MEMORANDUM

No 13/2017
December 2017

Identification, Instruments, Omitted Variables, and Rudimentary Models: Fallacies in the 'Experimental Approach' to Econometrics

Erik Biørn

ISSN: 0809-8786

Department of Economics
University of Oslo

This series is published by the
University of Oslo
Department of Economics

P. O.Box 1095 Blindern
N-0317 OSLO Norway
Telephone: + 47 22855127
Fax: + 47 22855035
Internet: http://www.sv.uio.no/econ
e-mail: econdep@econ.uio.no

In co-operation with
The Frisch Centre for Economic Research

Gaustadalleén 21
N-0371 OSLO Norway
Telephone: +47 22 95 88 20
Fax: +47 22 95 88 25
Internet: http://www.frisch.uio.no
e-mail: frisch@frisch.uio.no

Last 10 Memoranda

No 12/2017	Tapas Kundu and Tore Nilssen Delegation of Regulation*
No 11/2017	Pedro Brinca, Miguel H. Ferreira, Francesco Franco, Hans A. Holter and Laurence Malafry Fiscal Consolidation Programs and Income Inequality*
No 10/2017	Geir B. Asheim & Andrés Perea Algorithms for cautious reasoning in games*
No 09/2017	Finn Førsund Pollution Meets Efficiency: Multi-equation modelling of generation of pollution and related efficiency measures*
No 08/2017	John K. Dagsvik Invariance Axioms and Functional Form Restrictions in Structural Models
No 07/2017	Eva Kløve and Halvor Mehlum Positive illusions and the temptation to borrow
No 06/17	Eva Kløve and Halvor Mehlum The Firm and the self-enforcing dynamics of crime and protection
No 05/17	Halvor Mehlum A polar confidence curve applied to Fieller's ratios
No 04/17	Erik Biørn Revisiting, from a Frischian point of view, the relationship between elasticities of intratemporal and intertemporal substitution
No 03/17	Jon Vislie Resource Extraction and Uncertain Tipping Points

Previous issues of the memo-series are available in a PDF® format at:
http://www.sv.uio.no/econ/english/research/unpublished-works/working-papers/

Identification, Instruments, Omitted Variables, and Rudimentary Models: Fallacies in the 'Experimental Approach' to Econometrics

Erik Biørn

Department of Economics, University of Oslo,
P.O. Box 1095 Blindern, 0317 Oslo, Norway
E-mail: erik.biorn@econ.uio.no

Abstract: Since identification, instrumental variables and variables exclusion, core concepts in econometrics, are entwined, several questions arise: How is identification related to the existence of IVs? How are identification criteria related to omitted variables? Is omission/inclusion of variables from a model's equations part of the definition of IVs? Is exogeneity a critical claim to an IV? Is 'omitted variables' a meaningful term for a single equation when its 'environment' is incompletely described? Which are the borderlines between omitted, observable variables, omitted non-modeled variables, latent variables represented by proxies or measurement error mechanisms? These are among the questions addressed in this paper, partly with reference to the conflict between 'experimentalists' and 'structuralists', specifically relating to: (i) the contrast between 'rudimentary models' and models for 'limited information inference', (ii) the distinction between exogeneity of variables and the orthogonality claim for IVs and disturbances or errors, (iii) the role of predetermined variables in selecting IVs, and (iv) the 'omitted variables' concept and the role of IVs in 'handling' such variables, when considering the 'origin' of the omission.

Keywords: Identification. Instrumental variables. Omitted variables. Limited information. Experimental approach.

JEL classification: B23, C21, C26, C31, C36, C51

Acknowledgements: Comments from and discussions with Olav Bjerkholt, John K. Dagsvik, Ragnar Nymoen, Asbjørn Rødseth, Terje Skjerpen, and Yngve Willassen are gratefully acknowledged. The usual disclaimer applies.

1 Introduction

Identification and instrumental variables (IVs or instruments, for short) are core concepts in classical and modern econometrics. The former relates to *model properties* and is a 'pre-observation' concept concerning, briefly, *whether it is possible* from knowledge of the distribution of a model's observable endogenous variables (conditional on the exogenous variables) to 'uncover' the value of a specific parameter of interest, θ. The latter are elements in several procedures for estimating from data the value of θ, given that identifiability is assured, including Two-Stage and Three-Stage Least Squares (2SLS, 3SLS) and the General Method of Moments (GMM). The sharp distinction that in general should be made between 'model' and 'method' in empirical research, implies that it is essential to separate the parts of an investigation that relates to model description and identification from those relating to the use of IVs in statistical procedures. Among the conclusions that follow from this are that 'identification by using IVs' and 'an IV model', which mix the two, are expressions that should be avoided.[1] Likewise, 'IV-estimand' is a misnomer, since in common language, an estimand is that which is to be estimated in a statistical analysis, irrespective of method, usually a parameter.[2]

On the other hand, from the conditions for identification of the coefficients of a certain equation may, under certain conditions, be *derived* a requirement that a minimal number of IVs for its variables should exist *if an analyst considers using procedures that involve IVs*. Likewise, an examination of properties of the model outside of the equation will often suggest IV candidates for the latter and hence give *a model for which IVs exist*. But this should not be called 'an IV model'.

Koopmans and Reiersøl, in a classical article on identifiability and identification problems, state:

> "One might regard problems of identifiability as a necessary part of the specification problem ..... Identification problems are not problems of statistical inference in a strict sense, since the study of identifiability proceeds from a hypothetical exact knowledge of the probability distribution of observed variables rather than from a finite sample of observations. However ..... the study of identifiability is undertaken in order to explore the limitations of statistical inference." [Koopmans and Reiersøl (1950, pp.169–170)]

Aldrich summarizes the 'history of IVs' by:

> "The method's rise to prominence can be divided into three phases. In the 1940s the method was introduced for use with errors in variables model. In the 1950s it was related to methods devised for the errors in equations model. More recently it has been transformed into an organising principle underlying many apparently distinct models ..... According to Reiersøl (1950, p. 378), the "idea of using instrumental variables" was introduced independently by himself in 1941 and by Geary in 1943. .... As Morgan (1990, p. 226) remarks, Reiersøl's contribution was "hidden in amongst a number of extensions to Frisch's confluence analysis." ..... Frisch held that a combination of measurement error and multicollinearity makes economic data difficult to analyse ..... his influence on Reiersøl's work of 1941 and 1945 was overwhelming." [Aldrich (1993, p. 247–249)]

The third and fourth elements in the title, omitted variables and rudimentary models, are related to the first two, and the four elements are entwined. It is for example well known, and usually spelt out in econometric textbooks, that criteria for identifiability of the coefficients of an equation in a linear multi-equation model relate to properties of variables which occur *somewhere* in the model, but *do not enter the equation of interest*. An important matter for empirical research then becomes in which detail should the rest of the model (have to) be specified to draw conclusions on identifiability and availability of IVs for its equations. Can we do with models that are not specified in such a detail that the existence of valid IVs follows logically from the models' assumptions? *Rudimentary models* is the label we attach to such models, while model specifications containing theoretical underpinnings which directly motivate the IV-properties will be called *models for limited information inference*. The question above may be answered by a qualified yes, as it has been known for about sixty years that even from rudimentary models, consistent estimation of one (or a subset) of its equations is possible, if some additional assumptions, usually of an *ad hoc nature*, are met.

For some years there has been what Keane (2010, pp. 3–4) calls a conflict between 'the "structural" and "experimental" camps in econometrics', labels to be used also in this text. While both give core roles to IVs and omitted variables, there seems to be terminological disagreement along with disagreement about the status and requirements of formalized models in research. In particular, there are reasons to ask whether a changed use of the terms 'identification' and 'instrumental variables' has occurred.[3] My answer is in the affirmative, and when using the term identification, I stick to the definition often called *point identification* (as opposed to *e.g.*, 'set identification'), relating to the non-existence of observationally equivalent structures or parameter points, see Rothenberg (1971, p. 578) and to the concept named the *Haavelmo distribution*, by Spanos (1989); see also Aldrich (1994) and Hendry and Johansen (2015, Section 2). Signs of changed focus and distortion of terminology are the declining attention many practitioners of 'empirical economics' pay to identification problems in the classical sense and to the related concepts *structural parameter*, see Marschak (1953) and Koopmans (1953), and *autonomy*, see Frisch (1938, 2005) and Aldrich (1989), as well as the growing tendency to use 'identify' and 'estimate' interchangeably.

Since the concepts in the title are entwined, relevant questions become: How is identification related to the existence of IVs? How are criteria for identification related to omitted variables? Does omission/inclusion of variables from a model's equations belong to the definition of IVs? How draw the borderlines between omitted variables and disturbances and between omitted, non-modeled variables, latent variables modeled via proxies or measurement error mechanisms, and omitted, observable variables? Is exogeneity a critical claim to an IV? Can 'omitted variables' in a single equation be given a definite meaning when the equation's 'environment' is incompletely described? Is it possible to distinguish operationally between 'exactly identified' and 'overidentified' equations in incompletely specified models?

Having no ambition of fully answering these questions, I will address some of them, including: (i) The contrast between rudimentary models used by many 'experimentalists' and models for limited information inference used by most 'structuralists', the latter containing full lists of exogenous and endogenous variables, without listing all unknown coefficients by equation. (ii) The distinction between exogeneity of variables and the claim of orthogonality between an IV and the relevant disturbance or error. (iv) The way of motivating the requirement that IVs be correlated with the 'instrumented variables' (sometimes denoted as 'the full rank condition'), (iv) Predetermined variables in dynamic equations and their role in selecting IVs. (v) The role of IVs in 'handling' omitted variables, taking into account the 'reason' for the omission.

The conflict between the 'structuralists' (the 'structural approach') and the 'experimentalists' (the 'treatment effect approach') in economics and the way they invoke IVs can be illustrated by a few quotations. Pearl (2015, p.169), with address to two proponents of the 'experimentalist movement', for example asks:

> "..... why did the "experimentalists" end up with the primitive, single-equation exercises reported in Angrist and Pischke (2010)? ..... The answer usually given is that "experimentalists" are a priori skeptical about the assumptions embedded in structural models, and feel more comfortable with those involved in instrumental variables design. However, since the very choice of an instrument rests on the type of modeling assumptions that "experimentalists" attempt to avoid, namely, exclusion and exogeneity ..... why did "experimentalists" embrace the former and reject the latter?"

The following quotations indicate positions in this conflict more succinctly, the first distinguishing between natural experiments and *'natural natural experiments'*:[4]

> "The advantage of the natural natural experimental approach is that *the assumption of randomness for the instrumental variables employed is more credible than for those instruments used in almost all other studies.* But a weakness of many of the studies that adopt this approach is that the necessary additional ..... assumptions needed to justify the authors' interpretations of the estimates obtained are absent. The impression left ..... is that if one accepts that the instruments are perfectly random and plausibly affect the variable whose effect is of interest, then the instrumental-variables estimates are conclusive ..... However, the absence of models in the natural natural experiment literature does not mean that there are no important and implausible assumptions being implicitly used." (my italics) [Rosenzweig and Wolpin (2000, pp. 828-829)]

> "..... the natural experiments approach to instrumental variables is fundamentally grounded in theory, in the sense that there is usually a well-developed story or model motivating the choice of instruments ..... this approach contrasts favorably with *studies that provide detailed but abstract theoretical models, followed by identification based on implausible and unexamined choices about which variables to exclude from the model* and assumptions about what statistical distribution certain variables follow." (my italics) [Angrist and Krueger (2001, pp. 72–76)]

> "Evaluating the impacts of public policies, forecasting their effects in new environments, and predicting the effects of policies never tried are three central tasks of economics ..... The structural approach emphasizes clearly articulated economic models that can be used to accomplish all three tasks under the exogeneity and parameter policy invariance assumptions presented in that literature. Economic theory is used to guide the construction of models and to suggest included and excluded variables ...... The treatment effect literature ..... focuses on ..... evaluating the impact of a policy ..... in the special case where there is a "treatment group" and a "comparison group" ..... The literature on treatment effects has given rise to a new language ..... *where the link to economic theory is often obscure and the economic policy questions being addressed are not always clearly stated. Different instruments answer different economic questions that typically are not clearly stated.* Relationships among the policy parameters implicitly defined by alternative choices of instruments are not articulated." (my italics) [Heckman and Vytlacil (2005, p, 669–670)]

The rest of the paper, containing at several places selected commented examples and quotations, proceeds as follows: In Section 2, the connection between identification and IVs is considered with reference to five examples of different complexity and to the mentioned conflict between 'structuralists' and 'experimentalists'. The discussion of the examples is expanded in Section 3 with supplementary remarks, and in Section 4, with comments on the way 'IV-regression' and endogeneity, relative to exogeneity, externality, and orthogonality, are dealt with in some recent texts. In Section 5, connections between omitted and unobserved variables on the one hand and coefficient identification on the other are illustrated by examples. Section 6, moves attention from models having only observed variables, to examples containing some latent (structural) variables, emphasizing the role of the latter in motivating IV candidates. This section contains examples showing (in contrast to not uncommon belief) that an equation may well have as IVs variables that are endogenous in the model to which it belongs, *inter alia*, by reference to the distinction between interdependent and recursive models and between fully exogenous and conditionally exogenous variables.[5] Interesting in this connection is Koopmans' (1950, p. 393) distinction between exogeneity according to the "departemental principle" and the "causal principle". Section 7 switches attention from static to dynamic models and exemplifies model-based IV-selection when autoregressive structures interact with disturbance memory. Concluding remarks follow in Section 8.

2 Instrumental variables and identification: Five example models

This section exposes, by five examples, A through E, the definition of an IV and its relation to identification. An important distinction between them, for a specific equation of interest, is whether potential IVs belong to the equation together with other variables (C and E) or do not belong to it (A, B and D). In this respect there seems to be some confusion and disagreement in the literature, which may reflect that many applications, not least in the 'natural experiment' literature and in some elementary expositions in textbooks, are concerned with only the very simplest case where one variable in need of an IV is inside the equation and only one or a few IV candidates are available outside of it. Another distinction is that Examples A, B and C have one equation and additional assumptions from which IVs can be defined, these supplements being tacit about the form and properties of the other equations, whiles Examples D and E have one equation containing several endogenous variables needing IVs, and as a supplement a list of the full set of endogenous and exogenous variables in the model to which the equation belongs. The latter are *models for limited information inference*, 'limited' indicating that they lack detailed descriptions of these supplementary equations beyond what is required to establish the identification status of the equation of interest.[6]

A. One investigational variable,[7] *one instrument.* Consider the simple equation,

$$y = \alpha + x\beta + u, \tag{1}$$

with (y, x) assumed observable and $\text{cov}(x, u) \neq 0$. An IV, z, for x, with coefficient β, should, according to the classical definition, satisfy:

(2) z is observable, $\quad \text{cov}(z,x)\neq 0, \quad \text{cov}(z,u)=0.$

This prescribes neither how such a z could be found, nor, if found, how it should be used.[8] The above definition restricts the joint *theoretical* distribution of y, x, z, u, relative to (1), which should be founded on theory or some theory element. *'Exogenous' is not a label attached to z, nor is 'endogenous' attached to y and x.*[9] If, at one extreme $z = x$, $\text{cov}(x,u) = 0\,(\Longrightarrow\, \text{cov}(y,u) \neq 0)$, *$x$ is a potential IV for itself.* This is the situation assumed in regression analysis with regressor x. If, at the other extreme, $z=y$, $\text{cov}(y,u) = 0\,(\Longrightarrow\, \text{cov}(x,u) \neq 0)$, *$y$ is a potential IV for itself.* This corresponds to regression analysis of the inverse relation with regressor y. From (1) and (2) it follows that

$$\text{(3)}\qquad \begin{aligned} &\text{cov}(z,y) = \text{cov}(z,x)\beta \neq 0 \implies \beta = \frac{\text{cov}(z,y)}{\text{cov}(z,x)} \equiv \frac{\rho_{zy}}{\rho_{zx}}\frac{\sigma_y}{\sigma_x} \iff \\ &\text{cov}(z,x) = \text{cov}(z,y)\beta^{-1} \neq 0 \implies \beta^{-1} = \frac{\text{cov}(z,x)}{\text{cov}(z,y)} \equiv \frac{\rho_{zx}}{\rho_{zy}}\frac{\sigma_x}{\sigma_y}, \end{aligned}$$

where ρ and σ, with appropriate subscripts, denote, respectively, theoretical (population) correlation coefficients and standard deviations. It follows that *z is an IV for y as well*, with coefficient β^*, in the reverse relationship

$$\text{(4)}\qquad x = \alpha^* + y\beta^* + u^*, \quad \alpha^* = -\alpha/\beta,\ \beta^* = 1/\beta,\ u^* = -u/\beta.$$

Therefore: [1] z, satisfying (2), is a potential IV for both x and y in both the direct and reverse relationships, with coefficients β and β^*, respectively. [2] x can never be an IV for y, and y can never be an IV for x. [3] The *IV-estimators* motivated by expressions like those for β or β^{-1} in (3) follow by replacing, in the expressions for these parameters, theoretical second-order moments with their empirical counterparts, say $\widehat{\beta}_{IV} = S_{ZY}/S_{ZX}$, where S, with appropriate subscripts, denotes second-order moments obtained from the sample.

B. One investigational variable, several IVs, none in the equation. Now let $z_1, \dots, z_K$ be IVs for x in (1), assumed to have the same form as in Example A, with coefficient β, $\text{cov}(x,u) \neq 0$. Then (2) and (3) are extended to

$$\text{(5)}\qquad z_k \text{ observable}, \quad \text{cov}(z_k,x)\neq 0, \quad \text{cov}(z_k,u)=0, \quad \beta = \frac{\text{cov}(z_k,y)}{\text{cov}(z_k,x)}, \quad k = 1,\dots,K.$$

Defining in general

$$\beta_k = \frac{\text{cov}(z_k,y)}{\text{cov}(z_k,x)} \equiv \frac{\rho_{zk,y}}{\rho_{zk,x}}\frac{\sigma_y}{\sigma_x}, \qquad k = 1,\dots,K,$$

it follows that (5), by imposing that $\rho_{zk,y}/\rho_{zk,x}$ is k-invariant, imposes

$$\beta = \beta_1 = \cdots = \beta_K.$$

Here a multitude of IVs for x exist: any $w = \sum_{k=1}^{K} a_k z_k$, with $a_1, \dots, a_K$ prescribed, satisfying (5) and $\sum_k a_k \text{cov}(z_k,x)\neq 0$, is such an IV, and implies

$$\text{cov}(w,x)\neq 0, \qquad \text{cov}(w,u)=0,$$
$$\beta = \frac{\text{cov}(w,y)}{\text{cov}(w,x)} \equiv \frac{\sum_k a_k \text{cov}(z_k,x)\beta_k}{\sum_k a_k \text{cov}(z_k,x)}.$$

The IV-estimator of β follows by replacing, in the last expression for the chosen $a_1,\dots,a_K$, theoretical second-order moments by their empirical counterparts.

C. One investigational variable, two IV candidates, one in the equation, one outside. Extend (1) to

(6) $$y=\alpha+x_1\beta_1+x_2\beta_2+u,$$

with (y,x_1,x_2) assumed observable, x_1, with $\text{cov}(x_1,u)\neq 0$, corresponding to x in (1), x_2 being an additional explanatory variable, uncorrelated with the disturbance u, $\text{cov}(x_2,u)=0$, with coefficient β_2. An IV, z, for x_1, with coefficient β_1, satisfies the following modification of (2):

(7) z is observable, $\qquad \text{cov}(z,x_1|x_2)\neq 0, \qquad \text{cov}(z,u)=0.$

From (6) and (7) it follows that

(8) $$\begin{aligned}\text{cov}(z,y)&=\text{cov}(z,x_1)\beta_1+\text{cov}(z,x_2)\beta_2,\\ \text{cov}(x_2,y)&=\text{cov}(x_2,x_1)\beta_1+\text{var}(x_2)\beta_2,\end{aligned} \implies$$
$$\begin{bmatrix}\beta_1\\ \beta_2\end{bmatrix}=\begin{bmatrix}\text{cov}(z,x_1) & \text{cov}(z,x_2)\\ \text{cov}(x_2,x_1) & \text{var}(x_2)\end{bmatrix}^{-1}\begin{bmatrix}\text{cov}(z,y)\\ \text{cov}(x_2,y)\end{bmatrix},$$

as $\text{cov}(z,x_1|x_2)\neq 0 \implies \text{cov}(z,x_1)\text{var}(x_2)\neq \text{cov}(x_2,x_1)\text{cov}(z,x_2) \iff \rho_{z,x1}\neq\rho_{z,x2}\rho_{x2,x1}$, ensures that a solution exists.[10] It also follows that any linear combination of z and x_2,

(9) $$z^* = za_1+x_2a_2, \quad a_1\neq 0, \quad a_2/a_1\neq -\text{cov}(z,x_1)/\text{cov}(x_2,x_1),$$

is a potential IV for x_1 since $\text{cov}(z^*,x_1)\neq 0$ and $\text{cov}(z^*,u)=0$. Therefore, also in this example, a multitude of potential IVs exist, *all of which 'contain'* x_2, exist[11] and hence (8) can be generalized to

$$\begin{aligned}\text{cov}(z^*,y)&=\text{cov}(z^*,x_1)\beta_1+\text{cov}(z^*,x_2)\beta_2,\\ \text{cov}(x_2,y)&=\text{cov}(x_2,x_1)\beta_1+\text{var}(x_2)\beta_2,\end{aligned} \implies$$
$$\begin{bmatrix}\beta_1\\ \beta_2\end{bmatrix}=\begin{bmatrix}\text{cov}(z^*,x_1) & \text{cov}(z^*,x_2)\\ \text{cov}(x_2,x_1) & \text{var}(x_2)\end{bmatrix}^{-1}\begin{bmatrix}\text{cov}(z^*,y)\\ \text{cov}(x_2,y)\end{bmatrix}$$

Further, z^* *is an IV for* y *as well*, with coefficient β_1^*, in the reverse relationship

$$x_1=\alpha^*+y\beta_1^*+x_2\beta_2^*+u^*, \quad \alpha^*=-\alpha/\beta_1,\ \beta_1^*=1/\beta_1,\ \beta_2^*=-\beta_2/\beta_1,\ u^*=-u/\beta_1,$$

and hence,

$$\begin{aligned}\text{cov}(z^*,x_1)&=\text{cov}(z^*,y)\beta_1^*+\text{cov}(z^*,x_2)\beta_2^*,\\ \text{cov}(x_2,x_1)&=\text{cov}(x_2,y)\beta_1^*+\text{var}(x_2)\beta_2^*,\end{aligned} \implies$$
$$\begin{bmatrix}\beta_1^*\\ \beta_2^*\end{bmatrix}=\begin{bmatrix}\text{cov}(z^*,y) & \text{cov}(z^*,x_2)\\ \text{cov}(x_2,y) & \text{var}(x_2)\end{bmatrix}^{-1}\begin{bmatrix}\text{cov}(z^*,x_1)\\ \text{cov}(x_2,x_1)\end{bmatrix}.$$

The distinctive property of this example is that any z^* (with $a_2 \neq 0$) contains the equation's exogenous explanatory variable, so that x_2 serves the double role as an exogenous variable in the equation and a component of the IV for x_1. The IV-estimators of (β_1, β_2) and (β_1^*, β_2^*) follow by replacing, in the above expressions, for the chosen a_1 and a_2 theoretical moments by their empirical counterparts.

Examples A–C are simple, containing only one investigational variable in need of an IV. The last two examples increase applicability by allowing for a vector of such variables.

D. Several investigational variables and several IVs, none in the equation. Let the equation of interest be the following generalization of (1):

$$y = \boldsymbol{x}_A \boldsymbol{\beta}_A + u_A, \tag{10}$$

where y is the variable to be 'explained', the 'left-hand side (LHS) variable', $\boldsymbol{x}_A$ is the $(1 \times N)$ vector of 'explanatory variables', 'right-hand side (RHS) variables', including a one entry for the intercept, and $\boldsymbol{\beta}_A$ is the $(N \times 1)$ vector of coefficients, including the intercept. The equation belongs to a multi-equation model where: (i) y and $\boldsymbol{x}_A$ are jointly endogenous, (ii) $\boldsymbol{z}_B$ is an observable *exogenous* $(1 \times M)$-vector, which does not enter (10), and (iii) u_A is the disturbance, the first element of the model's full disturbance vector $\boldsymbol{u} = (u_A, \boldsymbol{u}'_B)'$. This setup describes the equation's 'environment' in more detail than Examples A–C, by specifying *properties* of the full model to which (10) belongs without specifying the model in full detail. Exogeneity of $\boldsymbol{z}_B$ may be defined in several ways, see Engle *et al.* (1983) and Hendry (1995, Ch. 5). The definition to be used here is $\mathsf{E}(\boldsymbol{z}'_B \boldsymbol{u}) = \boldsymbol{0}$, where $\boldsymbol{0}$ is a zero vector. This *is a claim to the disturbance vector of the full model*, that *implies* orthogonality between the IVs, $\boldsymbol{z}_B$, and the disturbance in the equation of interest, *i.e.*, $\mathsf{E}(\boldsymbol{z}'_B u_A) = \boldsymbol{0}$. These assumptions *imply* that $\boldsymbol{z}_B$ is an IV-vector for $\boldsymbol{x}_A$, *relative to (10)*, satisfying a generalization of (2):

$$\boldsymbol{z}_B \text{ is observable}, \qquad \text{rank}[\mathsf{E}(\boldsymbol{z}'_B \boldsymbol{x}_A)] = N, \qquad \mathsf{E}(\boldsymbol{z}'_B u_A) = \boldsymbol{0}. \tag{11}$$

Postulating (11) as *separate* requirements would be superfluous, for two reasons. First, since *exogeneity is a stronger claim than IV-disturbance orthogonality*, it follows that

$$\mathsf{E}(\boldsymbol{z}'_B y) = \mathsf{E}(\boldsymbol{z}'_B \boldsymbol{x}_A) \boldsymbol{\beta}_A. \tag{12}$$

Second, the specification of the 'environment' of (10) allows us to establish the reduced form (RF) equation for $\boldsymbol{x}_A$ (a similar equation exists for y), which has the form

$$\boldsymbol{x}_A = \boldsymbol{z}_B \boldsymbol{\Pi}_{AB} + \boldsymbol{\epsilon}_A, \qquad \mathsf{E}(\boldsymbol{z}'_B \boldsymbol{\epsilon}_A) = \boldsymbol{0}, \tag{13}$$

where in normal cases $\boldsymbol{\Pi}_{AB}$, the $(M \times N)$-matrix of coefficients in this RF-equation, has rank N, and $\mathsf{E}(\boldsymbol{z}'_B \boldsymbol{\epsilon}_A) = \boldsymbol{0}$ is implied by the exogeneity of $\boldsymbol{z}_B$ in combination with the RF status of (13) because $\boldsymbol{\epsilon}_A$ is a linear transformation of $\boldsymbol{u} = (u_A, \boldsymbol{u}'_B)'$. This *guarantees that* $\mathsf{E}(\boldsymbol{z}'_B \boldsymbol{x}_A) = \mathsf{E}(\boldsymbol{z}'_B \boldsymbol{z}_B) \boldsymbol{\Pi}_{AB}$ *has rank* N *provided that* $\mathsf{E}(\boldsymbol{z}'_B \boldsymbol{z}_B)$ *has full rank* N. *Such an element is missing in typical 'natural experiment' studies* and other studies being content with using rudimentary models. Using $\boldsymbol{z}_B \boldsymbol{\Pi}_{AB} \, [\equiv \mathsf{E}(\boldsymbol{x}_A | \boldsymbol{z}_B) \equiv \boldsymbol{x}_A - \boldsymbol{\epsilon}_A]$ as 'IV vector' for $\boldsymbol{x}_A$ in (10) (pretending for a moment that $\boldsymbol{\Pi}_{AB}$ is known, and therefore using quotation marks), gives

$$\boldsymbol{\Pi}'_{AB}\mathsf{E}(\boldsymbol{z}'_B y) = \boldsymbol{\Pi}'_{AB}\mathsf{E}(\boldsymbol{z}'_B \boldsymbol{x}_B)\boldsymbol{\beta}_A.$$

Since (13) implies

$$\mathsf{E}(\boldsymbol{z}'_B\boldsymbol{x}_A) = \mathsf{E}(\boldsymbol{z}'_B\boldsymbol{z}_B)\boldsymbol{\Pi}_{AB} \iff \boldsymbol{\Pi}_{AB} = [\mathsf{E}(\boldsymbol{z}'_B\boldsymbol{z}_B)]^{-1}[\mathsf{E}(\boldsymbol{z}'_B\boldsymbol{x}_A)],$$

the 'IV-vector' for $\boldsymbol{x}_A$ can be written as $\boldsymbol{z}_B\boldsymbol{\Pi}_{AB} = \boldsymbol{z}_B[\mathsf{E}(\boldsymbol{z}'_B\boldsymbol{z}_B)]^{-1}[\mathsf{E}(\boldsymbol{z}'_B\boldsymbol{x}_A)]$, so that the coefficient vector of interest can be written as

$$(14) \quad \boldsymbol{\beta}_A = \left\{[\mathsf{E}(\boldsymbol{x}'_A\boldsymbol{z}_B)][\mathsf{E}(\boldsymbol{z}'_B\boldsymbol{z}_B)]^{-1}[\mathsf{E}(\boldsymbol{z}'_B\boldsymbol{x}_A)]\right\}^{-1}\left\{[\mathsf{E}(\boldsymbol{x}'_A\boldsymbol{z}_B)][\mathsf{E}(\boldsymbol{z}'_B\boldsymbol{z}_B)]^{-1}[\mathsf{E}(\boldsymbol{z}'_B y)]\right\}.$$

If (10) is exactly identified, *i.e.*, $M = N$, $\boldsymbol{x}'_A\boldsymbol{z}_B$ quadratic, $\boldsymbol{\Pi}_{AB}$ is quadratic and immaterial, and (14) specializes to

$$(15) \qquad \boldsymbol{\beta}_A = [\mathsf{E}(\boldsymbol{z}'_B\boldsymbol{x}_A)]^{-1}[\mathsf{E}(\boldsymbol{z}'_B y)].$$

Replacing in (14) $\mathsf{E}(\boldsymbol{x}'_A\boldsymbol{z}_B)$, etc., by their empirical counterparts $\boldsymbol{X}'_A\boldsymbol{Z}_B/n$, etc., where n is the number of observations and $\boldsymbol{X}_A$, $\boldsymbol{Z}_B$, etc., are the observation matrices (with n rows) corresponding to the row vectors $\boldsymbol{z}_B$, $\boldsymbol{x}_A$, etc. gives the IV-estimator, applicable under exact identification as well as overidentification,

$$(16) \qquad \widehat{\boldsymbol{\beta}}_A^{IV} = \left\{[\boldsymbol{X}'_A\boldsymbol{Z}_B][\boldsymbol{Z}'_B\boldsymbol{Z}_B]^{-1}[\boldsymbol{Z}'_B\boldsymbol{X}_A]\right\}^{-1}\left\{[\boldsymbol{X}'_A\boldsymbol{Z}_B][\boldsymbol{Z}'_B\boldsymbol{Z}_B]^{-1}[\boldsymbol{Z}'_B y]\right\}.$$

E. Several investigational variables and IVs. Exogenous variables in the equation. Let the equation of interest be the following generalization of (10):

$$(17) \qquad y = \boldsymbol{z}_A\boldsymbol{\gamma}_A + \boldsymbol{x}_A\boldsymbol{\beta}_A + u_A,$$

where $\boldsymbol{z}_A$ is a $(1 \times K)$ vector of exogenous RHS variables, and $\boldsymbol{\gamma}_A$ its $(K \times 1)$ coefficient vector. The equation belongs to a multi-equation model where y, $\boldsymbol{x}_A$ *and* $\boldsymbol{x}_B$ are jointly endogenous. Now both $\boldsymbol{z}_A$ and $\boldsymbol{z}_B$ are assumed exogenous, and both $\boldsymbol{x}_B$ and $\boldsymbol{z}_B$ are excluded from the equation. The exogeneity claim is stronger than the corresponding claim in Example D, as more variables are involved. These assumptions *imply* that the $[1 \times (K+M)]$-vector $(\boldsymbol{z}_A \vdots \boldsymbol{z}_B)$ is a valid IV-vector for $\boldsymbol{x}_A$ relative to (17), satisfying the following generalization of (5) and (11):

$$(18) \quad (\boldsymbol{z}_A \vdots \boldsymbol{z}_B) \text{ is observable}, \quad \text{rank}\{\mathsf{E}[(\boldsymbol{z}_A \vdots \boldsymbol{z}_B)'(\boldsymbol{z}_A \vdots \boldsymbol{x}_A)]\} = K+N, \quad \mathsf{E}[(\boldsymbol{z}_A \vdots \boldsymbol{z}_B)'u_A] = \boldsymbol{0}.$$

First, exogeneity of $(\boldsymbol{z}_A \vdots \boldsymbol{z}_B)$ is a stronger claim than IV-disturbance orthogonality relative to (17). The former implies $\mathsf{E}(\boldsymbol{z}'_A\boldsymbol{u}) = \boldsymbol{0}$, $\mathsf{E}(\boldsymbol{z}'_B\boldsymbol{u}) = \boldsymbol{0}$, while the latter expresses that $\mathsf{E}(\boldsymbol{z}'_A u_A) = \boldsymbol{0}$ and $\mathsf{E}(\boldsymbol{z}'_B u_A) = \boldsymbol{0}$ (the dimensions of the zero vectors and matrices are not indicated and usually differ). Hence, (12) is generalized to

$$(19) \qquad \begin{bmatrix} \mathsf{E}(\boldsymbol{z}'_A\boldsymbol{z}_A) & \mathsf{E}(\boldsymbol{z}'_A\boldsymbol{x}_A) \\ \mathsf{E}(\boldsymbol{z}'_B\boldsymbol{z}_A) & \mathsf{E}(\boldsymbol{z}'_B\boldsymbol{x}_A) \end{bmatrix} \begin{bmatrix} \boldsymbol{\gamma}_A \\ \boldsymbol{\beta}_A \end{bmatrix} = \begin{bmatrix} \mathsf{E}(\boldsymbol{z}'_A y) \\ \mathsf{E}(\boldsymbol{z}'_B y) \end{bmatrix}.$$

Second, the specification of the 'environment' of the equation of interest, (17), allows us to establish the RF equation vector for $\boldsymbol{x}_A$ (similar equations exist for y and $\boldsymbol{x}_B$) which now has the form

$$\text{(20)} \qquad \boldsymbol{x}_A = (\boldsymbol{z}_A \,\vdots\, \boldsymbol{z}_B) \begin{bmatrix} \boldsymbol{\Pi}_{AA} \\ \boldsymbol{\Pi}_{AB} \end{bmatrix} + \boldsymbol{\epsilon}_A, \qquad \mathsf{E}[(\boldsymbol{z}_A \,\vdots\, \boldsymbol{z}_B)'\boldsymbol{\epsilon}_A] = \mathbf{0},$$

where $\boldsymbol{\Pi}_{AA}$ and $\boldsymbol{\Pi}_{AB}$ are $(K\times N)$- and $(M\times N)$-matrices, respectively, and $\boldsymbol{\epsilon}_A$ has the same interpretation as in Example D.[12] Using in (17) $\boldsymbol{z}_A\boldsymbol{\Pi}_{AA} + \boldsymbol{z}_B\boldsymbol{\Pi}_{AB}\,[\equiv \mathsf{E}(\boldsymbol{x}_A|\boldsymbol{z}_A, \boldsymbol{z}_B) \equiv \boldsymbol{x}_A - \boldsymbol{\epsilon}_A]$ as an 'IV vector' for $\boldsymbol{x}_A$ (pretending for a moment that $\boldsymbol{\Pi}_{AA}$ and $\boldsymbol{\Pi}_{AB}$ are known, and therefore again using quotation marks), while letting $\boldsymbol{z}_A$ serve as IV vector for itself, or equivalently, using

$$(\boldsymbol{z}_A \,\vdots\, \boldsymbol{z}_B) \begin{bmatrix} \boldsymbol{I} & \boldsymbol{\Pi}_{AA} \\ \mathbf{0} & \boldsymbol{\Pi}_{AB} \end{bmatrix} \quad \text{as IV vector for} \quad (\boldsymbol{z}_A \,\vdots\, \boldsymbol{x}_A),$$

and hence,

$$\begin{bmatrix} \boldsymbol{I} & \mathbf{0} \\ \boldsymbol{\Pi}'_{AA} & \boldsymbol{\Pi}'_{AB} \end{bmatrix} \begin{bmatrix} \mathsf{E}(\boldsymbol{z}'_A\boldsymbol{z}_A) & \mathsf{E}(\boldsymbol{z}'_A\boldsymbol{x}_A) \\ \mathsf{E}(\boldsymbol{z}'_B\boldsymbol{z}_A) & \mathsf{E}(\boldsymbol{z}'_B\boldsymbol{x}_A) \end{bmatrix} \begin{bmatrix} \boldsymbol{\gamma}_A \\ \boldsymbol{\beta}_A \end{bmatrix} = \begin{bmatrix} \boldsymbol{I} & \mathbf{0} \\ \boldsymbol{\Pi}'_{AA} & \boldsymbol{\Pi}'_{AB} \end{bmatrix} \begin{bmatrix} \mathsf{E}(\boldsymbol{z}'_A y) \\ \mathsf{E}(\boldsymbol{z}'_B y) \end{bmatrix},$$

it finally follows, as a generalization of (14),

$$\text{(21)} \qquad \begin{bmatrix} \boldsymbol{\gamma}_A \\ \boldsymbol{\beta}_A \end{bmatrix} = \left\{ \begin{bmatrix} \boldsymbol{I} & \mathbf{0} \\ \boldsymbol{\Pi}'_{AA} & \boldsymbol{\Pi}'_{AB} \end{bmatrix} \begin{bmatrix} \mathsf{E}(\boldsymbol{z}'_A\boldsymbol{z}_A) & \mathsf{E}(\boldsymbol{z}'_A\boldsymbol{x}_A) \\ \mathsf{E}(\boldsymbol{z}'_B\boldsymbol{z}_A) & \mathsf{E}(\boldsymbol{z}'_B\boldsymbol{x}_A) \end{bmatrix} \right\}^{-1} \times \left\{ \begin{bmatrix} \boldsymbol{I} & \mathbf{0} \\ \boldsymbol{\Pi}'_{AA} & \boldsymbol{\Pi}'_{AB} \end{bmatrix} \begin{bmatrix} \mathsf{E}(\boldsymbol{z}'_A y) \\ \mathsf{E}(\boldsymbol{z}'_B y) \end{bmatrix} \right\}.$$

If (17) is exactly identified, *i.e.*, $M = N$, $\boldsymbol{x}'_A\boldsymbol{z}_B$ quadratic, we know that $\boldsymbol{\Pi}_{AB}$ is quadratic and that $\boldsymbol{\Pi}_{AA}$ and $\boldsymbol{\Pi}_{AB}$ are immaterial, since the partitioned matrices occurring as the first factor in the expressions in curly brackets are block-diagonal and cancel. Then (21) simplifies to the following generalization of (15):

$$\text{(22)} \qquad \begin{bmatrix} \boldsymbol{\gamma}_A \\ \boldsymbol{\beta}_A \end{bmatrix} = \begin{bmatrix} \mathsf{E}(\boldsymbol{z}'_A\boldsymbol{z}_A) & \mathsf{E}(\boldsymbol{z}'_A\boldsymbol{x}_A) \\ \mathsf{E}(\boldsymbol{z}'_B\boldsymbol{z}_A) & \mathsf{E}(\boldsymbol{z}'_B\boldsymbol{x}_A) \end{bmatrix}^{-1} \begin{bmatrix} \mathsf{E}(\boldsymbol{z}'_A y) \\ \mathsf{E}(\boldsymbol{z}'_B y) \end{bmatrix}.$$

Replacing in (21) $\mathsf{E}(\boldsymbol{x}'_A\boldsymbol{z}_B)$, etc., by their empirical counterparts $\boldsymbol{X}'_A\boldsymbol{Z}_B/n$, etc., *and replacing* $\boldsymbol{\Pi}_{AA}$ *and* $\boldsymbol{\Pi}_{AB}$ *with their OLS estimates* obtained from (20) gives the *IV-estimators* of $\boldsymbol{\gamma}_A$ and $\boldsymbol{\beta}_A$, which is applicable under exact identification as well as overidentification.

As is well known, the procedure for IV-estimation just described is equivalent to applying *Two-Stage Least Squares (2SLS)* in estimating a single equation of interest. The equivalence relies on the orthogonality between OLS residuals and OLS fitted values. In general, 2SLS prescribes an OLS estimation of the RF equations for the relevant endogenous variables (first stage) followed by a modified OLS estimation of the equation on interest (second stage), as proposed by Theil (1953, 1954) and Basmann (1957, pp. 80–81) (see Eqs. (20) and (31)–(35) in the latter). This procedure is also strongly related to

Anderson and Rubin (1949, 1950) and Sargan (1958), as well as to Anderson (1974, 2005) for formal expansions on the 'limited information' concept in this context and relationship between 2SLS and Limited Information Maximum Likelihood (LIML). These pioneers, more than 60 years ago, gave 'first-stage' and 'second-stage' a definite meaning relative to an interdependent model. I will later explain how, unfortunately, this usage has been perverted in recent years by proponents of 'natural experiments', contributing to confusion in the applied literature.

3 Examples A–E: Supplementary remarks

The above examples, contrasting identification criteria and IV-definitions, will in this section be expanded by five remarks.

1. Variable exclusion restrictions mixed into IV-definitions: The IV-definitions (2) and (7), related to equations (1) and (6), respectively, are in conflict with definitions used (or implied) by proponents of 'natural experiments' (and in many other studies as well). An example is:[13]

> "A good instrument is correlated with the endogenous regressor for reasons the researcher can verify and explain, but uncorrelated with the outcome variable for reasons *beyond its effect* on the endogenous regressor." (my italics) [Angrist and Krueger (2001, p. 73)]

First, contrast this quotation with Example A, associating "instrument", "endogenous regressor" and "outcome variable" with z, x and y, respectively. Now, *the classical definition of an IV contains no requirement that z be uncorrelated with y "beyond its effect on" x.* Maybe the quotation suggests that a model of an equation of interest, specified as $\mathsf{E}(y|z) = \alpha + \mathsf{E}(x|z)\beta$, is supplemented by, say, $\mathsf{E}(x|z) = \gamma + z\delta$? Then it is true that the order condition (OC) for identification unequivocally says that the absence of z in the expression for $\mathsf{E}(y|z)$, combined with $\delta \neq 0$, ensures identification of β and δ. However, $\delta \neq 0$ *is not part of the definition of z being an IV for x with coefficient β*, because variable inclusion/exclusion is a property of the full model. Next, contrast the quotation with Example C, associating in the quotation "instrument", "endogenous regressor" and "outcome variable" with, respectively, z^*, x_1 and y, the former satisfying (9), in the model $\mathsf{E}(y|z^*, x_2) = \alpha + \mathsf{E}(x_1|z^*, x_2)\beta_1 + x_2\beta_2$ and $\mathsf{E}(x_1|z^*, x_2) = \gamma + z^*\delta_1 + x_2\delta_2$. The OC then unequivocally says that the absence of z^* in the equation for $\mathsf{E}(y|z^*, x_2)$, combined with $\delta_1 \neq 0$, ensures identification of β and δ. However, y is correlated with x_1 'corrected for the impact of z^*', $x_1 - z^*\delta_1$, when β_2 and δ_2 are non-zero. Again, requiring that IV z be uncorrelated with y "beyond the effect of z on the investigational variable" x_1, ("endogenous regressor" in the Angrist-Krueger vocabulary) is not to the the point.[14]

A passage from the same text, Angrist and Krueger (2001, p. 70): "certain "curve shifters", *what we would now call instrumental variables*" (my italics), there quoted with reference to Wright (1928, p. 312),[15] confuses claims to IV and criteria for identification of equations. Other examples are:

> "A valid instrumental variable, which helps determine whether an individual is treated, *but does not determine other factors that affect the outcomes* of interest..." (my italics) [Oreopoulos (2006, p. 152)]

“..... instrumental variables change the incentives for agents to choose a particular level of the treatment *without affecting the potential outcomes* associated with the treatment” (my italics) [Imbens (2014, p. 328)]

“One definition of instrumental variable estimation is the use of *additional “instrumental” variables, not contained in the equation of interest*, to estimate the unknown parameters of that equation ..... two derivations of the instrumental variable *estimator as the solution to the identification problem* .....” (my italics) [Stock and Trebbi (2003, pp. 179-180)]

The Stock-Trebbi quotation – whose last passage moreover perverts the distinction between IV-estimation and identification – is from a paper, “Who invented instrumental variable regression?”, setting out to scrutinize evidence in P. Wright (1928) and S. Wright (1934) (father and son) concerning the origin of the IV approach. However, their findings may equally well be interpreted as answering the question “Who clarified identification as a requirement for decent estimation?” P. Wright (1928) has the following lucid passage on the statistical analysis of simple market models, which obviously reveals his awareness of ‘curve shifting’ as a pathway to identification:[16]

“In the absence of intimate knowledge of demand and supply conditions, statistical methods for imputing fixity to one of the curves while the other changes its position must be based on the introduction of additional factors. Such additional factors may be factors which (A) affect demand conditions without affecting cost conditions or which (B) affect cost conditions, without affecting demand conditions.” [P. Wright (1928, pp. 311–312)]

2. Exclusion restrictions mixed into IV-definitions, examples from medicine: Regarding the IV claims exemplified by (2) and (7), it is of interest to notice misconceptions about the IV requirements in *medical* research studies from the last two decades. This is a relevant digression as strong indications exist that medical applications, sometimes referring to viewpoints from genetics, have formed part of the inspiration of proponents of the ‘natural experiment’ (‘treatment research’) movement in empirical *economics* in the last decades. Even the use of the term ‘treatment’ gives associations to medical doctors’ and nurses’ treatment of patients and their diseases.[17] Typical examples are:

“Instrumental variables estimation uses one or more IVs – observable factors that influence treatment *but do not directly affect patient outcomes to mimic a randomization* of patients to different likelihoods of receiving alternative treatments.” (my italics) [McClellan *et al.* (1994, p. 860)]

“An instrumental variable is a factor that is correlated with the exposure ..... but is not associated with any confounder of the exposure–outcome association, *nor is there any pathway by which the instrumental variable can influence the outcome other than via the exposure of interest*...... an instrumental variable is in some way external to the relationship between the exposure and outcome..... ” (my italics) [Burgess *et al.* (2017, p. 2333)]

“In epidemiological research, the causal effect of a modifiable phenotype or exposure on a disease is often of public health interest. Randomized controlled trials to investigate this effect are not always possible and inferences ..... can be confounded. However, if we know of a gene closely linked to the phenotype without direct effect on the disease, it can often be reasonably assumed that the gene is not itself associated with any confounding factors – a phenomenon called Mendelian randomization. These properties define an instrumental variable and allow estimation of the causal effect, despite the confounding, under certain model restrictions.” [Didelez and Sheehan (2007, p. 309)]

The first quotation exemplifies, once again, the unfortunate mixing of IV requirements and variable exclusion as part of identification criteria. In the second quotation, "exposure-outcome", "confounder" and "instrumental factor" can be associated with, respectively, y, x and z in Example A. However, the lack of both specified model elements and awareness of identification issues is worrying, an impression strengthened by the vague passuses "any pathway" and "in some way external to the relationship". The third quotation, explicitly refers to *Mendelian randomization*, whose origin is in genetics;[18] see also Smith (2010) and Hinke *et al* (2016, Section 2.1).

3. Experimentalists' requirements to IVs: The third remark supplementing the discussion of the examples relates to the way IVs are applied and motivated by adherents to 'natural experiments'. Imbens (2014, pp. 338–339) explains the primary assumptions as follows:

> "There are four key ..... assumptions underlying instrumental variables methods ..... The first assumption concerns the assignment to the instrument ..... *requires that the instrument is as good as randomly assigned* ..... The second ..... *limits or rules out completely direct effects of the assignment* ..... on the outcome, other than through the effect of the assignment on the receipt of the treatment of interest..... This is the most critical ..... and most controversial assumption ..... *sometimes viewed as the defining characteristic of instruments* ..... A third assumption ..... monotonicity ..... rules out the presence of units who always do the opposite of their assignment ..... Finally, we *need the instrument to be correlated with the treatment.*" (my italics)

The "random assignment" and "direct effects ..... completely ruled out" assumptions, are notably stronger accentuated than is the "IV-treatment correlation" assumption, which is mentioned almost as a brief addendum.

What precisely, should be meant by recovering from a "random assignment design" agents' responses to 'stimuli'? Some clarification is provided by:[19]

> "*Random selection* refers to how sample members (study participants) are selected from the population for inclusion in the study. *Random assignment* is an aspect of experimental design in which study participants are assigned to the treatment or control group using a random procedure. Random selection requires the use of some form of random sampling ..... *Random assignment takes place following the selection of participants for the study.* ...... A study using only random assignment could ask ..... to select the [potential respondents which] are most likely to enjoy participating in the study, and the researcher could then randomly assign this sample ..... to the treatment and control groups. In such a design the researcher could draw conclusions about the effect of the intervention *but couldn't make any inference about whether the effect would likely to be found in the population.* ....." (my italics)

This quotation also explains the critical concepts *"internal validity"* and *"external validity"*, crucial in judging how far the applicability of 'natural experiment' results may be stretched:

> "..... the lack of random assignment to be in the treatment or control group would make it impossible to conclude whether the intervention had any effect. *Random selection is thus essential to external validity*, or the extent to which the researcher can use the results of the study to generalize to the larger population. *Random assignment is central to internal validity*, which allows the researcher to make causal claims about the effect of the treatment." (my italics)

Heckman and Urzua (2010) clarify basic differences between the questions posed and answers given by 'structuralists' and 'experimentalists', when using IVs, as follows:[20]

> "The problem that plagues the IV approach is that *the questions it answers are usually defined as probability limits of estimators and not by well-formulated economic problems.* Unspecified "effects" replace clearly defined economic parameters as the objects of empirical interest. ..... for many problems of policy analysis, it is not necessary to identify fully specified structural models with parameters that are invariant to classes of policy modifications ..... All that is required to conduct many policy analyses or to answer many well-posed economic questions are combinations of the structural parameters that are often much easier to identify than the individual parameters themselves. ..... Most IV studies do not clearly formulate the economic question being answered by the IV analysis. *The probability limit of the IV-estimator is defined to be the object of interest.*" (pp. 28, 34, my italics)

This passage, in terms of stringency, contrasts with the rather vague description of the two 'perspectives' on modeling and IV-use in Stock and Watson (2011, p. 448–449):

> "There are two main approaches, which reflect two different perspectives on econometric and statistical modeling. The first approach is to use economic theory to suggest instruments .... The second approach to constructing instruments is to *look for some exogenous source of variation in X arising from what is, in effect, a random phenomenon* that induces shifts in the endogenous regressor" (my italics)

Finally, the reminders and warnings of Rubin, expressed more than 40 years ago, related to a much simpler situation than typical "natural experiment" cases in economics and *with no mention of instrumental variables*, should be observed:

> "..... estimating the typical causal effect of one treatment versus another is a difficult task unless we understand the actual process well enough to (a) assign most of the variability in Y to specific causes and (b) ignore associated but causally irrelevant variables ..... Almost never do we have a random sample from the target population of trials, and thus we must generally rely on the belief in subjective random sampling, that is, there is no important variable that differs in the sample and the target population ..... In both randomized and nonrandomized studies, the investigator should think hard about variables besides the treatment that may causally affect Y and plan in advance..." [Rubin (1974, p. 699-700)]

4. Rudimentary models versus models for limited information inference: While Examples A–C belong to what I call rudimentary models, Examples D and E are models for limited information inference. Example E, like C, has exogenous variables in the equations defining IVs. A claim to models for limited information inference, originating from formal claims in economic theory, known by any student of economics, is that the number of variables to be explained (endogenous variables) should equal the number of equations. On the other hand, in such specifications and unlike most 'theory models', no claim is made that the form of *all* equations should be specified in detail.[21] Nevertheless, in his review, Imbens (2014, p. 329) asserts about simultaneous equations and IVs that:

> "Simultaneous equations are both at the core of the econometrics canon and at the core of the confusion concerning instrumental variables methods in the statistics literature."

This statement, remarkably, follows one page after a statement on IVs, incentives, treatment level and potential outcomes (p. 328)), see also Imbens (2010, p. 403), on the "almost complete lack of instrumental variables in the statistical literature". Angrist and

Pischke (2008, p. 84) express their view by

> "Simultaneous equations models (SEMs) have been enormously important in the history of econometric thought. At the same time, few of today's most influential applied papers rely on an orthodox SEM framework, though the technical language used to discuss IV still comes from this framework.

The motivation for these controversial statements is far from clear. Regarding the first, it has to be said a remarkable number of 'natural experiment' studies are sloppy in counting equations and endogenous variables, refraining from convincing readers that the models intended for data confrontation satisfy the fundamental claim in economic theory of being 'determined'; see Koopmans (1950, p. 393). In the second statement, the passages "today's most influential applied papers", "an orthodox SEM framework" and "technical language" strongly need precise explanations. As it stands, without any references to the multifaceted contributions of founders of classical econometrics and their impact, this statement contains assertions with no basis.

5. Misuse of the term '2SLS' in experimental studies: In the 'natural experiment' literature, often considering setups with only two equations, there seems to be a growing tendency to refer to *two separate model blocks (most often single equations)* – denoted as the "first-stage" and "second-stage model" ("first-stage" and "second-stage equation"), although sometimes knit together via correlated disturbances – handled by separate OLS/IV regressions. This is an unfortunate 'practice', terminologically as well as methodologically, far away from the succinct exposition made by the pioneers of 2SLS (Theil, Basmann, Anderson, Rubin, Sargan) referred at the end of Section 2, in relation to Example E.

A typical case occurs in Angrist *et al.* (1996, p. 445), with binary, endogenous "treatment indicator" for unit i, D_i, determined by an underlying latent, continuous variable, D_i^*. The model having the elements (1) $Y_i = \beta_0 + \beta_1 D_i + \epsilon_i$, (2) $D_i^* = \alpha_0 + \alpha_1 Z_i + \nu_i$, (3) $D_i = 1$ and 0 for $D_i^* > 0$ and $D_i^* \leq 0$, respectively, (4) $\mathsf{E}(Z_i\epsilon_i) = \mathsf{E}(Z_i\nu_i) = 0$, (5) $\text{cov}(D_i, Z_i) \neq 0$. The authors describe it as follows:

> "β_1 represents the causal effect of D on Y..... The assumption that the correlation between ϵ and Z is zero and the absence of Z in Equation (1) captures the notion that any effect of Z on Y must be through an effect of Z on D. This is a key assumption in econometric applications of instrumental variables.... In general D_i ..... is potentially correlated with ϵ_i because ..... ϵ_i and ν_i are potentially correlated. This implies that the receipt of treatment D_i is not ignorable ..... and, in econometric terminology, not exogenous."

The following four quotations, from studies using two-equation applications similar to this prototype and relying on the "key" IV-assumption "that any effect of Z on Y must be through an effect of Z on D", are marred by the same misuse of IV terminology.[22] The first three studies relate to economics (respectively, disability pension influenced by social interaction, disability pension influenced by within-family contagion, females' career influenced by births):

"Operationally, we implement a *two-stage linear probability model (2SLS)*..... The *first-stage equation* predicts the disability participation ..... The *second-stage equation* determines the likelihood that a worker who is employed ..... draws disability....." (my italics) [Rege *et al.* (2012, pp. 1215–1216)]

"We perform two-stage least squares (2SLS) *with equation ..... as the first stage and equation ..... as the second stage*, with the goal of consistently estimating the parameter ......" (my italics) [Dahl *et al.* (2014, p.1723)]

"... we use a *two-stage least squares model* to estimate ..... The first and second-stage regressions are ..... " (my italics) [Lundborg *et al.* (2017, p. 1619)]

In neither of these examples (and several others not quoted here) the reader is informed of how the two "stages" are supposed to be related within a common structural model. As explained in Section 2 in presenting Example E, the two 'model elements' under estimation in the two stages *of the classical form of 2SLS*, *(i) have as their origin the same structural model, and (ii) do not form a recursive or block-recursive system.* The non-recursivity follows from the definitions of the the disturbances involved in the first stage (system of RF equations of the relevant endogenous variable) and the second stage of the method (the structural equation of interest). These two properties are neglected in the above quotations, which invalidates application of the 2SLS label. The fourth study is from medicine and concerns the relationship between infants' birth weight, mothers' smoking and interventions and misuses the term 'recursive model':

"It is implicit in this formulation that intervention has no effect on birth weight other than through its effect on smoking. ..... ϵ_1, and ϵ_2 [disturbances in the equations explaining smoking and birth weight, respectively] *are likely to be correlated.* This is the mathematical formulation of the selection effects: Women who smoke more than others with the same intervention status may tend systematically to have heavier or lighter infants for reasons other than smoking. In this case [smoking] is a random variable correlated with ϵ_2 ..... This is *a simple example of what is known in the econometric literature as a recursive system of simultaneous equations.*" (my italics) [Permutt and Hebel (1989, p. 620)]

4 'IV-regression': remarks on terminology and on some recent texts

This section completes the rather detailed, somewhat technical, discussion of Examples D and E in Section 2, and parts of Section 3, by giving commented quotations from recent texts which describe applications and interpretations of IV and 2SLS procedures, also outside of the 'natural experiment' field.

Stock and Watson (2011, p. 419) explain the idea behind IVs for the case where one endogenous RHS-variable, X, is in an equation with disturbance u, as follows:

"..... think of the variation in X as having two parts: one part that, *for whatever reason*, is correlated with u ..... and a second part that is uncorrelated with u. If you had information that allowed you to isolate the second part, you could focus on those variations in X that are uncorrelated with u and disregard the variations in X that bias the OLS estimates...." (my italics).

This passage might have been quite meaningful if the authors had supplemented the text by the essential information that the equation belongs to a simultaneous model, had introduced its reduced form as an element resembling (13) and had explained the role of the latter in "isolating" the two parts of X. Later (p. 421) they assert:

> "If an instrument is relevant, then the instrument is related to the variation in X. If in addition the instrument is exogenous, then that part of the variation of X captured by the instrumental variable is exogenous. Thus *an instrument that is relevant and exogenous* can capture movements in X that are exogenous. This exogenous variation can in turn be used to estimate the population coefficient.... *The first stage begins with a population regression linking X and Z.*" (my italics).

Four remarks are in order. First, the meaning of "population coefficient" hangs in the air when the equation's status relative to a full model, in structural form, remains unexplained. Second, "capturing exogenous movements", "exogenous variation" and the indicated distinction between IVs that are "relevant and exogenous" and those that are "relevant", but not "exogenous" are equally void of content as long as the text is non-informative about the rest of the model. What about IVs that are "relevant", but not "exogenous"? Third, the passage neglects that classical IV-requirements may well be satisfied for endogenous variables, as will be shown by examples in Section 6 and Appendix A. Fourth, the passage is silent about the status of the "population regression linking X and Z". The meaning of "population" here obviously differs from its meaning in "population coefficient". Both with respect to style and precision of argument this text represents a substantial step backwards relative to the succinct (and far from technically advanced) expositions, in, *e.g.*, Koopmans (1953) and Marschak (1953), which build on the work of the pioneers of econometrics. Angrist and Pischke (2008, p. 94) have a similar passage,[23] which can best be described as exemplifying 'obscurum per obscurius', distinguishing between dependent and "independent endogenous variables":

> "..... "2SLS aficionados live in a world of *mutually exclusive labels*: in any empirical study involving instrumental variables, the random variables to be studied are either dependent variables, *independent endogenous variables*, instrumental variables, or exogenous covariates." (my italics).

Stock and Watson (2011, p. 420) explain terminology by:

> "Instrumental variables regression has some specialized terminology to distinguish variables that are correlated with the population error term u from those that are not. Variables correlated with the error term are called *endogenous variables*, while variables uncorrelated with the error term are called *exogenous variables*. The historical source of these terms traces to models with multiple equations...." (my italics)

while they later (p. 430) describe types of variables in an 'IV regression model' (Y, X, W, Z apparently corresponding to y, $\boldsymbol{z}_A$, $\boldsymbol{z}_A$, $\boldsymbol{z}_B$ in Example E) by:

> "The general IV regression model has four types of variables: the dependent variable, Y; problematic endogenous regressors ..... which are correlated with the error term and which we will label X; additional regressors, called included exogenous variables, which we will label W; and instrumental variables, Z ..... The *relationship* between the number of endogenous variables and the number of endogenous regressors *has its own terminology.* The regression coefficients *are said to be* exactly identified if the number of instruments ..... equals the number of endogenous regressors ..... The coefficients are overidentified if the number of instruments exceed the number of endogenous regressors ..... " (my italics).

These two passages are strongly misconceived, the addendum in the first, "historical source of these terms traces to ....." being almost void of content. First, since exogeneity/endogeneity are model properties, related to a full setup with an equal number of endogenous variables and equations, invoking specific estimation methods in defining them (confer "are said to be") can bring nothing but confusion. Second, suggesting identification, a pre-observation concept, to be related to a particular estimation method is equally misleading. They continue (pp. 433–434) by:

> "When there is one included endogenous variable but multiple instruments, the condition for instrument relevance *is that at least one Z is useful for predicting X given W*. When there are multiple included endogenous variables, this condition is more complicated because we must *rule out perfect multicollinearity in the second-stage population regression.* Intuitively when there are multiple included endogenous variables, the instruments must *provide enough information about the exogenous movements in these variables* to sort out their separate effects on Y."

Even when taking into account that this text is most likely intended for technically non-advanced readers, the lack of stringency of phrases like "perfect multicollinearity in the second-stage population regression" and "enough information about the exogenous movements to sort out....." is remarkable. And why should IV relevance require "at least one Z useful for predicting X given W"?

Angrist and Pischke (2010, p. 12), in the same vein, assert:

> "When using instrumental variables, for example, *it's no longer enough to mechanically invoke a simultaneous equations framework, labeling some variables endogenous and others exogenous*, without substantially justifying the exclusion restrictions and as-good-as-randomly-assigned assumptions that make instruments valid." (my italics)

This raises the questions: What is meant by "mechanically" and what should a "non-mechanically invoked simultaneous equations framework" be? How define exclusion restrictions without specifying a complete model? How circumscribe an IV set for a model's equations without drawing a borderline between its exogenous and endogenous variables?

Deaton (2010, pp.430-431) brings in the term 'externality' (and use a narrower definition of 'exogeneity' than the present one), in a paper concerned with 'natural experiments' as basis for policy recommendations in development economics, and clarifies by:

> "According to Merriam Webster's dictionary, "exogenous" means "caused by factors or an agent from outside the organism or system," ..... However, the consistency of IV-estimation requires that the instrument be orthogonal to the error term ..... in the equation of interest... Heckman (2000) suggests using the term "external" ..... for variables whose values are not set or caused by the variables in the model and keeping "exogenous" for the orthogonality condition that is required for consistent estimation in this instrumental variable context. ..... The main issue, however, is ..... *that we can see when the argument being offered is a justification for externality when what is required is a justification for exogeneity* ..... Failure to separate externality and exogeneity or to build a case for the validity of the exclusion restrictions has caused, and continues to cause, endless confusion....... Whether any of these instruments is exogenous ..... depends on the specification of the equation of interest, and is not guaranteed by its externality. And because exogeneity is an identifying assumption that must be made prior to analysis of the data, empirical tests cannot settle the question." (my italics)

5 Omitted and unobserved variables and identification

This section, organized around three examples that resemble typical setups in the 'natural experiment' literature, has as its primary aim to illustrate how variable exclusion and parameter restrictions may affect a model's identification status and why 'omitted variables' is an elusive concept.[24] Elimination of unobserved variables by combining equations is common in econometric practice, which makes it interesting also to explore the impact of such dimension-reducing changes on coefficient identification and availability of IVs.

Since there are several reasons why variables may disappear from an equation, it seems futile to believe that one remedy may ensure consistent estimation of parameters of interest irrespective of which mechanism is in effect. Even the 'simple' problem of 'omitted regressors', discussed in most econometric textbooks, may in many cases raise identification problems that may not be 'cured' by invoking IVs.[25] Nevertheless, three of the proponents of 'natural experiments' cited above assert, in an optimistic tone:

> "..... a flowering of recent work *uses instrumental variables to overcome omitted variables problems* ..... The instrumental variables methods allow us to estimate the coefficient of interest consistently ..... *without actually having data on the omitted variables or even knowing what they are*." (my italics) [Angrist and Krueger (2001, pp. 72–76)]

> "IV *solves the problem of missing or unknown control variables*, much as a randomized trial obviates the need for extensive controls in a regression." (my italics) [Angrist and Pischke (2008, p. 84)

Reasons why a variable, q, is missing may be: (i) q is genuinely latent, and even the strongest efforts possible cannot make it observable. (ii) q is latent, but we have a theory connecting it to observable variables by an equation. (iii) Theory says that q is irrelevant in the actual equation and should be omitted, while it is relevant in other equations. (iv) Not insisting on having q represented in the model, we decide to eliminate it through dimension-reducing operations. (v) q is endogenous, and rather than specifying a full system of structural equations, we are satisfied by operating on the reduced form (RF) equations of the remaining endogenous variables.

The discussion of the following three example models (intercepts for simplicity omitted throughout), with exogenous, endogenous, and excluded variables indicated – and not least their variants – illustrate why the categorical statements about IVs remedies for handling omitted variables in the last two quotations are problematic.

Model 1: Simplified two-equation model with two exogenous variables.

The first is a five-variables model:

$$
(23)\qquad
\begin{array}{lll}
(a) & y = \beta x + \gamma z + u, & w \text{ excluded}, \\
(b) & z = \mu w + v, & (y,x) \text{ excluded}, \\
 & (x,w) \perp (u,v). & \\
 & \text{exogenous: } x, w; \text{ endogenous: } y, z. &
\end{array}
$$

It satisfies $\text{cov}(z,u)=\sigma_{uv}$, while $y=\beta x+\gamma\mu w+(u+\gamma v)$ and Eq. (b) are the RF equations of y and z. The five variants below exemplify (i) z or w observed versus unobserved,

(ii) (u, v) correlated versus uncorrelated, and (iii) μ being given versus unknown:

Variant 1: y, x, z, w observed, $\sigma_{uv} \neq 0$:
The OC is necessary and implies that β, γ, μ are identified. Using w as IV for z in Eq. (a) is a feasible, consistent procedure. OLS on Eq. (b) is consistent for μ.

Variant 2: y, x, z, w observed, $\sigma_{uv} = 0 \Longrightarrow \mathrm{cov}(z, u) = 0$:
The last restriction, making the model recursive, changes the status of z to becoming conditionally exogenous in Eq. (a). No IV is needed. OLS on Eq. (a) and on Eq. (b) are consistent. Although using w as IV for z is feasible and consistent, OLS on Eq. (a), exploiting the stronger restrictions, is more efficient. Moreover, overidentification prevails, as it follows that $\mathrm{var}(y|x, w) = \sigma_u^2 + \gamma^2\sigma_v^2$ and $\mathrm{var}(z|x, w) = \sigma_v^2$, $\mathrm{cov}(y, z|x, w) = \gamma\sigma_v^2$, and hence γ, can alternatively be obtained as $\gamma = \mathrm{cov}(y, z|x, w)/\mathrm{var}(z|x, w)$.

Variant 3: y, x, w observed, z unobserved:
Since no proxy for z in Eq. (a) exists, as Eq. (b) is uninformative, β and $\gamma\mu$ are identified from the RF of y. Therefore, if μ is known, γ is identified, and vice versa. The intuitive explanation for the former is that μ known and w observed implies $\mathsf{E}(z|w)$ known.

Variant 4: y, x, z observed, w unobserved, $\sigma_{uv} \neq 0$:
This variant, relative to variant 3, switches the observation status of z and w. The lack of w-observations implies that there will be no IV for z in Eq. (a), still Eq. (b) is uninformative. As a consequence, β, γ, μ are unidentified according to the OC, and because $\mathrm{cov}(z, u) = \sigma_{uv} \neq 0$, OLS on Eq. (a) is inconsistent.

Variant 5: y, x, z observed, w unobserved, $\sigma_{uv} = 0$:
The last restriction makes the model recursive and, relative to variant 4, changes the status of z to becoming conditionally exogenous in Eq. (a). This makes β and γ identified, while μ is unidentified. No IV for z is needed, the recursivity makes OLS on Eq. (a) consistent. This is a case where lack of observations on a potential IV is (partly) compensated by the covariance restriction.

Model 2: Three-equation model with one exogenous variable.

This is a five-variables model which can be viewed as a modification of Model 1 with Eq. (b) replaced by two equations, Eqs. (b) and (c):

$$
(24)\qquad
\begin{array}{lll}
(a) & y = \beta x + \gamma z + u, & w \text{ excluded}, \\
(b) & w = bz + \epsilon, & (x, y) \text{ excluded}, \\
(c) & z = \lambda x + v, & (y, w) \text{ excluded}, \\
 & x \perp (u, v, \epsilon). & \\
 & \text{exogenous: } x, \text{ endogenous: } y, z, w. &
\end{array}
$$

The RF equations of y, w and z are: $y = (\beta + \gamma\lambda)x + (u + \gamma v)$, $w = b\lambda x + (\epsilon + bv)$, and Eq. (c). The four variants below exemplify (i) z observed versus unobserved, (ii) w observed versus unobserved, (iii) disturbances correlated versus uncorrelated, and (iv) λ being given versus unknown.

Variant 1: y, x, z, w observed:
The OC is necessary and implies that $\beta, \gamma, \lambda, b$ are identified. OLS on Eq. (c) is consistent for λ, and using x as IV for z in Eq. (b) is consistent for b, while OLS on Eq. (b) is inconsistent when $\sigma_{v\epsilon} \neq 0$.

Variant 2: y, x, z observed, w unobserved, $\sigma_{uv} \neq 0$:

With observations on w unavailable, which makes Eq. (b) uninformative, the OC implies that β, γ, b are unidentified. On the other hand, λ is identified.

Variant 3: y, x, z observed, w unobserved, $\sigma_{uv} = 0$:

The last restriction makes the model recursive and, relative to variant 2, changes the status of z to becoming conditionally exogenous in Eq. (a), while Eq. (b) uninformative. Therefore, β, γ, λ are identified and can be consistently estimated by OLS, while b is unidentified. Briefly, the lack of w observations is (partly) compensated by the covariance restriction.

Variant 4: y, x, w observed, z unobserved:

This variant, relative to variant 2, switches the observation status of z and w. Since no proxy for z in Eq. (a) exists, Eq. (c) being uninformative, it follows that $\beta + \gamma\lambda$ and $b\lambda$ are identified from the RF of y and w. Hence, λ known makes b identified and vice versa. The intuitive explanation for the former is that λ known and w observed implies $\mathsf{E}(z|x)$ known. Moreover, (b, β) known makes (λ, γ) identified, and vice versa.

Model 3: Two-equation model with two exogenous variables

This is another two-equation, five-variables model:

$$\text{(25)} \qquad \begin{array}{lll} (a) & y = \beta x + \gamma z + u, & w \text{ excluded}, \\ (b) & z = \lambda x + \mu w + v, & y \text{ excluded}. \\ & (x, w) \perp (u, v). & \\ & \text{exogenous: } x, w, \text{ endogenous: } y, z. & \end{array}$$

It satisfies $\text{cov}(z, u) = \sigma_{uv}$, while $y = (\beta + \gamma\lambda)x + \gamma\mu w + (u + \gamma v)$ and Eq. (b) are the RF equations of y and z, respectively. The three variants below exemplify: (i) z observed versus unobserved, (ii) disturbances (u, v) correlated versus uncorrelated, and (iii) μ prescribed versus unknown:

Variant 1: y, x, z, w observed, $\sigma_{uv} \neq 0$:

The OC is necessary and implies that $\beta, \gamma, \lambda, \mu$ are identified. Using w as IV for z in Eq. (a) is a feasible, consistent procedure. OLS on Eq. (b) is consistent.

Variant 2: y, x, z, w observed, $\sigma_{uv} = 0 \Longrightarrow \text{cov}(z, u) = 0$:

The model is recursive. z is conditionally exogenous in Eq. (a). OLS on Eq. (a) and on Eq. (b) is consistent. No IV is needed. Overidentification prevails, as it follows that $\text{var}(y|x, w) = \sigma_u^2 + \gamma^2\sigma_v^2$ and $\text{var}(z|x, w) = \sigma_v^2$, $\text{cov}(y, z|x, w) = \gamma\sigma_v^2$. Hence, γ can alternatively be obtained as $\gamma = \text{cov}(y, z|x, w)/\text{var}(z|x, w)$.

Variant 3: y, x, w observed, z unobserved:

Eq. (b) is uninformative, $\beta + \gamma\lambda$ and $\gamma\mu$ are identified, z being eliminated by inserting Eq. (b). $\beta, \gamma, \lambda, \mu$ are unidentified. μ known makes γ identified from RF of y, because with observations on w available, lack of z-observations is compensated by knowledge of coefficients. Further, (λ, μ) known makes (β, γ) identified from RFs of y and of z. The intuitive explanation is that (λ, μ) known and (x, w) observed imply $\mathsf{E}(z|x, w)$ known.

6 Latent variables, omitted variables and IV-motivating models

Next, three versions of regression models will be considered, for which error-free measures of some variables are unavailable while latent variables connected to observed ones through relationships may be a recourse. Omitted variables may well be latent variables or variables represented by proxies. For each model and model transformation, the excluded variables, and hence their possible connection to exogeneity and implied IVs, will be indicated.[26]

Model 4: Simple EIV model with a proxy variable equation: Consider a *structural* errors-in-variables (EIV) model where y is connected to x (error-ridden) via the latent variable, ξ, where ξ is connected to z (observable) via a 'proxy-variable-equation', in the terminology of Leamer (1978, pp. 245, 251). The model is[27]

$$\begin{aligned}
&y = \beta\xi + u, && (x, z, \nu, \epsilon) \text{ excluded},\\
&x = \xi + \nu, && (y, x, u, \epsilon) \text{ excluded},\\
&z = b\xi + \epsilon, && (y, x, u, \nu) \text{ excluded},\\
&u \perp \nu \perp \epsilon \perp \xi, &&\\
&\text{observed, endogenous: } y, x, z, &&\\
&\text{latent, exogenous: } \xi. &&
\end{aligned} \tag{26}$$

All observed variables are endogenous, the only exogenous variable, ξ, is latent. All equations are in RF format, with uncorrelated disturbances. Elimination of ξ leads to the two-equation 'reduced form counterpart'[28] in observed variables:

$$\begin{aligned}
&y = \beta x + (u - \beta\nu), \quad && z \perp (u-\beta\nu), \quad && (z, \epsilon) \text{ excluded},\\
&z = bx + (\epsilon - b\nu), && y \perp (\epsilon - b\nu), && (y, u) \text{ excluded}.
\end{aligned} \tag{27}$$

While the disturbances in (26) are uncorrelated, the disturbances in (27), which contain the errors ν and ϵ, are correlated and are also correlated with the respective regressors. *The status of (27) with respect to excluded variables differs from that of (26).* The error/disturbance variances apart, the model has 3 parameters, (β, b, σ_ξ^2) and 3 *implied covariance equations* between the observable variables:

$$\frac{\text{cov}(y, x)}{\beta} = \frac{\text{cov}(z, x)}{b} = \frac{\text{cov}(y, z)}{\beta b} = \sigma_\xi^2. \tag{28}$$

Identification is ensured, as

$$\begin{aligned}
&\beta = \text{cov}(z, y)/\text{cov}(z, x), \quad \text{motivating } z \text{ as IV for } x \text{ to estimate } \beta,\\
&b = \text{cov}(y, z)/\text{cov}(y, x), \quad \text{motivating } y \text{ as IV for } x \text{ to estimate } b.\\
&\sigma_\xi^2 = \text{cov}(y, x)/\beta = \text{cov}(z, x)/b = \text{cov}(y, x)\text{cov}(x, z)/\text{cov}(y, z).
\end{aligned} \tag{29}$$

The IVs valid for the equations in (27) are different, endogenous and omitted from the respective equations. There are no 'exogenous IVs'. *The validity of the IVs and the variables exclusion in (27) are implied by the full model (26).* These properties do not emerge if its equations are examined separately. Two variants are of interest.

Variant 4a: Replace in (26) $x = \xi + \nu$ by $x = d\xi + \nu$ (d unknown), *i.e.*, replace a simple measurement error equation for ξ by a less restrictive proxy variable equation. Then (28) changes to:

$$\frac{\text{cov}(y, x)}{\beta d} = \frac{\text{cov}(z, x)}{bd} = \frac{\text{cov}(y, z)}{\beta b} = \sigma_\xi^2,$$

which is equivalent to

$$\frac{\text{cov}(y,x)}{\beta^*} = \frac{\text{cov}(z,x)}{b^*} = \frac{\text{cov}(y,z)}{\beta^* b^*} = \sigma_\xi^{*2},$$

with $\beta^* = \beta/d$, $b^* = b/d$ and $\sigma_\xi^* = d\sigma_\xi$, the latter expression having the same form as (28). Defining $\xi^* = d\xi$, a model of the same form as (26) with (ξ, β, b) replaced by (ξ^*, β^*, b^*) in (29) emerges, in which only (β^*, b^*) can be identified. The IVs and the exclusion properties are the same as in the main model.

Variant 4b: Allow for correlation between the error elements by replacing $u \perp \nu \perp \epsilon \perp \xi$ by $(u, \nu, \epsilon) \perp \xi$, denoting the covariances by $\sigma_{u\nu}, \sigma_{u\epsilon}, \sigma_{\nu\epsilon}$, respectively. This makes the reduced form (RF) disturbances correlated both when the RF is expressed in terms of latent variables and in terms of observed variables. It follows that $\text{cov}(z,x) = b\sigma_\xi^2 + \sigma_{\nu\epsilon}$ and $\text{cov}(z, u-\beta\nu) = \sigma_{u\epsilon} - \beta\sigma_{\nu\epsilon}$, while (28) changes to:

$$\frac{\text{cov}(y,x) - \sigma_{u\nu}}{\beta} = \frac{\text{cov}(z,x) - \sigma_{\nu\epsilon}}{b} = \frac{\text{cov}(y,z) - \sigma_{u\epsilon}}{\beta b} = \sigma_\xi^2.$$

Hence, $b \neq 0$ (and $b \neq -\sigma_{\nu\epsilon}/\sigma_\xi^2$) suffices for satisfying the IV-requirement $\text{cov}(z,x) \neq 0$, while $\beta\sigma_{\nu\epsilon} \neq 0$ (and $\sigma_{u\epsilon} \neq \beta\sigma_{\nu\epsilon}$) violates $\text{cov}(z, u-\beta\nu) = 0$. Therefore, allowing for correlated errors violates (27) and (28), as the model now has 6 parameters and only 3 implied covariance equations. This destroys identification of β and b and destroys the validity, respectively, of z as IV for x in estimating β, and y as IV for x in estimating b.

Model 5: Two-equation EIV model with no proxy variable equation[29]

$$(30) \quad \begin{array}{ll} y_1 = \beta_1\xi + u_1, & (y_2, x, u_2, \nu) \text{ excluded}, \\ y_2 = \beta_2\xi + u_2, & (y_1, x, u_1, \nu) \text{ excluded}, \\ x = \xi + \nu, & (y_1, y_2, u_1, u_2) \text{ excluded}, \\ u_1 \perp u_2 \perp \nu \perp \xi, & \\ \text{observed, endogenous: } y_1, y_2, x, & \\ \text{latent, exogenous: } \xi. & \end{array}$$

Eliminating ξ gives:

$$(31) \quad \begin{array}{lll} y_1 = \beta_1 x + (u_1 - \beta_1\nu), & y_2 \perp (u_1 - \beta_1\nu), & (y_2, u_2) \text{ excluded}, \\ y_2 = \beta_2 x + (u_2 - \beta_2\nu), & y_1 \perp (u_2 - \beta_2\nu), & (y_1, u_1) \text{ excluded}. \end{array}$$

While the disturbances in (30) are uncorrelated, the (composite) disturbances in (31) are correlated and correlated with the respective regressors. *The status of (31) with respect to excluded variables differs from that of (30).* The error/disturbance variances apart, the model has 3 parameters, $(\beta_1, \beta_2, \sigma_\xi^2)$, and 3 implied covariance equations between the observable variables:

$$(32) \quad \frac{\text{cov}(x,y_1)}{\beta_1} = \frac{\text{cov}(x,y_2)}{\beta_2} = \frac{\text{cov}(y_1,y_2)}{\beta_1\beta_2} = \sigma_\xi^2.$$

It follows that identification is ensured, as

$$\begin{aligned}
&\beta_1 = \mathrm{cov}(y_1, y_2)/\mathrm{cov}(x, y_2), \quad \text{motivating } y_2 \text{ as IV for } x \text{ to estimate } \beta_1,\\
(33) \qquad &\beta_2 = \mathrm{cov}(y_2, y_1)/\mathrm{cov}(x, y_1), \quad \text{motivating } y_1 \text{ as IV for } x \text{ to estimate } \beta_2,\\
&\sigma_\xi^2 = \mathrm{cov}(x, y_1)/\beta_1 = \mathrm{cov}(x, y_2)/\beta_2 = \mathrm{cov}(x, y_1)\mathrm{cov}(x, y_2)/\mathrm{cov}(y_1, y_2).
\end{aligned}$$

Two variants are of interest.

Variant 5a: Assume that we know that $\beta_1+\beta_2=1$. This replaces the second equation of (30) by $y_2=(1-\beta_1)\xi+u_2$ and implies $y_1+y_2=\xi+u_1+u_2$. Consequently, *not only x, but also* y_1+y_2 emerges as an error-ridden measure of ξ. Then (32) becomes

$$\frac{\mathrm{cov}(y_1, x)}{\beta_1}=\frac{\mathrm{cov}(y_2, x)}{1-\beta_1}=\frac{\mathrm{cov}(y_1, y_2)}{\beta_1(1-\beta_1)}=\sigma_\xi^2,$$

implying

$$\begin{aligned}
\beta_1 &= \frac{\mathrm{cov}(y_1, y_2)}{\mathrm{cov}(x, y_2)} = \frac{\mathrm{cov}(x, y_1)}{\mathrm{cov}(x, y_1+y_2)},\\
\beta_2 &= 1-\beta_1 = \frac{\mathrm{cov}(y_1, y_2)}{\mathrm{cov}(x, y_1)} = \frac{\mathrm{cov}(x, y_2)}{\mathrm{cov}(x, y_1+y_2)},\\
\sigma_\xi^2 &= \mathrm{cov}(x, y_1+y_2) = \frac{\mathrm{cov}(x, y_1)\mathrm{cov}(x, y_2)}{\mathrm{cov}(y_2, y_1)}.
\end{aligned}$$

The interpretation is the following: Utilizing the restriction $\beta_1+\beta_2=1$ and inserting $\xi=x-v$ and $\xi=y_1+y_2-(u_1+u_2)$ in the first equation of (30) give, respectively,

$$\begin{aligned}
y_1 &= \beta_1 x + (u_1-\beta_1\nu),\\
y_1 &= \beta_1(y_1+y_2) + (1-\beta_1)u_1-\beta_1 u_2.
\end{aligned}$$

The implied overidentification changes the 'match' between IV and investigational variable: IVs for estimating β_1 are, respectively, y_2 for x and x for y_1+y_2 (or equivalently, IVs for estimating β_2 are, respectively, y_1 for x and x for y_1+y_2).

Variant 5b: Again, impose $\beta_1+\beta_2=1$ and in addition replace $x=\xi+v$ by $x=d\xi+v$. Then the overidentification disappears, and (32) changes to

$$\frac{\mathrm{cov}(x, y_1)}{\beta_1 d}=\frac{\mathrm{cov}(x, y_2)}{(1-\beta_1)d}=\frac{\mathrm{cov}(y_1, y_2)}{\beta_1(1-\beta_1)}=\sigma_\xi^2,$$

and hence,

$$\begin{aligned}
&d\sigma_\xi^2 = \mathrm{cov}(x, y_1+y_2), \quad \frac{\beta_1}{d} = \frac{\mathrm{cov}(y_2, y_1)}{\mathrm{cov}(y_2, x)},\\
&\beta_1 = \frac{\mathrm{cov}(x, y_1)}{\mathrm{cov}(x, y_1+y_2)}, \quad \beta_2 = 1-\beta_1 = \frac{\mathrm{cov}(x, y_2)}{\mathrm{cov}(x, y_1+y_2)},\\
&d = \frac{\mathrm{cov}(x, y_1)\mathrm{cov}(x, y_2)}{\mathrm{cov}(x, y_1+y_2)\mathrm{cov}(y_1, y_2)},\\
&\sigma_\xi^2 = \frac{[\mathrm{cov}(x, y_1+y_2)]^2\mathrm{cov}(y_1, y_2)}{\mathrm{cov}(x, y_1)\mathrm{cov}(x, y_2)}.
\end{aligned}$$

The interpretation is the following: Eliminating ξ from the first equation of (31) by inserting $\xi=(x-\nu)/d$ and $\xi=y_1+y_2-(u_1+u_2)$ gives, respectively,

$$y_1 = (\beta_1/d)x + [u_1 - (\beta_1/d)\nu],$$
$$y_1 = \beta_1(y_1 + y_2) + [(1-\beta_1)u_1 - \beta_1 u_2].$$

Once again, the 'match' between IV, investigational variable and relevant parameter, as well as the exclusion properties changes. Now y_2 is an IV for x in estimating β_1/d, x is an IV for y_1+y_2 in estimating β_1. Equivalently, y_2 is an IV for y_1 in estimating d/β_1 from the (inverted) first equation, x is an IV for y_1 in estimating $1/\beta_1$ from the (inverted) second equation. Hence, although d cannot be estimated by an IV-estimator directly, it can be estimated as the ratio between two IV-estimators one for β_1, one for β_1/d, or equivalently, as the ratio between the IV-estimators for d/β_1 and $1/\beta_1$.

Model 6: Two-equation EIV model with a proxy variable equation: We augment Model 5 by an IV, z, and its 'IV equation', introducing z as a proxy for ξ, to obtain

$$\begin{aligned}
&y_1 = \beta_1\xi + u_1, && (y_2, x, z, u_2, \nu, \epsilon) \text{ excluded},\\
&y_2 = \beta_2\xi + u_2, && (y_1, x, z, u_1, \nu, \epsilon) \text{ excluded},\\
&x = \xi + \nu, && (y_1, y_2, z, u_1, u_2, \epsilon) \text{ excluded},\\
&z = b\xi + \epsilon, && (y_1, y_2, x, u_1, u_2, \nu) \text{ excluded},\\
&u_1 \perp u_2 \perp \epsilon \perp \nu \perp \xi, &&\\
&\text{observed, endogenous: } y_1, y_2, x, z, &&\\
&\text{latent, exogenous: } \xi. &&
\end{aligned} \tag{34}$$

Eliminating ξ, we obtain the system, extending (31):

$$\begin{aligned}
&y_1 = \beta_1 x + (u_1 - \beta_1\nu), && (y_2, z) \perp (u_1 - \beta_1\nu), && (y_2, z, u_2, \epsilon) \text{ excluded},\\
&y_2 = \beta_2 x + (u_2 - \beta_2\nu), && (y_1, z) \perp (u_2 - \beta_2\nu), && (y_1, z, u_1, \epsilon) \text{ excluded},\\
&z = bx + (\epsilon - b\nu), && (y_1, y_2) \perp (\epsilon - b\nu), && (y_1, y_2, u_1, u_2) \text{ excluded}.
\end{aligned} \tag{35}$$

The model has 4 parameters, $(\beta_1, \beta_2, b, \sigma_\xi^2)$, 6 implied covariance equations, which signalize overidentification, and imply that (32) is extended to

$$\begin{aligned}
&\frac{\text{cov}(y_1, x)}{\beta_1} = \frac{\text{cov}(y_2, x)}{\beta_2} = \frac{\text{cov}(y_1, y_2)}{\beta_1\beta_2} =\\
&\frac{\text{cov}(y_1, z)}{\beta_1 b} = \frac{\text{cov}(y_2, z)}{\beta_2 b} = \frac{\text{cov}(x, z)}{b} = \sigma_\xi^2.
\end{aligned} \tag{36}$$

It follows that

$$\begin{aligned}
\beta_1 &= \text{cov}(y_1, y_2)/\text{cov}(x, y_2) = \text{cov}(y_1, z)/\text{cov}(x, z),\\
\beta_2 &= \text{cov}(y_2, y_1)/\text{cov}(x, y_1) = \text{cov}(y_2, z)/\text{cov}(x, z),\\
b &= \text{cov}(z, y_1)/\text{cov}(x, y_1) = \text{cov}(z, y_2)/\text{cov}(x, y_2),\\
\sigma_\xi^2 &= \text{cov}(y_1, x)/\beta_1 = \text{cov}(y_2, x)/\beta_2 = \text{cov}(x, z)/b.
\end{aligned} \tag{37}$$

We now proceed by estimating β_1 by using y_2 or z (both endogenous) as IV for x, estimating β_2 by using y_1 or z (both endogenous) as IV for x, and estimating b by using y_1 or y_2 (both endogenous) as IV for x. Two variants are of interest.

Variant 6a: Replacing the measurement error equation $x = \xi + \nu$ by the proxy variable equation $x = d\xi + \nu$ leads to 6 covariance equations in 5 parameters, $(\beta_1, \beta_2, b, d, \sigma_\xi^2)$, extending (36) to:[30]

$$\frac{\text{cov}(x,y_1)}{\beta_1 d} = \frac{\text{cov}(y_2,x)}{\beta_2 d} = \frac{\text{cov}(y_1,y_2)}{\beta_1\beta_2} =$$
$$\frac{\text{cov}(z,y_1)}{\beta_1 b} = \frac{\text{cov}(y_2,z)}{\beta_2 b} = \frac{\text{cov}(x,z)}{bd} = \sigma_\xi^2.$$

'Full overidentification' of $\beta_1^* = \beta_1/d$, $\beta_2^* = \beta_2/d$ and $b^* = b/d$ is obtained since:

$\beta_1/d = \text{cov}(y_1,y_2)/\text{cov}(x,y_2) = \text{cov}(y_1,z)/\text{cov}(x,z)$,
$\beta_2/d = \text{cov}(y_2,y_1)/\text{cov}(x,y_1) = \text{cov}(y_2,z)/\text{cov}(x,z)$,
$b/d = \text{cov}(z,y_1)/\text{cov}(x,y_1) = \text{cov}(z,y_2)/\text{cov}(x,z)$.

When d is unknown, β_1, β_2 and b are unidentified.

Variant 6b: Imposing $\beta_1+\beta_2=1$ while replacing the measurement error equation $x = \xi+v$ with the proxy variable equation $x=d\xi+v$ results in 6 covariance equations in 4 parameters $\beta_1, b, d, \sigma_\xi^2$, changing (36) to:

$$\frac{\text{cov}(x,y_1)}{\beta_1 d} = \frac{\text{cov}(x,y_2)}{(1-\beta_1)d} = \frac{\text{cov}(y_1,y_2)}{\beta_1(1-\beta_1)} =$$
$$\frac{\text{cov}(z,y_1)}{\beta_1 b} = \frac{\text{cov}(z,y_2)}{(1-\beta_1)b} = \frac{\text{cov}(x,z)}{bd} = \sigma_\xi^2.$$

It follows that

$$\sigma_\xi^2 = \frac{\text{cov}(x,y_1+y_2)}{d} = \frac{\text{cov}(z,y_1+y_2)}{b},$$
$$\frac{\beta_1}{d} = \frac{\text{cov}(y_2,y_1)}{\text{cov}(y_2,x)}, \quad \frac{\beta_1}{b} = \frac{\text{cov}(y_2,y_1)}{\text{cov}(y_2,z)},$$

and therefore,

$$b = \frac{\text{cov}(x,z)}{\text{cov}(x,y_1+y_2)},$$
$$d = \frac{\text{cov}(z,x)}{\text{cov}(z,y_1+y_2)},$$
$$\beta_1 = \frac{\text{cov}(x,y_1)}{\text{cov}(x,y_1+y_2)} = \frac{\text{cov}(z,y_1)}{\text{cov}(z,y_1+y_2)},$$
$$\sigma_\xi^2 = \frac{\text{cov}(x,y_1+y_2)\text{cov}(z,y_1+y_2)}{\text{cov}(x,z)}.$$

The interpretation is the following: Inserting $\xi=y_1+y_2-(u_1+u_2)$ in the first, third and fourth equations of (34) gives, respectively,

$$\begin{aligned} y_1 &= \beta_1(y_1+y_2) + (1-\beta_1)u_1 - \beta_1 u_2, \\ z &= b(y_1+y_2) + (1-b)u_1 - b\nu, \\ x &= d(y_1+y_2) + (1-d)u_1 - d\nu, \end{aligned}$$

for which admissible IVs for y_1+y_2 for estimating β_1, b and d are, respectively, (x,z), (y_2,x), and (y_2,z). The restriction $\beta_1+\beta_2=1$ ensures identification of β_1, β_2 and b.

This discussion of Models 4–6, with variants, motivates important *conclusions*: [1] Endogenous variables may serve as IVs for equations in a model with no error correlation across equations. [2] 'Omitted variables' is an elusive concept, as for example, its interpretation changes when latent structural variables are eliminated and the equations written in terms of their manifest counterparts or when coefficients are subject to transformations. The meaning of 'omitted' is clearly context-specific. [3] Whether cross-equation error correlation is present or absent, crucially affects identification and availability of valid IVs, as is illustrated by contrasting Model 4 with variant 4b. [5] The parameters or parameter combinations for which an IV is valid may not be invariant to the model's identification status, compare Models 5 and 6 with variants 5a–5b, and 6a–6b. [6] Imposition of parameter restrictions may crucially affect the set of valid IVs, as illustrated by comparing Models 5 and 6 with variants 5a–5b and 6a–6b. [7] Imposing parameter restrictions that make some parameters overidentified, may well leave other parameters unidentified, as illustrated by comparing Models 4 and 6 with variants 4a and 6a. [8] Replacing a measurement error equation with a 'proxy variable equation', or vice versa, will change the 'metric' of some variables' and affect the identification status of coefficients and the way IVs and coefficients are connected. This is illustrated by contrasting Examples 4, 5 and 6 with variants 4a, 5b and 6a–6b.

Two extensions of Model 6 will be described in Appendix A, one with a latent exogenous variable added, and one with two latent exogenous variables and with both measurement error mechanisms and proxy variables mechanisms represented. An interdependent two-equation model, fully specified, with latent variables, which may also be considered an extension of the models in the present section, is considered in Appendix B.

7 Instrumental variables in dynamic models

In the previous sections, the core topics identification, IVs, omitted variables, and their interconnection, have been discussed within static contexts. Since their interest also extends to dynamic models, three versions of an autoregressive model in observed variables are well worth considering. Dynamic models were among the first examples for which the usefulness of IVs were demonstrated, see Reiersøl (1941, 1945).[31] Therefore it is more than remarkable that Angrist and Krueger (2001, p.76) ignore this literature, with the following tirade:

> "Indeed, one of the most mechanical and naive, yet common, approaches to the choice of instruments *uses atheoretical and hard-to-assess assumptions about dynamic relationships* to construct instruments from lagged variables in time series or panel data. The use of lagged endogenous variables as instruments is problematic if the equation error or omitted variables are serially correlated." (my italics)

The phrases "naive, yet common, approaches" and "atheoretical and hard-to-assess assumptions" definitely miss the point. This can be seen from the following models:

$$(38)\qquad y_t = \beta x_t + \gamma y_{t-1} + u_t, \quad |\gamma| < 1, \quad (u_t|X) \sim \mathsf{IID}(0, \sigma_u^2),$$

$$(39)\qquad y_t = \beta x_t + \gamma y_{t-1} + u_t, \quad |\gamma| < 1, \quad u_t = v_t + \theta v_{t-1},\ (v_t|X) \sim \mathsf{IID}(0, \sigma_v^2),$$

$$(40)\qquad y_t = \beta x_t + \gamma y_{t-1} + u_t, \quad |\gamma| < 1, \quad u_t = \rho u_{t-1} + v_t,\ |\rho| < 1,\ (v_t|X) \sim \mathsf{IID}(0, \sigma_v^2),$$

whose disturbances are, respectively, white noise, MA(1) and AR(1) processes, and X denotes the vector of x_ts observed. Symbolizing by L the lag-operator, it follows that

$$(41)\qquad y_t = \sum_{i=0}^{\infty} \gamma^i(\beta x_{t-i}+u_{t-i}) = \frac{\beta x_t+u_t}{1-\gamma\mathsf{L}},$$

$$(42)\qquad y_t = \sum_{i=0}^{\infty} \gamma^i(\beta x_{t-i} + v_t+\theta v_{t-1}) = \frac{\beta x_t + (1+\theta\mathsf{L})v_t}{1-\gamma\mathsf{L}},$$

$$(43)\qquad y_t = \sum_{i=0}^{\infty} \gamma^i(\beta x_{t-i} + \sum_{j=0}^{\infty} \rho^j v_{t-i-j}) = \frac{\beta x_t}{1-\gamma\mathsf{L}} + \frac{v_t}{(1-\gamma\mathsf{L})(1-\rho\mathsf{L})}.$$

Consistent estimation of the coefficients of Model (38) requires no IV; OLS ensures consistency, even though $\text{cov}(y_{t-1}, u_{t-s}) \neq 0$ for $s=1,2,\ldots$. This follows, according to the famous theorem of Mann and Wald (1943), because y_t remembers x_t and u_t infinitely long back in time, while u_t has no memory.

In Model (39) u_t has a one period memory. It follows from (42) that IV-candidates for y_{t-1} are $x_{t-1}, x_{t-2}, \ldots$ and $y_{t-2}, y_{t-3}, \ldots$. They are all omitted from the equation, are correlated with y_{t-1}, and are uncorrelated with u_t. Linear combinations of $x_{t-1}, x_{t-2}, \ldots$ and $y_{t-2}, y_{t-3}, \ldots$ are also valid IVs for y_{t-1}. The longer the lag, the smaller is the number of effective observations (and degrees of freedom) available for application of such IV procedures. The set $x_{t-1}, x_{t-2}, y_{t-2}$ includes the IVs which give the smallest loss of degrees of freedom.

In Model (40) u_t has an infinite memory. It follows from (43) that IVs-candidates for y_{t-1} are $x_{t-1}, x_{t-2}, \ldots$. They are all omitted from the equation, are correlated with y_{t-1}, and uncorrelated with u_t. Linear combinations of $x_{t-1}, x_{t-2}, \ldots$ are also valid IVs for y_{t-1}. The longer the lag, the smaller is the number of effective observations (and degrees of freedom) available for application of such IV procedures. The set x_{t-1}, x_{t-2} includes the IVs which give the smallest loss of degrees of freedom.

These conclusions clearly refute the assertion in the Angrist-Krueger quotation at the beginning of the section. The suggestions for IV candidates for the two last cases could not have been obtained if the model of the disturbance processes had not been specified. The relevant orthogonality conditions are testable.

A latent variables extension of Model (40) is considered in Appendix C.

8 Concluding remarks

The primary conclusions, or reminders, can be summarized as follows:

First, identification is a model concept, IV-procedures belong to methods for statistical inference. The existence of IVs should not be invoked in *defining* identification criteria.

Second, an IV is a variable, which is *theoretically* correlated with some variables in an equation and uncorrelated with others. This theoretical correlation should be founded on some economic theory cast in econometric terms. The classical IV-definition, applicable to a class of linear multi-equation models, has a rank condition and an orthogonality condition, of equal importance.

Third, exclusion restrictions belong to classical criteria for identification. Considering them as part of IV-definitions, as has seemingly become common in parts of the 'natural experiment' literature, is not recommendable because its disguises the crucial distinction between model elements and elements in methods, which is basic to empirical research.

Fourth, the contrast between 'rudimentary models' and models for 'limited information inference', referring to specifications containing full lists of exogenous and endogenous variables, without requiring all unknown coefficients to be listed equation by equation, is important when motivating the use of IVs. For the choice to have a theoretical underpinning, it is required that the full set of exogenous variables in the model to which the relevant equation belongs be fully specified. The limited information format of many classical linear multi-equation structural models departs essentially from the rudimentary setups in typical 'natural experiment' studies.

Fifth, omitted variables is not a precise term. It may refer to a latent variable for which a proxy exists, a variable observed with a random error, a variable which theory says should not be in a specific equation, a variable pretended to be absorbed by equations' disturbances, or a more general, irrelevant nuisance variable to be modeled or eliminated from the equation system. Claiming the IV technique to be a general remedy to cope with omitted variables is farfetched.

Sixth, the parameters or parameter combinations for which an IV is valid may not be invariant to the model's identification status. Imposition of parameter restrictions may crucially affect the set of valid IVs. Replacing a measurement error equation with a 'proxy variable equation' changes the 'metric' of some variables' and affect the identification status of coefficients and the way IVs and coefficients are connected.

Seventh, distinguishing between the statements 'X is exogenous' and 'X is orthogonal to the disturbance' when considering X an an IV candidate relative to an equation, is crucial. While exogeneity in both single- and multi-equation contexts is a model property, IV-requirements relate to one or a few of the model's equations. A valid IV may be endogenous, while if the model is recursive, such an IV may be conditionally exogenous.

Eighth, terms like 'IV-estimand', 'IV-Model', 'first-stage' and 'second-stage equations' violate classical econometric terminology and should be avoided. Like OLS, the term 2SLS (Two-stage least squares) has a definite meaning in the classical literature, which should not be disguised or obscured by using it in contexts alien to those assumed by the method's constructors.

Ninth, classical econometrics is not alien to incorporating potential 'experiments' as part of the model, if it can be convincingly argued that certain model elements, *e.g.*, fiscal policy instruments or institutional rules, vary randomly relative to unexplained elements. However, if a kind of a 'natural experiment' setup involving economic variables is considered, it should not be handled in isolation. The (often rather few) equations which describe the 'experiments' and their set of orthogonality conditions should be embedded in a more comprehensive system of stochastic equations describing the economy. The latter should include equations that are not intended to 'directly' answer the economic-policy or economic-behavioural question(s) of primary interest. Relying solely on rudimentary models can be 'dangerous'.

Tenth, certain proponents of 'natural experiment' approaches in empirical economics call their results 'credible causal inference', some even using the label 'the credibility revolution' for the ideas they promote. Interesting in this connection is the observation that Ioannidis (2012) and Ioannidis and Doucouliagos (2013), in a survey of key parameters that may influence credibility of research findings in general, and in economics in particular, ask "why science is not necessarily self-correcting?", and remain sceptical. Anyway, a legitimate question is: Can 'loaded' words as 'credibility' and 'revolution' adequately characterize a scientific approach with as many loose ends as described in the previous sections? My answer is: definitely not.

Appendix A: Two extensions of Model 6

In this appendix, two extensions of Model 6 are briefly considered.

Two-equation EIV model with observed regressor added: We extend (30) to

$$\text{(a.1)}\quad \begin{array}{ll} y_1 = \beta_1\xi + \gamma_1 z + u_1, & (y_2, x, u_2, \nu) \text{ excluded},\\ y_2 = \beta_2\xi + \gamma_2 z + u_2, & (y_1, x, u_1, \nu) \text{ excluded},\\ x = \xi + \nu, & (y_1, y_2, z, u_1, u_2) \text{ excluded},\\ u_1 \perp u_2, \quad (u_1, u_2, \nu) \perp (\xi, z), & \\ \text{observed, endogenous: } y_1, y_2, x, & \\ \text{observed, exogenous: } z, & \\ \text{latent, exogenous: } \xi. & \end{array}$$

Eliminating ξ, we obtain

$$\text{(a.2)}\quad \begin{array}{lll} y_1 = \beta_1 x + \gamma_1 z + (u_1 - \beta_1\nu), & (y_2, z) \perp (u_1 - \beta_1\nu), & (y_2, u_2) \text{ excluded},\\ y_2 = \beta_2 x + \gamma_2 z + (u_2 - \beta_2\nu), & (y_1, z) \perp (u_2 - \beta_2\nu), & (y_1, u_1) \text{ excluded}, \end{array}$$

and hence,

$$\begin{aligned} \text{cov}(y_1, y_2) &= \beta_1 \text{cov}(x, y_2) + \gamma_1 \text{cov}(z, y_2),\\ \text{cov}(y_1, z) &= \beta_1 \text{cov}(x, z) + \gamma_1 \text{var}(z), \end{aligned}$$

$$\begin{aligned} \text{cov}(y_2, y_1) &= \beta_2 \text{cov}(x, y_1) + \gamma_2 \text{cov}(z, y_1),\\ \text{cov}(y_2, z) &= \beta_2 \text{cov}(x, z) + \gamma_2 \text{var}(z), \end{aligned}$$

which motivates estimating (β_1, γ_1) and (β_2, γ_2) by using, respectively, (y_2, z) as IVs for (x, z), and (y_1, z) as IVs for (x, z).

Generalizing the measurement error equation $x = \xi + v$ to the proxy variable equation $x = d\xi + v$, again changing the 'metric' of x, leads to a model with the same covariance equations as above except that β_1 and β_2 are replaced with β_1/d and β_2/d. This suggests estimating $(\beta_1/d, \gamma_1)$ by using (y_2, z) as IVs for (x, z), and estimating $(\beta_2/d, \gamma_2)$ by using (y_1, z) as IVs for (x, z), which leaves β_1 and β_2 unidentified. Relative to (a.2), z is both an IV for itself and serves as IV for x together with y_2 in the first equation and for x together with y_1 in the second equation, and z is not excluded from any of these two equations. This model therefore also emerges as an extension of the Example C model in Section 2.

Two-equation model, with two latent explanatory variables and error and proxy mechanisms represented: We further extend (30) and (34) to a model with two latent variables which enter both equations and having six observed variables, two of which being proxies and two being standard error-ridden:

$$\text{(a.3)}\quad \begin{array}{ll} y_1 = \beta_1\xi + \gamma_1\kappa + u_1, & (y_2, z, s, x, q, u_2, \nu, \delta, \epsilon_z, \epsilon_s) \text{ excluded}, \\ y_2 = \beta_2\xi + \gamma_2\kappa + u_2, & (y_1, z, s, x, q, u_1, \nu, \delta, \epsilon_z, \epsilon_s) \text{ excluded}, \\ x = \xi + \nu, & (y_1, y_2, q, z, s, u_1, u_2, \delta, \epsilon_z, \epsilon_s) \text{ excluded}, \\ q = \kappa + \delta, & (y_1, y_2, x, z, s, u_1, u_2, \nu, \epsilon_z, \epsilon_s) \text{ excluded}, \\ z = b_z\xi + \epsilon_z, & (y_1, y_2, s, x, q, u_1, u_2, \nu, \delta, \epsilon_s) \text{ excluded}, \\ s = b_s\kappa + \epsilon_s, & (y_1, y_2, z, x, q, u_1, u_2, \nu, \delta, \epsilon_z) \text{ excluded}, \\ u_1 \perp u_2 \perp \nu \perp \delta \perp \epsilon_s \perp \epsilon_z \perp \xi \perp \kappa, & \\ \text{observed, endogenous: } y_1, y_2, x, q, z, s, & \\ \text{latent, exogenous: } \xi, \kappa. & \end{array}$$

This system, in RF format, after elimination of (ξ, κ), becomes:

$$\text{(a.4)}\quad \begin{array}{lll} y_1 = \beta_1 x + \gamma_1 q + (u_1 - \beta_1\nu - \gamma_1\delta), & (y_2, z, s) \perp (u_1 - \beta_1\nu - \gamma_1\delta), & (y_2, z, s, u_2, \epsilon_s, \epsilon_z) \text{ excluded}, \\ y_2 = \beta_2 x + \gamma_2 q + (u_2 - \beta_2\nu - \gamma_2\delta), & (y_1, z, s) \perp (u_2 - \beta_2\nu - \gamma_2\delta), & (y_1, z, s, u_1, \epsilon_s, \epsilon_z) \text{ excluded}. \end{array}$$

Hence, there is overidentification and it follows that:

$$\begin{aligned} \text{cov}(y_1, y_2) &= \beta_1 \text{cov}(x, y_2) + \gamma_1 \text{cov}(q, y_2), \\ \text{cov}(y_1, z) &= \beta_1 \text{cov}(x, z) + \gamma_1 \text{cov}(q, z), \\ \text{cov}(y_1, s) &= \beta_1 \text{cov}(x, s) + \gamma_1 \text{cov}(q, s), \end{aligned}$$

$$\begin{aligned} \text{cov}(y_2, y_1) &= \beta_2 \text{cov}(x, y_1) + \gamma_2 \text{cov}(q, y_1), \\ \text{cov}(y_2, z) &= \beta_2 \text{cov}(x, z) + \gamma_2 \text{cov}(q, z), \\ \text{cov}(y_2, s) &= \beta_2 \text{cov}(x, s) + \gamma_2 \text{cov}(q, s). \end{aligned}$$

This suggests estimating (β_1, γ_1) by using (y_2, z, s) as (endogenous) IVs for (x, q) and estimating (β_2, γ_2) by using (y_1, z, s) as (endogenous) IVs for (x, q). *This model exemplifies a case where only endogenous variables serve as IVs,* which crucially hinges on the last, orthogonality assumption in (a.3). Since also

$$\text{(a.5)}\quad \begin{array}{lll} z = b_z x + (\epsilon_z - b_z\nu), & (y_1, y_2, s, q) \perp (\epsilon_z - b_z\nu), & (y_1, y_2, s, q, u_1, u_2, \delta, \epsilon_s) \text{ excluded}, \\ s = b_s q + (\epsilon_s - b_s\delta), & (y_1, y_2, z, x) \perp (\epsilon_s - b_s\delta), & (y_1, y_2, z, x, u_1, u_2, \nu, \epsilon_z) \text{ excluded}, \end{array}$$

and hence,

$$\begin{aligned} b_z &= \frac{\text{cov}(z, y_1)}{\text{cov}(x, y_1)} = \frac{\text{cov}(z, y_2)}{\text{cov}(x, y_2)} = \frac{\text{cov}(z, s)}{\text{cov}(x, s)} = \frac{\text{cov}(z, q)}{\text{cov}(x, q)}, \\ b_s &= \frac{\text{cov}(s, y_1)}{\text{cov}(q, y_1)} = \frac{\text{cov}(s, y_2)}{\text{cov}(q, y_2)} = \frac{\text{cov}(s, z)}{\text{cov}(q, z)} = \frac{\text{cov}(s, x)}{\text{cov}(q, x)}. \end{aligned}$$

a suggested way of estimating b_z is by using (y_1, y_2, s, q) (endogenous) as IV for x and estimating b_s by using (y_1, y_2, z, x) (endogenous) as IV for q.

Appendix B: Interdependence, measurement error, and IVs

In this appendix an interdependent, two-equation model involving both 'simultaneity' (feedback between endogenous variables) and mismeasured variables, and ways of using IVs, is considered. Like the models in Appendix A, it may be considered an extension of the models in Section 6.

$$\text{(b.1)}\quad \begin{aligned} \eta &= \alpha\xi + \gamma z + u, \\ \xi &= \beta\eta + \delta q + v, \\ y &= \eta + \epsilon_y, \\ x &= \xi + \epsilon_x, \\ (u, v) &\perp (\epsilon_y, \epsilon_x) \perp (z, q). \end{aligned}$$

Here (η, ξ) are the endogenous latent variables, (y, x) their observed counterparts, (z, q) are the exogenous variables, assumed free of measurement errors, (u, v) are the structural form disturbances, and

(ϵ_y, ϵ_x) are the measurement errors. Correlation between the two measurement errors and between the two disturbances, as well as correlation between the exogenous variables, are allowed for. Eliminating the latent endogenous variables, the structural equations expressed in observed variables become:

$$\text{(b.2)} \qquad \begin{aligned} y &= \alpha x + \gamma z + \epsilon_y - \alpha\epsilon_x + u, \quad q \perp (\epsilon_y - \alpha\epsilon_x + u), \quad (q, v) \text{ excluded} \\ x &= \beta y + \delta q + \epsilon_x - \beta\epsilon_y + v, \quad z \perp (\epsilon_x - \beta\epsilon_y + v), \quad (z, u) \text{ excluded} \end{aligned}$$

This suggests (i) using q as an IV for x in the first equation (z serving as IV for itself), when estimating α and γ, or, equivalently, as an IV for y in the reverse equation when estimating $1/\alpha$ and γ/α and (ii) using z is an IV for x in second equation (q serving as IV for itself), when estimating β and δ, or, equivalently, as an IV for y in the reverse equation when estimating $1/\beta$ and δ/β.

Since (b.1)–(b.2) lead to the following RF equations for the observed variables:

$$\text{(b.3)} \qquad \begin{aligned} y &= \frac{\gamma z + \alpha\delta q}{1-\alpha\beta} + \frac{u + \alpha v}{1-\alpha\beta} + \epsilon_y, \\ x &= \frac{\delta q + \beta\gamma z}{1-\alpha\beta} + \frac{v + \beta u}{1-\alpha\beta} + \epsilon_x, \end{aligned}$$

the marginal covariances between the observed endogenous variables and the exogenous variables and their conditional counterparts are, respectively,

$$\begin{aligned} \text{cov}(y,z) &= \frac{\gamma\text{var}(z) + \alpha\delta\text{cov}(q,z)}{1-\alpha\beta}, & \text{cov}(y,z|q) &= \frac{\gamma\text{var}(z|q)}{1-\alpha\beta}, \\ \text{cov}(y,q) &= \frac{\gamma\text{cov}(z,q) + \alpha\delta\text{var}(q)}{1-\alpha\beta}, & \text{cov}(y,q|z) &= \frac{\alpha\delta\text{var}(q|z)}{1-\alpha\beta}, \\ \text{cov}(x,z) &= \frac{\delta\text{cov}(q,z) + \beta\gamma\text{var}(z)}{1-\alpha\beta}, & \text{cov}(x,z|q) &= \frac{\beta\gamma\text{var}(z|q)}{1-\alpha\beta}, \\ \text{cov}(x,q) &= \frac{\delta\text{var}(q) + \beta\gamma\text{cov}(q,z)}{1-\alpha\beta}, & \text{cov}(x,q|z) &= \frac{\delta\text{var}(q|z)}{1-\alpha\beta}. \end{aligned}$$

This implies that the coefficient of the latent endogenous variables can be expressed by

$$\begin{aligned} \alpha &= \frac{\text{cov}(y,q|z)}{\text{cov}(x,q|z)} \quad \text{or} \quad \alpha = \frac{\text{cov}(y,q)}{\text{cov}(x,q)} \quad \text{when} \quad \gamma = 0, \\ \beta &= \frac{\text{cov}(x,z|q)}{\text{cov}(y,z|q)} \quad \text{or} \quad \beta = \frac{\text{cov}(x,z)}{\text{cov}(y,z)} \quad \text{when} \quad \delta = 0. \end{aligned}$$

From (b.3) we further have

$$\begin{aligned} \text{var}(y|z,q) &= \frac{\sigma_u^2 + 2\alpha\sigma_{uv} + \alpha^2\sigma_v^2}{(1-\alpha\beta)^2} + \sigma_{\epsilon y}^2, \\ \text{var}(x|z,q) &= \frac{\sigma_v^2 + 2\beta\sigma_{uv} + \beta^2\sigma_v^2}{(1-\alpha\beta)^2} + \sigma_{\epsilon x}^2, \\ \text{cov}(y,x|z,q) &= \frac{\beta\sigma_u^2 + (1+\alpha\beta)\sigma_{uv} + \alpha\sigma_v^2}{(1-\alpha\beta)^2} + \sigma_{\epsilon y \epsilon x}. \end{aligned}$$

If z is excluded from (b.1) (and hence vanishes from the model) ($\gamma = 0$, $\delta \neq 0$), q is the IV-candidate for x in estimating α in its first equation, while β and δ in its second equation are unidentified. Symmetrically, if q is excluded from (b.1) (and hence vanishes from the model) ($\delta = 0$, $\gamma \neq 0$), z is the IV-candidate for y in estimating β in its second equation, while α and γ in its first equation are unidentified. Why it is necessary to have specified a full two-equation EIV model, is obvious.

The first and second equation in (b.2) satisfy, respectively:

$$\begin{aligned} \text{cov}(y,z) &= \alpha\,\text{cov}(x,z) + \gamma\,\text{var}(z), \\ \text{cov}(y,q) &= \alpha\,\text{cov}(x,q) + \gamma\,\text{cov}(z,q), \end{aligned}$$

$$\begin{aligned} \text{cov}(x,z) &= \beta\,\text{cov}(y,z) + \delta\,\text{cov}(q,z), \\ \text{cov}(x,q) &= \beta\,\text{cov}(y,q) + \delta\,\text{var}(q), \end{aligned}$$

expressing, respectively, that (q, z) are IVs for (x, z), while (z, q) are IVs for (y, q). It $\gamma \neq 0,\ \delta \neq 0$,

$$\alpha = \frac{\text{var}(z)\text{cov}(y,q) - \text{cov}(z,q)\text{cov}(y,z)}{\text{var}(z)\text{cov}(x,q) - \text{cov}(x,z)\text{cov}(z,q)},$$
$$\gamma = \frac{\text{cov}(x,q)\text{cov}(y,z) - \text{cov}(x,z)\text{cov}(y,q)}{\text{var}(z)\text{cov}(x,q) - \text{cov}(x,z)\text{cov}(z,q)},$$
$$\beta = \frac{\text{var}(q)\text{cov}(x,z) - \text{cov}(q,z)\text{cov}(x,q)}{\text{var}(q)\text{cov}(y,z) - \text{cov}(y,q)\text{cov}(q,z)},$$
$$\delta = \frac{\text{cov}(y,z)\text{cov}(x,q) - \text{cov}(y,q)\text{cov}(x,z)}{\text{var}(q)\text{cov}(y,z) - \text{cov}(y,q)\text{cov}(q,z)}.$$

If the two exogenous variables (IVs) are uncorrelated ($\text{cov}(z, q)=0$), the expressions simplify to

$$\alpha = \frac{\text{cov}(y,q)}{\text{cov}(x,q)},$$
$$\beta = \frac{\text{cov}(x,z)}{\text{cov}(y,z)},$$
$$\gamma = \frac{\text{cov}(x,q)\text{cov}(y,z) - \text{cov}(x,z)\text{cov}(y,q)}{\text{var}(z)\text{cov}(x,q)} = \frac{\text{cov}(y,z)}{\text{var}(z)}(1-\alpha\beta),$$
$$\delta = \frac{\text{cov}(y,z)\text{cov}(x,q) - \text{cov}(y,q)\text{cov}(x,z)}{\text{var}(q)\text{cov}(y,z)} = \frac{\text{cov}(x,q)}{\text{var}(q)}(1-\beta\alpha).$$

Appendix C: IV use in a dynamic model in error-ridden variables

This appendix extends one of the dynamic models in Section 7 to a latent variables variant of (38):

$$\text{(c.1)}\quad \begin{aligned} \eta_t &= \beta\xi_t + \gamma\eta_{t-1} + u_t, \qquad |\gamma| < 1,\\ y_t &= \eta_t + \epsilon_t,\\ x_t &= \xi_t + \delta_t,\\ u_t &\perp \epsilon_t \perp \delta_t \perp \xi_t, \end{aligned}$$

where ξ_t, η_t are latent variables. The form of the processes for $u_t, \epsilon_t, \delta_t, \xi_t$ will influence the set of IV for x_t. It follows that:

$$\text{(c.2)}\quad y_t = \beta x_t + \gamma y_{t-1} + v_t,$$
$$\text{(c.3)}\quad v_t = u_t + \epsilon_t - \gamma\epsilon_{t-1} - \beta\delta_t,$$

and that

$$\text{(c.4)}\quad \begin{aligned} \text{cov}(x_t, x_{t-s}) &= \text{cov}(\xi_t, \xi_{t-s}) + \text{cov}(\delta_t, \delta_{t-s}),\\ \text{cov}(x_{t-s}, v_t) &= -\beta\text{cov}(\delta_{t-s}, \delta_t). \end{aligned}$$

This can be used to define the valid IVs for x_t in (c.2). It follows that $\text{cov}(x_t, x_{t-s}) \neq 0$ and $\text{cov}(x_{t-s}, v_t) = 0$ when $\text{cov}(\xi_t, \xi_{t-s}) \neq 0$ and $\text{cov}(\delta_t, \delta_{t-s}) = 0$ hold jointly. These conditions therefore ensure that x_{t-s} is a valid IV for x_t. This is a specific, theoretically founded, statement about the validity of s-period lagged observed regressors x_{t-s} as IV for the one period lagged one x_{t-1}, contingent on the relative memory of ξ_t and δ_t and of the signal and noise components of x_t.

Notes

[1]Stock and Watson (2011), in their textbook, remarkably, use the term 'IV Model' (pp. 420,430–433, etc.). See also Chesher (2010), Chesher and Smolinski (2012) and Heckman *et al.* (2006, Section III).

[2]In empirical studies in, *e.g.*, economics there is an increasing tendency to use 'IV-estimand' as synonymous with the probability limit of a suggested statistic involving one or more variables declared as IVs, without ensuring (and convincing readers) that this probability limit is an interesting interpretable parameter in a statistical model declared for a certain purpose; cf. footnote 20. See, for example, Angrist *et al.* (1996), Angrist *et al.* (2000), Heckman and Vytlacil (2005, 2007), and Imbens (2014).

[3]Lewbel (2017) starts a recent survey of the 'Identification Zoo' by saying: "Econometric identification really means just one thing: model parameters being uniquely determined from the observable population that data are drawn from. Yet well over two dozen different terms for identification now appear in the econometrics literature", and in the concluding section remarks: "Unlike statistical inference, there is not a large body of general tools or techniques that exist for proving identification. As a result, identification proofs are often highly model specific and idiosyncratic."

[4]Rosenzweig and Wolpin let the label, with the word 'natural' repeated, indicate experiments where random "treatments ... have arisen serendipitously".

[5]On conditional exogeneity and its relation to recursivity, see *e.g.*, White and Chalak (2010) and White and Pettenuzzo (2014).

[6]A full information specification would contain detailed descriptions of all equations, including omitted and included variables, coefficients, as well as exogeneity and orthogonality assumptions.

[7]This term is used by, *e.g.*, Geary (1949, p. 30). *Instrumented variable* is more often used today.

[8]On the origin of the *name Instrumental Variable* in the 1940s and its utilization by Reiersøl in the early history of econometrics, see Willassen (2000, pp. 118–119).

[9]Although x and z cannot be latent, they may be connected to common latent variables, as will be demonstrated by examples in Section 6.

[10]Reminder: $\rho_{z,x1|x2} = [\rho_{z,x1} - \rho_{z,x2}\rho_{x2,x1}]/[(1-\rho^2_{z,x2})(1-\rho^2_{z,x1})]^{1/2}$ is the (partial) correlation coefficient between z and x_1, given x_2.

[11]$a_1 = 0$ would have given an infeasible IV, with x_2 serving both an IV for x_1 and being an exogenous variable in the equation. Although no IV is needed for x_2, we might, for symmetry reasons, consider x_2 as being an IV for itself, or, more generally, consider the pair $z^* = za_1 + x_2a_2$ and $z^{**} = zb_1 + x_2b_2$, representing a one-to-one transformation from (z, x_2) to (z^*, z^{**}), as an IV pair for (x_1, x_2) in (6).

[12]Bowden and Turkington (1984, Sec. 4.3), see also Turkington (2013, Sec. 6.6), denote such a 'model', combining an equation in an interdependent model, (17), with the RF-equations of its endogenous RHS-variables, (20) (in my view not very aptly) a 'limited-information model'.

[13]See also Pearl (2009, Section 7.4.5).

[14]Certain texts even denote the primary claims to IVs as 'instrument relevance' and 'instrument exogeneity', the former meaning that the IV is 'correlated with the endogenous regressor' and the latter meaning that the IV is 'uncorrelated with the error term of the actual equation and has no direct effect on the variable it explains'.

[15]Angrist and Krueger (2001, p. 72) also remark: "The observed association between the outcome and explanatory variable ..... is likely to be misleading in the sense that it partly reflects omitted factors that are related to both variables. If these factors *could be measured and held constant in a regression*, the omitted variables bias would be eliminated. In practice, however, economic theory typically does not specify all of *the variables that should be held constant while estimating a relationship*, and it is difficult to measure all of the relevant variables accurately even if they are specified" (my italics). Do the authors really consider an analyst of economic data an acting person who, when running regressions, can force variables to 'be held constant'? I hope not.

[16]Passages from S. Wright (1934, pp. 161,175), on the utilization of 'path coefficients' in exploring causal relations, point in the same direction: "..... the method of path coeffcients ..... was developed primarily as a means of combining the quantitative information given by a system of correlation coefficients *with such information as may be at hand with regard to causal relations ...... and thus making quantitative an interpretation which would otherwise be merely qualitative ..... The setting up of a qual-*

itative scheme depends primarily on information outside of the numerical data and the judgement as to its validity must rest primarily on this outside information..." (my italics). Confer also the passages on shifts in the supply curve to 'see' the demand curve, and vice versa, from a text published in 1906, referred to in Kærgaard (1984, p. 441) – an obvious early illustration of the identification problem.

[17]Laudably, to prevent confusion, Hinke *et al.* (2016, p. 193) states "Depending on the discipline, the terms 'treatment', 'risk factor', 'exposure', 'predictor', or 'intermediate phenotype' have all been used to denote the variable of interest that potentially causes the outcome".

[18]By Wikipedia described as the "method of using measured variation in genes of known function to examine the causal effect of a modifiable exposure on disease in observational studies".

[19]Extract from `https://www.socialresearchmethods.net/kb/random.htm`.

[20]Adherents to 'natural experiments' often consider the mentioned probability limits of suggested 'IV-estimators' the 'research objects', replacing the parameters of the 'structuralists' and (inaptly) labeled *'IV-estimands'*; cf. footnote 2. Labeling a declared coefficient, β an 'IV-estimand' must be judged superfluous and confusing; β might equally well (and equally confusing) be labeled an 'OLS-estimand', a 'GLS-estimand', an 'ML-estimand' and much else. A Google search gave more than 1600 entries using this term (and 240 entries using 'OLS-estimand', but only 3 entries using 'ML-estimand')! There is a strong discrepancy between having θ and having $\text{plim}(\widehat{\theta}^{IV})$ (or $\text{plim}(\widehat{\theta}^{OLS})$ or $\text{plim}(\widehat{\theta}^{ML})$) in the focus of a study.

[21]For extensions and modifications of this setup, mostly strengthening assumptions, see Wegge (1965), Brundy and Jorgenson (1971, 1974), Hausman (1974), Hausman *et al.* (1987), and Schmidt (1990).

[22]Their setups resemble that in Models 1–3 to be discussed in Section 5.

[23]A passus on p. 133 violates classical terminology twice (confer italics): "Consider the *2SLS estimand* based on the *first-stage equation*". In the same vein, Paxton *et al.* (2011), in a textbook on "nonrecursive models, endogeneity and reciprocal relationships", misguide readers by explaining the first stage of 2SLS by: "*In most cases* the first-stage equation will be the reduced-form equation." (p. 49, my italics; see also p. 60).

[24]Recall that the classical OC is necessary (but not sufficient) for identification of an equation in a certain kind of linear simultaneous equations models, the corresponding rank condition being both necessary and sufficient. Neither are applicable to rudimentary models as in Examples A, B and C in Section 2 and by typical setups in studies following 'natural experiment' ideas. Neither are these order and rank conditions useful for models containing latent structural variables. Geraci (1976) discusses the identification criteria in interdependent models with measurement error.

[25]See also the discussion in Clarke (2005) of the 'Phantom Menace' of omitted variables and (relevant and irrelevant) control variables.

[26]All models, except Model 4, variant 4b, assume error-disturbance orthogonality, as in Leamer (1987, Sections 2 and 4).

[27]In considering this model and Models 5 and 6 (as well as the extended models in Appendices A and B) I, for simplicity, (i) omit intercepts, say α and a in the first and third equation of (26), written explicitly as $y = \alpha + \beta\xi + u$ and $z = a + b\xi + \epsilon$, and (ii) omit equations expressing the expectations and variances of the observed variables, say in (28) $\mathsf{E}(y)$, $\mathsf{E}(x)$, $\mathsf{E}(z)$, $\text{var}(y) = \beta^2\sigma_\xi^2 + \sigma_u^2$, $\text{var}(x) = \sigma_\xi^2 + \sigma_\nu^2$, and $\text{var}(z) = b^2\sigma_\xi^2 + \sigma_\epsilon^2$, whose only function is to identify $\mathsf{E}(\xi)$, α, a, σ_u^2, σ_ν^2 and σ_ϵ^2. Then we in (28) are left with the equations for the covariances.

[28]I hesitate to call it an RF, because of its endogenous RHS-variables. A 'hybrid model form' is a more appropriate label.

[29]Although this model can be interpreted as a two-equation EIV model, (30) is *formally* a reparameterization of (26) with (y_1, y_2, u_1, u_2) corresponding to (y, z, u, ϵ), the second structural equation in (30) corresponding to the proxy variable equation in (26).

[30]Letting $\xi^* = d\xi$ we get, after reparameterization, a model of the same form as the first model in Appendix A, (a.1), with ξ replaced by ξ^*, and β_1, β_2 and b replaced by β_1^*, β_2^* and b^* in (35).

[31]Frisch also showed interest in IVs in connection with his work on Confluence Analysis, Frisch (1934), according to unpublished material from the Frisch-Archives of the University of Oslo. I thank Olav Bjerkholt for turning my attention to, and giving me access to, this material. It is also worth noting that the familiar theorem of Frisch and Waugh (1933) may be considered a particular application of what was later to become known as the IV-technique.

References

Aldrich, J. (1989): Autonomy. *Oxford Economic Papers* **41**, 15–34.

Aldrich, J. (1993): Reiersøl, Geary and the Idea of Instrumental Variables. *Economic and Social Review* **24**, 247–273.

Aldrich, J. (1994): Haavelmo's Identification Theory. *Econometric Theory* **10**, 198–219.

Anderson, T.W. (1974): An Asymptotic Expansion of the Distribution of the Limited Information Maximum Likelihood Estimate of a Coefficient in a Simultaneous Equation System. *Journal of the American Statistical Association* **69**, 565–573.

Anderson, T.W. (2005): Origins of the Limited Information Maximum Likelihood and Two-stage Least Squares Estimators. *Journal of Econometrics* **127**, 1–16.

Anderson, T.W. and Rubin, H. (1949): Estimation of the Parameters of a Single Equation in a Complete System of Stochastic Equations. *Annals of Mathematical Statistics* **20**, 46–63.

Anderson, T.W. and Rubin, H. (1950): The Asymptotic Properties of Estimates of the Parameters of a Single Equation in a Complete System of Stochastic Equations. *Annals of Mathematical Statistics* **21**, 570–582.

Angrist, J.D., Graddy, K. and Imbens, G.W. (2000): The Interpretation of Instrumental Variables Estimators in Simultaneous Equations Models with an Application to the Demand for Fish. *Review of Economic Studies* **67**, 499-527.

Angrist, J.D., Imbens, G.W. and Rubin, D.B. (1996): Identification of Causal Effects Using Instrumental Variables. *Journal of the American Statistical Association* **91**, 444–455.

Angrist, J.D. and Krueger, A.B. (2001): Instrumental Variables and the Search for Identification: From Supply and Demand to Natural Experiments. *Journal of Economic Perspectives* **15**, 69–85.

Angrist, J.D. and Pischke, J.S. (2008): *Mostly Harmless Econometrics: An Empiricist's Companion.* Princeton: Princeton University Press.

Angrist, J.D. and Pischke, J.S. (2010): The Credibility Revolution in Empirical Economics: How Better Research Design is Taking the Con out of Econometrics. *Journal of Economic Perspectives* **24**, 3–30.

Basmann, R.L. (1957): A Generalized Classical Method of Linear Estimation of Coefficients in a Structural Equation. *Econometrica*, **25**, 77–83.

Bowden, R.J. and Turkington, D.A. (1984): *Instrumental Variables.* Cambridge: Cambridge University Press.

Brundy, J.M. and Jorgenson, D.W. (1971): Efficient Estimation of Simultaneous Equations by Instrumental Variables. *Review of Economics and Statistics* **53**, 207–224.

Brundy, J.M. and Jorgenson, D.W. (1974): The Relative Efficiency of Instrumental Variables Estimators of Simultaneous Equations. *Annals of Economic and Social Measurement* **3/4**, 679–700.

Burgess, S., Small, D.S., and Thompson, S.G. (2017): A Review of Instrumental Variable Estimators for Mendelian Randomization. *Statistical Methods in Medical Research* **26**, 2333–2355.

Chesher, A. (2010): Instrumental Variable Models for Discrete Outcomes. *Econometrica* **78**, 575–601.

Chesher, A. and Smolinski, K. (2012): IV Models of Ordered Choice. *Journal of Econometrics* **166**, 33–48.

Clarke, K.A. (2005): The Phantom Menace: Omitted Variable Bias in Econometric Research. *Conflict Management and Peace Science* **22**, 341-352.

Dahl, G.B., Kostøl, A.R. and Mogstad, M. (2014): Family Welfare Cultures. *Quarterly Journal of Economics* **129**, 1711-1752.

Deaton, A. (2010): Instruments, Randomization, and Learning about Development. *Journal of Economic Literature* **48**, 424–455.

Didelez, V. and Sheehan, N. (2007): Mendelian Randomization as an Instrumental Variable Approach to Causal Inference. *Statistical Methods in Medical Research* **16**, 309–330.

Engle, R.F., Hendry, D.F. and Richard, J.-F. (1983): Exogeneity. *Econometrica* **51**, 277–304.

Frisch, R. (1934): *Statistical Confluence Analysis by Means of Complete Regression Systems.* Oslo: Universitetets Økonomiske Institutt.

Frisch, R. (2005): Autonomy of Economic Relations (1938). Statistical versus Theoretical Relations in Economic Macrodynamics. Chapter 35 in: Hendry, D.F. and Morgan, M.S. (eds.): *The Foundations of Econometric Analysis.* Cambridge: Cambridge University Press.

Frisch, R. and Waugh, F.V. (1933): Partial Time Regressions as Compared with Individual Trends. *Eonometrica* **1**, 387–401.

Geary, R.C. (1949): Determination of Linear Relations between Systematic Parts of Variables with Errors of Observation the Variances of Which Are Unknown. *Econometrica* **17**, 30–58.

Geraci, V.J. (1976): Identification of Simultaneous Equation Models with Measurement Error. *Journal of Econometrics* **4**, 263–283.

Hausman, J.A. (1974): Full Information Instrumental Variables Estimation of Simultaneous Systems. *Annals of Economic and Social Measurement.* **3/4**, 641–652.

Hausman, J.A., Newey. W.K. and Taylor, W.E. (1987): Efficient Estimation and Identification of Simultaneous Equation Models with Covariance Restrictions. *Econometrica* **55**, 849–874.

Heckman, J.J. and Urzua, S. (2010): Comparing IV with Structural Models: What Simple IV Can and Cannot Identify. *Journal of Econometrics* **156**, 27-37.

Heckman, J.J., Urzua, S. and Vytlacil, E. (2006): Understanding Instrumental Variables in Models with Essential Heterogeneity. *Review of Economics and Statistics* **88**, 389–432.

Heckman, J.J. and Vytlacil, E.J. (2005): Structural Equations, Treatment Effects, and Econometric Policy Evaluation. *Econometrica* **73**, 669–738.

Heckman, J.J. and Vytlacil, E.J. (2007): Econometric Evaluation of Social Programs, Part II: Distributional Treatment Effects, Dynamic Treatment Effects, Dynamic Discrete Choice, and General Equilibrium Policy Evaluation. Chapter 71 in: Heckman, J.J., and Leamer, E.E. (eds.): *Handbook of Econometrics, vol 6B*

Hendry, D.F. (1995): *Dynamic Econometrics.* Oxford: Oxford University Press.

Hendry, D.F. and Johansen, S. (2015): Model Discovery and Trygve Haavelmo's Legacy. *Econometric Theory* **31**, 93–114.

Hinke, S. von, Smith, G.D., Lawlor, D.A., Propper, C. and Windmeijer, F. (2016): Genetic Markers as Instrumental Variables. *Journal of Health Economics* **45**, 131–148.

Imbens, G.W. (2010): Better LATE Than Nothing: Some Comments on Deaton (2009) and Heckman and Urzua (2009). *Journal of Economic Literature* **48**, 399–423.

Imbens, G.W. (2014): Instrumental Variables: An Econometricians's Perspective. *Statistical Science*, **29**, 323–358.

Ioannidis, J.P.A. (2012): Why Science Is Not Necessarily Self-Correcting. *Perspectives on Psychological Science* **7**, 645-654.

Ioannidis, J.P.A. and Doucouliagos, C. (2013): What's to Know About the Credibility of Empirical Economics? *Journal of Economic Surveys* **27**, 997–1004.

Keane, M.P. (2010): Structural vs. Atheoretic Approaches to Econometrics. *Journal of Econometrics* **156**, 3–20.

Koopmans, T.C. (1950): When is an Equation System Complete for Statistical Purposes? Chapter XVII in: Koopmans, T.C. (ed.): *Statistical Inference in Dynamic Economic Models.* New York: Wiley.

Koopmans, T.C. (1953): Identification Problems in Economic Model Construction. Chapter II in: W.C. Hood & T.C. Koopmans (eds.): *Studies in Econometric Methods*, New York: Wiley.

Koopmans, T.C. and Reiersøl, O. (1950): The Identification of Structural Characteristics. *Annals of Mathematical Statistics* **21**, 165–181.

Kærgaard, N. (1984): The Earliest History of Econometrics: Some Neglected Danish Contributions. *History of Political Economy* **16**, 437–444.

Leamer, E.E. (1978): *Specification Searches: Ad Hoc Inference with Nonexperimental Data.* New York: Wiley.

Leamer, E.E. (1987): Errors in Variables in Linear Models. *Econometrica* **55**, 893–909.

Lewbel, A. (2017): The Identification Zoo – Meanings of Identification in Econometrics. *Journal of Economic Literature*, forthcoming.

Lundborg, P., Plug, E. and Rasmussen, A.W. (2017): Can Women Have Children and a Career? IV Evidence from IVF Treatments. *American Economic Review* **107**, 1611–1637.

Mann, H.B. and Wald, A. (1943): On Stochastic Limit and Order Relationships. *Annals of Mathematical Statistics* **14**, 217–226.

Marschak, J. (1953): Economic Measurements for Policy and Prediction. Chapter I in: Hood, W.C. and Koopmans, T.C. (eds.): *Studies in Econometric Methods.* New York: Wiley.

McClellan, M., McNeal, B.J. and Newhouse, J.P. (1994): Does More Intensive Treatment of Acute Myocardial Infarction in the Elderly Reduce Mortality? Analysis Using Instrumental Variables. *Journal of the American Medical Association* **272**, 859–866.

Morgan, M.S. (1990): *The History of Econometric Ideas: Historical Perspectives on Modern Economics.* Cambridge: Cambridge University Press.

Oreopoulos, P. (2006): Estimating Average and Local Average Treatment Effects of Education when Compulsory Schooling Laws Really Matter. *American Economic Review* **96**, 152–175.

Paxton, P., Hipp, J.R. and Marquart-Pyatt, S. (2011): *Nonrecursive Models. Endogeneity, Reciprocal Relationships, and Feedback Loops.* Thousand Oaks, CA: SAGE Publications

Pearl, J. (2009): *Causality. Models, Reasoning and Inference. Second Edition.* Cambridge: Cambridge University Press.

Pearl, J. (2015): Trygve Haavelmo and the Emergence of Causal Calculus. *Econometric Theory* **31**, 152–179.

Permutt, T. and Hebel, J.R. (1989): Simultaneous-Equation Estimation in a Clinical Trial of the Effect of Smoking on Birth Weight. *Biometrics* **45**, 619–622.

Rege, M., Telle, K. and Votruba, M. (2012): Social Interaction Effects in Disability Pension Participation: Evidence from Plant Downsizing. *Scandinavian Journal of Economics* **114**, 1208-1239.

Reiersøl, O. (1941): Confluence Analysis by Means of Lag Moments and Other Methods of Confluence Analysis. *Econometrica* **9**, 1–23.

Reiersøl, O. (1945): *Confluence Analysis by Means of Sets of Instrumental Variables.* Stockholm: Almquist & Wiksells.

Reiersøl, O. (1950): Identifiability of a Linear Relation Between Variables Which are Subject to Error. *Econometrica* **18**, 375-389.

Rosenzweig, M.R. and Wolpin, K.I. (2000): Natural "Natural Experiments" in Economics. *Journal of Economic Literature* **38**, 827–874.

Rothenberg, T.J. (1971): Identification in Parametric Models. *Econometrica* **39**, 577-591.

Rubin, D.B. (1974): Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies. *Journal of Educational Psychology* **66**, 688–701.

Sargan, J.D. (1958): Estimation of Economic Relationships Using Instrumental Variables. *Econometrica* **67**, 557-586.

Schmidt, P. (1990): Three-Stage Least Squares with Different Instruments for Different Equations. *Journal of Econometrics* **43**, 389–394.

Smith, G.D. (2010): Mendelian Randomization for Strengthening Causal Inference in Observational Studies: Application to Gene × Environment Interactions. *Perspectives on Psychological Science* **5**, 527-545.

Spanos, A. (1989): On Rereading Haavelmo: A Retrospective View of Econometric Modeling. *Econometric Theory* **5**, 405–429.

Stock, J.H. and Trebbi, F. (2003): Who Invented Instrumental Variable Regression? *Journal of Economic Perspectives* **17**, 177–194.

Stock, J.H. and Watson, M.W. (2011): *Introduction to Econometrics.* 3rd ed. Boston: AddisonWesley.

Theil, H. (1953): Estimation and Simultaneous Correlation in Complete Equation Systems. Centraal Planbureau, Memorandum. Reprinted In: Raj, B. and Koerts, J. (eds.) (1992): *Henri Theil's Contributions to Economics and Econometrics, Vol. 1.* Dordrecht: Kluwer.

Theil, H. (1954): Estimation of Parameters in Econometric Models. *Bulletin of the International Statistical Institute* **34**, 122-129.

Turkington, D.A. (2013): *Generalized Vectorization, Cross-products, and Matrix Calculus.* Cambridge: Cambridge University Press

Wegge, L.L. (1965): Identifiability Criteria for a System of Equations as a Whole. *Australian Journal of Statistics* **7**, 67-77.

White, H. and Chalak, H. (2010): Testing a Conditional Form of Exogeneity. *Economics Letters* **109**, 88–90.

White, H. and Pettenuzzo, D. (2014): Granger Causality, Exogeneity, Cointegration, and Economic Policy analysis. *Journal of Econometrics* **178**, 316–330.

Willassen, Y. (2000): The ET Interview: Professor Olav Reiersøl. *Econometric Theory* **16**, 113–125.

Wright, P.G. (1928): *The Tariff on Animal and Vegetable Oils.* New York: MacMillan.

Wright, S. (1934): The Method of Path Coefficients. *Annals of Mathematical Statistics* **5**, 162–215.