{ then } \mathbf{y}' \in \mathbf{Y}$$

- Once again we are interested in when there exists a production set which rationalises a set of observations on a firm

Definition: A production set $\mathbf{Y}$ *rationalises* the data $\{\mathbf{p}_t, \mathbf{y}_t\}_{t=1,...,T}$ if $\mathbf{p}_t'\mathbf{y}_t \geq \mathbf{p}_t'\mathbf{y}$ for all $\mathbf{y} \in \mathbf{Y}$ for all $t = 1, ..., T$.

Firms

Theorem[‡]. *The following statements are equivalent:*

A. *there exists a production set which rationalises the data* $\{\mathbf{p}_t, \mathbf{y}_t\}_{t=1,...,T}$.

B. $\mathbf{p}_t'\mathbf{y}_t \geq \mathbf{p}_t'\mathbf{y}_s$ for all $s, t = 1, ..., T$

C. *there exists a close convex negative monotonic production set which rationalises the data* $\{\mathbf{p}_t, \mathbf{y}_t\}_{t=1,...,T}$.

[‡]Hanoch and Rothschild (1972), Diewert and Parkan (1979), Varian (1984).

Firms

- Statement (B) is known as the weak axiom of profit maximisation (WAPM).

- Firstly it is immediate that if (A) is true then (B) follows because the production set must contain the observations $\{\mathbf{y}_t\}_{t=1,\ldots,T}$

- To show that if WAPM (B) holds, then that implies the existence of a well-behaved production set (C) we construct one.

Firms

- Let $\mathbf{Y} = com^{-}\{\mathbf{y}_t\}$. This is the negative convex monotonic hull of $\mathbf{y}_t$.

- To show that this rationalises the data consider an arbitrary point $\mathbf{y}$ constructed as
$$\mathbf{y} = \sum_{s=1}^{T} \lambda_s(\mathbf{y}_s + \mathbf{e}_s)$$
where $\mathbf{e}_s \leq 0$ and $\sum_{s=1}^{T} \lambda_s = 1$.

- So $\mathbf{y}$ is any point which lies somewhere inside the constructed set. We need to show that it gives a lower profit than the chosen point.

Firms

- We know that condition (B) means that $\mathbf{p}_t'\mathbf{y}_t \geq \mathbf{p}_t'\mathbf{y}_s$ so

$$\lambda_s \mathbf{p}_t'\mathbf{y}_t \geq \lambda_s \mathbf{p}_t'(\mathbf{y}_s + \mathbf{e}_s)$$

  for all s, t

- Summing over s gives

$$\begin{aligned}
\sum_{s=1}^{T} \lambda_s \mathbf{p}_t'\mathbf{y}_t &\geq \left[\sum_{s=1}^{T} \lambda_s \mathbf{p}_t'(\mathbf{y}_s + \mathbf{e}_s)\right] \\
\mathbf{p}_t'\mathbf{y}_t &\geq \mathbf{p}_t' \left[\sum_{s=1}^{T} \lambda_s(\mathbf{y}_s + \mathbf{e}_s)\right] \\
\mathbf{p}_t'\mathbf{y}_t &\geq \mathbf{p}_t'\mathbf{y}
\end{aligned}$$

Firms

- When there is a single output there is a very straightforward version of WAPM.

$$\mathbf{p}_t'\mathbf{y}_t \geq \mathbf{p}_t'\mathbf{y}_s \Rightarrow p_t y_t - \mathbf{w}_t'\mathbf{x}_t \geq p_t y_s - \mathbf{w}_t'\mathbf{x}_s$$

- This can be rearrange to give

$$\begin{aligned} p_t y_s &\leq p_t y_t + \mathbf{w}_t'\mathbf{x}_s - \mathbf{w}_t'\mathbf{x}_t \\ y_s &\leq y_t + \frac{1}{p_t}\mathbf{w}_t'(\mathbf{x}_s - \mathbf{x}_t) \end{aligned}$$

which is an Afriat Inequality composed of observables.

Firms

Theorem. *The following statements are equivalent:*

A. *there exists a production function which rationalises the data* $\{\mathbf{w}_t, \mathbf{x}_t, y_t, p_t\}_{t=1,...,T}$.

B. *the data satisfy the Strong Axiom of Profit Minimisation:*

$$y_s \leq y_t + \frac{1}{p_t}\mathbf{w}_t' (\mathbf{x}_s - \mathbf{x}_t) \qquad \forall s, t \in \{1, ..., T\}$$

C. *there exists a continuous, monotonic and concave production which rationalises the data* $\{\mathbf{w}_t, \mathbf{x}_t, y_t, p_t\}_{t=1,...,T}$.

Firms

All of the sort of questions which we looked at in the context of individual choices can be studies in the firm context.

For example we can forecast conditional factor demands given new input prices and output levels using WACM, or predict netputs given a change in the prices of inputs and or final goods.

We can also test for imperfect competition and measure (in)efficiency.

Statistical Issues

- Revealed preference "test" are deterministic: either the subject passes in which case her behaviour can (heuristically) be regarded as being that of a utility-maximiser, or she doesn't in which case it cannot.

- This reflects the idea that the DGP is the deterministic model

$$\max_{\mathbf{q}} u(\mathbf{q}) \text{ subject to } \mathbf{p}_t'\mathbf{q} = y_t$$

  rather than a stochastic process.

- Important statistical considerations are present nonetheless.

Statistical Issues

1. we might only get to see sample of individuals from a larger population.

2. the data may very well be subject to measurement errors.

3. the individual may make optimisation errors which are stochastic in nature.

- There is also an inferential question which arises even if we set aside these other issues.

Afriat's Theorem - A Bayesian View

- *Under-determinism* is a concept from the philosophy of science about the relationship between theory and data.

"Most thinkers of any degree of sobriety allow, that an hypothesis ... is not to be received as probably true because it accounts for all the known phenomena, since this is a condition sometimes fulfilled tolerably well by two conflicting hypotheses...while there are probably a thousand more which are equally possible, but which, for want of anything analogous in our experience, our minds are unfitted to conceive."

J.S. Mill in *A System of Logic* ([1867] 1900, p328)

Afriat's Theorem - A Bayesian View

- Suppose we observe a repeated observations on a single individual: $\{\mathbf{p}_t, \mathbf{q}_t\}_{t=1,\ldots,T}$.

- Suppose they pass GARP. What are we to make of that?

- How justified might we be in thinking that this individual is, heuristically at any rate, really a utility-maximiser?

Afriat's Theorem - A Bayesian View

- Clearly our assessment of this will depend on

1. the number of observations

2. the ability of the GARP test to detect non-rational behaviour.

- If we have few observations, or constraints which do not cross often, then the evidence is probably weak and we should be unwilling to conclude simply that since this person has passed GARP, she must be a utility-maximiser.

Afriat's Theorem - A Bayesian View

- We are interested in whether or not the individual is a utility-maximiser (denoted U), given the data satisfy GARP (denoted G).

- Bayes' Theorem gives

$$
\begin{aligned}
P(U|G) &= \frac{P(G|U)P(U)}{P(G)} = \\
&= \frac{P(G|U)P(U)}{P(G|U)P(U) + P(G|\neg U)[1 - P(U)]}
\end{aligned}
$$

Afriat's Theorem - A Bayesian View

- Assuming no optimisation or measurement error then $P(G|U) = 1$ because a utility-maximiser will certainly pass GARP.

- So this gives us

$$P(U|G) = \frac{P(U)}{P(U) + P(G|\neg U)[1 - P(U)]}$$

where $P(U)$ is the prior.

- Bayes' Theorem tells us how to weigh the evidence.

Afriat's Theorem - A Bayesian View

$$P(U|G) = \frac{P(U)}{P(U) + P(G|\neg U)[1 - P(U)]}$$

- If the GARP test is not very well able to detect non-rational behaviour very well, then $P(G|\neg U) \approx 1$ and $P(U|G) \to P(U)$.

- This means that the evidence of the successful GARP test should not impress us much and should do little to shift our prior beliefs.

Afriat's Theorem - A Bayesian View

$$P(U|G) = \frac{P(U)}{P(U) + P(G|\neg U)[1 - P(U)]}$$

- If the GARP test is in fact very sensitive (for example, we have many observations) and $P(G|\neg U) \approx 0$, then $P(U|G) \to 1$

- Consequently the GARP test gives us rational grounds to become very confident that the individual is, in fact, a utility-maximiser.

Afriat's Theorem - A Bayesian View

- The term $P(G|\neg U)$ is therefore centrally important to the way in which we interpret a successful empirical GARP test.

- Its value depends on the alternative DGP (i.e. what ever $\neg U$ is).

- The difficulty is that there are many alternatives to rational choice models.

Afriat's Theorem - A Bayesian View

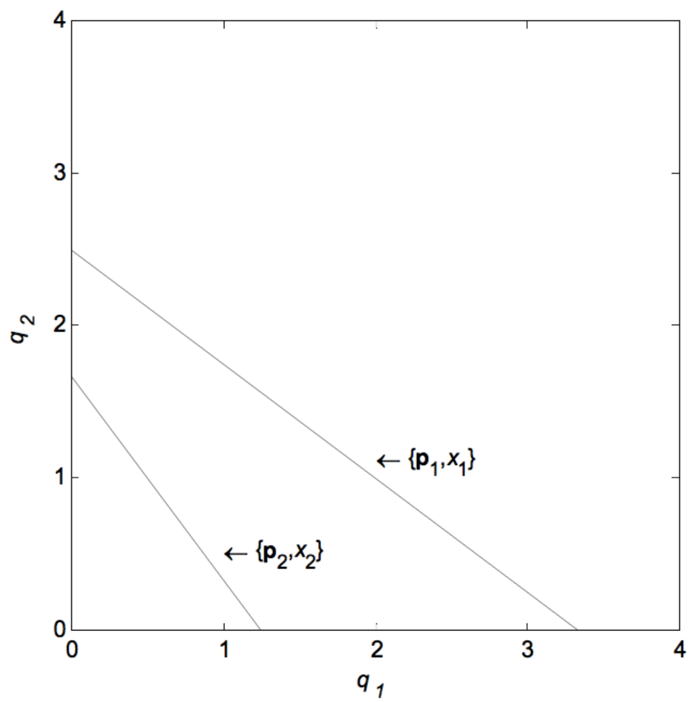
- One important, non-rational alternative considered by Becker (1962) was a probabilistic DGP: uniform random choice.

- Bronars (1987) applied this in an RP context by calculating the probability of observing a violation of GARP with this DGP operating on the observed constraints.

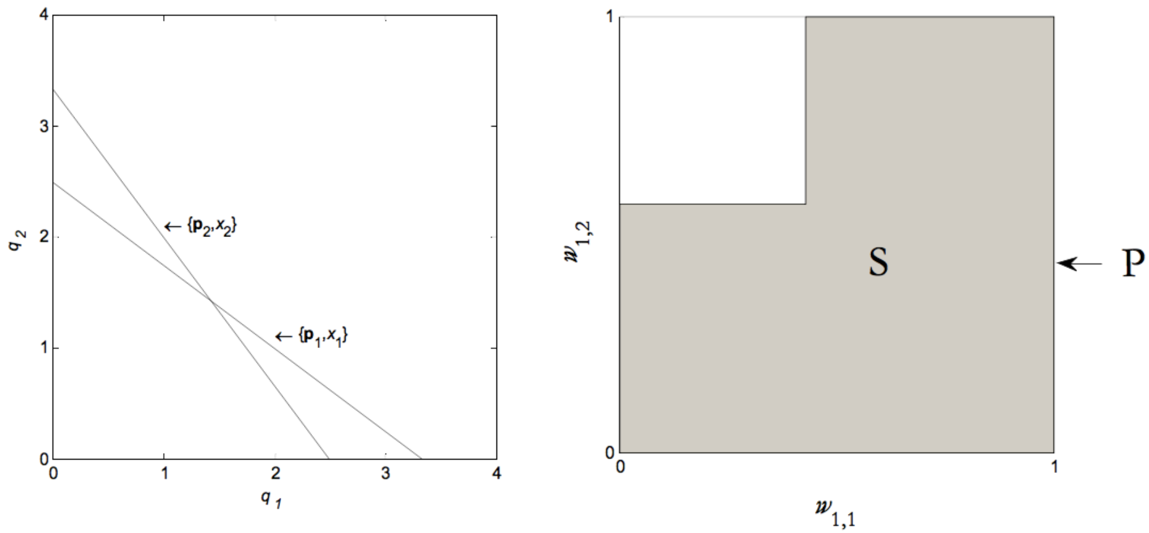
- Bronars' approach remains the most popular method but more recent contributions (notably Andreoni, Gillen and Harbaugh (2013))), whilst sticking with the idea of a probabilistic alternative DGP, consider more data-driven alternatives to uniform random choice - they suggest drawing from the empirical distribution of observed choices to allow for a more realistic alternative.

Predictive Success

"... lack of variation in the price data limits the power of these methods"
Hal Varian (*Econometrica*, 1982, pp 966-7)



Predictive Success



$P =$ the set of possible choices which satisfy the budget constraints.
$S =$ the set of choices which also satisfy GARP.

Predictive Success

When we check RP conditions for an individual we look to see whether their choices fall within the areas allowed by the restrictions.

The size of the target area provided by the restrictions is a sensible measure of how demanding the restrictions are.

$r \in \{0, 1\}$: the pass/fail indicator.

$a \in [0, 1]$: the relative *area* of the predicted subset compared to the outcome space.

Predictive Success

The simple hit/miss rate should not be the sole measure of the performance of the theory (if it was, then nothing could do better than "anything goes").

- good theories combine good hit rates (high pass rates) with demanding predictions (small areas);

- poor theories make imprecise predictions (large areas) which the data fail to satisfy (low pass rates).

Suggestion: take account of both r and a.

This idea is due to Reinhard Selten (*J. of Math Soc Sci*, 1991) who developed it in the context of experimental game theory.

Predictive Success

Some suggested properties of a measure of predictive success $m(r, a)$:

Monotonicity: $m(1, 0) > m(0, 1)$.

Equivalence of trivial theories: $m(0, 0) = m(1, 1)$.

Aggregability:. For every $\lambda \in [0, 1]$

$$m(\lambda r_1 + (1 - \lambda) r_2, \lambda a_1 + (1 - \lambda) a_2) = \lambda m(r_1, a_1) + (1 - \lambda) m(r_2, a_2).$$

Predictive Success

Selten's Theorem. *The function,* $m = r - a$ *satisfies the axioms. If* $\widetilde{m}(r,a)$ *also satisfies these axioms, then there exist real numbers* $\{\gamma, \delta > 0\}$ *such that* $\widetilde{m}(r,a) = \gamma + \delta m$.

Remarks on the Theorem

Selten (*J. Math. Soc. Sci.*, 1991) provides an ordinal characterisation of $r-a$ where he replaces aggregability with a continuity axiom and an axiom which says that the difference between theories should be a function of the differences between r's and a's. He uses stronger monotonicity axioms.

Predictive Success

$m \to 1$: demanding restrictions and data which satisfy them.

$m \to -1$: undemanding restrictions and yet the data fail to conform.

$m \to 0$: the apparent accuracy of the data simply mirrors the size of the target.

Predictive Success

An additional interpretation of $m \approx 0$ is that the theory performs about as well as a uniform random number generator

This interpretation provides a link between the measure discussed here and the investigation of *statistical power* conducted by Bronars (*Econometrica*,1987).

The alternative hypothesis is uniform random behaviour (as *per* Becker *J.Pol.E.*, 1962).

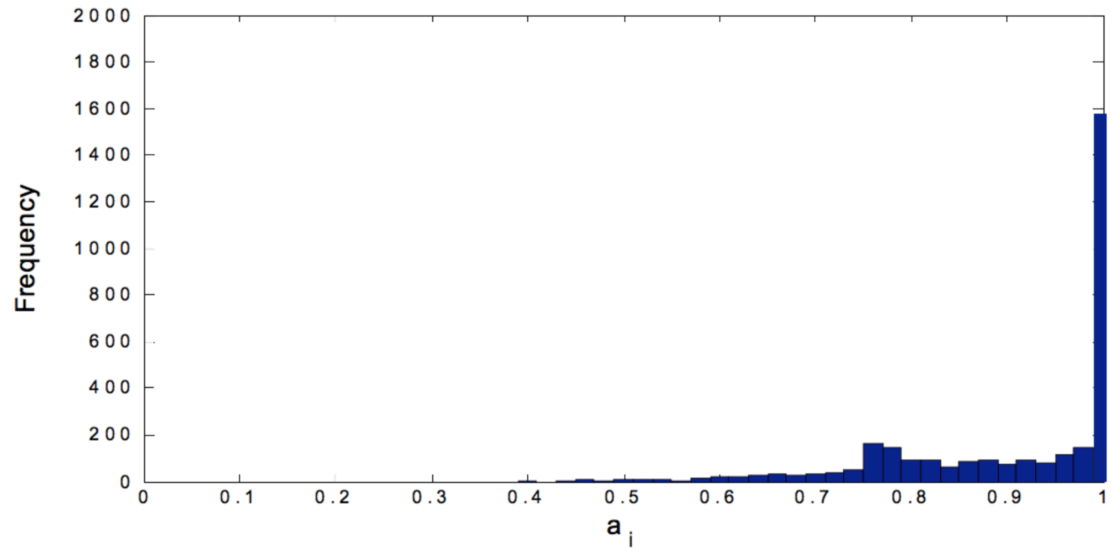
$(1 - a)$ can be interpreted as $P\left(\neg G \mid \text{uniform random behaviour}\right)$

Predictive Success

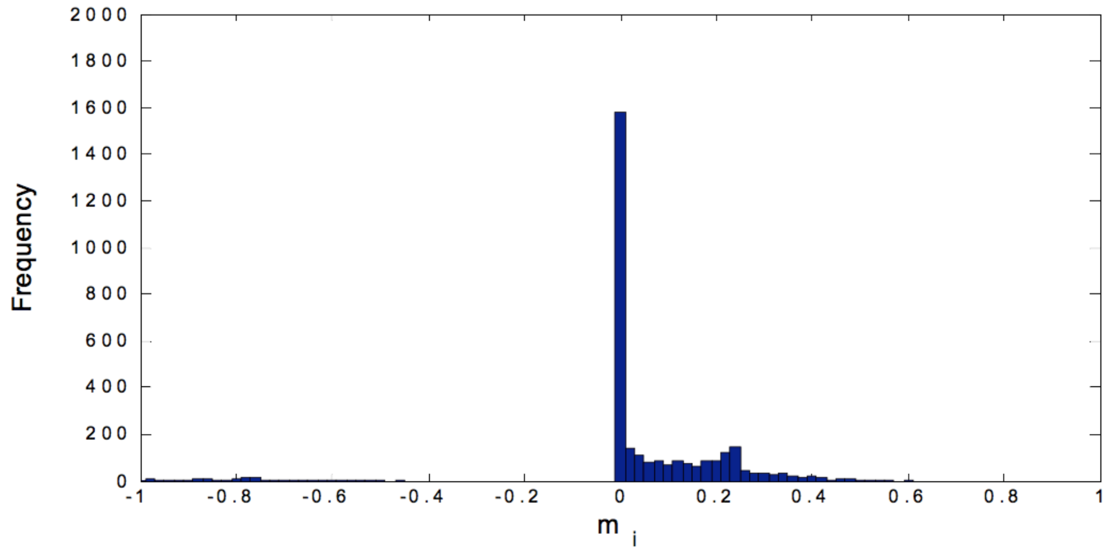
We return to the Spanish Continuous Family Expenditure Survey data (the *Encuesta Continua de Presupuestos Familiares* - ECPF) we saw last week.

$$
\begin{aligned}
r &= 0.957 \\
a &= 0.912 \\
m &= 0.045
\end{aligned}
$$

The distribution of a_i



The distribution of m_i



Predictive Success

- I'm not, of course, claiming that these particular results apply more widely than the dataset/conditions studied here (more restrictive models, e.g. intertemporal models or HARP, seem to provide a great deal of discipline on the data).

- But I am claiming that presenting results using these measures sheds a great deal more light on the empirical performance of a theory than does the uncorrected aggregate pass rate which is often reported in the empirical literature.

Inference

- If the data involved are a random panel sample of households and demands are measured without error, then inference about objects like the proportion of households in the population which satisfy RP restrictions is straightforward.

- A sample proportion can be viewed as the fraction of “successes” in N independent Bernoulli trials with the same success probability p.

- The central limit theorem implies that for large N, the sample proportion $\widehat{p}$ is normally distributed with mean p and standard deviation $\sqrt{p\,(1-p)\,N}$ so the statistic

$$z = (\widehat{p} - p) / \sqrt{p\,(1-p)\,N} \sim N\,(0, 1)$$

Inference

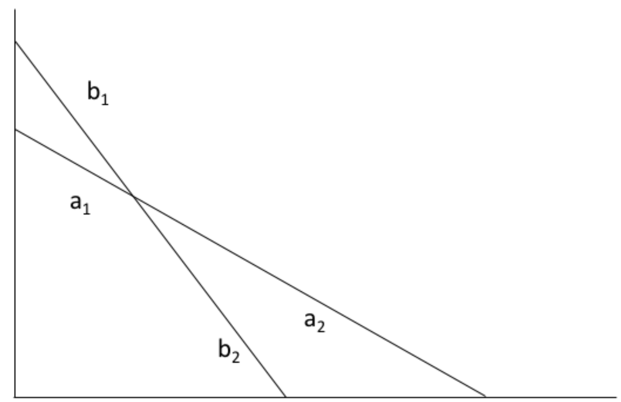
- Inference with repeated cross-sections from a heterogeneous population is more difficult.

- The issue here is that we do not see the same consumer twice, so we cannot proceed on an consumer-by-consumer basis, checking the RP conditions for each one as before.

- The object of interest remains the population proportion of consumers who satisfy the RP conditions.

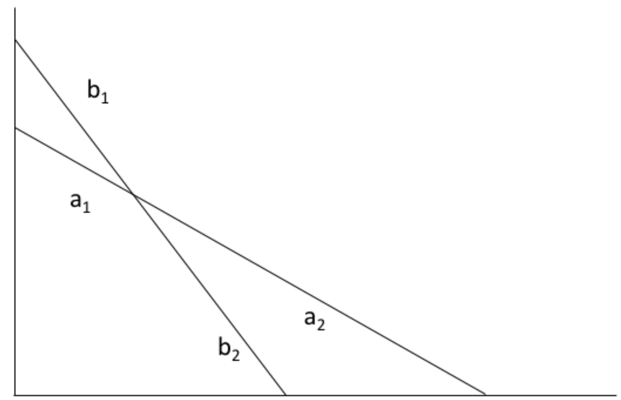
- However, this parameter depends on the joint distribution of choices over different budget sets and repeated cross-sectional data do not reveal this: only its marginal distributions can be observed.

Inference



Suppose we have a fixed population observed twice. In the first observation they are distributed along budget constraint a and in the second on b. We observe the two distributions but not the joint.

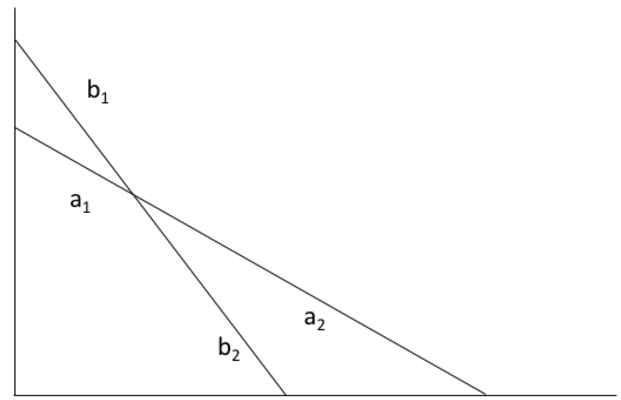
Inference



Let the proportion of the population on each segment be given by $a_1+a_2 = 1$ and $b_1 + b_2 = 1$. The population parameter of interest is the proportion of people who behave rationally (pass GARP).

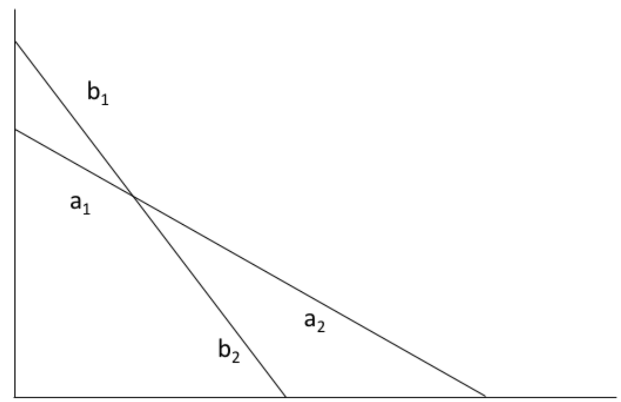
Because we cannot track individuals (we don't observe the joint distribution) we have to think about best- and worst-case scenarios.

Inference



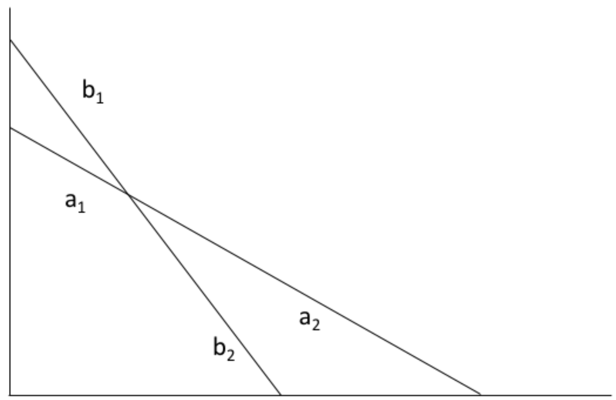
To make things easier suppose that the population consists of 100 people. And that there are 50 people in a_1, 50 in a_2, 40 in b_1 and 60 in b_2 How many people fail GARP?

Inference



Worst case: There were 50 people in a_1. Suppose all of them moved to b_2. There are 60 in total in b_2 so the other ten would have had to have come from a_2. What about the 50 on a_2? They can go anywhere they like on b and none of them will violate GARP. The 50 from a_1 are all violating GARP so the proportion who are irrational is at most 0.5.

Inference



Best case: There are 60 people on b_2. At most only 50 could have come from a_2 so at least 10 must have come from a_1. That means that 10 individual violated rationality and so the proportion who are irrational is at least 0.1.

Inference

- Under these circumstances, the population parameter of interest is not point identified. But we can bound the proportion of the population which behave irrationally within the interval $[0.1, 0.5]$ and hence the rational proportion in $[0.5, 0.9]$.

- Note that the actual distributions n the budget constraints don't matter much – just the proportions in each "patch".

- Hoderlein and Stoye (2013) show how to do inference on the sample analogue of this in the context of the Weak Axiom of Revealed Preference (i.e. without transitivity).

Measurement Errors

- An important difference between structural econometrics and empirical revealed preference lies in the absence of an error term in the latter.

- Certainly error terms rarely appear in revealed preference theory: there is no mention of an error term in Afriat's Theorem.

- But as soon as we attempt to take those revealed preference conditions to data, errors can no longer necessarily be ignored.

Measurement Errors

- The most obvious situation arises when we consider measurement errors, but identical issues arise when revealed preferences are applied to statistical objects (like estimates of aggregate consumption as in Browning (*International Economic Review,* 1989) or nonparametric Engel curves as in Blundell, Browning and Crawford (*Econometrica*, 2003 and 2008)).

- In these cases the price-quantity data we observe is a function of a random variable.

- This introduces a statistical element to empirical revealed preference and forms an important link between revealed preference with structural econometrics.

Measurement Errors

- To illustrate the case for classical additive measurement error consider the model

$$\mathbf{q}_t = \mathbf{q}_t^* + \mathbf{e}_t$$

where $\mathbf{q}_t^*$ denote the true values of demands and $\mathbf{e}_t$ is a vector of classical measurement errors.

- Now the DGP is the deterministic economic model plus a stochastic model.

- Suppose that we are interested in the null hypothesis that the true data $\{\mathbf{p}_t, \mathbf{q}_t^*\}_{t\in T}$ satisfy GARP.

Measurement Errors

- If the RP conditions fail for the observed demands $\mathbf{q}_t$, it is possible to generate a restricted estimator, $\widehat{\mathbf{q}}_t$ using the following Gaussian quasi-likelihood ratio or minimum distance criterion function:

$$L = \min_{\{\widehat{\mathbf{q}}_t\}_{t \in T}} \sum_{t=1}^{T} (\mathbf{q}_t - \widehat{\mathbf{q}}_t)' \, \Omega_t^{-1} \, (\mathbf{q}_t - \widehat{\mathbf{q}}_t)$$

  subject to the restriction that $\{\mathbf{p_t}, \widehat{\mathbf{q}}_t\}_{t \in T}$ satisfies GARP and where the weight matrix Ω_t^{-1} is the inverse of the covariance matrix of the demands.

- The solution to this problem leads to demands $\widehat{\mathbf{q}}_t$, which satisfy the RP restrictions and which are unique almost everywhere.

Measurement Errors

- Evaluated at the restricted demands, the above distance function also provides a test statistic for the RP conditions.

- It can be used for conducting conservative inference.

- This test falls within the general class of misspecification tests investigated in Andrews and Guggenberger (2007, Section 7).

Optimisation Errors

- Instead of asking whether the outcome of an empirical RP test represents a *statistically* significant departure from a DGP in to which a stochastic element has been introduced, we can also ask whether the results of the test represent an *economically* significant departure from rational choice.

- The key to this is to see that when a consumer violates RP conditions, that consumer appears to waste money by buying a consumption bundle when a cheaper bundle is available and also revealed preferred to it.

Optimisation Errors

- The cost-efficiency measure suggested in Afriat (1967) is the smallest amount of this wastage (as a fraction of the overall budget) consistent with the given demand data.

- This index provides a simple way of measuring the size of a violation of GARP and does so in units which are easy to understand and to interpret economically.

Optimisation Errors

- The idea is to modify the revealed preference relation R, essentially relaxing it to allow for some inefficency by the consumer.

- Normally we define "directly revealed preferred to" using $\mathbf{p}_t'\mathbf{q}_t \geq \mathbf{p}_t'\mathbf{q}_s \Leftrightarrow \mathbf{q}_t R^0 \mathbf{q}_s$ and the transitive closure of R^0 by R in the usual way.

- Instead we say that $\mathbf{q}_t$ is directly revealed preferred to $\mathbf{q}_s$ *at efficiency level e* using $e\mathbf{p}_t'\mathbf{q}_t \geq \mathbf{p}_t'\mathbf{q}_s \Leftrightarrow \mathbf{q}_t R_e^0 \mathbf{q}_s$ and define the transitive closure of this relation as R_e in the usual way.

- Then we have GARP$_e$ is $\mathbf{q}_t R_e \mathbf{q}_s$ implies not $\mathbf{q}_t P_e^0 \mathbf{q}_s$.

Optimisation Errors

- The number e can be thought of as how much less the potential expenditure on a bundle has to be before we will consider it worse than the observed choice.

- If e is 0.95, for example, we will only count bundles whose cost is less than 95% of an observed choice as being revealed worse than that choice.

- Said another way: if e is 0.95 and $\mathbf{q}_s$ would cost only 2% less than $\mathbf{q}_t$ at $\mathbf{p}_t$ prices we would not consider this a significant enough difference to conclude that $\mathbf{q}_t$ was preferred by the consumer to.$\mathbf{q}_s$

Optimisation Errors

- We are allowing the consumer a 'margin of error' of $(1 - e)$.

- Afriat's Critical Cost Efficiency Index, or the Afriat Efficiency Index for short, is the largest value of $e \in [0, 1]$ such that there are no violations of GARP_e.

Optimisation Errors

- If $e = 1$ then there are no violations of GARP in the original data, but for $e < 1$ there are violations.

- Traditionally, researchers begin their analysis of consumer behavior by setting some critical level of e , say e^*; such that they would consider any $e > e^*$ a small or tolerable violation of GARP.

- Varian (1991), for instance, suggests a value of $e = 0.95$

Adding Structure

- In the basic model of rational demand that we discussed above, we are considering any type of (well-behaved) utility function.

- But many models in economics depend critically on more particular functional assumptions.

- For example: *additive separability* is essentially the defining characteristic of expected utility theory.

- If you want to investigate with particular structures or particular models using RP we need more than just Afriat's Theorem.

Adding Structure: Separability

"Separability is about the structure we are to impose on our model: what to investigate in detail, what can be sketched in with broad strokes without violence to the facts."
W.M. Gorman (1987

- Separability is the most important restriction used in applicable theory.

- It refers to certain restrictions on functional representations of preferences or technologies which add structure to the decision making tasks undertaken by economic agents.

- These restrictions also allow the economic researcher to study the behavior of these agents in a more effective manner.)

Adding Structure: Separability

- Partition our data into two sets of goods and prices

$$\left\{\left\{\mathbf{p}_t^1, \mathbf{q}_t^1\right\}, \left\{\mathbf{p}_t^2, \mathbf{q}_t^2\right\}\right\}_{t=1,\ldots,T}$$

- A utility function is separable in the group 1 goods, if

$$\{\mathbf{q}^1, \mathbf{q}^2\} \succeq \left\{\mathbf{q}_*^1, \mathbf{q}^2\right\} \Leftrightarrow \{\mathbf{q}^1, \mathbf{q}_\#^2\} \succeq \left\{\mathbf{q}_*^1, \mathbf{q}_\#^2\right\}$$

for all $\mathbf{q}^1, \mathbf{q}_*^1, \mathbf{q}^2$ and $\mathbf{q}_\#^2$.

- That is preferences within group 1 are independent of the composition of group 2.

Adding Structure: Separability

- The functional representation is that a utility function u is (weakly) separable in the the group 1 goods if we can find a "subutility function" $v\left(\mathbf{q}^1\right)$ and a "macro function" $w(v, \mathbf{q}^2)$ with $w(v, \mathbf{q}^2)$ strictly increasing in v such that:

$$u\left(\mathbf{q}^1, \mathbf{q}^2\right) = w(v(\mathbf{q}^1), \mathbf{q}^2)$$

- Separability confers two major simplifying benefits:

  1. the ability to ignore certain things,

  2. dimension reduction

Adding Structure: Separability

- What revealed preference conditions would reflect this structure?

- Recall we need both necessary and sufficient conditions.

- Clearly the entire data set must satisfy GARP since it comes from maximisation of $u(\mathbf{q}^1, \mathbf{q}^2)$.

Adding Structure: Separability

- Weak separability is also necessary and sufficient for the second (lower) stage of two-stage budgetting.

- The sub-dataset must satisfy GARP since each $\mathbf{q}^1$ must solve the problem:

$$\max_{\mathbf{q}^1} v\left(\mathbf{q}^1\right) \text{ subject to } \mathbf{p}_t^{1\prime}\mathbf{q}^1 = \mathbf{p}_t^{1\prime}\mathbf{q}_t^1$$

- To see why suppose that $\mathbf{q}_*^1$ satisfied the budget constraint and yielded higher subutility. Then $w(v\left(\mathbf{q}_*^1\right), \mathbf{q}_t^2) > w(v\left(\mathbf{q}_t^1\right), \mathbf{q}_t^2)$ and $\mathbf{p}_t^1\mathbf{q}_*^1 + \mathbf{p}_t^2\mathbf{q}_t^2 \leq \mathbf{p}_t^1\mathbf{q}_t^1 + \mathbf{p}_t^2\mathbf{q}_t^2$ contradicting maximisation.

Adding Structure: Separability

- So the pooled data $\left\{\mathbf{p}_t^1, \mathbf{p}_t^2, \mathbf{q}_t^1, \mathbf{q}_t^2\right\}_{t=1,\ldots,T}$ and the sub-group data $\left\{\mathbf{p}_t^1, \mathbf{q}_t^1\right\}_{t=1,\ldots,T}$ must satisfy GARP but we also need to allow for the aggregating/dimension-reducing aspect of separability.

- Concavity conditions for the macro and the sub-utility functions are

$$
\begin{aligned}
u\left(\mathbf{q}_s^1, \mathbf{q}_s^2\right) &\leq u\left(\mathbf{q}_t^1, \mathbf{q}_t^2\right) + \nabla_{\mathbf{q}_t^1} u\left(\mathbf{q}_t^1, \mathbf{q}_t^2\right)'\left(\mathbf{q}_s^1 - \mathbf{q}_t^1\right) \\
&\quad + \nabla_{\mathbf{q}_t^2} u\left(\mathbf{q}_t^1, \mathbf{q}_t^2\right)'\left(\mathbf{q}_s^2 - \mathbf{q}_t^2\right) \\
v\left(\mathbf{q}_s^1\right) &\leq v\left(\mathbf{q}_t^1\right) + \nabla v\left(\mathbf{q}_t^1\right)'\left(\mathbf{q}_s^1 - \mathbf{q}_t^1\right)
\end{aligned}
$$

Adding Structure: Separability

- Denote $v_t = v\left(\mathbf{q}_t^1\right)$.

- Then the concavity condition for $w\left(v, \mathbf{q}^2\right)$ is is

$$w\left(v_s, \mathbf{q}_s^2\right) \leq w\left(v_t, \mathbf{q}_t^2\right) + \frac{\partial w}{\partial v_t}\left(v_s - v_t\right) + \nabla w\left(v_t, \mathbf{q}_t^2\right)'\left(\mathbf{q}_s^2 - \mathbf{q}_t^2\right)$$

Adding Structure: Separability

- Optimising behaviour (first order conditions) gives us

$$
\begin{aligned}
\nabla_{\mathbf{q}_t^1} u\left(\mathbf{q}_t^1, \mathbf{q}_t^2\right) &\leq \lambda_t \mathbf{p}_t^1 \\
\nabla v\left(\mathbf{q}_t^1\right)' &\leq \mu_t \mathbf{p}_t^1 \\
\nabla_{\mathbf{q}_t^2} u\left(\mathbf{q}_t^1, \mathbf{q}_t^2\right) &\leq \lambda_t \mathbf{p}_t^2 \\
\nabla_{\mathbf{q}_t^2} w\left(v_t, \mathbf{q}_t^2\right) &\leq \lambda_t \mathbf{p}_t^2
\end{aligned}
$$

where λ_t is the marginal utility of income and μ_t is the marginal utility of income allocated to the $\mathbf{q}^1$ group, that is the Legrange multiplier on the problem

$$\max_{\mathbf{q}^1} v\left(\mathbf{q}^1\right) \text{ subject to } \mathbf{p}_t^{1\prime}\mathbf{q}^1 = \mathbf{p}_t^{1\prime}\mathbf{q}_t^1$$

Adding Structure: Separability

- Define

$$\rho_t = \frac{\partial w}{\partial v_t}$$

- Then the concavity conditions become

$$\begin{aligned}
u\left(\mathbf{q}_s^1, \mathbf{q}_s^2\right) &\leq u\left(\mathbf{q}_t^1, \mathbf{q}_t^2\right) + \lambda_t \mathbf{p}_t^{1\prime}\left(\mathbf{q}_s^1 - \mathbf{q}_t^1\right) + \lambda_t \mathbf{p}_t^{2\prime}\left(\mathbf{q}_s^2 - \mathbf{q}_t^2\right) \\
v\left(\mathbf{q}_s^1\right) &\leq v\left(\mathbf{q}_t^1\right) + \mu_t \mathbf{p}_t^{1\prime}\left(\mathbf{q}_s^1 - \mathbf{q}_t^1\right) \\
w\left(v_s, \mathbf{q}_s^2\right) &\leq w\left(v_t, \mathbf{q}_t^2\right) + \rho_t\left(v_s - v_t\right) + \lambda_t \mathbf{p}_t^{2\prime}\left(\mathbf{q}_s^2 - \mathbf{q}_t^2\right)
\end{aligned}$$

Adding Structure: Separability

- From the chain rule we have

$$\nabla_{\mathbf{q}_t^1} u\left(\mathbf{q}_t^1, \mathbf{q}_t^2\right) = \frac{\partial w}{\partial v_t} \nabla v\left(\mathbf{q}_t^1\right) = \rho_t \mu_t \mathbf{p}_t^1$$

- We also had

$$\nabla_{\mathbf{q}_t^1} u\left(\mathbf{q}_t^1, \mathbf{q}_t^2\right) \leq \lambda_t \mathbf{p}_t^1$$

- So

$$\frac{\lambda_t}{\mu_t} = \rho_t$$

Adding Structure: Separability

- The concavity/optimality conditions are therefore

$$
\begin{aligned}
u\left(\mathbf{q}_s^1, \mathbf{q}_s^2\right) &\leq u\left(\mathbf{q}_t^1, \mathbf{q}_t^2\right) + \lambda_t \mathbf{p}_t^{1\prime}\left(\mathbf{q}_s^1 - \mathbf{q}_t^1\right) + \lambda_t \mathbf{p}_t^{2\prime}\left(\mathbf{q}_s^2 - \mathbf{q}_t^2\right) \\
v\left(\mathbf{q}_s^1\right) &\leq v\left(\mathbf{q}_t^1\right) + \mu_t \mathbf{p}_t^{1\prime}\left(\mathbf{q}_s^1 - \mathbf{q}_t^1\right) \\
w\left(v_s, \mathbf{q}_s^2\right) &\leq w\left(v_t, \mathbf{q}_t^2\right) + \frac{\lambda_t}{\mu_t}\left(v_s - v_t\right) + \lambda_t \mathbf{p}_t^{2\prime}\left(\mathbf{q}_s^2 - \mathbf{q}_t^2\right)
\end{aligned}
$$

- The final step just replaces the values of these real-valued functions with real numbers.

Adding Structure: Separability

Theorem (Varian (1982), Afriat (1967)). The following conditions are equivalent:
(1) there exists a weakly separable, concave, monotonic, continuous non-satiated utility function that rationalises the data;
(2) there exist numbers $\{V_t, W_t, \lambda_t > 0, \mu_t > 0\}_{t=1,\ldots,T}$ that satisfy:

$$\begin{aligned} V_s &\leq V_t + \mu_t \mathbf{p}_t^{1\prime} \left(\mathbf{q}_s^1 - \mathbf{q}_t^1\right) \\ W_s &\leq W_t + \frac{\lambda_t}{\mu_t}(V_s - V_t) + \lambda_t \mathbf{p}_t^{2\prime} \left(\mathbf{q}_s^2 - \mathbf{q}_t^2\right) \end{aligned}$$

(3) the data $\left\{\mathbf{p}_t^1, \mathbf{q}_t^1\right\}_{t=1,\ldots,T}$ and $\left\{1/\mu_t, \mathbf{p}_t^2, V_t, \mathbf{q}_t^2\right\}_{t=1,\ldots,T}$ satisfy GARP for some choice of $\{1/\mu_t, V_t\}_{t=1,\ldots,T}$ that satisfies the Afriat inequalities.

Adding Structure: Separability

- There is no explicit mention of a condition corresponding to an Afriat condition for the entire dataset or the statement "the entire dataset satisfies GARP".

- This is because that condition is implied by the other two [hint: add the other inequalities up].

- There is also a new computational problem. This can be seen in two ways which are equivalent.

Adding Structure: Separability

- The first is that it is necessary to find a set of Afriat numbers $\{1/\mu_t, V_t\}_{t=1,\ldots,T}$ which represent a group quantity and price index such that

1. they satisfy the Afriat inequality $V_s \leq V_t + \mu_t \mathbf{p}_t^{1\prime} \left(\mathbf{q}_s^1 - \mathbf{q}_t^1\right)$

2. $\left\{1/\mu_t, \mathbf{p}_t^2, V_t, \mathbf{q}_t^2\right\}_{t=1,\ldots,T}$ satisfies GARP

- The difficulty is that the numbers which satisfy (1), if they exist, are not unique.

Adding Structure: Separability

- The second way to see the problem is to look at the Afriat inequality

$$W_s \leq W_t + \frac{\lambda_t}{\mu_t}(V_s - V_t) + \lambda_t \mathbf{p}_t^{2\prime}\left(\mathbf{q}_s^2 - \mathbf{q}_t^2\right)$$

- The issue is whether there exists a solution, or not.

- This is known as a "certification" problem.

Adding Structure: Separability

- For *linear* inequalities certification is possible thanks (in theory) to a result call *Farkas Lemma* and thanks (in practice) to the simplex algorithm

- But these inequalities are *non-linear* in unknowns.

- There exists no certification method which is guaranteed to converge in a finite number of steps.

Adding Structure: Separability

- It is therefore possible in principle to determine whether or not data are consistent with utility maximisation and a weakly separable preference relation.

- But it is not always possible to show that data are not consistent with that structure.

- Of course failure of just one of the necessary conditions would be enough to rule it out.

More structure: additive separability

- A stronger assumption than weak separability, and one which is almost as frequently made is *additive* separability.

- We say a utility function $u(\mathbf{q}_t^1, \mathbf{q}_t^2)$ is additively separable if we can write it as

$$u(\mathbf{q}_t^1, \mathbf{q}_t^2) = v(\mathbf{q}_t^1) + w(\mathbf{q}_t^2)$$

for some utility functions $v(\mathbf{q}_t^1)$ and $w(\mathbf{q}_t^2)$.

More structure: additive separability

- Since additive separability implies weak separability we know immediately that one condition is that there must exist some "Afriat numbers" $\{V_t, W_t, \lambda_t > 0, \mu_t > 0\}_{t=1,\ldots,T}$ that satisfy:

$$\begin{aligned} V_s &\leq V_t + \lambda_t \mathbf{p}_t^{1\prime}\left(\mathbf{q}_s^1 - \mathbf{q}_t^1\right) \\ W_s &\leq W_t + \mu_t \mathbf{p}_t^{2\prime}\left(\mathbf{q}_s^2 - \mathbf{q}_t^2\right) \end{aligned}$$

More structure: additive separability

- The first order conditions for overall utility maximisation imply that

$$\begin{aligned}\nabla_{\mathbf{q}_t^1} u\left(\mathbf{q}_t^1, \mathbf{q}_t^2\right) &= \nabla w\left(\mathbf{q}_t^1\right)' \leq \lambda_t \mathbf{p}_t^1 \\ \nabla_{\mathbf{q}_t^2} u\left(\mathbf{q}_t^1, \mathbf{q}_t^2\right) &= \nabla v\left(\mathbf{q}_t^2\right)' \leq \lambda_t \mathbf{p}_t^2\end{aligned}$$

- This means that the marginal utility of income is equalised across the groups. Hence we can take $\lambda_t = \mu_t$

More structure: additive separability

Theorem. The following two conditions are equivalent:
(1) there exist two concave, monotonic, continuous utility functions whose sum rationalises the data;
(2) there exist numbers $\{V_t, W_t, \lambda_t > 0, \mu_t > 0\}_{t=1,...,T}$ that satisfy:

$$\begin{aligned} V_s &\leq V_t + \lambda_t \mathbf{p}_t^{1\prime}\left(\mathbf{q}_s^1 - \mathbf{q}_t^1\right) \\ W_s &\leq W_t + \lambda_t \mathbf{p}_t^{2\prime}\left(\mathbf{q}_s^2 - \mathbf{q}_t^2\right) \end{aligned}$$

for all $s, t \in \{1, ..., T\}$.

More structure: additive separability

- Once again there is no *explicit* condition which applies to the entire dataset but such a condition is implied by the two conditions stated.

- There is no GARP-like condition for this model: you *have to* work with the Afriat Inequalities.

- The Afriat conditions for additive separability are linear in unknowns this is easy to implement - unlike the condition for weak separability.

Adding Structure: Returns to scale

- Constant returns to scale is another example of a functional restriction (on a production function in this case) which is often useful to know about.

- The basic result for profit maximisation with a single-output production function was SAPM:

$$y_s \leq y_t + \frac{\mathbf{1}}{p_t}\mathbf{w}_t'(\mathbf{x}_s - \mathbf{x}_t) \qquad \forall s, t \in \{\mathbf{1}, ..., T\}$$

- Obviously constant returns to scale is a special case of this so the data will have to satisfy SAPM plus another condition.

Adding Structure: Returns to scale

- To briefly recap the SAPM theorem we suppose that the production function is concave: $f(\mathbf{x}_s) \leq f(\mathbf{x}_t) + \nabla f(\mathbf{x}_t)' (\mathbf{x}_s - \mathbf{x}_t)$ where $y_s = f(\mathbf{x}_s)$ and $y_t = f(\mathbf{x}_t)$.

- Optimising price-taking behaviour says that the firm should use inputs until the marginal revenue product is equal to the price of each input:

$$p_t \nabla f(\mathbf{x}_t) = \mathbf{w}_t$$

- So combining these gives SAPM:

$$y_s \leq y_t + \frac{1}{p_t} \mathbf{w}_t' (\mathbf{x}_s - \mathbf{x}_t)$$

Adding Structure: Returns to scale

- A production function has constant returns to scale if it is homogeneous of degree one:

$$f(\lambda \mathbf{x}) = \lambda f(\mathbf{x})$$

- Another very useful property of hod1 functions is that

$$f(\mathbf{x}) = \nabla f(\mathbf{x})'\mathbf{x}$$

which is due (on the balance of probabilities if nothing else) to Euler.

Adding Structure: Returns to scale

- Using the optimality $(p_t \nabla f(\mathbf{x}_t) = \mathbf{w}_t)$ implies that

$$\nabla f(\mathbf{x}_t) = \frac{1}{p_t}\mathbf{w}_t$$

- Therefore using the hod1 property we get

$$y_t = f(\mathbf{x}_t) = \nabla f(\mathbf{x}_t)'\mathbf{x}_t = \frac{1}{p_t}\mathbf{w}_t'\mathbf{x}_t$$

Adding Structure: Returns to scale

Theorem. *The following statements are equivalent:*

A. *there exists a constant returns to scale production function which rationalises the data* $\{\mathbf{w}_t, \mathbf{x}_t, y_t, p_t\}_{t=1,...,T}$.

B. *the data satisfy the conditions*

$$\begin{aligned} y_s &\leq y_t + \frac{1}{p_t}\mathbf{w}_t'\left(\mathbf{x}_s - \mathbf{x}_t\right) \qquad \forall s,t \in \{1,...,T\} \\ y_t &= \frac{1}{p_t}\mathbf{w}_t'\mathbf{x}_t \end{aligned}$$

C. *there exists a continuous, monotonic and concave constant returns to scale production which rationalises the data* $\{\mathbf{w}_t, \mathbf{x}_t, y_t, p_t\}_{t=1,...,T}$.

Adding Structure: Returns to scale

- This way of presenting the result accentuates the fact that the CRS structural assumption adds a condition to the empirical requirements: the data has to satisfy SAPM *plus* another condition.

- You can (of course!) combine the two conditions

$$
\begin{aligned}
y_s &\leq y_t + \frac{1}{p_t}\mathbf{w}_t'\mathbf{x}_s - \frac{1}{p_t}\mathbf{w}_t'\mathbf{x}_t \qquad \forall s,t \in \{1,...,T\} \\
y_s &\leq \frac{1}{p_t}\mathbf{w}_t'\mathbf{x}_s \qquad \forall s,t \in \{1,...,T\}
\end{aligned}
$$

Adding Structure: Returns to scale

- Combining this with $y_t = \frac{1}{p_t}\mathbf{w}_t$ Varian (1984) takes ratios to get the condition::

$$\frac{y_s}{y_t} \leq \frac{\mathbf{w}_t'\mathbf{x}_s}{\mathbf{w}_t'\mathbf{x}_t} \qquad \forall s, t \in \{1, ..., T\}$$

- This is less clear economically, but computationally it is simpler.

- The RP test for homothetic utility functions is basically identical - work it out as an exercise.

Adding Structure: Characteristics

- The linear characteristics model is due to Gorman (1956).

$$\begin{array}{r} K \text{ market goods} : \mathbf{q} \\ J \text{ characteristics} : \mathbf{z} \end{array} \qquad J < K$$

$$\max_{\mathbf{q}} v(\mathbf{z}) \text{ subject to } \mathbf{z} = \mathbf{A}'\mathbf{q} \text{ and } \mathbf{p}_t'\mathbf{q} \leq x_t$$

- The structure is akin to separability but with complete overlap across groups rather than a partition.

Adding Structure: Characteristics

Definition: A utility function $v(\mathbf{z})$ *rationalises* the data $\{\mathbf{p}_t, \mathbf{q}_t\}_{t=1,...,T}$ for the technology $\mathbf{A}$ if $v(\mathbf{z}_t) \geq v(\mathbf{z})$ for all $\mathbf{z}$ such that $\mathbf{z}_t = \mathbf{A}'\mathbf{q}_t$, $\mathbf{z} = \mathbf{A}'\mathbf{q}$ and $\mathbf{p}_t'\mathbf{q}_t \geq \mathbf{p}_t'\mathbf{q}$.

Adding Structure: Characteristics

Theorem. *The following statements are equivalent.*

(P) *there exists a utility function* $v(\mathbf{z})$ *which is non-satiated, continuous and concave in characteristics which rationalises the data* $\{\mathbf{p}_t, \mathbf{q}_t\}_{t=1,\ldots,T}$ *for given* $\mathbf{A}$.

(A) *there exist numbers* $\{V_t, \lambda_t > 0\}_{t=1,\ldots,T}$ *and vectors* $\{\boldsymbol{\pi}_t\}_{t=1,..,T}$ *such that*

$$V_s \leq V_t + \lambda_t \boldsymbol{\pi}_t' \left(\mathbf{A}'\mathbf{q}_s - \mathbf{A}'\mathbf{q}_t\right) \qquad \text{(A1)}$$

$$\mathbf{p}_t \geq \mathbf{A}\boldsymbol{\pi}_t \qquad \left(\text{with equality if } q_t^k > 0\right) \qquad \text{(A2)}$$

Adding Structure: Characteristics

- The model is

$$\max_{\mathbf{q}} v(\mathbf{z}) \text{ subject to } \mathbf{z} = \mathbf{A}'\mathbf{q} \text{ and } \mathbf{p}_t'\mathbf{q} \leq x_t$$

- Maximising behaviour and linear structure

$$\lambda_t \mathbf{p}_t \geq \mathbf{A} \boldsymbol{\nabla} v(\mathbf{z}_t) = \boldsymbol{\nabla} u(\mathbf{q}_t)$$

$$\boldsymbol{\pi}_t = \frac{1}{\lambda_t} \boldsymbol{\nabla} v(\mathbf{z}_t)$$

$$\mathbf{p}_t \geq \mathbf{A} \boldsymbol{\pi}_t \tag{A2}$$

Adding Structure: Characteristics

- Using the standard property of concave functions

$$v(\mathbf{z}_s) \leq v(\mathbf{z}_t) + \boldsymbol{\nabla} v(\mathbf{z}_t)' (\mathbf{z}_s - \mathbf{z}_t)$$

- Given $\mathbf{p}_t \geq \mathbf{A}\boldsymbol{\pi}_t$ and $\mathbf{z}_t = \mathbf{A}'\mathbf{q}_t$

$$v(\mathbf{z}_s) \leq v(\mathbf{z}_t) + \lambda_t \boldsymbol{\pi}_t' \left(\mathbf{A}'\mathbf{q}_s - \mathbf{A}'\mathbf{q}_t\right)$$

$$V_s \leq V_t + \lambda_t \boldsymbol{\pi}_t' \left(\mathbf{A}'\mathbf{q}_s - \mathbf{A}'\mathbf{q}_t\right) \tag{A1}$$

Adding Structure: Characteristics

- We have T overestimates of the utility of an arbitrary bundle

$$V(\mathbf{z}) \leq \left\{ V_s + \lambda_s \boldsymbol{\pi}'_s (\mathbf{z} - \mathbf{z}_s) \right\}_{s=1,\ldots,T}$$

so take

$$V(\mathbf{z}) = \min_s \left\{ V_s + \lambda_s \boldsymbol{\pi}'_s (\mathbf{z} - \mathbf{z}_s) \right\}_{s=1,\ldots,T}$$

as a utility function (it's piecewise linear, non-satiated, and concave).

Adding Structure: Characteristics

- Suppose we have some $\mathbf{z} = \mathbf{A}'\mathbf{q}$ with $\mathbf{p}_t'\mathbf{q}_t \geq \mathbf{p}_t'\mathbf{q}$.

$$V(\mathbf{z}) = \min_s \left\{ V_s + \lambda_s \boldsymbol{\pi}_s' (\mathbf{z} - \mathbf{z}_s) \right\} \leq V_t + \lambda_t \boldsymbol{\pi}_t' (\mathbf{z} - \mathbf{z}_t)$$

- Since $\mathbf{p}_t' \geq \boldsymbol{\pi}_t' \mathbf{A}'$ with equality when $q_t^k > 0$

$$\mathbf{p}_t'\mathbf{q}_t \geq \mathbf{p}_t'\mathbf{q} \;\; \Rightarrow \quad \boldsymbol{\pi}_t'\mathbf{z}_t \geq \boldsymbol{\pi}_t'\mathbf{z}$$

- Since $\lambda_t > 0$ this means $\lambda_t \boldsymbol{\pi}_t' (\mathbf{z} - \mathbf{z}_t) \leq 0$. So $V(\mathbf{z}) \leq V_t$.

Adding structure - summary

- Separability and returns to scale are two examples in which the precise structure of preferences or technology is of interest.

- The basic condition remains, but this is augmented with further inequality restrictions reflecting the extra assumption of interest.

- By switching on/off this additional condition we can conduct a "specification" search for particular properties of the model.

Adding structure - summary

- The additional restriction also helps with prediction - the prediction has to statisfy the basic conditions (e.g. GARP) but also the further restrictions created by the additional structure.

- They therefore sharpen the bounds by reducing the size of the set of theory-consistent observations.

- This emphasises the general point that : stronger theoretical restrictions give you stronger predictions.

Empirical Revealed Preference - summary

1. A “elementary” way of combining theory and data.

2. We have looked at

   (a) The basic ideas/results

   (b) Practical issues in implementation

   (c) Adding structure to models

Empirical Revealed Preference - summary

The good things about RP include:

1. It is simple and clean from a theoretical point of view

2. It focuses analysis on the behaviour of the individual, i.e. the level at which the theory applies, and not the behaviour of statistics.

3. It does not rely on or indeed require errors.

4. It introduces loss functions to empirical work which are more economically meaningful than the sum of squared residuals.

Empirical Revealed Preference - summary

The bad things about RP include (but are not limited to):

1. Bounds on objects of interest can be so wide that it is close to useless for practical day-to-day empirical work.

2. It is hard to satisfactorily connect RP to standard econometric practice.

3. When a model fails you cannot just throw in an error

4. When a model doesn't fail what should you make of that when "all models are wrong"?

Thanks

Comments suggestions and queries to

ian.crawford@economics.ox.ac.uk