Article

The Status of Bridge Principles in Applied Econometrics

Bernt P. Stigum

Department of Economics, University of Oslo, Pb. 1095 Blindern, 0317 Oslo, Norway; bstigum@econ.uio.no; Tel.: +47-21998259

Abstract: The paper begins with a figurative representation of the contrast between present-day and formal applied econometrics. An explication of the status of bridge principles in applied econometrics follows. To illustrate the concepts used in the explication, the paper presents a simultaneous-equation model of the equilibrium configurations of a perfectly competitive commodity market. With artificially generated data I carry out two empirical analyses of such a market that contrast the prescriptions of formal econometrics in the tradition of Ragnar Frisch with the commands of present-day econometrics in the tradition of Trygve Haavelmo. At the end I demonstrate that the bridge principles I use in the formal-econometric analysis are valid in the Real World—that is in the world in which my data reside.

1. Introduction

Econometrics is a study of good and bad ways to measure economic relations. This paper is about the use of economic theory in such measurements, and about the need for bridge principles in applied econometrics.[1]

A long time ago in Fahlbeckska Stiftelsen's *Statsvetenskaplig Tidskrift* (1931) [2] (p. 281), Ragnar Frisch insisted (in my words) that economic statistics and parameter estimates tell us nothing about pertinent economic phenomena unless they are produced and analyzed within an economic-theoretic conceptualization. A few years later (cf. Haavelmo (1944) [3] (p. 1)), Trygve Haavelmo professed the same idea: "Theoretical models are necessary tools in our attempts to understand and 'explain' events in real life. In fact, even a simple description and classification of real phenomena would probably not be possible or feasible without viewing reality through the framework of some scheme conceived a priori." I grew up with this idea, and I took it to be valid not just in the paradigms of Frisch and Haavelmo's econometrics but in all of econometrics. Now I am not so sure. The younger generation of econometricians seems to think differently.

---

[1] The present paper is based on ideas that I developed and presented in chapters 1, 2, 3, and 10 of my 2015 book, *Econometrics in a Formal Science of Economics* [1]. It gives a novel explication of the status of bridge principles in applied econometrics, and develops necessary and sufficient conditions that their use in an empirical analysis be valid. A simultaneous-equation example serves to illustrate both the empirical relevance of the explication and the usefulness of the axiomatic method in theory-data confrontations. In writing the paper I have benefitted from questions and comments by Jørgen Aasnes, Leif Andreassen, Ådne Cappelen, Taran Fæhn, Roger Hammersland, Eilev Jansen, Tom Kornstad, and Terje Skjerpen in a seminar in Statistics Norway. I have also benefitted from discussions with Tony Hall, Edwin Leuven, and Timo Terasvirta. Finally, I have benefitted from interesting and constructive criticisms by two referees and from helpful comments by the Academic Editor of *Econometrics*.

1.1. Three Scenarios for Applied Econometrics

The younger generation of econometricians and the members of Frisch and Haavelmo's paradigms carry out the theory-data confrontation in different ways. Thus, one may envision three abstract scenarios for the empirical analysis of an economic theory. The scenarios differ in the way the researchers in charge view theory and use it in the measurement of economic relations.

In the first scenario the researcher in charge is an econometrician who, like Frisch, has one leg in mathematical economics and the other leg in mathematical statistics. The theory he confronts with data is an axiomatized theory about a few undefined terms that live and function in an imaginary model world. The researcher acknowledges that there is a definite divide between his model world and the world of observations, and that his data need not provide accurate measurements of the undefined terms of his theory.[2] He formulates his theory in a theory universe and uses bridge principles to describe how his theoretical variables are related to his data variables. The theory-data confrontation, then, becomes a triple of two disjoint universes—one for theory and one for data—and a bridge between them. This scenario I take to depict the way *formal econometrics* in the tradition of Frisch functions.

In the second scenario the researcher in charge is an econometrician who confronts an abstract theory with data. The theory may be axiomatized or formulated in accordance with the semantic view of theories. In Haavelmo's treatise the theory is taken to comprise a system of ordinary or functional equations that express definitional identities, technical relations, or relations depicting behavior (Haavelmo (1944) [3] (p. 2)). The researcher acknowledges the divide between his model world and the world in which his data resides. To overcome the divide, he identifies his theory variables with *true* variables in the real world, and he allows the possibility that he might not have accurate observations of all the relevant true variables.[3] Thus in the second scenario there is no need for a bridge between theory and data. The researcher formulates his theory, describes his data, and carries out the empirical analysis in a data universe. To me this scenario pictures the way *present-day econometrics* in the tradition of Haavelmo is carried out.

In the third scenario the researcher uses his knowledge of economic theory and the functioning of economic institutions to formulate his econometric model. The model concerns the behaviour of data variables and may account for both errors of observations and errors in equations. Then, the theory-data confrontation is carried out in a data universe, where the researcher formulates his econometric model and carries out the empirical analysis. It differs from the second scenario in that the model is not derived from the theorems of a pertinent economic theory but springs from the imagination of the researcher in charge. I think this must be a description of the way present-day econometrics is carried out by many young econometricians.

1.2. A Significant Contrast

The statistical analyses in the three scenarios differ, and the inferences about social reality that the respective researchers gain from their statistical results also differ. One way to describe the contrast between the empirical analyses in the three scenarios is as follows: The data variables constitute a vector-valued random process, $Y = \{y(t,\omega_P); t \ \varepsilon \ N\}$, on a probability space, $(\Omega_P, \aleph_P, P_P(\cdot))$, where $N = \{0, 1, 2, \ldots \}$, Ω_P is a subset of a vector space, $\aleph_P$ is a σ-field of subsets of Ω_P, and $P_P(\cdot):\aleph_P \to [0, 1]$ is a probability measure. The family of finite-dimensional probability distributions of the members of Y relative to $P_P(\cdot)$ is the *true probability distribution* of the data variables. I denote it by the acronym TPD in which T is short for true, P for probability, and D for distribution. In the

---

[2]   Frisch describes his model world in Bjerkholt and Qin (2010) [4] (p. 32).
[3]   In Haavelmo's Treatise [3] (p. 8) the *true* variables are variables in the real world with which a researcher identifies his theory variables.

last two scenarios the true values of the estimated parameters are taken to be values of pertinent TPD parameters.

The probability distribution that the researcher in the first scenario derives from the probability distribution of his theory variables and the bridge principles is, also, a family of finite-dimensional probability distributions of the members of Y. I denote it by the acronym MPD in which M is short for marginal, P for probability, and D for distribution. With an MPD I can associate a probability measure, $P_M(\cdot):\aleph_P \rightarrow [0, 1]$, such that Y has the probability distribution MPD relative to $P_M(\cdot)$, and such that Y can be thought of as a vector-valued random process on the probability space, $(\Omega_P, \aleph_P, P_M(\cdot))$. In a given theory-data confrontation there may be many MPD distributions. They vary with the models of the theory and the models of the bridge principles. Consequently, parameter estimates in the first scenario are estimates of the parameters of any one of the possible MPD distributions.

1.3. A Figurative Representation of the Contrast

The probability measures, $P_P(\cdot)$ and $P_M(\cdot)$, are probability measures on one and the same measure space, $(\Omega_P, \aleph_P)$. For pedagogical reasons it is useful to think of $(\Omega_P, \aleph_P, P_P(\cdot))$ and $(\Omega_P, \aleph_P, P_M(\cdot))$ as forming part of two different data universes—an upper data universe for $(\Omega_P, \aleph_P, P_P(\cdot))$ and a lower data universe for $(\Omega_P, \aleph_P, P_M(\cdot))$. The picture I have in mind looks roughly like Figure 1. Here, the bridge connects the theory universe with the lower data universe. The two data universes share the pair, $(\Omega_P, \aleph_P)$ and the random process, Y. The probability distribution of Y is TPD in the upper data universe and MPD in the lower data universe. I picture in my mind that present-day econometric analyses in accord with the second or third scenario take place in the upper data universe, while formal-econometric analyses in accord with the first scenario happen in the lower data universe.
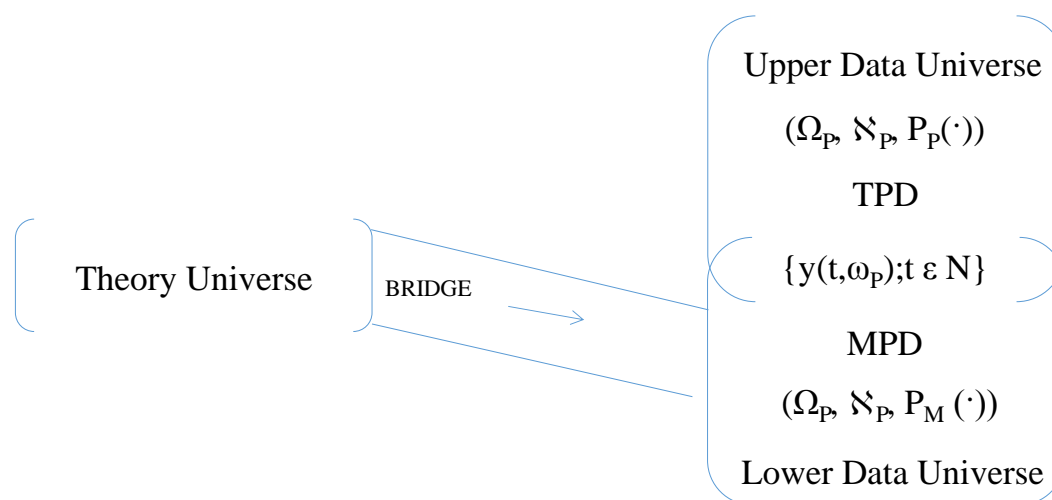

Figure 1. A Theory Universe and two Data Universes for an Empirical Contrast. TDP: true probability distribution; MDP: marginal probability distribution.

1.4. The Preferred Scenario

I believe that Frisch and Haavelmo were right in insisting that economic theory is a necessary tool in a researcher's attempts to understand and explain the economic phenomena that he observes in social reality. An economic theory is a family of models of a finite number of assertions—axioms—concerning a few undefined terms that live and function in a Frischian model world. As such the theory does not describe behavior and events in real life. Instead it delineates characteristic features of objects and relations in social reality that the originator of the theory has observed. The researcher is to determine when and where the given characteristics have empirical relevance.

The researchers in the three scenarios differ in the way they use theory in their empirical analysis. I have misgivings with the third scenario. The researcher's cavalier use of theory hampers his ability to give a well-founded analysis of his empirical results. Different theories lead to different interpretations of the results, and the researcher ought to reveal on which theory he bases his analysis. Without such information, meaningful discussions of the researcher's results become difficult.

I have misgivings with the second scenario, too. An economic theory about behavior or events is a family of models of a finite number of axioms. By identifying theoretical variables with true variables in the real world, and by using his assumptions about the true variables and the data variables to describe properties of the TPD, the researcher in effect treats his theory as if it has only one model. The theory, then, becomes a theory about the characteristics, say, of a representative individual or of a typical strike in the public sector of the economy. Such a theory restricts the kind of inference about social reality that a researcher's empirical results may confer.

I like the first scenario. In it the researcher's theory is about the characteristics of behavior and not about the characteristics of a representative individual's behavior. The statistical analysis confers information about how different the individuals may be for whom the theory is empirically relevant. Similarly, the researcher's theory about characteristics of events; e.g., strikes in the public sector, is about strikes and not about the typical strike. The statistical analysis confers information about how different the groups that conduct the strikes may be and for which the theory is empirically relevant. Such an empirical analysis is novel both in the way theory is used in the statistical analysis and in the way bridge principles are used to form the inference about social reality that the researcher can derive from his statistical results.

The first scenario is the one I prefer. It ought to appeal to applied econometricians once I have convinced them that there is a way to establish the validity of the bridge principles that are used in an empirical analysis. I hope to do that in Section 2 of the paper.

2. The Status of Bridge Principles in Applied Econometrics

The bridge principles play three roles in the empirical analysis in the lower data universe. In one role, the bridge principles translate the theory so that it becomes a statement about relations among variables in the data universe. In the second role, the bridge principles convert the probability distributions of variables in the theory universe into a family of probability distributions of variables in the data universe—the MPD—that plays the role of the data generating process in the empirical analysis. In the third role the bridge principles provide the means by which the researcher can identify the empirically relevant models of his theory that his statistical results determine. The three roles of the bridge principles in the lower data universe depict the way theory is incorporated in the empirical analysis of formal econometrics.

Since the MPD may be very different from the pertinent data generating process, the meaningfulness of the empirical analysis in the lower data universe hinges on the validity of the bridge principles. For that reason it is important to search for a general explication of the status of bridge principles in applied econometrics. I will set out to provide such an explication in this section of the paper.

To explicate the status of bridge principles in applied econometrics, I need to explain what I mean by several new concepts: empirical context, empirical relevance, encompassing, congruence, and a data admissible confidence region.

First, empirical context: In the upper data universe in Figure 1, *the empirical context* has two components: (1) axioms and loosely formulated assumptions that the pertinent researcher uses to delineate the characteristics of the data generating process; and (2) the prescriptions that underlie the statistical analysis. The second component details, for example, the way the researcher is to search for a data admissible econometric model that parsimoniously encompasses the so-called local data generating process (cf. Bontemps and Mizon (2003) [5] (p. 356,359,366) and Hendry and Krolzig (2003) [6] (p. 380,386–391)).

In the lower data universe in Figure 1, *the empirical context* has, also, two components: (1) the prescriptions that underlie the statistical analysis and (2) a description of the characteristics of a data admissible mathematical model of the MPD. The prescriptions that underlie the statistical analysis detail the way the researcher is to obtain a consistent estimate of a mathematical model of the MPD. Also, a mathematical model of MPD is *data admissible* if (1) the estimated mathematical model of MPD satisfies the strictures on which the axioms of the data universe and the prescriptions that underlie the statistical analysis insist; (2) the values that the given model assigns to the parameters of MPD satisfy the strictures on which the theory-axioms, the data-axioms, and the axioms of the bridge insist; and (3) the model lies within a 95% confidence region of the estimated mathematical model of MPD.

Next, empirical relevance: in the upper data universe in Figure 1 *the empirical relevance* of a theoretical hypothesis is determined the way the validity of a null hypothesis is decided in mathematical statistics. Specifically, the theoretical hypothesis is judged to be empirically relevant in a given empirical context if and only if it cannot be rejected. In contrast, in the lower data universe in Figure 1 a family of models of the theory axioms is *empirically relevant* if and only if there is a member of the given family which—together with a model of the data axioms and the bridge principles—induces a model of the MPD that belongs to a 95% confidence band of a meaningful statistical estimate of the MPD.

Then, three different concepts of encompassing: exact encompassing, encompassing in the upper data universe, and encompassing in the lower data universe.

Roughly speaking, an econometric model is said to *encompass* another econometric model if it can account for the results obtained by the latter (cf. Mizon (1984) [7] and Hendry and Richard (1989) [8]). My discussion of encompassing is designed to exhibit the characteristics of the three concepts of encompassing as simply as possible. I consider only cases of parametric encompassing. Also, I adopt the simplest version of the many binding functions that appear in articles on encompassing; e.g., in Gourieroux et al. (1983) [9], Mizon (1984) [7], and Florens et al. (1996) [10]. Finally, since I have had little to say about Bayesian econometrics in the paper, I do not discuss Bayesian ideas of encompassing here. For a comprehensive survey of the development of the encompassing principle in econometrics, the interested reader is referred to Bontemps and Mizon's (2008) [11] article on "Encompassing: Concepts and Implementation".

I will discuss the problem of encompassing within the confines of some given formal theory data confrontation as depicted in Figure 1. Hence, I am discussing cases in which there is one measurable space, $(\Omega_P, \aleph_P)$, with many different probability measures. I also presume that I am considering the probability distribution of a random vector, $y(\cdot)$, relative to the given probability measures. The distribution of $y(\cdot)$ relative to $P_P(\cdot)$ is the TPD distribution of $y(\cdot)$. Relative to a particular $P_M(\cdot)$ the distribution of $y(\cdot)$ is the MPD distribution corresponding to $P_M(\cdot)$. I will denote a vector of parameters of the TPD distribution of $y(\cdot)$ by θ_P and any one of the theoretically possible vectors of parameters of the MPD distribution of $y(\cdot)$ by θ_M. Also, I will denote an econometric model of MPD by a pair, $(M1, \theta_M{}^{\circ}(\cdot))$ and an econometric model of TPD by a pair, $(M2, \theta_P{}^{\circ}(\cdot))$, where, M1 and M2, and $\theta_M{}^{\circ}(\cdot)$, and $\theta_P{}^{\circ}(\cdot)$ are two probability distributions and two estimators of θ_M and θ_P, respectively. Finally, I will let s_n denote a finite sample, $y_1, \ldots, y_n$, of observations of $y(\cdot)$.

The fact that there is one true TPD and many "true" MPDs has one particularly important implication: The estimator in an econometric model of the TPD is an estimator of the "true" value of a pertinent TPD parameter. In contrast, the estimator in an econometric model of the MPD is an estimator of the value of an MPD parameter of any one of the possible mathematical models of the MPD. In different words: In an econometric model of the TPD, $(M, \theta_P{}^{\circ}(\cdot))$, the parameter of M, θ_P, has a true value. In an econometric model of the MPD, $(M, \theta_M{}^{\circ}(\cdot))$, the parameter of M, θ_M, may assume any one of a number of "true" values.

With these remarks in mind I can define the three concepts of encompassing formally as follows:

Definition 1. *Suppose that there are three subsets of R^k, Φ_P, Φ_M, and Φ_S, such that $\theta_P \ \varepsilon \ \Phi_P$, $\theta_M \ \varepsilon \ \Phi_M$, and $\theta_S \ \varepsilon \ \Phi_S$. Suppose, also, that I have obtained n independently and identically distributed observations of y, y_1, ... , y_n, and let $\theta_P{}^\circ(\cdot)$, $\theta_M{}^\circ(\cdot)$, and $\theta_S{}^\circ(\cdot)$, respectively, be estimators of θ_P, θ_M, and θ_S in the given sample. Finally, suppose that (M1, $\theta_S{}^\circ(\cdot)$) and (M2, $\theta_P{}^\circ(\cdot)$) are two different econometric models of the TPD, that (M3, $\theta_M{}^\circ(\cdot)$) is an econometric model of the MPD, and that the three estimators are consistent in their respective models. Then,*

A.   *(M1, $\theta_S{}^\circ(\cdot)$) exactly encompasses (M2, $\theta_P{}^\circ(\cdot)$) in the upper data universe in Figure 1 if and only if there is a binding function, $\Gamma(\cdot):\Phi_S \to \Phi_P$, such that in M1, the estimates of θ_P and θ_S satisfy the equation : $\theta_P{}^\circ(s_n) = \Gamma(\theta_S{}^\circ(s_n))$ a.e..*

B.   *(M1, $\theta_S{}^\circ(\cdot)$) encompasses (M2, $\theta_P{}^\circ(\cdot)$) in the upper data universe in Figure 1 if and only if there is a binding function, $\Gamma(\cdot):\Phi_S \to \Phi_P$, such that in M1, $\Gamma(\theta_{S0}) = plim \ \theta_P{}^\circ(s_n)$, and in M2, $\theta_{P0} = plim \ \theta_P{}^\circ(s_n)$, $\theta_{S0} = plim \theta_S{}^\circ(s_n)$, and $\theta_{P0} = \Gamma(\theta_{S0})$.*

C.   *(M3, $\theta_M{}^\circ(\cdot)$) encompasses (M2, $\theta_P{}^\circ(\cdot)$) in the lower data universe in Figure 1 if and only if there is a value of $\theta_M \ \varepsilon \ \Phi_M$ and a binding function, $\Gamma(\cdot):\Phi_M \to \Phi_P$, such that in M3, $\Gamma(\theta_M) = plim \ \theta_P{}^\circ(s_n)$, and in M2, $\theta_{P0} = plim \ \theta_P{}^\circ(s_n)$, $\theta_{M0} = plim \ \theta_M{}^\circ(s_n)$, and $\theta_{P0} = \Gamma(\theta_{M0})$.*

In reading this definition there are several things to notice. The definition of exact encompassing is in accord with Definition 1 in Bontemps and Mizon (2008) [11] (p. 6). Bontemps and Mizon attribute the definition to Florence et al. (1996) [10]. I have problems with using the definition in the lower data universe for several reasons. Firstly, there need not be a true value of θ_M that I can use to describe the distribution in M1. Secondly, there is no reference to the limiting distributions of the two estimates.

The definition of encompassing in the upper data universe, B, is a rewording of Bontemps and Mizon's (2003) definition of encompassing (cf. Bontems and Mizon, 2003 [5] (p. 359)). It is interesting in this context that the definition of encompassing in the lower data universe, C, is a natural extension of definition B. To see why, observe that in M3, $\theta_M = plim \ \theta_M{}^\circ(s_n)$. The required θ_M need not satisfy the equation $\theta_M = \theta_{M0}$. However, if $\theta_M = \theta_{M0}$, (M3, $\theta_M{}^\circ(\cdot)$) also encompasses (M2, $\theta_P{}^\circ(\cdot)$) in the sense of B. In that sense, C is the natural extension of B to the lower data universe in Figure 1.

At last I must say what the meanings of congruence and data admissible confidence sets are in the present context. The concept of congruence plays a pivotal role in the LSE methodology. According to Mizon (1995) [12], in the LSE methodology an econometric model is congruent if it is coherent with a priori theory, coherent with observed sample information, coherent with the measurement system, and encompasses all rival models. For present purposes I will insist on the following modified explication of congruence.

Definition 2. *Suppose that there are two subsets of R^k, Φ_P and Φ_M, such that $\theta_P \ \varepsilon \ \Phi_P$ and $\theta_M \ \varepsilon \ \Phi_M$. Suppose, also, that I have n independently and identically distributed observations of y, y_1, ... , y_n, and let $\theta_P{}^\circ(\cdot)$ and $\theta_M{}^\circ(\cdot)$ be estimators of θ_P and θ_M, respectively. Finally, suppose that (M2, $\theta_P{}^\circ(\cdot)$) and (M1, $\theta_M{}^\circ(\cdot)$) are different econometric models of the TPD and the MPD with M2 actually being the TPD, and suppose that the estimators, $\theta_P{}^\circ(\cdot)$ and $\theta_M{}^\circ(\cdot)$, are consistent in their respective models. Then, (M1, $\theta_M{}^\circ(\cdot)$) is a congruent econometric model of TPD if and only if M1 is coherent with the pertinent a priori theory and (M1, $\theta_M{}^\circ(\cdot)$) encompasses (M2, $\theta_P{}^\circ(\cdot)$) in accord with Definition 1C. The M1 in a congruent econometric model of TPD is termed a congruent model of TPD.*

In this definition the a priori theory is a family of models of the axioms of the data universe in a pertinent formal theory-data confrontation. It is, therefore, important to observe that the parameters of M1 in a congruent econometric model of TPD, (M1, $\theta_M{}^\circ(\cdot)$), need not satisfy the conditions of a data admissible mathematical model of the MPD.

Definition 3. *Consider a theory-data confrontation in which the sampling scheme that generates the data is adequate. The 95% confidence region around the estimated mathematical model of the MPD is data admissible if it contains both a data admissible mathematical model of the MPD and a congruent model of the TPD.*

Theorem 1. *Consider one subset of R^k, Φ_M, and suppose that θ_{M1} ε Φ_M, θ_{M2} ε Φ_M, and θ_{M3} ε Φ_M are, respectively, parameter vectors of three different models of MPD. Suppose, also, that I have n independently and identically distributed observations of y, $y_1, \ldots, y_n$, and let $\theta_M{}^\circ(\cdot)$ be an estimator of θ_{M1}, θ_{M2}, and θ_{M3}. Finally, suppose that $(M1, \theta_M{}^\circ(\cdot))$, $(M2, \theta_M{}^\circ(\cdot))$, and $(M3, \theta_M{}^\circ(\cdot))$ are econometric models of the MPD with θ_{M1} being the parameter of M1, θ_{M2} being the parameter of M2, and θ_{M3} being the parameter of M3, and suppose that the estimator, $\theta_M{}^\circ(\cdot)$, is consistent in the respective models. If M2 and M3 are data admissible, $(M2, \theta_M{}^\circ(\cdot))$ and $(M3, \theta_M{}^\circ(\cdot))$ encompass each other in accord with Definition 1C. Moreover, if $(M1, \theta_M{}^\circ(\cdot))$ is a congruent econometric model of TPD and M1 is a data admissible model of MPD, $(M1, \theta_M{}^\circ(\cdot))$ encompasses all the econometric models of the MPD, $(M, \theta_M{}^\circ(\cdot))$, with a data admissible mathematical model of MPD, M, and is encompassed by them. However, $(M2, \theta_M{}^\circ(\cdot))$ and $(M3, \theta_M{}^\circ(\cdot))$ need not be congruent econometric models of TPD.*

I will sketch a proof of the theorem and leave the missing details to the reader. First let Φ_P and Φ_M be subsets of R^k, and suppose that the parameter of TPD, θ_P, belongs to Φ_P, and that Φ_M, θ_{M1}, θ_{M2}, and θ_{M3} are as described in the theorem. Also, let $\theta_P{}^\circ(\cdot)$ be an estimator of θ_P, let $\theta_M{}^\circ(\cdot)$ be an estimator of the three θ_{Mi}, i = 1, 2, 3, and assume that the estimators are consistent in their respective models. Finally, let $(MP, \theta_P{}^\circ(\cdot))$ be an econometric model of the TPD. Then observe that if $(M1, \theta_M{}^\circ(\cdot))$ is a congruent econometric model of the TPD, there is a binding function, $\Gamma(\cdot):\Phi_M \to \Phi_P$, such that in M1 $\Gamma(\theta_{M1}) = \text{plim } \theta_P{}^\circ(s_n)$, and in MP, $\theta_{P0} = \text{plim } \theta_P{}^\circ(s_n)$, $\theta_{M0} = \text{plim } \theta_M{}^\circ(s_n)$, and $\theta_{P0} = \Gamma(\theta_{M0})$. Next, observe that if M2 and M3 are data admissible, in M2 $\theta_{M2} = \text{plim } \theta_M{}^\circ(s_n)$, and in M3, $\theta_{M3} = \text{plim } \theta_M{}^\circ(s_n)$. Hence with $\Gamma(\theta) = \theta$, $(M2, \theta_M{}^\circ(\cdot))$ encompasses $(M3, \theta_M{}^\circ(\cdot))$ in the sense of Definition 1C. Finally, observe that if $(M1, \theta_M{}^\circ(\cdot))$ is a congruent econometric model of TPD and M1 is data admissible, then $(M1, \theta_M{}^\circ(\cdot))$ encompasses both $(M2, \theta_M{}^\circ(\cdot))$ and $(M3, \theta_M{}^\circ(\cdot))$ in the sense of Definition 1C. However, even though the converse is true, neither $(M2, \theta_M{}^\circ(\cdot))$ nor $(M3, \theta_M{}^\circ(\cdot))$ need be congruent econometric models of TPD.

The preceding theorem allows me to explicate the status of bridge principles in applied econometrics as follows:

The Status of Bridge Principles in Applied Econometrics. *Consider a theory-data confrontation like the one pictured in Figure 1. The theory is a non-empty family of models of a finite number of assertions concerning a few undefined terms. The bridge comprises a non-empty family of models of a finite number of assertions—the bridge principles—that relate variables in the theory to variables in the data universe. The data universe is as described in a few assertions that delineate the characteristics of the data generating process—the TPD. Assume (1) that there is an empirically relevant model of the theory, and (2) that each data-admissible model of the MPD singles out an empirically relevant model of the theory. Then, the bridge principles are valid in the Real World if and only if all the econometric models of the MPD, $(M, \theta_M{}^\circ(\cdot))$, with a data-admissible M are congruent econometric models of the TPD.*[4]

I will exemplify the preceding definitions and the status of bridge principles in an empirical analysis of a simultaneous-equations model the idea of which I learned from reading Aris Spanos' (1989) [14] and (2012) [15] deliberations about the legacy of Trygve Haavelmo's (1944) [3] treatise "The Probability Approach in Econometrics".

---

[4] It is relevant here that the ideas underlying my explication of the status of bridge principles in applied econometrics are akin to David Hendry's idea that in model evaluation the requirement that an econometric model be congruent and encompassing can substitute for truth as a final decision criterion in empirical modeling (Hendry (1995) [13], Chapter 9).

3. A Simultaneous-Equation Example

The present simultaneous-equation example considers the equilibrium configurations of a perfectly competitive market for one commodity. I generate artificial data for such a market, describe the way an applied econometrician in present-day econometrics analyses the data, and contrast it with the way an applied econometrician in formal econometrics analyses the same data. The present-day econometrician in this example is taken to be a present-day econometrician in the tradition of Haavelmo.

3.1. Formal Aspects of the Present Theory-Data Confrontation

I begin by formulating the axioms that underlie the formal-econometric analysis of the given market. They are axioms that delineate the characteristics of the theory universe, the two data universes, the Bridge, and the MPD that I need for the two empirical analyses of my data.

3.1.1. Assumptions Concerning the Theory Universe

The theory universe is a triple, $(\Omega_T, \Gamma_T, (\Omega T, \aleph_T, P_T(\cdot)))$, where Ω_T is a subset of a vector space, Γ_T is a finite set of assertions concerning the properties of the vectors in Ω_T, and $(\Omega T, \aleph_T, P_T(\cdot))$ is a probability space, where $\aleph_T$ is a σ field of subsets of the given Ω_T, and $P_T(\cdot)):\aleph_T \to [0, 1]$ is a probability measure.

The assertions in Γ_T comprise six axioms, A1–A6.

A1 $\Omega_T \subset R^5 \times R^2$. Thus $\omega_T \in \Omega_T$ only if $\omega_T = (x, u)$ for some $x \in R^5$, $u \in R^2$, and $(x, u) \in R^5 \times R^2$.

A2 $(x_1, x_2, x_3) \in R_+^3$, $x_4 \in \{1, 2\}$, $x_5 \in \{1, 2\}$, $u_1 \in \{-1, 1\}$, and $u_2 \in \{-1, 1\}$.

In the intended interpretation of the components of x and u, x_1 and x_2 denote, respectively, the market participants' intended demand for the commodity and the market participants' intended supply of the commodity. Also, x_3 denotes the price of the commodity, x_4 and x_5 denote two auxiliary variables, and u_1, and u_2 are two error variables.

In an aggregate sense, the market participants' intended demand for the commodity is a linear function of its price and the auxiliary variable, x_4. Similarly, the market participants' intended supply of the commodity is a linear function of its price and the auxiliary variable, x_5. The values of the price variable and the auxiliary variables are such that the demand and supply of the commodity are always equal. Specifically,

A3 $x_1 = A + Bx_3 + Cx_4$; $x_2 = D + Ex_3 + Fx_5$; and $x_1 = x_2$.

The coefficients in the equations in A3 are taken to satisfy the conditions in A4.

A4 $A \in [8, 10]$, $B \in [-2, -1]$, $C = 1$, $D = 0$, $E \in [1, 2]$, and $F = 1$

Relative to $P_T(\cdot)$, the components of x and u are random variables. A5 and A6 bear witness to that.

A5 Let $x(\cdot):\Omega_T \to R^5$ and $u(\cdot):\Omega_T \to R^2$ be defined by the equations, $(x(\omega_T), u(\omega_T)) = \omega_T$ and $\omega_T \in \Omega_T$. The vector-valued function, $(x, u)(\cdot)$, is measurable with respect to $\aleph_T$ and has, subject to the conditions on which Γ_T insists, a well-defined probability distribution relative to $P_T(\cdot)$, the RPD, where R is short for researcher, P for probability, and D for distribution.[5]

---

[5] The Ω_T that appears in two places in the triple that represents the theory universe is one and the same subset of a vector space. Moreover, for the vector-valued function $(x, u)(\cdot)$, the equation $(x(\omega_T), u(\omega_T)) = \omega_T$ in axiom A5 is short for $(x_1(\omega_T), \ldots, x_h(\omega_T), u_1(\omega_T), \ldots, u_k(\omega_T)) = \omega_T$. In the present theory-data confrontation (x, u) plays the role of a vector in Ω_T and also the role of a vector-valued random variable on $(\Omega_T, \aleph_T)$. The components of (x, u) are theoretical variables and their values are not observable. Also, the probability measure, $P_T(\cdot)$, exists in the mind of the pertinent researcher and cannot, even ideally, be calculated by an outsider.

A6　Relative to $P_T(\cdot)$, the components of x have finite means and positive variances, and the covariance matrix of x_4, and x_5 is invertible. Also u_1, and u_2 have means zero, positive variances, and are distributed independently of x_4 and x_5.

In the axioms I insist on strict limits on the values of the constants; e.g., D = 0, C = 1, and B ∈ [−2, −1], to help me display the contrast between present-day econometrics and formal econometrics. I will, also, insist that a model of A1–A4 is a model of the axioms that fixes the values of the constants in A4 but leaves the ranges of values of x and u unrestricted as described in A2. I have specified the range of values of x_4, x_5, u_1, and u_2, but not their probability distribution. The RPD distribution of the four variables has many models. That is an interesting aspect of the example.

The three equations in A3 constitute the theory of the present simultaneous-equation example. I will refer to these equations as the structural equations of the model and to the constants as the structural parameters. Theorem 2 shows how the structural equations can be solved for x_1 and x_3 so that the two variables become linear functions of the two auxiliary variables, x_4 and x_5. I will refer to the resulting system of equations, (1), as the reduced form of the present simultaneous-equation example. The reduced form plays a pivotal role in my deliberations.

Theorem 2. *Suppose that the theory axioms are valid, and let*

$$\alpha = [(A - D)/(E - B)], \beta = [C/(E - B)], \gamma = -[F/(E - B)];$$
$$a = [(AE - BD)/(E - B)], b = [CE/(E - B)], \text{ and } c = -[BF/(E - B)].$$ (1)

Then, it is the case that $x_1 = a + bx_4 + cx_5$; and $x_3 = \alpha + \beta x_4 + \gamma x_5$.

From the preceding theorem it follows that in each model of A1–A4 the components of x can assume only four values. For example,

if A = 10, B = −2, C = 1, D = 0, E = 1, and F = 1, the four values of x are:

$$(15/3, 15/3, 9/3, 1, 2); (16/3, 16/3, 10/3, 2, 2);$$
$$(13/3, 13/3, 10/3, 1, 1); \text{ and } (14/3, 14/3, 11/3, 2, 1).$$ (2)

Similarly, if A = 8, B = −1, C = 1, D = 0, E = 2, and F = 1, the four values of x are:

$$(20/3, 20/3, 7/3, 1, 2); (22/3, 22/3, 8/3, 2, 2),$$
$$(19/3, 19/3, 8/3, 1, 1); \text{ and } (21/3, 21/3, 9/3, 2, 1).$$ (3)

Equations (2) and (3) give an indication of the way the RPD distribution of x varies with the models of Γ_T.

From Equations (2) and (3) it, also, follows that in each model of axioms A1–A4 the (x, u) vector, can assume sixteen values. For example, in the model of Equation (2)

$$(16/3, 16/3, 10/3, 2, 2, -1, -1); (16/3, 16/3, 10/3, 2, 2, -1, 1);$$

$$(16/3, 16/3, 10/3, 2, 2, 1, -1); \text{ and } (16/3, 16/3, 10/3, 2, 2, 1, 1)$$

are four of the 16 possible values of (x, u). In the model of Equation (3)

$$(21/3, 21/3, 9/3, 2, 1, -1, -1); (21/3, 21/3, 9/3, 2, 1, -1, 1);$$

$$(21/3, 21/3, 9/3, 2, 1, 1, -1) \text{ and } (21/3, 21/3, 9/3, 2, 1, 1, 1)$$

are four of the 16 possible values of (x, u). Since, $(x, u)(\omega_T) = \omega_T$, these vectors are at the same time possible values of ω_T.

According to Theorem 2, the values of the components of (A, B, C, D, E, F) determine the values of the components of (a, b, c, α, β, γ). It is, therefore, interesting that a model of the RPD distribution of x, also, determines the values of the components of (a, b, c, α, β, γ). The two sets of values of a, b, c, α, β, and γ must accord with each other.

Example 1: Consider the first model of Γ_T, where (A, B, C, D, E, F) = (10, −2, 1, 0, 1, 1). According to Theorem 2, the corresponding values of the components of (a, b, c, α, β, γ) are (10/3, 1/3, 2/3, 10/3, 1/3, −1/3). Suppose, now, that the RPD distribution of x_4 and x_5 is as described in the following equations:

$$P_T(\{x_4 = 1\}) = 0.25; P_T(\{x_4 = 2\}) = 0.75; P_T(\{x_5 = 1\}) = 2/3; P_T(\{x_5 = 2\}) = 1/3.$$

From these equations and Equations (2) and (3), we see that (cov. x_1,x_4, cov. x_1,x_5) = (0.0625, 0.14815), (cov. x_3,x_4, cov. x_3,x_5) = (0.0625, −0.07407), and (var.x_4, cov. x_4,x_5, var. x_5) = (0.1875, 0, 0.2222). Also, (mean (x_1, x_2, x_4, x_5)) = (4.8056, 3.4722, 1.75, 1.3333). Consequently, with G being the covariance matrix of x_4 and x_5, I can show that

$$(b, c) = (0.0625, 0.14815)G^{-1} = (0.3334, 0.6668),$$

$$(\beta, \gamma) = (0.0625, -0.07407)G^{-1} = (0.3334, -0.3334),$$

$$a = 4.8056 - 0.3334 \times 1.75 - 0.6668 \times 1.3333 = 3.3331; \text{ and}$$

$$\alpha = 3.4722 - 0.3334 \times 1.75 + 0.3334 \times 1.3333 = 3.3333.$$

In this case the two sets of values of a, b, c, α , β, and γ are equal.[6]

3.1.2. Assumptions Concerning the Two Data Universes

There is a basic data universe in which the data for the present analysis were generated. The basic data universe is a triple, (Ω_P, Γ_P, (Ω_P, $\aleph_P$, $P_P(\cdot)$),where Ω_P is a subset of a vector space, Γ_P is a finite set of assertions concerning the properties of the vectors in Ω_P, and (Ω_P, $\aleph_P$, $P_P(\cdot)$)) is a probability space, where $\aleph_P$ is a σ field of subsets of the given Ω_P, and $P_P(\cdot)$):$\aleph_P \to [0, 1]$ is a probability measure.

The assertions in Γ_P comprise three axioms, D1–D3.

D1　$\Omega_P \subset R^4$. Thus, $\omega_P \in \Omega_P$ only if $\omega_P = y$ for some $y \in R^4$.

The four components of the vectors in the data universe, y_1, y_2, y_3, and y_4, are taken to record, respectively, the actual sales and purchases of the given commodity, the price at which the sales were carried out, and the actual values of the two auxiliary variables that we met in the theory universe.

Relative to $P_P(\cdot)$, the components of y are random variables. D2 and D3 attest to that.

D2　Let $y(\cdot):\Omega_P \to R^4$ be defined by the equations, $y(\omega_P) = \omega_P$ and $\omega_P \in \Omega_P$. The vector-valued function, $y(\cdot)$, is measurable with respect to $\aleph_P$ and has, subject to the conditions on which Γ_P insists, a well-defined probability distribution, the TPD.[7]

D3　Relative to $P_P(\cdot)$, $y(\cdot)$ has finite means and a finite covariance matrix with positive diagonal elements. Finally, the covariance matrix of y_3 and y_4 is invertible.

---

[6]　I discuss the consistency conditions that the constants in A4 and the RPD distribution must satisfy in Stigum (2015) [1] (pp. 54–55).

[7]　The Ω_P that appears in two places in the triple that represents the basic data universe is one and the same subset of a vector space. Moreover, the equation $y(\omega_P) = \omega_P$ in axiom D2 is short for $(y_1(\omega_P), \ldots ,y_k(\omega_P)) = \omega_P$. In the present theory-data confrontation y plays two roles: that of a vector in Ω_P and that of a vector-valued random variable on (Ω_P, $\aleph_P$). The components of y can be observed, and the values that $P_P(\cdot)$ assigns to the members of $\aleph_P$ can (ideally, at least) be estimated.

Here, it is important to observe that, in the intended interpretation of the assumptions about this basic data universe, the TPD is taken to have one "true" model. The researcher does not know the "true" model of TPD.

In a formal theory-data confrontation the axioms of the basic data universe, Γ_P, play two roles—one as part of the empirical analysis in the lower data universe of Figure 1 and another as the basic axioms for the empirical analysis in the upper data universe. For the upper data universe these basic axioms delineate an identifiable and low-level-theory-consistent structure of the empirical analysis. The low-level theory—here the Γ_P—is taken to "provide a framework within which the high-level theory can be tested" (Bontemps and Mizon (2003) [5] (p. 365)).

In the present case, the researcher in the upper data universe accepts the Γ_P axioms of the basic data universe, and observes that they imply that there exists a pair of random variables, ξ_1 and ξ_2, with means 0 and positive variances such that y satisfies two linear equations,

$$y_1(\omega_P) = d + ey_3(\omega_P) + fy_4(\omega_P) + \xi_1(\omega_P); \text{ and} \tag{4}$$

$$y_2(\omega_P) = \lambda + \pi y_3(\omega_P) + \varrho y_4(\omega_P) + \xi_2(\omega_P). \tag{5}$$

The researcher, then, adds to the assertions in Γ_P the following three assumptions about the TPD distribution of $y(\cdot)$:

(1) The observations of y need not be accurate observations of the "true" variables with which the researcher identifies his theory. The "true" values of the constants in Equations (4) and (5) are such that the "true" components of $y(\cdot)$ satisfy the two equations without error terms.
(2) There are constants, A, B, C, D, E, and F whose values satisfy the conditions in A4 and the equations in Theorem 2 with the "true" TPD values of d, e, f, λ, π, and ϱ in place of a, b, c, α, β, and γ.
(3) The "true" components of y satisfy the relations:

$$y_1 = A + By_2 + Cy_3; \text{ and } y_1 = D + Ey_2 + Fy_4$$

The assumptions that the researcher in the upper data universe adds to Γ_P are assumptions about the characteristics of the TPD. Consequently, they have one "true" model. That is a fact that plays an important role in the empirical analysis of present-day econometrics.

3.1.3. Assumptions Concerning the Bridge and the MPD

The Bridge is a pair, $(\Omega, \Gamma_{T,P})$, where Ω is a subset of $\Omega_T \times \Omega_P$, and $\Gamma_{T,P}$ is a finite set of assertions about the vectors in Ω. It is understood that a researcher's observations consist of pairs, (ω_T, ω_P), where $\omega_T \in \Omega_T$, $\omega_P \in \Omega_P$, and $(\omega_T, \omega_P) \in \Omega$. The components of ω_T are unobservable, while the components of ω_P are observable. For example, in the present example two of the components of ω_T may record the intended demand and intended supply of the commodity in a pertinent market while one of the components of ω_P will record the actual sales and purchases of the commodity in the same market.

The assertions in $\Gamma_{T,P}$ comprise five axioms, B1–B5.

B1 $\Omega \subset \Omega_T \times \Omega_P$. Thus $\omega \in \Omega$ only if $\omega = (\omega_T, \omega_P)$ for some $\omega_T \in \Omega_T$, $\omega_P \in \Omega_P$, and $(\omega_T, \omega_P) \in \Omega_T \times \Omega$; i.e., $\omega \in \Omega$ only if $\omega = ((x,u), y)$ for some $(x,u) \in \Omega_T$, $y \in \Omega_P$, and $((x,u), y) \in \Omega_T \times \Omega_P$.

B2 Ω_T and Ω_P are disjoint and $\aleph_T$ and $\aleph_P$ are stochastically independent.[8]

---

[8] When $\aleph_T$ and $\aleph_P$ are stochastically independent, the probability spaces in the theory universe and the data universe induce a uniquely determined probability measure on $\Omega_T \times \Omega_P$. To see how, let $\aleph$ denote the family of all sets in $\Omega_T \times \Omega_P$ of the form $E_T \times E_P$ with $E_T \in \aleph_T$ and $E_P \in \aleph_P$, and let $\aleph$ denote the smallest σ field in $\Omega_T \times \Omega_P$ containing $\aleph$. There is a uniquely determined probability measure, $P(\cdot):\aleph \to [0,1]$, such that for all $E \in \aleph$, $P(E) = P_T(E_T)P_P(E_P)$ (Dunford and Schwartz (1957) [16] (pp. 183–189)).

B3　In the probability space, $(\Omega_T \times \Omega_P, \aleph, P(\cdot))$, that the probability spaces in the theory universe and the data universe determine, $\Omega \in \aleph$, and $P(\Omega) > 0$.

B4　$\Omega_T \subset \{(x, u) \in \Omega_T$ for which there is a $y \in \Omega_P$ such that $((x, u), y) \in \Omega\}$.

B5　For all $(\omega_T, \omega_P) \in \Omega$,

$$y_1(\omega_P) = x_1(\omega_T) + u_1(\omega_T); y_2(\omega_P) = x_3(\omega_T) + u_2(\omega_T);$$

$$y_3(\omega_P) = x_4(\omega_T); \text{ and } y_4(\omega_P) = x_5(\omega_T).$$

B5 is an interesting axiom that claims several things that it is important to keep in mind. First, if I substitute the reduced form equations in (1) for x_1 and x_3, the first two equations in B5 insist that

$$y_1(\omega_P) = a + bx_4(\omega_T) + cx_5(\omega_T) + u_1(\omega_T); \text{ and} \tag{6}$$

$$y_2(\omega_P) = \alpha + \beta x_4(\omega_T) + \gamma x_5(\omega_T) + u_2(\omega_T). \tag{7}$$

Next, the last two equations in B5 can be read to insist that on a bridge y_3 and y_4 can assume only the values 1 and 2.

A superficial look at axioms B4 and B5 gives the impression that the bridge has only one model. That is not true. So an example is called for. In reading the example, keep in mind that $x_i(\omega_T) = \omega_{Ti}$, $i = 1, \ldots, 5$, $u_j(\omega_T) = \omega_{T5+j}$, $j = 1, 2$.

Example 2: Consider the two models of the A1–A4 axioms, where (1) A = 10, B = -2, C = 1, D = 0, E = 1, and F = 1; and (2) A = 8, B = -1, C = 1, D = 0, E = 2, and F = 1. I will record two vectors in the Ω of each model—see Table 1—and describe the possible values of y_1 and y_2 in the support of the MPD distribution of $y(\cdot)$ in each model. First, the vectors in the two Ωs:

Table 1. Pairs of vectors in two different Bridges.

	x_1	x_2	x_3	x_4	x_5	u_1	u_2	y_1	y_2	y_3	y_4
(1)	15/3	15/3	9/3	1	2	-1	-1	4	2	1	2
	15/3	15/3	9/3	1	2	1	-1	6	2	1	2
(2)	20/3	20/3	7/3	1	2	-1	-1	17/3	4/3	1	2
	20/3	20/3	7/3	1	2	1	-1	23/3	4/3	1	2

Then, the possible values of y_1 and y_2 that in the respective models are members of a pair $(\omega_T, \omega_P) \in \Omega$:

$$y_1 \in \{10/3, 11/3, 12/3, 13/3, 16/3, 17/3, 18/3, 19/3\}$$
$$y_2 \in \{6/3, 7/3, 8/3, 12/3, 13/3, 14/3\}$$
$$y_1 \in \{16/3, 17/3, 18/3, 19/3, 22/3, 23/3, 24/3, 25/3\}$$
$$y_2 \in \{4/3, 5/3, 6/3, 10/3, 11/3, 12/3\}$$

The preceding example demonstrates that a bridge comprises no more than sixteen different vectors. Also, each model of A1–A4 determines one well defined bridge, and different models of A1–A4 determine different bridges. Finally, the relation between a model of A1–A4 and the bridge which it determines is independent of the RPD.

3.1.4. The MPD Distribution of $y(\cdot)$

The probability distribution of the data variables that is derived from the RPD and the bridge principles I have designated by the acronym MPD, where M is short for marginal, P stands for

probability, and D is short for distribution. I refer to it as the marginal probability distribution of the data variables.[9]

Formally, the MPD is the probability distribution of the components of ω_P that is induced by the bridge principles and the probability distribution of the components of ω_T that $P_T(\cdot)$ determines. In the present context, MPD is the probability distribution of $y(\cdot)$ induced by the RPD distribution of $(x,u)(\cdot)$ and the bridge principles.

To describe the MPD distribution of $y(\cdot)$ I need to introduce two new variables:

$$\text{For all } (\omega_T, \omega_P) \, \epsilon \, \Omega, \, v_1(\omega_P) = u_1(\omega_T), \text{ and } v_2(\omega_P) = u_2(\omega_T).$$

It follows from the bridge principles and from Theorem 2 that $v_1(\cdot)$ and $v_2(\cdot)$ are well defined in a given Ω, and satisfy the Equations in (8) and (9) in Theorem 3.

Theorem 3. *Suppose that the theory assumptions and the bridge principles are valid. Then it is the case that, for all ω_P in the support of a given MPD,*

$$y_1(\omega_P) = a + by_3(\omega_P) + cy_4(\omega_P) + v_1(\omega_P); \text{ and} \tag{8}$$

$$y_2(\omega_P) = \alpha + \beta y_3(\omega_P) + \gamma y_4(\omega_P) + v_2(\omega_P) \tag{9}$$

The two vs obviously vary with Ω. Thus in the first model of the theory universe, $v_1(\omega_P) = 1$ in the subset of Ω_P, in which, $y_1(\omega_P) \, \epsilon \, \{16/3, 17/3, 18/3, 19/3\}$; $v_1(\omega_P) = -1$ in the subset of Ω_P, in which, $y_1(\omega_P) \, \epsilon \, \{10/3, 11/3, 12/3, 13/3\}$; $v_2(\omega_P) = 1$ in the subset of Ω_P, in which, $y_2(\omega_P) \, \epsilon \, \{12/3, 13/3, 14/3\}$; and $v_2(\omega_P) = -1$ in the subset of Ω_P, in which, $y_2(\omega_P) \, \epsilon \, \{6/3, 7/3, 8/3\}$.

In the second model of the theory universe, $v_1(\omega_P) = 1$ in the subset of Ω_P, in which, $y_1(\omega_P) \, \epsilon \, \{22/3, 23/3, 24/3, 25/3\}$; $v_1(\omega_P) = -1$ in the subset of Ω_P, in which, $y_1(\omega_P) \, \epsilon \, \{16/3, 17/3, 18/3, 19/3\}$; $v_2(\omega_P) = 1$ in the subset of Ω_P, in which, $y_2(\omega_P) \, \epsilon \, \{10/3, 11/3, 12/3\}$; and $v_2(\omega_P) = -1$ in the subset of Ω_P, in which, $y_2(\omega_P) \, \epsilon \, \{4/3, 5/3, 6/3\}$.

For the interpretation of the coming empirical analysis I need two theorems. The validity of the first theorem is an immediate consequence of the definition of the vs and the definition of the MPD that the statement of the theorem gives. The validity of the second theorem follows from the first and from Kolmogorov's consistency theorem (cf. Stigum (1990) [17] (pp. 346–347)).

Theorem 4. *Suppose that the axioms of the theory universe and the bridge are valid, and let MPD($\cdot$) be an MPD that a model of the theory universe and the bridge determine. Then it is a fact that for any $A \subset R^2$ and $B \subset R^2$,*

$$MPD(\{(y_3, y_4, v_1, v_2)(\omega_P) \, \epsilon \, A \times B\}) =$$
$$\frac{P_T(\{\omega_T \, \epsilon \, \Omega_T : (x_4, x_5, u_1, u_2)(\omega_T) \epsilon \, A \times B\} \cap \{\omega_T \, \epsilon \, \Omega_T \text{ with an } \omega_P \, \epsilon \, \Omega_P \text{ and } (\omega_T, \omega_P) \, \epsilon \, \Omega\})}{P_T(\{\omega_T \, \epsilon \, \Omega_T \text{ with an } \omega_P \, \epsilon \, \Omega_P \text{ and } (\omega_T, \omega_P) \, \epsilon \, \Omega\})}$$

which, since $\Omega_T \subset \{\omega_T \in \Omega_T$ for which there is an $\omega_P \in \Omega_P$ with $(\omega_T, \omega_P) \in \Omega\}$,

$$= P_T(\{(x_4, x_5, u_1, u_2)(\omega_T) \, \epsilon \, A \times B\}).$$

From this theorem and the bridge principles it follows that the MPD distribution of (y_3, y_4, v_1, v_2) equals the RPD distribution of (x_4, x_5, u_1, u_2). This fact and Equations (6)–(9) show how the RPD distribution of (x, u) and the bridge principles determine the MPD distribution of y.

---

[9]   I describe properties of the MPD and its relationship to the conditional probability measure on Ω, $P(\cdot \mid \Omega)$, in Stigum (2015) [1] (pp. 47–52).

Theorem 5. *Suppose that the theory axioms and the axioms of the Bridge are valid. For each model of the theory universe and the Bridge there exists a probability measure $P_M(\cdot){:}\aleph_P \to [0, 1]$ and two random variables v_1 and v_2 with two important properties. Relative to $P_M(\cdot)$, (y, v_1, v_2) has the MPD distribution which the model of the theory universe and the Bridge determine. Also, relative to $P_M(\cdot)$, v_1 and v_2 have the same means and variances as u_1 and u_2, are distributed independently of y_3 an y_4, and satisfy the Equations (10) and (11) for all $\omega_P \epsilon \Omega_P$; i.e., for all $\omega_P \epsilon \Omega_P$,*

$$y_1(\omega_P) = a + by_3(\omega_P) + cy_4(\omega_P) + v_1(\omega_P); \text{ and} \tag{10}$$

$$y_2(\omega_P) = \alpha + \beta y_3(\omega_P) + \gamma y_4(\omega_P) + v_2(\omega_P). \tag{11}$$

In Equations (10) and (11) the values of $(a, b, c, \alpha, \beta, \gamma)$ are the values that the pertinent model of the A1–A4 axioms determines. Also, in the support of the pertinent $P_M(\cdot)$ the values of the v_1 and v_2 in Equations (10) and (11) coincide with the values of the v_1 and v_2 in Equations (8) and (9). Finally, in the support of the pertinent $P_M(\cdot)$ y_3 and y_4 can assume only the values 1 and 2.

The last two theorems call for several clarifying remarks. Note first that I derived Equations (8) and (9) from A1–A4 and the bridge principles without involving the RPD distribution of (x, u). The RPD distribution of (x_4, x_5, u_1, u_2) is independent of the models of A1–A4. Hence, for each model of A1–A4 and the bridge principles there are as many MPD distributions of (y_3, y_4, v_1, v_2) as there are RPD distributions of (x_4, x_5, u_1, u_2). Also, for each model of A1–A4 there are as many MPD distributions of (y, v_1, v_2) as there are MPD distributions of (y_3, y_4, v_1, v_2). Finally, even if there is just one RPD distribution of (x_4, x_5, u), there are many MPD distributions of (y, v_1, v_2) since they vary with the models of A1–A4.

4. The Two Empirical Analyses

For the two empirical analyses I am to generate 1249 "observations" of $((x, u), y)$. The components of (x, u) are unobservable. Hence only the sample of ys will be observed. To generate the sample of ys I proceed as follows. I presume that the components of y in the TPD distribution satisfy the equations,

$$y_1 = 16/3 + 2/3y_3 + 1/3y_4 + \xi_1; \text{ and } y_2 = 8/3 + 1/3y_3 - 1/3y_4 + \xi_2 \tag{12}$$

I presume, also, that y_3, y_4, ξ_1, and ξ_2 are distributed independently of each other, and that their distribution is given by the equations,

$$P_P(\{y_3 = 1\}) = 0.25; P_P(\{y_3 = 2\}) = 0.75; P_P(\{y_4 = 1\}) = 2/3; P_P(\{y_4 = 2\}) = 1/3; \text{ and}$$

$$P_P(\{\xi_i = -1\}) = P_P(\{\xi_i = 1\}) = 0.5, i = 1,2.$$

From these equations and the equations in (12) it follows that the joint probability distribution of y_1 and y_2 is as described in Table 2. But if that is so, then the means and co-variances of the components of y satisfy the equations:

$$\text{Mean}(y_1) = 6.94; \text{Mean}(y_2) = 2.806; \text{Mean}(y_3) = 7/4; \text{Mean}(y_4) = 4/3;$$

$$\text{Variance}(y_1) = 1.108; \text{Variance}(y_2) = 1.046; \text{Variance}(y_3) = 0.188; \text{Variance}(y_4) = 0.2222;$$

$$\text{Covariance}(y_1,y_2) = 0; \text{ and Covariance}(y_3,y_4) = 0.$$

The corresponding values of ξ_1, and ξ_2 are

$$\text{Mean}(\xi_1) = \text{Mean}(\xi_2) = 0: \text{Variance}(\xi_1) = \text{Variance}(\xi_2) = 1; \text{ and Covariance}(\xi_1,\xi_2) = 0.$$

Table 2. The Joint Probability Distribution of y1 and y2.

y2 \\ y1	16/3	17/3	18/3	19/3	22/3	23/3	24/3	25/3	Py2
4/3	2/576	1/576	6/576	3/576	2/576	1/576	6/576	3/576	1/24
5/3	10/576	5/576	30/576	15/576	10/576	5/576	30/576	15/576	5/24
6/3	12/576	6/576	36/576	18/576	12/576	6/576	36/576	18/576	6/24
10/3	2/576	1/576	6/576	3/576	2/576	1/576	6/576	3/576	1/24
11/3	10/576	5/576	30/576	15/576	10/576	5/576	30/576	15/576	5/24
12/3	12/576	6/576	36/576	18/576	12/576	6/576	36/576	18/576	6/24
Py1	2/24	1/24	6/24	3/24	2/24	1/24	6/24	3/24	1

With these probabilities and moments in mind I use the PcGive's uniform random number generator, ranu(), and PcGive to produce the required data and to carry out the statistical analysis (cf. Dornik (2009) [18] (pp. 158–162) and Dornik and Hendry (2009) [19]). Specifically, I generate 1249 observations of y_1, y_2, y_3, y_4, ξ_1, and ξ_2 with OxMetrics' random number generator for the uniform distribution, ranu(). The algebra codes in Appendix A detail how the observations were generated. In the printout, $y_1 = y6$, $y_2 = y7$, $y_3 = y8$, $y_4 = y9$, $\xi_1 = u6$, and $\xi_2 = u7$. The results of the statistical analysis are displayed in Table 3, Figures 2 and 3, and Tables A1 and A2 in Appendix A. They demonstrate that the probability distribution of (y, ξ) described above is statistically adequate in the sense that it accounts for the chance regularities in the data (cf. Spanos (2012) [15] p. 14).

Table 3. Means, Standard Deviations, and Correlations of y and u.

Means						
	y6	y7	y8	y9	u6	u7
	6.9933	2.7537	1.7406	1.3379	0.053643	−0.047238

Standard Deviations (Using T − 1)						
	y6	y7	y8	y9	u6	u7
	1.0569	1.0273	0.43849	0.47317	0.99896	0.99928

Correlation Matrix						
	y6	y7	y8	y9	u6	u7
y6	1.0000	−0.0084852	0.28097	0.15833	0.95078	−0.024830
y7	−0.0084852	1.0000	0.15925	−0.17482	−0.027976	0.97712
y8	0.28097	0.15925	1.0000	−0.032943	0.0098421	0.012243
y9	0.15833	−0.17482	−0.032943	1.0000	0.019262	−0.017058
u6	0.95078	−0.027976	0.0098421	0.019262	1.0000	−0.027159
u7	−0.024830	0.97712	0.012243	−0.017058	−0.027159	1.0000
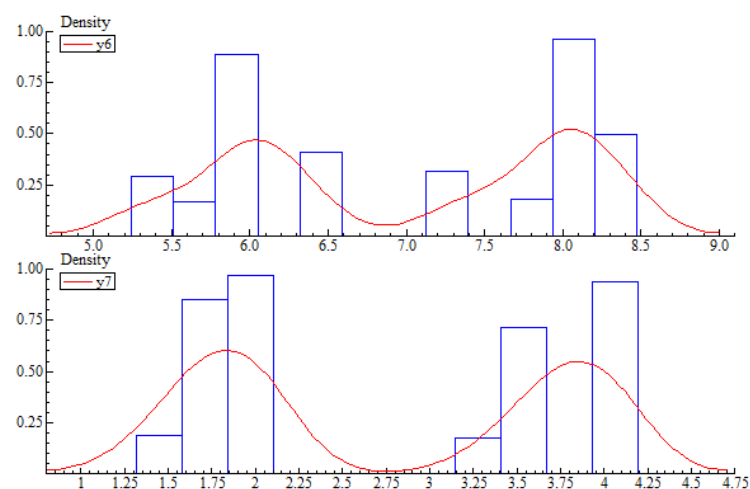
Figure 2. Estimates of the Probability Distributions of y6 and y7.

4.1. An Empirical Analysis in Accord with the Commands of Present-Day Econometrics

In the upper data universe in Figure 1 the researcher begins by estimating the parameters in Equations (4) and (5). His statistical analysis of the data is centered on three basic ideas: (1) In accord with the commands of present-day econometrics he must obtain consistent estimates of the pertinent TPD parameters; (2) An estimate of the parameters in Equations (4) and (5) is theoretically meaningful if it and Theorem 2 determine an acceptable structural model; (3) An economic theory is taken to be empirically relevant if it—in a standard statistical sense—cannot be rejected. In the present case the theory cannot be rejected only if all the vectors in the confidence region around a given estimate of TPD contains just theoretically meaningful parameter vectors.

The results of the researcher's statistical analysis are recorded in Appendix A. The estimated versions of Equations (4) and (5) are as follows:

$$y_1 = 5.29001 + 0.690561y_3 + 0.374728y_4 + eu6; \text{ and} \tag{13}$$

$$y_2 = 2.62015 + 0.359983y_3 - 0.368543y_4 + eu7. \tag{14}$$

Except for the normality tests, the diagnostic tests have acceptable results. Also, the sample of ys was obtained with a purely random sampling scheme. This and the fact that the probability distribution of (y, ξ) in (12) is statistically adequate allow the researcher to claim that his parameter estimates are consistent and asymptotically normally distributed.[10]
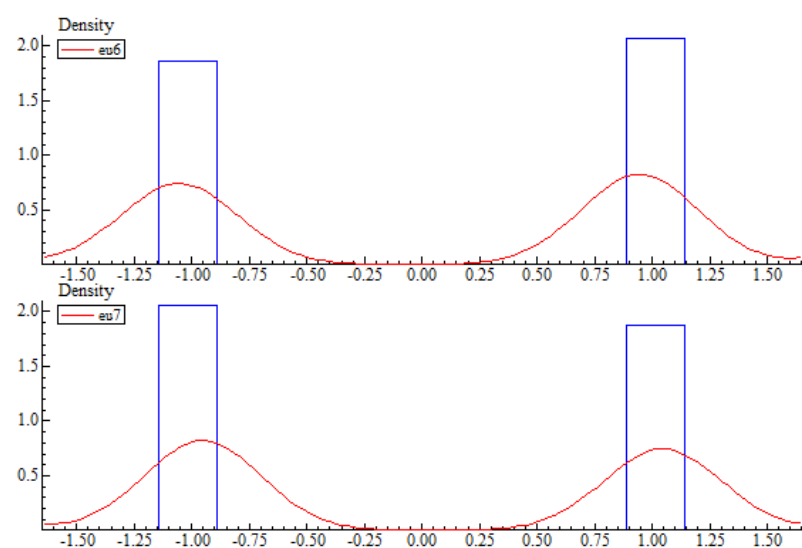

Figure 3. Estimates of the probability distributions of eu6 and eu7.

The acceptable diagnostic results and the consistency of the researcher's estimates need not mean that the estimates in (13) and (14) are theoretically meaningful. For an easy check of that, the researcher uses the following inverse of Theorem 2.

Theorem 6. *It follows from the equations in Theorem 2 that*

$$A = a - B\alpha; B = c/\gamma; \text{ and } C = [(E - B)/E]b; \tag{15}$$

---

[10] The error terms in Equations (13) and (14) are not normally distributed. Estimates of their distributions are displayed in Figure 3. The non-normality of these error terms imply that the Wald intervals of the parameters in Equations (13) and (14) do not constitute an accurate 95% confidence region for the estimates of the parameters in Equations (4) and (5).

$$D = a - E\alpha;\ E = b/\beta,\ and\ F = [(B - E)/B]c \tag{16}$$

From Equations (15) and (16) and the regression results he obtains the following indirect estimates of the structural parameters;

$$A = 7.95413;\ B = -1.01678,\ C = 1.05658 \tag{17}$$

$$D = 0.26372;\ E = 1.91832;\ F = 1.08171 \tag{18}$$

These estimates demonstrate that the parameters in the estimated equations do not constitute parameters of a theoretically meaningful estimate of the TPD.

Since his estimate of TPD is not theoretically meaningful, the researcher finds it interesting to check if there is a parameter vector in Wald's confidence band that constitutes a parameter vector of a theoretically meaningful estimate of the TPD.

Wald's confidence band is as described by the following intervals:

$$a \in (5.00421, 5.57581),\ b \in (0.561441, 0.819681),\ and\ c \in (0.255068, 0.494388) \tag{19}$$

$$\alpha \in (2.33415, 2.90615),\ \beta \in (0.230823, 0.489143),\ and\ \gamma \in (-0.488243, -0.248843) \tag{20}$$

In this confidence region the constants in Equations (21) and (22) below constitute, within an admissible margin of error, parameters of a theoretically meaningful estimate of the TPD:

$$a = 5.00422,\ b = 0.62545,\ c = 0.37455 \tag{21}$$

$$\alpha = 2.90614,\ \beta = 0.36322,\ \gamma = -0.36322 \tag{22}$$

In fact, these two equations and Equations (15) and (16) determine the following values of the structural parameters

$$A = 8.001,\ B = -1.03119,\ and\ C = 1.00000; \tag{23}$$

$$D = 0.00001,\ E = 1.72195,\ and\ F = 1.00000. \tag{24}$$

The results of the statistical analysis in the upper data universe in Figure 1 carry three kinds of information: (1) The estimates do not determine theoretically meaningful values of the structural parameters; (2) With 95% probability there exist parameter vectors that satisfy the researcher's theoretical hypothesis; and (3) Wald's confidence region contains parameter vectors that do not satisfy the researcher's theoretical hypothesis. These three observations leave the researcher in the upper data universe with no other choice than to reject the hypothesis which he formulated in the three assumptions he added to Γ_P. i.e., he must conclude that his theory is not empirically relevant.

4.2. A Formal-Econometric Empirical Analysis

In the lower data universe in Figure 1 the researcher's empirical deliberations center on three fundamental ideas that describe the prescriptions underlying the statistical analysis, delineate the salient characteristics of a data admissible model of the MPD, and describe what it means for a theory to be empirically relevant. In the present context the prescriptions that underlie the statistical analysis delineate the way the researcher is to obtain a consistent estimate of a mathematical model of the MPD. Also, a mathematical model of the MPD is data admissible if (1) the estimated mathematical model of MPD satisfies the strictures on which the basic-data axioms and the prescriptions that underlie the statistical analysis insist; (2) the values that the given model assigns to the parameters of MPD satisfy the strictures on which the theory axioms and the bridge principles insist; and (3) the model lies within a 95% confidence region of the estimated mathematical model of MPD. Finally, a family of mathematical models of Γ_T is empirically relevant if and only if there is at least one member of the family which—with a corresponding model of the axioms of the basic data universe and the Bridge—determines a data-admissible model of the MPD.

The researcher in the lower data universe begins by estimating the parameters in Equations (10) and (11). Since these equations are, formally, like Equations (4) and (5), his estimates are the same as the ones given in Equations (13) and (14). These estimates are not theoretically meaningful. Hence, they cannot be data admissible. Now, the arguments that demonstrate that the parameter values in Equations (21) and (22) are theoretically meaningful suffice to demonstrate that these parameter values are also parameters of a data admissible MPD. Theorem 7 attests to that:

Theorem 7. *Suppose that the theory axioms and the axioms of the Bridge are valid. In the present empirical analysis a theoretically meaningful parameter vector whose components belong in Wald's confidence intervals is the parameter vector of a data admissible MPD distribution of y.*

A simple sketch of a proof runs as follows. Let G be a theoretically meaningful parameter vector whose components belong in Wald's confidence intervals, let H be the structural parameter vector that it determines, and let $\Omega(H)$ be the bridge that H determines. Then, define v_1 and v_2 on $\Omega(H)$, formulate the pertinent versions of Equations (8) and (9), and pick an RPD distribution of (x_4, x_5, u_1, u_2). The Equations (8) and (9), Theorem 4, and the given RPD distribution of (x_4, x_5, u_1, u_2) suffice to construct an MPD distribution of (y, v_1, v_2). This MPD distribution is data admissible, and G is its parameter vector.

The preceding theorem makes it possible to establish the next theorem. Theorem 8 is the fundamental theorem of the formal-econometric analysis of my data.

Theorem 8 *Suppose that the theory axioms and the axioms of the Bridge are valid. Then a parameter vector whose components belong in Wald's confidence intervals determines an empirically relevant model of Γ_T if and only if it is a theoretically meaningful parameter vector.*

It is obvious that the parameter vector of an MPD which an empirically relevant model of Γ_T determines is theoretically meaningful and has components that belong to Wald's confidence intervals. Hence to establish the theorem it suffices to show that a theoretically meaningful parameter vector whose components belong in Wald's confidence intervals determines an empirically relevant model of Γ_T. A theoretically meaningful parameter vector determines a model of axioms A1–A4. Also it follows from Theorem 7 that a theoretically meaningful parameter vector whose components belong in Wald's confidence intervals is the parameter vector of a data admissible MPD. Hence I can conclude the proof by showing that the pertinent data admissible MPD determines a model of an RPD distribution of (x, u). Let the parameter vector in Equations (21) and (22) be the theoretically meaningful vector in question, and let v_1 and v_2 be the random variables that satisfy the equations

$$y_1(\omega_P) - 5.00422 - 0.62545y_3(\omega_P) - 0.37455y_4(\omega_P) = v_1(\omega_P) \tag{25}$$

$$y_2(\omega_P) - 2.90614 - 0.36322y_3(\omega_P) + 0.36322y_4(\omega_P) = v_2(\omega_P) \tag{26}$$

Then, pick an $\omega_P{}^\circ$ in the support of the MPD distribution of v_1 and v_2, and let

$$(u_1, u_2) = (v_1, v_2)(\omega_P{}^\circ),$$

$$x_1 = 5.00422 + 0.62545y_3(\omega_P{}^\circ) + 0.37455y_4(\omega_P{}^\circ),$$

$$x_3 = 2.90614 + 0.36322y_3(\omega_P{}^\circ) - 0.36322y_4(\omega_P{}^\circ),$$

$$x_4 = y_3(\omega_P{}^\circ),$$

and

$$x_5 = y_4(\omega_P{}^\circ).$$

The MPD distribution of the right-hand side of these equations determines an RPD distribution of the left-hand side that has all the required properties; i.e., a probability distribution that is a model of the RPD distribution of $(x_1, x_3, x_4, x_5, u_1, u_2)$. That and the equation $x_1 = x_2$ conclude the proof that the theory universe has an empirically relevant model.

The preceding deliberations show that there is at least one empirically relevant model of Γ_T. There may be many more. To find out how many more, the researcher will need two more theorems.

Theorem 9. *Suppose that the theory axioms and the axioms of the Bridge are valid. Then the subsets of Wald's confidence intervals that contain the data admissible parameter vectors are as described below:*

$$a \in (5.00421, 5.57581), b \in (0.561441, 0.6667), \text{ and } c \in (0.3333, 0.494388) \tag{27}$$

$$\alpha \in (2.33415, 2.90615), \beta \in (0.25, 0.489143), \text{ and } \gamma \in (-0.488243, -0.25) \tag{28}$$

I have obtained these intervals in two steps. First I used the equations in A4 and Theorem 2 to determine the acceptable minimum and maximum values of the pertinent parameters; e.g., $2.666 < a < 6.6667$ and $0.3333 < b < 0.6667$. Then, I compared these maximum and minimum values with the Wald intervals in Equations (19) and (20) and determined the extent of the asked-for intervals in Equations (27) and (28).

Theorem 10. *Suppose that the theory axioms and the axioms of the Bridge are valid. Then the intervals in Equations (27) and (28) imply that the empirically relevant values of the components of (A, B, E) must satisfy the conditions:*

$$7.33831 < A < 11.32278; -1.977552 < B < -0.68265; \text{ and } 1.14781 < E < 2.667. \tag{29}$$

Consequently, in the present empirical context not all the models of the theory axioms are empirically relevant.

The intervals in Equation (29) put added restrictions on the values of B, and E. In doing that, Theorem 10 demonstrates that models of the theory with a $B \in [-2, -1.977552]$ and/or an $E \in [1, 1.14781]$ cannot be empirically relevant. From this the researcher concludes that the inequalities in Equation (29) demonstrate that all the models of the theory axioms with $A \in [8, 10]$, $B \in (\div 1.977552, \div 1]$, $C = 1$, $D = 0$, $E \in (1.14781, 2]$, and $F = 1$ are empirically relevant in the present empirical context. Theorem 11 bears witness to that:

Theorem 11. *Suppose that the theory axioms and the axioms of the Bridge are valid. Then all the models of the theory axioms with $A \in [8, 10]$, $B \in (-1.977552, -1]$, $C = 1$, $D = 0$, $E \in (1.14781, 2]$, and $F = 1$ are empirically relevant in the present empirical context.*

4.3. The Contrasts and the Status of B1–B5

In this Section, I will describe the contrasts between the empirical analyses in the two data universes in Figure 1 and conclude the example by establishing the status of the bridge principles, B1–B5, in the present theory-data confrontation.

4.3.1. The Contrasts

The three observations which in a nutshell concretize the results of the empirical analysis in the upper data universe, are also observations that the researcher in the lower data universe made. In the upper data universe the second assertion is to no avail for the empirical relevance of the researcher's hypotheses because he is unable to assign a probability to his observation of the vector in Equations (21) and (22). In the lower data universe the researcher uses the same vector to show that his theory is empirically relevant. In the upper data universe the third observation leaves the

researcher with no other choice than to declare that his theory is not empirically relevant. In the lower data universe the researcher uses a similar observation—Theorem 10—to delineate the subfamily of models of his theory that are empirically relevant.

4.3.2. The Status of B1–B5

The bridge principles play a central role in my deliberations in the lower data universe. It is, therefore, right to question the validity of the axioms of the Bridge. In Stigum (2015) [1], I claimed that there are good reasons to believe in the validity of the bridge principles in the Real World if the 95% confidence region around a meaningful estimate of the pertinent MPD is data admissible. A 95% confidence band is data admissible if it contains a congruent model of the TPD and a data admissible model of the MPD. In this paper, I insist that the bridge principles are valid in the Real World if and only if (1) the MPD in a congruent econometric model of the TPD is data admissible; and (2) all the pairs of econometric models of the MPD, $(M1, \theta_M^\circ(\cdot))$ with a data-admissible model of the MPD, are congruent.

Next I will show that the Wald intervals in the present example constitute a 95% data-admissible confidence region, and that B1–B5 are valid in the Real World. I begin by showing that the Wald intervals contains a model of MPD that is congruent. In this case there are two econometric models to consider, $(M1, \theta_M^\circ(\cdot))$ and $(M2, \theta_P^\circ(\cdot))$, where M1 is any one of the many MPD distributions, $\theta_M^\circ(\cdot)$ is the estimator of $(a, b, c, \alpha, \beta, \gamma, \sigma_{v1}^2, \sigma_{v2}^2)$ in Equations (10) and (11), M2 is the TPD distribution, and $\theta_P^\circ(\cdot)$ is the estimator of $(d, e, f, \lambda, \pi, \varrho, \sigma_{\xi1}^2, \sigma_{\xi2}^2)$ in Equations (4) and (5). The estimators are consistent, and in M1 it is a fact that $\theta_M^\circ(s_n) = \theta_P^\circ(s_n)$ a.e.. Hence, $(M1, \theta_M^\circ(\cdot))$ exactly encompasses $(M2, \theta_P^\circ(\cdot))$ no matter which one of the possible MPD distributions one chooses for M1. It is, also, the case that if I choose the components of θ_{M1} to have the values ascribed to them in Equations (21) and (22), and if I let θ_P equal θ_{P0}, where $\theta_{PO} = (16/3, 2/3, 1/3, 8/3, 1/3, -1/3, 1, 1)$, then in M2, plim $\theta_P^\circ(s_n) = \theta_{P0} = $ plim $\theta_M^\circ(s_n)$. Hence, with the given choice of M1, M1 is coherent with the pertinent a priori theory, and $(M1, \theta_M^\circ(\cdot))$ encompasses $(M2, \theta_P^\circ(\cdot))$ in the sense of Definition 1C. Since the sampling system in the present case is adequate, it follows that with the given choice of θ_M, $(M1, \theta_M^\circ(\cdot))$ is a congruent econometric model of the TPD.

At last, I must show that the Wald intervals constitute a data-admissible confidence region and that B1–B5 are valid in the real world. I have shown before that the M1 in the congruent model of TPD is data admissible. Hence, the Wald intervals constitute a data-admissible confidence region. It is, also, obvious from my deliberation above that I can substitute for M1 any data admissible model of MPD and obtain a congruent model of TPD. Hence all the data-admissible models of MPD are congruent. But, if that is true, then it is a fact that the bridge principles in the present example are valid in the real world.

Conflicts of Interest: The author declares no conflict of interest.

Appendix A. The Statistical Analysis

I have generated 1249 observations of y_1, y_2, y_3, y_4, ξ_1, and ξ_2 with OxMetrics' random number generator for the uniform distribution, ranu(). In the printout, $y_1 = y6$, $y_2 = y7$, $y_3 = y8$, $y_4 = y9$, $\xi_1 = u6$, and $\xi_2 = u7$. The algebra code shows how the observations were generated.

Algebra Code for generating 1249 observations on u6, u7, y6, y7, y8, and y9

```
eps6 = ranu();
u6 = eps6 < 0.5 ? 1 : −1;
eps7 = ranu();
u7 = eps7 < 0.5 ? 1 : −1;
eps8 = ranu();
y8 = eps8 < 0.25 ? 1 : 2;
```

```
eps9 = ranu();
y9 = eps9 < 0.6666 ? 1 : 2;
y6 = 16/3 + 2/3*y8 + 1/3*y9 + u6;
y7 = 8/3 + 1/3*y8 − 1/3*y9 + u7;
```

I have used PcGive's Autometrics programme to estimate Equations (4) and (5). The programme is described in Chapter 6 of Volume 1 of the 2009 version of Dornik and Hendry's PcGive 13 [19].

Table A1. Modelling y6 by OLS-CS with 1249 observations.

	Coefficient	Std.Error	t-Value	t-Prob	Part.R^2
Estimates of the constants in Equation (4) and their significance.					
Constant	5.29001	0.1429	37.0	0.0000	0.5237
y8	0.690561	0.06456	10.7	0.0000	0.0841
y9	0.374728	0.05983	6.26	0.0000	0.0305
Measures of the goodness of fit of the estimated equation.					
Sigma	0.999521	RSS	1244.80701		
R^2	0.107061	F(2,1246) =	74.7 [0.000]**		
Adj.R^2	0.105627	log-likelihood	−1770.15		
No. of observations	1249	No. of parameters	3		
mean(y6)	6.99333	se(y6)	1.0569		
Tests of certain probabilistic characteristics of the error term in Equation (13).					
Normality test:	Chi^2(2) = 1569.0 [0.0000]**				
Hetero test:	F(2,1246) = 0.42845 [0.6516]				
Hetero-X test:	F(3,1245) = 0.39513 [0.7565]				
RESET23 test:	F(1,1245) = 0.22032 [0.6389]				

** indicates that the pertinent null hypothesis is rejected at the one percentage level of significance.

Table A2. EQ(2) Modelling y7 by OLS-CS with 1249 observations.

	Coefficient	Std.Error	t-value	t-prob	Part.R^2
Estimates of the constants in Equation (5) and their significance.					
Constant	2.62015	0.1430	18.3	0.0000	0.2123
y8	0.359983	0.06458	5.57	0.0000	0.0243
y9	−0.368543	0.05985	−6.16	0.0000	0.0295
Measures of the goodness of fit of the estimated equation.					
Sigma	0.999872	RSS	1245.68014		
R^2	0.0541462	F(2,1246) =	35.66 [0.000]**		
Adj.R^2	0.052628	log-likelihood	−1770.59		
No. of observations	1249	no. of parameters	3		
mean(y7)	2.75367	se(y7)	1.02727		
Tests of certain probabilistic characteristics of the error term in Equation (14).					
Normality test:	Chi^2(2) = 1577.5 [0.0000]**				
Hetero test:	F(2,1246) = 0.097428 [0.9072]				
Hetero-X test:	F(3,1245) = 0.46596 [0.7061]				
RESET23 test:	F(1,1245) = 0.80716 [0.3691]				

** indicates that the pertinent null hypothesis is rejected at the one percentage level of significance.

Algebra Code for the error terms in the two regressions

```
eu6 = y6 − 5.29001 − 0.690561*y8 − 0.374728*y9;
eu7 = y7 − 2.62015 − 0.359983*y8 + 0.368543*y9;
```

References

1. Stigum, B.P. *Econometrics in a Formal Science of Economics*; MIT Press: Cambridge, MA, USA; London, UK, 2015.
2. Frisch, R. Bedømmelse av Jonas Åkerman: Om det Økonomiska Livets Rytmik. *Statsvetensk. Tidskr.* **1931**, *34*, 281–300.
3. Haavelmo, T. The Probability Approach in Econometrics. *Econometrica* **1944**, *12*, iii–vi and 1–115. [CrossRef]
4. Bjerkholt, O.; Qin, D. (Eds.) *A Dynamic Approach to Economic Theory: The Yale Lectures of Ragnar Frisch, 1930*; Routledge: London, UK, 2010.
5. Bontemps, C.; Mizon, G.E. Congruence and Encompassing. In *Econometrics and the Philosophy of Economics: Theory-Data Confrontations in Economics*; Stigum, B.P., Ed.; Princeton University Press: Princeton, NJ, USA; Woodstock, UK, 2003; pp. 354–378.
6. Hendry, D.F.; Krolzig, H.-M. New Developments in Automatic General-to-Specific Modeling. In *Econometrics and the Philosophy of Economics*; Stigum, B.P., Ed.; Princeton University Press: Princeton, NJ, USA; Woodstock, UK, 2003; pp. 379–419.
7. Mizon, G.E. The Encompassing Approach in Econometrics. In *Econometrics and Quantitative Economics*; Hendry, D.F., Wallis, K.E., Eds.; Blackwell: Oxord, UK; New York, NY, USA, 1984; pp. 135–172.
8. Hendry, D.F.; Richard, J.F. Recent Developments in the Theory of Encompassing. In *Contributions to Operations Research and Economics: The XXth Aniversary of CORE*; Cornet, B., Tulkens, H., Eds.; MIT Press: Cambridge, MA, USA; London, UK, 1989; pp. 393–440.
9. Gourieroux, C.; Montfort, A.; Trognon, A. Testing Nested or Non-Nested Hypotheses. *J. Econom.* **1983**, *21*, 83–115. [CrossRef]
10. Florence, J.P.; Hendry, D.F.; Richard, J.F. Encompassing and Specificity. *Econom. Theory* **1996**, *12*, 620–656. [CrossRef]
11. Bontemps, C.; Mizon, G.E. Encompassing: Concepts and Implementation. *Oxf. Bull. Econ. Stat.* **2008**, *70*, 721–750. [CrossRef]
12. Mizon, G.E. Progessive Modelling of Macroeconomic Time Series: The LSE Methodology. In *Macroeconometrics; Developments, Tensions and Prospects*; Hoover, K.D., Ed.; Kluver Academic Press: Dordrecht, Germany, 1995; pp. 107–169.
13. Hendry, D.F. *Dynamic Econometric*; Oxford University Press: Oxford, UK, 1995.
14. Spanos, A. On Rereading Haavelmo: A Retrospective View of Econometric Modeling. *Econom. Theory* **1989**, *5*, 405–429. [CrossRef]
15. Spanos, A. Revisiting Haavelmo's Structural Econometrics: Bridging the Gap between Theory and Data. In Presented at the Trygve Haavelmo Centennial Symposium, Oslo, Norway, 13–14 December 2011.
16. Dunford, N.; Schwartz, J.T. *Linear Operators, Part I: General Theory*; Interscience: New York, NY, USA, 1957.
17. Stigum, B.P. *Toward a Formal Science of Economics*; MIT Press: Cambridge, MA, USA, 1990.
18. Dornik, J.A. *An Introduction to OxMetrics 6*; Timberlake Consultants Press: London, UK, 2009.
19. Dornik, J.A.; Hendry, D.F. *PcGive 13, Volume I: Modelling Dynamic Systems*; Timberlake Consultants Press: London, UK, 2009.