Falsifiability of Economic Assumptions and Partial Identification

Jinhyun Lee*
University of St Andrews

February 2011

Abstract

This paper develops a logical framework of "structural" analysis allowing for partial identification. It discusses testability of a structural (economic) feature when it is set identified. Set identification and sharpness are formally defined by using the Hurwicz (1950) structure. Emphasis is put on sharpness with regards to meaningful inference. The power of a test based on a non-sharp identified set could be zero. Taking intersection does not guarantee sharpness. Set identified structural features are only *refutable* in Breusch's (1986) sense. Criteria for testability/falsifiability of an economic model or restrictions under set identification are proposed in the case where the one model is nested by the other model. Unless there exists an identification result specifying a one-to-one mapping between a specific economic model and a specific observed data, it is not possible to "**confirm**" that the observed data are generated by a certain economic model. Rather, all we can do with data is to "**falsify**" whether a specific economic model is the underlying economic model that generates the data we observe.

1 Introduction

> "*Most of economic intuition is expressed in terms of the structure, so the structure is often the object of interest for estimation and for testing. .... the reduced form is convenient theoretically, but to be most useful, facts about it have to be translated back into structural statements...* (Kadane and Anderson (1977), p1028, quoted by Breusch (1986)"

*This paper is the revised version of one chapter of my Ph D thesis. This is motivated by A. Chesher's unpublished lecture note entitled "Evidence in Economics" in which *falsifiability* of a model is discussed using point identification.

By a "structure" we mean an economic model. In order to link an economic model with data, we introduce the Hurwicz (1950) structure. Unless there exists an identification result specifying a one-to-one mapping between a specific economic model and a specific observed (realized) data, it is not possible to "**confirm**" that the observed data are generated by a certain economic model. Rather, all we can do with data is to "**falsify**" whether a specific economic model is the underlying economic model that generates the data we observe.

Structural data analysis often starts by taking a specific economic model and see if a specific feature of the economic model can be (possibly nonparametrically) identified and then we proceed to estimate the economic feature. We emphasize falsifiability of an economic model as a pre-step of such structural data analyses. Whether a model can be falsifiable or not would affect the interpretability of data analysis substantially. However, falsifying economic models/assumptions is not always possible. We discuss one of the possibilities of falsifiabilit in this paper.

Since it is such a challenging problem to identify an economic model, we may have to focus on identifiability of certain features of an economic model rather than the economic model. *Testability* of structural features is closely related to, but distinct from *identifiability.* Examining testability of structural features can provide ways to falsify economic or econometric assumptions, thereby, economic models. Following the spirit of partial identification of features of *probability distributions*, pioneered by C. Manski (see Manski (1989,1990), and Manski (1995) for economic examples that motivate partial identification, and Manski (2003) for a survey of recent developments), in this paper set identification and sharpness of an identified set are formally defined by using the nonparametric Hurwicz (1950) structure that can be applied to models with multiple equilibria. The Hurwicz (1950) structure, as a tuple of structural relations and the distribution of the unobservables, has been used in many nonparametric identification studies. The logic of testability of structural features is discussed by extending earlier results of Koopmans and Reiersol (1950) into a general nonparametric setup allowing for set identification. Jovanovic (1989) modified Koopmans and Reiersol (1950)'s framework so as to deal with models with multiple equilibria and offers a general framework for statistical inference in such models. Jovanovic (1989) noticed the possibility of set identification, however, "identification" in his paper is restricted to mean *point* identification.

The identifying power of a model comes from the restrictions imposed by an econometric model and credibility of restrictions should be discussed in each application of the model. However, not all restrictions have testable implications on the distribution of the observables (data). Even in such a case the restrictions can be tested when there exist two distinct models that identify the same structural features with the one model nested by the other. The criteria proposed for testability or falsifiability involve comparison of the two identified sets - the identified set defined by the nested model should be smaller. Galichon and Henry (2009) develop a test of nonidentifying restrictions under the Jovanovic (1989) setup by explicitly allowing for partial identification. Galichon and Henry (2009)'s test can be used to falsify an econometric model. This paper suggests an alternative framework of falsification of an econometric model, by which falsification of restrictions may be achieved.

Results from nonparametric *structural* data analysis may require a revisit of a theoretical model, which shows how data analyses can interact with theoretical developments. Deriving testable implications from economic models such as comparative statics and falsifying them could be crucial, particularly if different economic models implied different policy suggestions. For policy purposes experiments may have to be conducted on a reasonably *representative*, and possibly *large*, sample from the population that a policy is planned to be implemented on. Generalization or extrapolation of the results from a small sample experiment needs to be avoided without convincing justification, and it is especially so if there is any evidence of heterogeneity in response.

2 Set identification of Hurwicz (1950) Structural Features

2.1 Gaps between Economic Models and Data Analysis

Microeconomic models attempt to explain how individual agents make their decisions by which policy makers may predict individuals' (possibly heterogeneous) responses in the face of policy changes. While economic models may describe deterministic rules for each individual decision makers, data analysis needs to incorporate stochastic elements. The gaps between economic models and data analysis are caused by how to incorporate stochastic elements and how to interpret them.

We start with choice theory to motivate the use of Hurwicz (1950) structure for structural data analysis. There exist well known testable axioms from the choice theory under rationality. The revealed preference idea is adopted as an example of testable implications.

2.1.1 Data analysis based on choice theory

Let $\gamma(A) : \Sigma \to \mathcal{A}/\varnothing$, with $\gamma(A) \subset A$, $\forall A \in \Sigma$, where $\mathcal{A}$ is the set of all the alternatives to the agent, and $\Sigma \subset 2^{\mathcal{A}/\varnothing}$ is a certain set of some subsets of $\mathcal{A}$, called choice sets. $\gamma(A)$ is called choice function. An economic model may specify how an individual chooses among the alternative choice sets available. $\gamma(A)$ is a primitive of the model.

Individuals' choice data are generated by $\gamma(A)$ and let Y be the choice outcome, X denote a vector of variables defining choice sets, $A \in \Sigma$, and $F_{Y|X}(y|x)$ indicate the conditional distribution of Y given X. Data may be used to recover the underlying choice rule, $\gamma(A)$.

If experimental data were not available, and if observational data do not contain individuals' choice for rich choice alternatives, then we would have to discuss how to extend the axioms on individual's choices into those for the population.

Example (Blundell, Kristenssen, Matzkin (2010)) A is a budget set defined by a price vector (P) and income (M). X includes P and M and Y is demand for each good. Then the choice mapping can be written as

$$Y = \gamma(X), \text{ where } X = [P, M]'.$$

Suppose we observe data on X and Y, $(X_1, Y_1), ...(X_i, Y_i), ...(X_n, Y_n)$. The decision rule, γ may be varying across different individuals. Then individual i's choice data (X_i, Y_i) is generated by $Y_i = \gamma_i(X_i)$. However, it is impossible to recover $\gamma_i(\cdot)$, since we would observe only limited number of realization of (X, Y) for each individual, while $\gamma_i(\cdot)$ needs to specify how choices are made for all given possible available choice sets. Thus, Blundell, Kristenssen, Matzkin (2010) specify the choice function as the following

$$Y = \gamma(X, U), \text{ where } X = [P, M]'$$

where U indicates a scalar unobserved heterogeneity, which is generated by the probability law, $F_U(u)$. With this specification, individuals' choice can vary even if the same value of X is given, thus, generating a distribution of Y given a fixed value of X. Then the primitive that may be the focus of identification would be $\{\gamma(\cdot, \cdot)$, $F_U(u)\}$, rather than $\{\gamma_i(\cdot)\}$.

That is, while economic models may characterize individual-specific choice functions, individuals' data are not usually enough to recover them. Rather, data analyses attempt to recover a commonly applicable choice relation to the whole population, from the sample (that is, data from the population). Individual heterogeneity, then, will be explained by the unobserved stochastic term.

To deal with this gap we introduce Hurwicz (1950) structure.

2.2 Elements of Identification

Distribution functions are denoted by F_A indicating the distribution function of A. $F_{A|B}$ indicates the conditional distribution of A given B. The corresponding $\tau-$quantiles are denoted by $Q_A(\tau)$ and $Q_{A|B}(\tau|b)$[1].

We assume that economic processes are generated by individual's decision mechanisms.[2] The decision mechanisms are usually described as relationships between variables. We denote these underlying mechanisms as "structures" following Hurwicz (1950). Hurwicz (1950) assumed that the distribution of the observables is generated

[1] The quantiles are defined by $Q_{A|B}(\tau|b) = \inf\{a|F_{A|B}(a|b) \geq \tau\}$

[2] By "individual" we mean any economic decision unit of interest.

by a transformation $\mathcal{H}$ performed on the distribution of the unobservables, $F_{U|X}$ and defined the structure, S, as a tuple of the mapping ($\mathcal{H}$) and the distribution of the unobservables ($F_{U|X}$) where Y is a vector of endogenous[3] variables determined by the economic decision processes, X is a vector of covariates (exogenous variables) and U is a vector of unobserved elements to the analyst. From a random sample of observations on Y and X the conditional distribution of Y given X is identified[4]. Denote $F^S_{Y|X}$ to be the distribution of the observables that is generated by a structure S.

Let $\mathcal{S}$ be a set of all structures. Let Ψ^U be a set of all distributions of unobservables ($F_{U|X}$) and $\Psi^{\mathcal{S}}$ be the set of all distribution functions of observables generated by elements in $\mathcal{S}$, that is, $\Psi^{\mathcal{S}} = \{F^S_{Y|X} |\ S \in \mathcal{S}\}$. Then $\mathcal{H}$ is a one-to-one mapping from Ψ^U to $\Psi^{\mathcal{S}}$. That is, a Hurwicz (1950) structure[5], $S = \{\mathcal{H}, F_{U|X}\}$, and observed data ($F_{Y|X}$) have the following relation called "Hurwicz mapping"

$$\underbrace{F_{Y|X}}_{\text{Reduced form}} = \underbrace{\mathcal{H}(F_{U|X})}_{\text{Structure}}.$$

The Hurwicz mapping $\mathcal{H}(\cdot)$ is assumed to *uniquely*[6] determine the distribution of the observables. $\mathcal{H}(\cdot)$ can specify structural relations and can assume the existence of an equilibrium selection mechanism if there are multiple equilibria. However, this does not mean that the econometric model for any economic setup with multiple equilibria should specify the equilibrium selection mechanism. Only the existence of a selection mechanism is required.

Testable implications such as WARP/GARP from choice theory can be stated ex ante, that is, before data are realized. (Point or partial) identification strategies from econometric models can also be stated ex ante. Once data are realized, one can determine whether the observed data satisfy the testble implications and one can find the identified point or set. Testable implications are derived economic models, and econometric models may adopt such implications for identification purposes, or may use them to test theory. Typically econometric models based on specific economic

[3]Following Koopmans (1949, p133), endogenous variables are "observed variables which are not known, or assumed to be statistically *not* independent of the latent variables, and whose occurrence in one or more equations of the set of equations is necessary on grounds of "theory"".

[4]The exact knowledge of $F_{Y|X}$ cannot be derived from any finite number of observations. Such knowledge is the limit approachable but not attainable by increasing the number of observation (Koopmans and Reiersol (1950)).

[5]Note that this definition of a structure is different from that by Koopmans (1949) in which it is defined as "the combination of a distribution ofd latent variables and a complete set of structural equations." The definition in this paper does not require "complete" specification of each endogenous variable.

[6]Note that this uniqueness should be distinguished from the assumption that the *structural relations* uniquely determine values of endogenous variables given exogenous variables. When structural relations do not specify a one-to-one mapping between the endogenous variables and unobserved variables given exogenous variables, this setup assumes that there should be a mechanism that selects one point among many. This selection is possibly unknown thus, unspecified by the analysist since there may not be a well defined or a convincing way of doing it.

models need to impose additional restrictions to those imposed by the economic models to deal with the gaps between the economic models and econometric models.

2.2.1 Econometric models and identification

The econometric model, $\mathcal{M}$, is characterized by a priori information (restrictions) applied to the structure and the distribution of the observables. The model, $\mathcal{M}$, is defined to be the set of the structures that satisfy the restrictions. Let $\Psi^{\mathcal{M}}$ be a set of all possible distribution functions generated by $S \in \mathcal{M}$, $\Psi^{\mathcal{M}} = \{F^S_{Y|X} | \ S \in \mathcal{M}\}$.

Let S_0 be the true structure that generates the distribution of observables available to us, $F^0_{Y|X}$. The two structures S and S' are called **observationally equivalent to each other** if $F^S_{Y|X} = F^{S'}_{Y|X}$. Define $\Omega_0 = \{S : F^S_{Y|X} = F^0_{Y|X}\}$, a set of structures that are observationally equivalent to the true structure, S_0 (note that $S_0 \in \Omega_0$ by definition of Ω_0 . See Figure 1.)

Example from choice theory - observationally equivalent structures. Manzini and Mariotti (2009) report that choice data generated by their categorization then choose (CTC) model with salient categorization satisfy WARP. Thus, the two decision processes - the one under rationality and the other under bounded rationality - are observationlly equivalent and no amount of data can distinguish the one from the other.

Often the whole Hurwicz structure is hard to identify, thus, data analysis may focus on certain features of the structure, rather than the whole structure. The structural feature, $\theta(S)$ is defined as a real-valued functional of the structure, $\theta(\cdot) : \mathcal{S} \to \mathcal{R}^d$. It can be an economic object that is important in policy design such as elasticities, risk attitudes, time preferences, valuation of an auction object, or willingness to pay etc. One of the main objectives of specifying an econometric model is to recover the true economic data generating structure S_0, or some features of it, $\theta(S_0)$.

The true structure, S_0 is said to be point identifiable in $\mathcal{M}$ if there is no other member of $\mathcal{M}$ that is observationally equivalent to S_0. A structural feature $\theta(S_0)$ is said to be point identifiable if there is no variation in the values of the structural feature of the admitted structures. See Hurwicz (1950), Koopmans and Reiersol (1950), Roehrig (1988), and Matzkin (1994, 2007) for general discussion of point identification. Matzkin (2007) reviews recent developments of nonparametric identification.

2.2.2 Examples of Hurwicz's (1950) Structure

1. **Nonparametric Simultaneous Equations Models** of Roehrig (1988), Benkard and Berry (2006), or Matzkin (2008) : The mapping $\mathcal{H}$ is a system of structural functions, $U = G(Y, X)$, where Y, X, and U are defined as before. The dimension of Y needs to be the same as that of U. Their studies assume that the

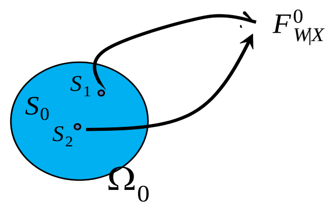


Figure 1: Note that $S_0 \in \Omega_0$, by definition of Ω_0. Suppose $S_1, S_2 \in \Omega_0$. Then $F^{S_1}_{Y|X} = F^{S_2}_{Y|X} = F^0_{Y|X}$. That is, S_1 and S_2 are observationally equivalent structures and they are indistinguishable from data - no amount of data can distinguish S_1 from S_2. Note that the true structure, S_0, that generates the distribution of the observables we have, is always in Ω_0.

structural functions are unique valued and one-to-one mapping between Y and U, excluding discrete endogeneous variables. The structural feature of interest, $\theta(S)$, can be the value of the structural function evaluated at a specific point, or partial derivatives of the structural functions.

2. **Treatment Effects** or missing data problem using Hurwicz's (1950) Structure : $\mathcal{H}$ can be a mapping from a set of joint distributions of the scalar potential outcomes, Y_1 and Y_0, $F_{Y_1Y_0|X}$, to Ψ, such that

$$F_{Y|X} = \mathcal{H}(F_{Y_1Y_0|X}).$$

   Note that $F_{Y_1Y_0|X}$ is unobserved. $\theta(S)$ can be an average, or quantiles of treatment effects, $\theta(S) = E(Y_1 - Y_0|X = x)$, or $\theta(S) = Q_{Y_1-Y_0|X}(\tau|x)$.

3. **Models for Oligopoly Entry Games** : The mapping $\mathcal{H}$ can be structural relations together with an equilibrium selection mechanism. Let the structural relations be specified by the threshold crossing structures as $Y_1 = 1(X_1\beta_1 + Y_2\Delta_1 + U_1 \geq 0)$ and $Y_2 = 1(X_2\beta_2 + Y_1\Delta_2 + U_2 \geq 0)$ as in Brehnahan and Reiss (1991). These structural relations do not predict unique outcomes. By assuming a specific eqilibrium selection rule, π, point identification can be achieved .

4. **Binary Choice Models** without endogeneity of Manski (1988) and Matzkin (1992) : $\mathcal{H}$ is a latent structural relation together with a threshold crossing

structure that transfoms the latent structural relation into the observed distribution. Manski (1988)'s case is $Y = 1(X\beta + U \geq 0)$, and Matzkin (1992)'s case is $Y = 1(h(X) + U \geq 0)$. The structural feature of interest in these papers are : $\theta(S) = \{\beta,\ F_{U|X}\}$ and $\theta(S) = \{h(x),\ F_{U|X}\}$, where x is a realized value of the random variable X. Under a set of restrictions, point identification of the structural features can be established. Note that neither of the restrictions of the two models are included by the other. One of the common restrictions in the two models is the existence of a continuous explanatory variable (can be called a special regressor as in Lewbel (2000)), which is relaxed in Magnac and Maurin (2007, 2008).

5. **Auction models - the interpretation of** $F_{U|X}$: The data available are individuals' bids, and the underlying structure is $S = \{h, F_{U|X}\}$, where $F_{U|X}$ is the joint distribution of latent individual valuation and private information (types) satisfying certain statistical properties (e.g. independence and symmetry etc.) and h is a true mapping from the true distribution of types (U) to a distribution of observable bids (Y) implied by the assumption of the Baysian Nash equilibrium. The structural relation h between bids and valuation is implied by the Bayesian Nash equilibrium, which is represented by a mapping, $\mathcal{H}(Y, X, U) = 0$. See Athey and Haile (2005).

6. Intrahoushold allocation

7. Insurnace market ; testable implication ?

2.2.3 The identification problem

> "Statistical inference, from observations to economic behavior parameters, can be made in two steps : inference from the observations to the parameters of the assumed joint distribution of the observations, and inference from that distribution to the parameters of the structural equations describing economic behavior. The latter problem of inference described by the term "identification problem"... (Koopmans (1949)"

The word "identification" has been used in different contexts meaning different things. As indicated in the above quote from Koopmans (1949) the identification problem in this paper means

Remarks on "local identification a la Rothenberg (1978)

2.3 Set Identification and Sharpness

Sometimes point identification is not achievable unless stronger restrictions on the structure are imposed. If such strong, often parametric restrictions are hard to justify in the context of an economic application, then we may try to obtain partial identification by imposing weaker restrictions instead of imposing unreasonable restrictions that guarantee point identification.

For a given econometric model, $\mathcal{M}$, defined in the previous section, let $\mathbf{\Theta}^{\mathcal{M}}(\cdot)$ be a mapping from $\Psi^{\mathcal{M}}$ to a class of sets in R^d, where d is the dimension of the structural feature, $\theta(S)$, specified in the economic example. We say a structure is *admissible* if the structure satisfies all the restrictions specified by the model.

Two concepts - admissibility and observational equivalence - are key in the discussion of identification. The mapping is written as $\mathbf{\Theta}^{\mathcal{M}}(F^S_{Y|X})$. We say the model, $\mathcal{M}$, set identifies the structural feature, $\theta(S_0)$, if we can *determine* a set $\mathbf{\Theta}^{\mathcal{M}}(F^0_{Y|X})$, given data, $F^0_{Y|X}$, such that for **any** admitted structure S that is observationally equivalent to S_0, $\theta(S) \in \mathbf{\Theta}^{\mathcal{M}}(F^0_{Y|X})$. $\mathbf{\Theta}^{\mathcal{M}}(F^0_{Y|X})$ can be defined either explicitly with the boundary explicitly specified (Manski (1990,1997), Manski and Pepper (2000), Chesher (2005), and Lee (2010)), or implicitly by some moment inequalities (some of the entry models (see Berry and Tamer (2007) for survey), Honore and Tamer (2006), Magnac and Maurin (2008), and Chesher (2010), for example).

Since several studies define their identified sets as those that may contain outer regions[7], an identified set is defined as a bigger set that contains a sharp identified set, which will be defined later. An identified set is defined as the following :

Definition 1 Set Identification : the model $\mathcal{M}$ **set identifies** the structural feature, $\theta(S_0)$ if $\exists\ \mathbf{\Theta}^{\mathcal{M}}(\cdot)$ s.t. $\forall S \in \mathcal{M} \cap \Omega_0$, $\theta(S) \in \mathbf{\Theta}^{\mathcal{M}}(F^0_{Y|X})$, where Ω_0 is defined as before.

If $\mathbf{\Theta}^{\mathcal{M}}(\cdot)$ is singleton, in other words, if every admitted and observationally equivalent structure generates the same value of the structural feature, then we say the structural feature is point identified by the model. See Figure 2.

Failure of point identification and set identification (examples continued)

1. **Nonparametric and Nonseparable Simultaneous Equations Models** : Identifiability results in Matzkin (2008) cannot be applied when any of endogenous variables are discrete since differentiability and the one-to-one mapping assumption between the unobservables and the discrete endogenous variable does not hold. Other identification strategies using a triangular system or single equation IV models have been proposed to show point identification. For a triangular system, see Chesher (2003) and Imbens and Newey (2009), and for single equation IV models see Chernozhukov and Hansen (2005). However, when the regressor is discrete under the triangular system (see Chesher (2005), Jun, Pinkse, and Xu (2010), and Lee (2010) for discrete endogenous regressor)

[7] An identified set may contain some *outer regions* as discussed in Beresteanu, Molchanov, and Molinari (2008).

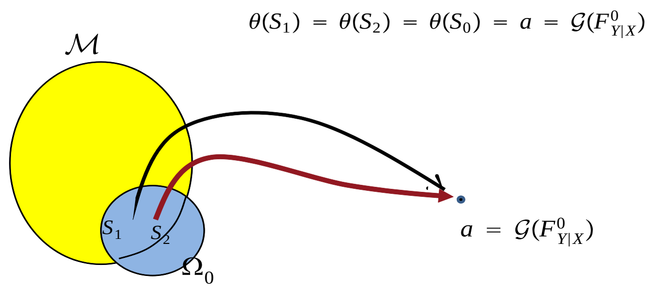


Figure 2: Point identification of a structural feature, $\theta(S)$ by Lemma 1 in Chesher (2007) : it is said to be point identified if there exists a unique valued functional $\mathcal{G}(\cdot)$ such that $\theta(S) = \mathcal{G}(F^0_{Y|X}), \forall S \in \mathcal{M} \cap \Omega_0$. In other words, if $S_1, S_2 \in \mathcal{M} \cap \Omega_0$, then $\theta(S_1) = \theta(S_2) = \theta(S_0) = a = \mathcal{G}(F^0_{Y|X})$. Thus, in this case, $\Theta^{\mathcal{M}}(F^0_{Y|X}) = \mathcal{G}(F^0_{Y|X}) = \{a\}$. The identification analysis provides a way to find out the form of unique-valued functional $\mathcal{G}(F^0_{Y|X})$ to achieve point identification, or the form of the set, $\Theta^{\mathcal{M}}(F^0_{Y|X})$.

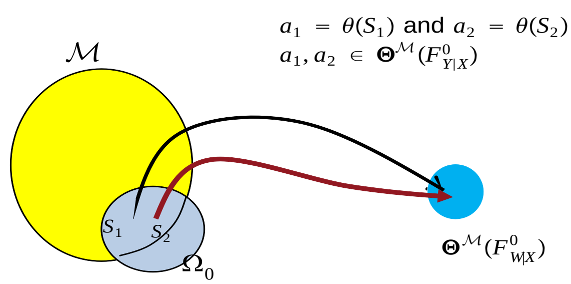


Figure 3: When the structural feature is set identified, and not point identified, then for any two admitted and observationally equivalent structures, $S_1, S_2 \in \mathcal{M} \cap \Omega_0$, with $a_1 = \theta(S_1)$ and $a_2 = \theta(S_2)$, all we can say is that $a_1 \in \Theta^{\mathcal{M}}(F^0_{Y|X})$ and $a_2 \in \Theta^{\mathcal{M}}(F^0_{Y|X})$, a_1 and a_2 can be distinct values.

and when the outcome is discrete in the single equation IV model (see Chesher (2010)), point identification fails.

2. **Treatment Effects** : When parametric assumptions on the distribution function are relaxed, strong restrictions such as identification at infinity (see Heckman (1990)) are required for point identification of average treatment effects. Several studies report partial identification results under weaker restrictions : see Manski (1990,1997) and Heckman and Vytlacil (2001), Manski and Pepper (2000), Shaikh and Vytlacil (2005), and Bhattacharya, Shaikh, Vytalcil (2008).

3. **Models for Oligopoly Entry Games :** without specifying an equilibrium selection mechanism point identification in the entry models is not achievable. See Tamer (2003), and Berry and Tamer (2007) for a recent survey. See also Pakes (2010).

4. **Binary Choice Models with Endogeneity** : With the large support condition in Lewbel's (2000) model with a special regressor, Magnac and Maurin (2008) show partial identification results using the moment conditions derived from their model restrictions.

If a set, $\mathbf{\Theta}^{\mathcal{M}}(F^0_{Y|X})$, includes **all** the values of a feature of structures that are admissible and observationally indistinguishable and if it contains **only** such values, then $\mathbf{\Theta}^{\mathcal{M}}(F^0_{Y|X})$ is called a sharp identified set.

Definition 2 A sharp identified set, $\mathbf{\Theta}^{\mathcal{M}}_{Sharp}(F^0_{Y|X})$ is defined as $\mathbf{\Theta}^{\mathcal{M}}_{Sharp}(F^0_{Y|X}) \equiv \{a : \theta(S) = a, \forall S \in \mathcal{M} \cap \Omega_0\}$.

To show set identification, it needs to be shown that an identified set, $\mathbf{\Theta}^{\mathcal{M}}(F^0_{Y|X})$, contains **all** the values of a feature of structures that are observationally equivalent and admitted by $\dot{\mathcal{M}}$. However, not every point in $\mathbf{\Theta}^{\mathcal{M}}(F^0_{Y|X})$ is necessarily generated by an admitted structure that is observationally equivalent. (e.g. Andrews, Berry, and Jia (2004), Ciliberto and Tamer (2009)). A set defined by an identification strategy may **not** include *all* the points that are generated by admitted and observationally equivalent structures.

Beresteanu, Molchanov, and Molinari (2008) define a sharp identified region as

> "... the region in the parameter space which includes *all* possible parameter values that (i) could generate the same distribution of observables for some data generation process (ii) *consistent* with the maintained modeling assumptions and no other parameter value, is called the sharp identified region. .." .

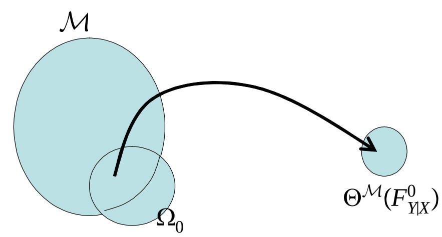


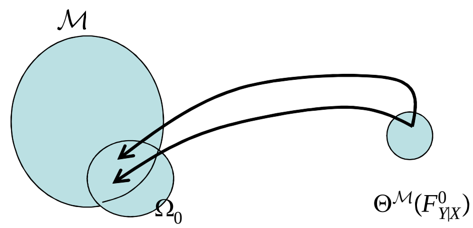
Figure 4: Set identification : the value of the structural feature $(\theta(S))$ generated by any structure that is admitted by the model and observationally equaivalent to the true structure should lie in the set $\Theta^{\mathcal{M}}(F^0_{Y|X})$. Two distinct points in the identified set may have been generated by two distinct structures, but they should be admitted (consistent with the model) as well as observationally equivalent to each other (consistent with the data). If the structural feature is point identified, that is, $\Theta^{\mathcal{M}}(F^0_{Y|X})$ is singleton, distinct structures that are admitted and observationally equivalent should generate the same value for the structural feature. Note that there can be some parts of the set, $\Theta^{\mathcal{M}}(F^0_{Y|X})$, where $\theta(S)$ never lies. Sharpness of the identified set guarantees that there will be no such parts, in which case, the set can be described as "the smallest set that exhausts all the information from the data and the model" as some authors define sharpness.

This is a descriptive definition of sharpness. This descriptive definition can be mapped into Definition 2 because $\mathbf{\Theta}^{\mathcal{M}}_{Sharp}(F^0_{Y|X})$ is the set of all values of the structural feature, $\theta(S)$, that are generated by an element in $S \in \mathcal{M} \cap \Omega_0$. "Consistent with the model ($S \in \mathcal{M}$)" and "generate the same the distribution of the observables ($S \in \Omega_0$)" can be guaranteed by the fact that $S \in \mathcal{M} \cap \Omega_0$.

If one can show that $\mathbf{\Theta}^{\mathcal{M}}(F^0_{Y|X}) = \mathbf{\Theta}^{\mathcal{M}}_{Sharp}(F^0_{Y|X})$, then sharpness of an identified set is shown. Another way of showing sharpness is to use the following lemma. Suppose that for every value in an identified set, there exists an admitted and observationally equivalent structure whose feature is that value, then the identified set is sharp.

Lemma 1 Suppose that $\forall\ a \in \mathbf{\Theta}^{\mathcal{M}}(F^0_{Y|X}), \exists S \in \mathcal{M} \cap \Omega_0$ with $\theta(S) = a$. Then $\mathbf{\Theta}^{\mathcal{M}}(F^0_{Y|X}) = \mathbf{\Theta}^{\mathcal{M}}_{Sharp}(F^0_{Y|X})$.

Proof. Suppose that $\forall\ a \in \mathbf{\Theta}^{\mathcal{M}}(F^0_{Y|X}), \exists S \in \mathcal{M} \cap \Omega_0$ with $\theta(S) = a$. First, note that $\mathbf{\Theta}^{\mathcal{M}}(F^0_{Y|X}) \subseteq \mathbf{\Theta}^{\mathcal{M}}_{Sharp}(F^0_{Y|X})$, since for any $a \in \mathbf{\Theta}^{\mathcal{M}}(F^0_{Y|X}), \exists S \in \mathcal{M} \cap \Omega_0$ with $\theta(S) = a$, it should be the case that $a \in \mathbf{\Theta}^{\mathcal{M}}_{Sharp}(F^0_{Y|X})$. Next, $\mathbf{\Theta}^{\mathcal{M}}(F^0_{Y|X}) \supseteq \mathbf{\Theta}^{\mathcal{M}}_{Sharp}(F^0_{Y|X})$



$$\forall a \in \Theta^{\mathcal{M}}(F^0_{Y|X}), \exists S \in \mathcal{M} \cap \Omega_0$$

Figure 5: Sharpness : showing sharpness involves showing that for each point in the set there exists at least one structure that is admitted (consistent with the model) and observationally equivalent (consistent with the data) to the true structure, S_0. Note that two distinct structures could generate the same value for the structural feature.

since $\mathbf{\Theta}^{\mathcal{M}}_{Sharp}(F^0_{Y|X})$ is the smallest identified set without any outer region. Then $\mathbf{\Theta}^{\mathcal{M}}(F^0_{Y|X}) = \mathbf{\Theta}^{\mathcal{M}}_{Sharp}(F^0_{Y|X})$ follows. ■

Discussion on Sharpness (examples continued)

1. **Nonparametric Structural analysis with Discrete Data** : Chesher (2010) and Lee (2010) show sharpness of their identified sets. The proofs involve showing that for every point in the identified set, there exists at least one admitted and observationally equivalent structure

2. **Treatment effects** : Heckman, Smith, and Clements (1997), Shaikh and Vytlacil (2005), Fan and Park (2010), and Firpo and Ridder (2009) study the distribution of treatment effects defined as the difference between the potential outcomes, $Y_1 - Y_0$. Since the Hurwicz's (1950) mapping can be considered to transform the distribution of $Y_1 - Y_0$ (unobservable) into the distribution of Y (observable), their sharpness proofs involve construction of $F_{Y_1 - Y_0|X}$ from the observed distribution, $F_{Y|X}$. For the constructed distribution, $F_{Y_1 - Y_0|X}$ to be legitimate, it has to satisfy the properties of distribution functions.

3. **Entry models** : Ciliberto and Tamer (2007) recognize that the inequality restrictions taken in the entry game do not generate sharp identified sets.

4. **Monotone Binary Choice Models :** Magnac and Maurin (2008)'s identified set is all the points that are observationally equivalent and that satisfy the moment restriction derived in their papers. This is enough since they showed

that their moment conditions equivalently express all the restrictions imposed by the model, thus, all the observationally equivalent structures that satisfy the moment conditions should be those that are admitted by the model.

2.4 Sharpness and Inference - the Zero Power Property

Sharpness of an identified set is a logically essential property for inference. When the identified set is not sharp, the confidence region is constructed "conservatively". However, if an identified set is not shown to be sharp, then the inference based on the non-sharp identified set can be *meaningless* - it could be the case that one never rejects a hypothesis (regarding the structural feature) even if the hypothesis were not true. For example, consider one is interested in $H_0 : \theta(S) = 0$. Without the information on the outer region, confidence region constructed on the non-sharp identified set might not be informative, since even if zero lied in the confidence region, one never knows whether zero is in the outer region or not. If zero lied in the outer region of the identified set, then the inference would fail.

When one fails to show sharpness of an identified set, one could proceed by constructing confidence regions for the non-sharp identified set. However, it may be the case that the inference based on this may not be very informative. Rather, one could serch for an alternative model to define a sharp set.

2.5 Over-identification and the Intersection of Sharp Identified Sets

A model defines an identified set and a sharp identified set *always* exists once a model is given. There can be many such sets (overidentification)[8], for example, different values of IVs define different identified sets. Not every identified set is sharp. Also, even though a model may define several identified sets, the intersection of them does not guarantee sharpness since every identified set may contain some common *outer regions.*

Lemma 2 Suppose that there exist two *sharp* identified sets. Then the intersection of the two cannot be an identified set.

Proof. Definition 1 and 2 are to be used. Suppose two distinct sets, $\mathbf{\Theta}_1^{\mathcal{M}}(F_{Y|X}^0)$ and $\mathbf{\Theta}_2^{\mathcal{M}}(F_{Y|X}^0)$ sharply identify a structural feature $\theta(S)$, for $S \in \mathcal{M}$. Let $\mathbf{\Theta}^{\cap}$ be the intersection of the two. Then $\mathbf{\Theta}^{\cap}$ cannot be an identified set for $\theta(S)$: consider a value $a \in \mathbf{\Theta}_1^{\mathcal{M}}/\mathbf{\Theta}^{\cap}$. Since $\mathbf{\Theta}_1^{\mathcal{M}}$ is a sharp identified set, there exists a structure, S^a such that $a = \theta(S^a)$ and $S^a \in \mathcal{M} \cap \Omega_0$. Thus, not all values that are generated by a

[8] This terminology, "overidentification" in the partial identification context was used in Chesher (2005) and Bontemps, Magnac and Maurin (2007).

structure in $\mathcal{M} \cap \Omega_0$ are included in $\mathbf{\Theta}^{\cap}$. By definition 1, $\mathbf{\Theta}^{\cap}$ can*not* be an identified set. ■

When there are several sharp identified sets, the lemma implies that the intersection of those sharp sets may lose information contained in the model meaning that the intersection of sharp identified sets should not be considered to be an identified set because some parts excluded by taking intersection may be those points that are admitted by the model and observationally equivalent to the data.

2.6 Over-identification and Specification Tests under Set Identification

When there is overidentification, tests regarding the specification of the model can be conducted.

Examples

1. **Treatment Effects** : Manski (1990) discusses testability (refutability) of "level-set" assumption - constant treatment effect asssumption across different observable characteristics - by taking intersection of each identified interval and seeing if the intersection is empty.

2. **Overidentification Tests under Set Identification** : Bontemps, Magnac, and Maurin (2007) develop a Sargan-type specification test of overidentifying restriction in the case of overidentification when the parameter is partially identified.

3 *Refutability* of Structural Features under Set Identification

> "...particularly where the model is to a large degree speculative, empirical confirmation of the validity or usefulness of the model is obtained only to the extent that *observationally restrictive* specifications are upheld by the data....(Koopmans and Reiersol (1950) p180, emphasis added)"

If a model (or the restrictions imposed by a model) can be confirmed by data, this will validate the usefulness and credibility of the model. Some features of an economic model may not be identifiable in which case no amount of empirical information will answer the questions regarding the features of the underlying economic decision processes. To be able to use data as evidence for or against any hypothesis regarding the underlying structure, identifiability of the structural feature which is the object of the hypothesis is essential. The general rule described in the above quote from Koopmans and Reiersol (1950) and Breusch (1986)'s observation on "testability" are adopted.

Let $\mathcal{F}$ be a set of all restrictions and $\mathcal{F}^{YX}$ be a set of restrictions on the distribution of the observables. Then $\mathcal{F}$ would be partitioned into the two sets, $\mathcal{F}^{YX}$ and $\mathcal{F}/\mathcal{F}^{YX}$. Let R denote an element of the set $\mathcal{F}$. R can be a statement regarding either a structure or a distribution of observed variables.

Examples of restrictions on the structure can be monotone treatment response or monotone selection restriction in Manski (1997) and Manski and Pepper (2000). They can be regarding the functional form such as additive separability, linearity, or regardingt he distribution of the unobservables such as the mean or quantile. In most econometric models restrictions on the structure are not enough for identification.

Often restrictions on the distribution of the observables should be required. Examples of restrictions on the distribution of the observables are various types of rank conditions, or no multicollinearity condition or completeness conditions in Newey and Powell (2003) or Chernozhukov and Hansen (2005). Sometimes existence of a continuous variable plays a key role in identification - identification at infinity, special regressor in Lewbel (2000) (which can be a special case of Manski (1988) and Matzkin (1992)'s conditions for identification). In principle, any such restrictions on observables could be checked whether they are satisfied by the data. That is, $R \in \mathcal{F}^{YX}$ is "directly testable", while $R \in \mathcal{F}/\mathcal{F}^{YX}$ is testable if we can derive an equivalent expression for the restriction in terms of the observable $F^{S}_{Y|X}$. Any $R \in \mathcal{F}^{YX}$ is directly testable, while $R \in \mathcal{F}/\mathcal{F}^{YX}$ is **testable (confirmable)** if and only if $\exists R' \in \mathcal{F}^{YX}$ s.t. $R \Leftrightarrow R'$. Note that some of the restrictions have no testable implications on data.

For the discussion of testability of $R \in \mathcal{F}/\mathcal{F}^{YX}$ and how to interpret the test results, we adopt Breusch's (1986) framework. From now on we are concerned with restrictions on structural features which do not have any implications on the distribution of observables.

3.1 Breusch's (1986) Framework of "Testability"

We introduce Breusch's (1986) framework to determine "testability" of hypotheses on S. Let $\mathcal{H}_0$ be the set of structures that satisfy the null hypothesis. Then $\mathcal{S}$, a set of all structures, is partitioned into two, $\mathcal{H}_0$ and $\mathcal{S}/\mathcal{H}_0$. The testability of the hypothesis is a decision problem regarding whether the true structure S_0 that generates the data is included in $\mathcal{H}_0$ or not using the data.

Determining how to "interpret" the test results structurally requires further clarification of ideas. We adopt the "refutability" and "confirmability" from Breusch (1986) and define them as the following in a general setup. A hypothesis is refutable if, when the true structure, S_0, is not included in $\mathcal{H}_0$, every observationally equivalent structure to S_0 is also not included in $\mathcal{H}_0$.

Definition 3 A hypothesis is called refutable if $S_0 \notin \mathcal{H}_0 \Longrightarrow \nexists S \in \mathcal{H}_0 \; s.t. F^0_{Y|X} = F^S_{Y|X}$.

A hypothesis is confirmable if, when the true structure, S_0 is included in $\mathcal{H}_0$, every observationally equivalent structure to S_0 is also included in $\mathcal{H}_0$.

Definition 4 A hypothesis is called confirmable if $S_0 \in \mathcal{H}_0 \Longrightarrow \nexists S \notin \mathcal{H}_0$ $s.t. F^0_{Y|X} = F^S_{Y|X}$.

Discussion :

1. A hypothesis is **refutable** if when it is rejected, we can use the data as evidence against the null hypothesis and conclude that the hypothesis is not true. A hypothesis is **confirmable** if when it is accepted, we can conclude that the model (hypothesis) is true.

2. However, if a hypothesis is **refutable, but not confirmable**, then we cannot conclude that the hypothesis is true even though the hypothesis is not rejected.

When the restriction is regarding a structural feature, $\theta(S)$, testability/interpretability of $H_0 : \theta(S) = a$ depends on identifiability of $\theta(S)$.

3.2 When H_0 : $\theta(S) = a$

Consider a hypothesis regarding a structural feature of the form $H_0 : \theta(S) = a$. Let $\mathcal{H}_0$ be the set of structures that satisfy the null hypothesis which is of the form $H_0 : \theta(S) = a$, for some functional $\theta(\cdot)$. We write $\mathcal{H}_0 = \{S : \theta(S) = a, S \in \mathcal{S}\}$. Typically, $a = 0$, as in the example in Chapter 5, one may be interested in whether heterogeneous causal effects, $\theta(S) = h(1, u^*) - h(0, u^*)$ has any nontrivial effects.

We state Breusch (1986)'s results regarding point-identification and testability under the general setup. The proof is direct from the definition of identification of the structural feature.

Lemma 1 If $\theta(S)$ is point-identified by a model $\mathcal{M}$, then the hypothesis is refutable as well as confirmable.

Proof. Let $a \in A \equiv \{a : \theta(S) = a, S \in \mathcal{M}\}$. Suppose that $\theta(S)$ is point-identified by $\mathcal{M}$, that is, there exists a single-valued functional $\mathcal{G}(\cdot)$ such that $\theta(S) = \mathcal{G}(F^S_{Y|X})$ by lemma in Chesher (2007). To show "refutability" suppose $S_0 \notin \mathcal{H}_0$. This means that $a \neq \theta(S_0) = \mathcal{G}(F^0_{Y|X})$. Then for all observationally equivalent structures S to S_0, $a \neq \theta(S) = \mathcal{G}(F^S_{Y|X}) = \mathcal{G}(F^0_{Y|X})$. Thus, we conclude that there can be no structure included in $\mathcal{H}_0$ among observationally equivalent structures to S_0. Next, to show "confirmability", suppose $S_0 \in \mathcal{H}_0$, that is, $\theta(S_0) = a$. Since $\theta(S)$ is identified, $\theta(S_0) = \mathcal{G}(F^0_{Y|X})$. Then any observationally equivalent structure S to S_0 should be in $\mathcal{H}_0$ because $\theta(S_0) = \mathcal{G}(F^0_{Y|X}) = \mathcal{G}(F^S_{Y|X}) = \theta(S) = a$. ■

The lemma says whether $\theta(S)$ is identified by the model or not is the key determinant to "testability" of a hypothesis and remove the ambiguity in interpretation. When $\theta(S)$ is set-identified, the hypothesis is only refutable. This shows set identification causes certain ambiguity in interpretation in test of structural features.

Lemma 2 If $\theta(S)$ is set-identified by a model $\mathcal{M}$, then the hypothesis is refutable, but not necessarily confirmable.

Proof. Suppose that $\theta(S)$ is set-identified by $\mathcal{M}$, that is, there exists a set-valued mapping $\mathbf{\Theta}^{\mathcal{M}}$ such that $\theta(S) \in \mathbf{\Theta}^{\mathcal{M}}(F^0_{Y|X})$, for $\forall S \in \mathcal{M} \cap \Omega_0$**.** To show "refutability" suppose $S_0 \notin \mathcal{H}_0$. This means that $a \neq \theta(S_0) \in \mathbf{\Theta}^{\mathcal{M}}(F^0_{Y|X})$, i.e. $a \notin \mathbf{\Theta}^{\mathcal{M}}(F^0_{Y|X})$. Then for all observationally equivalent structures S to S_0, $\theta(S) \in \mathbf{\Theta}^{\mathcal{M}}(F^S_{Y|X}) = \mathbf{\Theta}^{\mathcal{M}}(F^0_{Y|X})$. Since $a \neq \theta(S_0) \in \mathbf{\Theta}^{\mathcal{M}}(F^0_{Y|X})$, $a \notin \mathbf{\Theta}^{\mathcal{M}}(F^S_{Y|X})$, for $\forall S \in \mathcal{M} \cap \Omega_0$. Thus, we conclude that there can be no structure included in $\mathcal{H}_0$ among observationally equivalent structures to S_0 in $\mathcal{M}$. Next, for confirmability, suppose $S_0 \in \mathcal{H}_0$, that is, $\theta(S_0) = a$. Since $\theta(S)$ is set-identified, $a = \theta(S_0) \in \mathbf{\Theta}^{\mathcal{M}}(F^0_{Y|X})$. Then for any observationally equivalent structure S to S_0, by the fact that $\theta(S)$ is identified, we have $\theta(S) \in \mathbf{\Theta}^{\mathcal{M}}(F^S_{Y|X}) = \mathbf{\Theta}^{\mathcal{M}}(F^0_{Y|X})$. From the fact that $a \in \mathbf{\Theta}^{\mathcal{M}}(F^0_{Y|X})$, we know that $a \in \mathbf{\Theta}^{\mathcal{M}}(F^S_{Y|X})$, for $\forall S \in \mathcal{M} \cap \Omega_0$, but this does not guarantee $a = \theta(S)$ for all observationally equivalent S to S_0. Thus, confirmability is not guaranteed. ■

4 Falsifiability of a Model

Econometric models characterized by restrictions are used to infer certain information on the true data generating structure, $\theta(S_0)$. Identification analysis assumes that S_0 is in the model, i.e. S_0 satisfies all the restrictions imposed by the model. Otherwise, the identified set by the model would not be informative on $\theta(S_0)$. In this section testability of whether the true structure actually lies in the model ($S_0 \in \mathcal{M} \cap \Omega_0$) is considered. This is a problem of deciding whether $S_0 \in \mathcal{M}$, or $S_0 \in \mathcal{M}^C$.This can be restated as $S_0 \in \mathcal{M} \cap \Omega_0$, or, $S_0 \in \mathcal{M}^C \cap \Omega_0$, since $S_0 \in \Omega_0$ by definition of Ω_0. See Figure 6. In this section one way of falsifying a model is discussed.

Let $\mathcal{M}^1$ be a set of structures that satisfy the set of restrictions R^1 and $\mathcal{M}$ be a set of structures that satisfy the set of restrictions $R^{\mathcal{M}}$. Then a model $\mathcal{M}'$ imposing restrictions R^1 and R can be written as

$$\mathcal{M}' = \mathcal{M} \cap \mathcal{M}^1 \qquad (**)$$

Let $\Psi^{\mathcal{M}}$ be a set of distribution functions of observables generated by the struc-

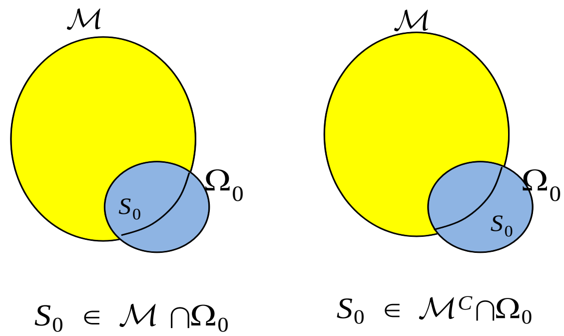


Figure 6: **Falsifiability of a model** : this is a problem of deciding whether $S_0 \in \mathcal{M}$, or $S_0 \in \mathcal{M}^C$.This can be restated as $S_0 \in \mathcal{M} \cap \Omega_0$, or, $S_0 \in \mathcal{M}^c \cap \Omega_0$, since $S_0 \in \Omega_0$ by definition of Ω_0.

tures in $\mathcal{M}$ and $\Psi^{\mathcal{M}'}$ be a set of distribution functions of observables generated by the structures in $\mathcal{M}'$.

Koopmans and Reiersol (1950) state that R^1 is "subject to test" if we can test $H_0 : F^0_{Y|X} \in \Psi^{\mathcal{M}'}$. From this we further develop the logic of testability of restrictions and discuss how identification results can be used in "falsifying" a model/restrictions.

Suppose that a model, $\mathcal{M}^1$, identifies a structural feature, $\theta(S)$, by a set $\mathbf{\Theta}^1(F^S_{Y|X})$, and another model, $\mathcal{M}^2$, identifies the same structural feature, $\theta(S)$, by $\mathbf{\Theta}^2(F^S_{Y|X})$. Recall that we define $\Psi^{\mathcal{M}^1} = \{F^S_{Y|X} : S \in \mathcal{M}^1\}$ and $\Psi^{\mathcal{M}^2} = \{F^S_{Y|X} : S \in \mathcal{M}^2\}$. Note that $\Psi^{\mathcal{M}^1 \cap \Omega_0} = \{F^0_{Y|X}\}$.

If $F^0_{Y|X} \notin \Psi^{\mathcal{M}}$, the model $\mathcal{M}$ should be falsified, since the true structure, S_0, that generates $F^0_{Y|X}$ cannot be in $\mathcal{M}$. Falsification of a model is not always possible.

The proofs of the following from set theory (see for example Pinter (1971)).

Lemma 3 If $\mathcal{M}^1 \subseteq \mathcal{M}^2$, then $\Psi^{\mathcal{M}^1} \subseteq \Psi^{\mathcal{M}^2}$.
Proof. Trivial by definition of $\Psi^{\mathcal{M}}$. ■

Lemma 4 Suppose $\Psi^{\mathcal{M}^1} \subseteq \Psi^{\mathcal{M}^2}$. Then $\mathbf{\Theta}^1(F^S_{Y|X}) \subseteq \mathbf{\Theta}^2(F^S_{Y|X})$, $\forall S \in \mathcal{M}^1 \cap \Omega_0$.
Proof. Consider an arbitrary $S^* \in \mathcal{M}^1 \cap \Omega_0$. Then by definition of $\Psi^{\mathcal{M}^1}$ and Ω_0, $F^{S^*}_{Y|X}(= F^0_{Y|X}) \in \Psi^{\mathcal{M}^1}$ and $\theta(S^*) \in \mathbf{\Theta}^1(F^0_{Y|X})$ from Definition 1. Suppose $\Psi^{\mathcal{M}^1} \subseteq \Psi^{\mathcal{M}^2}$. If $F^{S^*}_{Y|X} \in \Psi^{\mathcal{M}^1}$, then $F^{S^*}_{Y|X} \in \Psi^{\mathcal{M}^2}$. Then the definition of $\Psi^{\mathcal{M}^2}$ and set identification

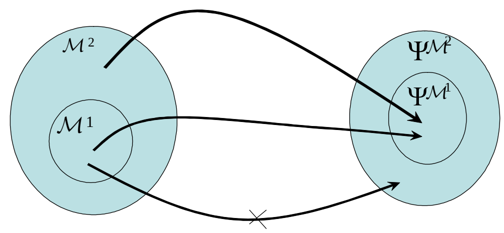


Figure 7: **Lemma 3**

(Definition 1) imply that $S^* \in \mathcal{M}^2 \cap \Omega_0$, and $\theta(S^*) \in \mathbf{\Theta}^2(F^0_{Y|X})$, leading to the conclusion that $\mathbf{\Theta}^1(F^0_{Y|X}) \subseteq \mathbf{\Theta}^2(F^0_{Y|X})$. ■

Theorem 1 If $\mathcal{M}^1 \subseteq \mathcal{M}^2$, then $\mathbf{\Theta}^1(F^S_{Y|X}) \subseteq \mathbf{\Theta}^2(F^S_{Y|X})$, $\forall S \in \mathcal{M}^1 \cap \Omega_0$.
Proof. The result follows from Lemma 3 and Lemma 4. ■

Theorem 1 is a natural and intuitive result. Consider the following examples.

Example 1 suppose that $\mathcal{M}^1$ imposes linearity with mean independence of the unobserved U, so that the structural relation admitted is of the form, $Y = X\beta + U$, and $\mathcal{M}^2$ admits the additively separable structural relation $Y = f(X) + U$,with mean independence of U. Then $\mathcal{M}^1 \subseteq \mathcal{M}^2$. If the true structure lies in $\mathcal{M}^1$, by **Theorem 1** for the structural feature of partial derivative, $\beta = f'(X)$, implying that $\mathbf{\Theta}^1(F^0_{Y|X}) = \mathbf{\Theta}^2(F^0_{Y|X})$ since both model point identifies the partial derivative. If the identified sets were different, the linearity restriction should be refuted.

Example 2 Consider Manski (1990), Manski's (1997) Monotone Treatment Response (MTR) model, and Manski and Pepper's (2000) Monotone Treatment Response and Monotone Treatment Selection (MTR-MTS) model. Let $\mathcal{M}^1, \mathcal{M}^2$, and $\mathcal{M}^3$ denote each model. Then $\mathcal{M}^1 \supseteq \mathcal{M}^2 \supseteq \mathcal{M}^3$. **Theorem 1** implies that if the true structure satisfies MTR-MTS restrictions, that is, the true structure lies in $\mathcal{M}^3$, we have

$$\mathbf{\Theta}^1(F^S_{Y|X}) \supseteq \mathbf{\Theta}^2(F^S_{Y|X}) \supseteq \mathbf{\Theta}^3(F^S_{Y|X}), \text{ for } \forall S \in \mathcal{M}^3 \cap \Omega_0$$

Note that both MTR and MTS are *not* "directly testable".

In **Theorem 1** at least one model - either $\mathcal{M}^1$ or $\mathcal{M}^2$ - is overidentifying. However,

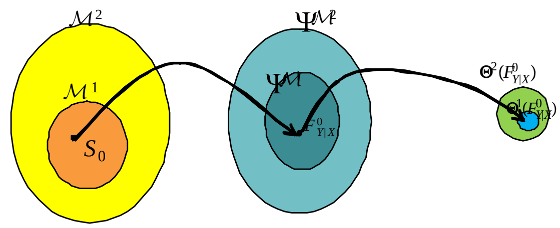


Figure 8: **Lemma 4**

existence of an overidentifying model is not required to falsify a model. As long as $\Psi^{\mathcal{M}^1} \subseteq \Psi^{\mathcal{M}^2}$ the criterion can be used to falsify $\mathcal{M}^1$ by **Lemma 4**. That is, although the two models are just-identifying, if $\Psi^{\mathcal{M}^1} \subseteq \Psi^{\mathcal{M}^2}$, we can falsify $\mathcal{M}^1$.

Corollary 1 Suppose $\mathcal{M}^1 \subseteq \mathcal{M}^2$. If $S \in \mathcal{M}^1 \cap \Omega_0$, $\mathbf{\Theta}^1(F^S_{Y|X}) \subseteq \mathbf{\Theta}^2(F^S_{Y|X})$.
Proof. Suppose $\mathcal{M}^1 \subseteq \mathcal{M}^2$ and $S \in \mathcal{M}^1 \cap \Omega_0$. Then $\mathbf{\Theta}^1(F^S_{Y|X}) \subseteq \mathbf{\Theta}^2(F^S_{Y|X})$ by Theorem 1 since for all $S \in \mathcal{M}^1 \cap \Omega_0$, if $\mathcal{M}^1 \subseteq \mathcal{M}^2$, then $\mathbf{\Theta}^1(F^S_{Y|X}) \subseteq \mathbf{\Theta}^2(F^S_{Y|X})$. ■

Corollary 2 Suppose $\mathcal{M}^1 \subseteq \mathcal{M}^2$. If $\mathbf{\Theta}^1(F^S_{Y|X}) \supset \mathbf{\Theta}^2(F^S_{Y|X})$, then $S \notin \mathcal{M}^1 \cap \Omega_0$.
Proof. This follows from Corollary 1 for the negation of $\mathbf{\Theta}^1(F^S_{Y|X}) \subseteq \mathbf{\Theta}^2(F^S_{Y|X})$. (Refer to, for example, Exercise 4(a) in Section 1.1 in Pinter (1971)). ■

4.1 Choice-based Models and Falisifiability of Rationality

4.1.1 Falsifiability of "Rationality"

There have been several attepts to test "rationality" using the revealed preference idea. Unless there exists any identification result of the model under the rationality, it is only "refutable" or falsifiable, rahter thab confirmable. That is, if data rejects the testable implications, then we can falsify the economic model. However, not rejecting the testable implications do not imply the economic model is true.

Blundell, Kristenssen and Matzkin (2010) Using the observational data it is not feasible to test "consistency" of individual choices since often data do not show much variation in relative prices. Therefore, instead of testing individual's rationality by using the axioms directly, Blundell, Browning and Crawford (2003,2008) and

Blundell, Kristenssen and Matzkin (2010) imposed GARP to find bounds on demand response for different income levels. Rationality is considered to be rejected if the unrestrictied expansion path is statistically different from the expantion path found by imposing rationality.

Choi, Fisman, Gale, and Kariv (2007) Choi, Fisman, Gale, and Kariv (2007) use experimental data that contain choice decisions for various sets of hypothetical price and income. With this data testing each individual's rationality is possible based on the axioms. Choi, Fisman, Gale, and Kariv (2007) found majority of individuals did not reject the rationality assumption. However, this does not confirm that individual's choices are rational. They fit the data to measure individual's degree of risk aversion under certain specification of the utility function and report heterogeneity in risk aversion.

Falsifiability of Rationality based on Manzini and Mariotti (2009) Let $\mathcal{M}^R$ and $\mathcal{M}^{CTC}$ be the choice models under rationality and Manzini and Mariotti (2009) each and let Ψ^R and Ψ^{CTC} be the sets of distribution of observables generated by the structures in $\mathcal{M}^R$ and $\mathcal{M}^{CTC}$ each. If $\Psi^R \subseteq \Psi^{CTC}$, then rationality may be falsifiable by adopting the idea in Blundell, Kristenssen and Matzkin (2010). The bounds under rationality should be smaller than the bounds by Manzini and Marrioti (2009) by Lemma 1. If $F^0_{Y|X} \in \Psi^{CTC}/\Psi^R$, then rationality would be falisified, and if $F^0_{Y|X} \in \Psi/\Psi^{CTC}$, then CTC would be falisified and this can be determined by comparing the bounds defined by two distinct models, $\mathbf{\Theta}^R(F^0_{Y|X})$ and $\mathbf{\Theta}^{CTC}(F^0_{Y|X})$. Note that $F^0_{Y|X} \in \Psi^R$ does not imply that rationality is *not* falsified due to the identification problem discussed in Manzini and Mariotti (2009) - consistent categorazation under CTC implies WARP.

5 Just-identifying Models and Falsifiabilty of Restrictions

We first define just-identification. A just-identifying model loses its identifying power if any of its restrictions is relaxed.

Definition 5 A model $\mathcal{M}$ characterized by a set of restrictions R^M **just-identifies** a structural feature $\theta(S)$ if $\nexists \mathcal{M}_1$ characterized by a set of restrictions R^1 s.t. (i) $R^M \supset R^1$ and (ii) $\theta(S) \in \mathbf{\Theta}^{\mathcal{M}_1}$, $\forall S \in \mathcal{M}_1 \cap \Omega_0$, where Ω_0 is defined as before.

The set of restrictions, R^M of a model $\mathcal{M}$ which is just-identifying is a minimal set that identifies the structural features.

If a model, $\mathcal{M}$, characterized by a set of restrictions R^M is **not** just-identifying, then there exists a less restrictive model $\mathcal{M}_1$ characterized by a set of restriction R^1 s.t. (i) $R^M \supset R^1$ and (ii) $\theta(S) = \theta(S_0)\ \forall S \in \mathcal{M}_1 \cap \Omega_0$.

Suppose two models, $\mathcal{M}_1$ and $\mathcal{M}_2$, with $\mathcal{M}_1 \neq \mathcal{M}_2$ are just-identifying for the same structural feature. Let the set of restrictions for $\mathcal{M}_1$ be R_1 and that for $\mathcal{M}_2$ be R_2. If $\Psi^{\mathcal{M}^1} \subset \Psi^{\mathcal{M}^2}$, then $\mathcal{M}^1$ can be falsified. In other words, if $\mathcal{M}^1$ is **observationally more restrictive**, and R_1/R_2 are **observationally relevant restrictions**. Then we have the following main result of this paper.

Definition 6 (Koopmans and Reiersol (1950)) $\mathcal{M}^1$ is called **observationally restrictive** if $\Psi^{\mathcal{M}^1} \subset \Psi^{\mathcal{M}^2}$.

Definition 7 R_1/R_2 is called **observationally relevant restrictions** if $\Psi^{\mathcal{M}^1} \subset \Psi^{\mathcal{M}^2}$.

Falsifiability of a model can be linked to refutabilty of restrictions under set identification since a model is characterized by restrictions. If a model is falsified, then some of the restrictions imposed by the model must not be the true description of the true underlying data generating structure. However, which restrictions among all the restrictions imposed by the model are not clear. This can be determined by the following proposition.

Proposition 1 If $\Psi^{\mathcal{M}_1} \subset \Psi^{\mathcal{M}_2}$, R_1/R_2, the observationally relevant restrictions, can be refuted.

Proof. The result follows from **Lemma 4**. ■

References

[1] Andrews, D., S. Berry and P. Jia (2004), "Confidence regions for parameters in discrete games with multiple equilibria, with an application to discount chain store," mimeo.

[2] Athey, S. and P. Haile (2005), "Nonparametric approaches to auctions," *Handbook of Econometrics.*

[3] Benkard, L. and S. Berry, (2006), "On the nonparametric identification of nonlinear simultaneous equations models : Comment on Brown (1983) and Roehrig (1988)," *Econometrica*, 74(5), 1429-1440.

[4] Beresteanu,A., I. Molchanov, and F. Molinari (2009), "Sharp identification regions in game," mimeo.

[5] Blundell, R., M. Browning, and I. Crawford (2003), "Nonparametric Engel curves and revealed preference," *Econometrica*, 71 (1), 205-240.

[6] Blundell, R., M. Browning, and I. Crawford (2008), "Best nonparametric bounds on demand responses," *Econometrica*, 76 (6), 1227-1262.

[7] Blundell, R., D. Kristenssen, and R. Matzkin (2010), "Stochastic demand and revealed preference," *mimeo.*

[8] Bontemps, C, T. Magnac, and E. Maurin (2007), "Set identified Linear models," mimeo.

[9] Bresnahan, R. and Reiss, P. (1991), "Empirical models of discrete games," *Journal of Econometrics*, 48, 57-81.

[10] Breusch (1986), "Hypothesis testing in unidentified models," *Review of Economic Studies*, 635-651.

[11] Chernozhukov, V. and C. Hansen (2005), "An IV Model of Quantile Treatment Effects," *Econometrica*, 73, 245-261.

[12] Chesher, A. (2005), "Nonparametric identification under discrete variation," *Econometrica*, 73(5), 1525-1550.

[13] Chesher, A. (2007), "Instrumental values," *Journal of Econometrics*, 139, 15-34.

[14] Chesher, A. (2010), "An instrumental variable model for discrete outcomes," *Econometrica*, forthcoming.

[15] Choi, S.,R. Fisman, D. Gale, and S. Kariv (2007), "Consistency and heterogeneity of individual behavior under uncertainty," *The American Economic Review*, 97 (5), 1921-1938.

[16] Fan, Y. and S. Park (2009), "Sharp bounds on the distribution of treatment effects and their statistical inderence," mimeo.

[17] Firpo. S. and G. Ridder (2008), "Bounds on functionals of the distribution of treatment effects," IEPR working paper 08.09.

[18] Galichon, A. and M. Henry (2009), "A test of non-identifying restrictions and confidence regions for partially identified parameters," *Journal of Econometrics*, 152(2), 186-196.

[19] Haavelmo, T. (1944), "The Probability Approach in Econometrics," *Econometrica*, 12.

[20] Heckman, J., J. (1990) "Variety of selection bias," *American Economics Review Papers and Proceedings,* 80, 313-318.

[21] Heckman, J., J. Smith, and N. Clements (1997), "Making the most out of program evaluations and social experiments accounting for heterogeneity in program impacts," *Review of Economic Studies*, 64, 487-535.

[22] Hurwicz, L. (1950) "Generalization of the concept of identification," in Statistical Inference in Dynamic Economic Models, Cowles Commision Monograph 10. Wiley, New York.

[23] Jovanovic, B. (1989), "Observable implications of models with multiple equilibria," *Econometrica*, 57. 1431-1437.

[24] Kadane , J. and T. Anderson (1977), "A comment on the test of overidentifying restrictions," *Econometrica*, 45. 1027-1031.

[25] Kitagawa, T. (2010), "Testing for Instrument Independence in the Selection Model," mimeo.

[26] Koopmans, T. (1949), "Identification Problems in Economic Model Construction," *Econometrica*, 17(2), 125-144

[27] Koopmans, T. and O. Reiersol (1950), "The identification of structural charactersitics," *Annals of Mathematical Statistics*, 21, 165- 181.

[28] Lee, J. (2010), "Sharp bounds on heterogeneous treatment responses," mimeo.

[29] Lewbel, A. (2000), "Semiparametric qualitative response model estimation with unknown heteroskedasticity or instrumental variables," *Journal of Econometrics*, 97, 145-177.

[30] Magnac, T. and E. Maurin (2007), "Identification and information in monotone binary models," *Journal of Econometrics*, 139, 76-104.

[31] Magnac, T. and E. Maurin (2008), "Partial identification in monotone binary models : discrete regressors and interval data," *Review of Economic Studies*, 75, 835-864.

[32] Manski, C. (1990), "Nonparametric bounds on treatment effects," *American Economic Review*, 65(6), 1311-1334.

[33] Manski, C. (1997), "Monotone Treatment Response," *Econometrica*, 65(6), 1311-1334.

[34] Manski, C. (2003), *Partial identification of probability distribution*, Springer-Verlag.

[35] Manski, C. and J. Pepper (2000), "Monotone instrumental variables : with an application to the returns to schooling," *Econometrica*, 68(4), 997-1010.

[36] Manzini, P and M. Mariotti (2009), "Categorization then choose : boundedly rational choice and welfare," mimeo.

[37] Matzkin, R. (1994), "Restrictions of Economic Theory in Nonparametric Methods," *Handbook of Econometrics.*

[38] Matzkin, R. (2007), "Nonparametric identification," *Handbook of Econometrics*, 5307-5368.

[39] Matzkin, R. (2008), "Identification of nonparametric simultaneous equations," *Econometrica*, 76, 945-978.

[40] Newey, W. and J. Powell (2003), "Instrumental variable estimation of nonparametric models," *Econometrica*, 71(5), 1565-1578.

[41] Pakes, A. (2010), "Alternative models for moment inequalities," mimeo.

[42] Pinter, C. (1971), *Set Thoory*, Addison-Wesley Publishing Company.

[43] Roehrig, C (1988), "Conditions for identification in nonparametric and parametric models," *Econometrica*, 56(2), 433-447.

[44] Tamer, E. (2003), "Incomplete simultaneous discrete response model with multiple equilibria," *Reveiw of Economic Studies,* 70, 147 - 165.