TECHNICAL REPORT
R-278
September 2000

Exogeneity and Superexogeneity: A No-tear Perspective

Judea Pearl
Cognitive Systems Laboratory
Computer Science Department
University of California, Los Angeles, CA 90024
judea@cs.ucla.edu

Abstract

There is hardly a concept in econometrics that is more enigmatic and controvertial than that of *exogeneity*. This report—an edited excerpt from (Pearl 2000)—claims that exogeneity is a rather simple concept, readily definable in terms of standard econometric models, and that the confusion stems primarily from improper usage of statistical vocabulary in a structural framework.

1 Introduction

Economics textbooks invariably warn readers that the distinction between exogenous and endogenous variables is, on the one hand, "most important for model building" (Darnell 1994, p. 127) and, on the other hand, "a subtle and sometimes controversial complication" (Greene 1997, p. 712). Economics students would naturally expect the concepts and tools of causal modeling (e.g., (Pearl 2000)) to shed some light on the subject, and rightly so. This paper offers a simple definition of exogeneity that captures the important nuances appearing in the literature and that is both palatable and precise,

It is fashionable today to distinguish three types of exogeneity: weak, strong, and super (Engle et al. 1983); the former two are statistical and the latter causal. However, the importance of exogeneity—and the reason for its controversial status—lies in its implications for policy interventions. Some economists believe, therefore, that only the causal aspect (i.e. superexogeneity) deserves the exogenous title and that the statistical versions are unwarranted intruders that tend to confuse issues of identification and interpretability with those of estimation efficiency (Ed Leamer, personal communication).[1] This paper will serve both camps by starting with a simple definition of causal exogeneity and then offering a more general definition, from which both the causal and the statistical aspects would follow as special cases. Thus, what we call "exogeneity" corresponds to what Engle et al. called "superexogeneity," a notion that captures economists' interest in the structural invariance of certain relationships under policy intervention.

[1] Similiar opinions have also been communicated by John Aldrich and James Heckman. See also Aldrich (1993).

2 Exogeneity: Motivation and Definition

Suppose that we consider intervening on a set of variables X and that we wish to characterize the statistical behavior of a set Y of outcome variables under the intervention $do(X = x)$. [2] Denote the postintervention distribution of Y by the usual expression $P(y|do(x))$. If we are interested in a set λ of parameters of that distribution, then our task is to estimate $\lambda[P(y|do(x)]$ from the available data. However, the data available is typically generated under a different set of conditions: X was not held constant but instead was allowed to vary with whatever economical pressures and expectations prompted decision makers to set X in the past. Denoting the process that generated data in the past by M and the probability distribution associated with M by $P_M(v)$, we ask whether $\lambda[P_M(y|do(x)]$ can be estimated consistently from samples drawn from $P_M(v)$, given our background knowledge T (connoting "theory") about M. This is essentially the problem of identification that familiar to most economists and that is extended to nonparametric causal quantities in (Pearl 1995a and 2000). There is one important difference though; we now ask whether $\lambda[P(y|do(x)]$ can be identified from the conditional distribution $P(y|x)$ alone, instead of from the entire joint distribution $P(v)$. When identification holds under this restricted condition, X is said to be *exogenous* relative to (Y, λ, T).

We may state this formally as follows.

Definition 1 (Exogeneity)
Let X and Y be two sets of variables, and let λ be any set of parameters of the postintervention probability $P(y|do(x))$. We say that X is exogenous relative to (Y, λ, T) if λ is identifiable from the conditional distribution $P(y|x)$, that is, if

$$P_{M_1}(y|x) = P_{M_2}(y|x) \implies \lambda[P_{M_1}(y|do(x))] = \lambda[P_{M_2}(y|do(x))] \tag{1}$$

for any two models, M_1 and M_2, satisfying theory T.

In the special case where λ constitutes a complete specification of the postintervention probabilities, (1) reduces to the implication

$$P_{M_1}(y|x) = P_{M_2}(y|x) \implies P_{M_1}(y|do(x)) = P_{M_2}(y|do(x)). \tag{2}$$

If we further assume that, for every $P(y|x)$, our theory T does not a priori exclude some model M_2 satisfying $P_{M_2}(y|do(x)) = P_{M_2}(y|x)$,[3] then (2) reduces to the equality

$$P(y|do(x)) = P(y|x), \tag{3}$$

a condition recognized as "no confounding" (see (Greenland et al. 1999; Pearl 2000, Sections 3.3 and 6.2)). Equation (3) follows (from (2)) because (2) must hold for all M_1 in T. Note that, since the theory T is not mentioned explicitly, (3) can be applied to any individual

[2]Precise semantics of the $do(X = x)$ operator, in terms of structural equation is given in (Pearl, 1995 and 2000), the novice reader may simply interpret $do(X = x)$ to mean a local intervention that replaces the equations that determine the variables in X by constants, $X = x$.

[3]For example, if T stands for all models possessing the same graph structure, then such M_2 is not a priori excluded.

model M and can be taken as yet another definition of exogeneity—albeit a stronger one than (1).

The motivation for insisting that λ be identifiable from the conditional distribution $P(y|x)$ alone, even though the marginal distribution $P(x)$ is available, lies in its ramification for the process of estimation. As stated in (3), discovering that X is exogenous permits us to predict the effect of interventions (in X) directly from passive observations, without even adjusting for confounding factors. Recent developments in causal modeling have led to a simple graphical tests of exogeneity: X is exogenous for Y if there is no unblocked back-door path from X to Y (Pearl, 1995). This test supplements the declarative definition of (3) with a procedural definition and thus completes the formalization of exogeneity. That the invariance properties usually attributable to superexogeneity are discernible from the topology of the causal diagram should come as no surprise, considering that each causal diagram represents a structural model and that each structural model already embodies the invariance assumptions necessary for policy predictions (see Pearl 2000, page 160).[4]

3 The Coarseness of Statistical Vocabulary

Leamer (1985) defined X to be exogenous if $P(y|x)$ remains invariant to changes in the *"process that generates"* X. This definition coincides[5] with (3) because $P(y|do(x))$ is governed by a structural model in which the equations determining X are wiped out; thus, $P(y|x)$ must be insensitive to the nature of those equations. In contrast, Engle et al. (1983) defined exogeneity (i.e., their superexogeneity) in terms of changes in the *"marginal density"* of X; as usual, the transition from process language to statistical terminology leads to ambiguities. According to Engle et al. (1983, p. 284), exogeneity requires that all the parameters of the conditional distribution $P(y|x)$ be "invariant for any change in the distribution of the conditioning variables"[6] (i.e. $P(x)$). This requirement of constancy under *any* change in $P(x)$ is too strong—changing conditions or new observations can easily alter both $P(x)$ and $P(y|x)$ even when X is perfectly exogenous. (To illustrate, consider a change that turns a randomized experiment, where X is indisputably exogenous, into a nonrandomized experiment; we should not insist on $P(y|x)$ remaining invariant under such change.) The class of changes considered must be restricted to local modification of the mechanisms (or equations) that determine X, as stated by Leamer, and this restriction must be incorporated into any definition of exogeneity. In order to make this restriction precise, however, the vocabulary

[4]Lack of attention to this fact is one of the most perplexing phenoenon in modern teaching of econometrics. I have asked dozens of economists (including authors of seminal papers and classical texts) whether super-exogeneity is a property that one can discern from a complete specification of a structural econometric model M. Not one has responded with an unqualified "yes". Some responded with noncommittal qualifications (e.g., "it depends if M has more that one equation" or "exogeneity requires change, which may contradict the claim that M is structural" or "the marginal model for z must be considered," or "conditioning on z may not be valid") but the general attitude was that superexogeneity requires some extra assumptions (regarding policy interventions), assumptions that are not part of standard econometric models. If the answer was a simple "yes," so the argument goes, economists would have defined exogeneity directly, in terms of the model's equations, and there would be no disagreement on whether the definition serves its purpose.

[5]Provided that changes are confined to modification of functions without expanding the set of arguments (i.e. parents) in each function.

[6]This requirement is repeated verbatim in Darnell (1994, p. 131) and Maddala (1992, p. 192).

of structural equations must be invoked as in the definition of $P(y|do(x))$; the vocabulary of marginal and conditional densities is far too coarse to properly define the changes against which $P(y|x)$ ought to remain invariant.

4 Exogeneity: A General Definition

We are now ready to define a more general notion of exogeneity, one that includes "weak" and "super" exogeneities under the same umbrella.[7] Toward that end, we remove from Definition 1 the restriction that λ must represent features of the postintervention distribution. Instead, we allow λ to represent *any* feature of the underlying model M, including structural features such as path coefficients, causal effects, and counterfactuals, and including statistical features (which could, of course, be ascertained from the joint distribution alone). With this generalization, we also obtain a simpler definition of exogeneity.

Definition 2 (General exogeneity)
Let X and Y be two sets of variables, and let λ be any set of parameters defined on a structural model M in a theory T. We say that X is exogenous relative to (Y, λ, T) if λ is identifiable from the conditional distribution $P(y|x)$, that is,

$$P_{M_1}(y|x) = P_{M_2}(y|x) \implies \lambda(M_1) = \lambda(M_2) \tag{4}$$

for any two models, M_1 and M_2, satisfying theory T.

When λ consists of structural parameters, such as path coefficients or causal effects, (4) expresses invariance to a variety of interventions, not merely $do(X = x)$. Although the interventions themselves are not mentioned explicitly in Eq. (4), the equality $\lambda(M_1) = \lambda(M_2)$ reflects such interventions through the structural character of λ. In particular, if λ stands for the values of the causal effect function $P(y|do(x))$ at selected points of x and y, then (4) reduces to the implication

$$P_{M_1}(y|x) = P_{M_2}(y|x) \implies P_{M_1}(y|do(x)) = P_{M_2}(y|do(x)), \tag{5}$$

which is identical to (2). Hence the causal properties of exogeneity follow.

When λ consists of strictly statistical parameters—such as means, modes, regression coefficients, or other distributional features—the structural features of M do not enter into consideration; we have $\lambda(M) = \lambda(P_M)$ and so (4) reduces to

$$P_1(y|x) = P_2(y|x) \Longrightarrow \lambda(P_1) = \lambda(P_2) \tag{6}$$

for any two probability distributions $P_1(x, y)$ and $P_2(x, y)$ that are consistent with T. We have thus obtained a statistical notion of exogeneity that permits us to ignore the marginal $P(x)$ in the estimation of λ and that we may call "weak exogeneity".[8]

[7] We leave out discussion of "strong" exogeneity, which is a slightly more involved version of weak exogeneity applicable to time-series analysis.

[8] Engle et al. (1983) further imposed a requirement called "variation-free," which is satisfied by default when dealing with genuinely structural models M in which mechanisms do not constrain one another.

Finally, if λ consists of causal effects among variables in Y (excluding X), we obtain a generalized definition of *instrumental variables.* For example, if our interest lies in the causal effect $\lambda = P(w|do(z))$, where W and Z are two sets of variables in Y, then the exogeneity of X relative to this parameter ensures the identification of $P(w|do(z))$ from the conditional probability $P(z, w|x)$. This is indeed the role of an instrumental variable—to assist in the identification of causal effects not involving the instrument. (See (Pearl 1995; Antrist et al. 1996; Balke and Pearl 1997))

5 A Second Statistical Intrusion

A word of caution regarding the language used in most textbooks: exogeneity is frequently defined by asking whether parameters "enter" into the expressions of the conditional or the marginal density. For example, Maddala (1992, p. 392) defined weak exogeneity as the requirement that the marginal distribution $P(x)$ "does not involve" λ. Such definitions are not unambiguous, because the question of whether a parameter "enters" a density or whether a density "involves" a parameter are syntax-dependent; different algebraic representations may make certain parameters explicit or obscure. For example, if X and Y are dichotomous, we can define parameters such as

$$\lambda_1 = P(x_0, y_0) + P(x_0, y_1),\ \ \lambda_2 = P(x_0, y_0),\ \text{ and } \lambda = \lambda_2/\lambda_1,$$

and write the marginal probability of X as

$$P(x_0) = \lambda_2/\lambda,\ \ P(x_1) = 1 - \lambda_2/\lambda.$$

The equation $P(x_0) = \lambda_2/\lambda$ shows that λ is certainly "involved" in the marginal probability $P(x_0)$, and one may be tempted to conclude that X is not exogenous relative to λ. Yet X *is* in fact exogenous relative to λ, because the ratio $\lambda = \lambda_2/\lambda_1$ is none other than $P(y_0|x_0)$; hence it is determined uniquely by $P(y_0|x_0)$ as required by (6).[9]

The advantage of the definition given in (4) is that it depends not on the syntactic representation of the density function but rather on its semantical content alone. Parameters are treated as quantities *computed from* a model, and not as mathematical symbols that *describe* a model. Consequently, the definition applies to both statistical and structural parameters and, in fact, to any quantity λ that can be computed from a structural model M, regardless of whether it serves (or may serve) in the description of the marginal or conditional densities.

6 The Mystical Error Term and Cowles Exogeneity

Historically, the definition of exogeneity that has evoked most controversy is the one expressed in terms of correlation between variables and errors (or distrubances). It reads as follows.

[9] Engle et al. (1983, p. 281) and Hendry (1995, pp. 162–3) attempted to overcome this ambiguity by using "reparameterization"—an unnecessary complication.

Definition 3 (Error-Based Exogeneity)
A variable X is exogenous (relative to $\lambda = P(y|do(x))$) if X is independent of all errors that influence Y, except those mediated by X.

This definition, which Hendry and Morgan (1995) trace to Orcutt (1952), became standard in the econometric literature between 1950 and 1970 (e.g. Christ 1966, p. 156; Dhrymes 1970, p. 169) and still serves to guide the thoughts of most econometricians (as in the selection of instrumental variables; Bowden and Turkington 1984). However, it came under criticism in the early 1980s when the distinction between structural errors and regression errors became obscured (Richard 1980). (Regression errors, by definition, are orthogonal to the regressors.) The Cowles Commission logic of structural equations has not reached full mathematical maturity and—by denying notational distinction between structural and regressional parameters—has left all notions based on error terms suspect of ambiguity. The prospect of establishing an entirely new foundation of exogeneity—seemingly free of theoretical terms such as "errors" and "structure" (Engle et al. 1983)—has further dissuaded economists from tidying up the Cowles Commission logic, and criticism of the error-based definition of exogeneity has become increasingly fashionable. For example, Hendry and Morgan (1995) wrote that "the concept of exogeneity rapidly evolved into a loose notion as a property of an observable variable being uncorrelated with an unobserved error," and Imbens (1997) readily agreed that this notion "is inadequate."[10]

These critics are hardly justified if we consider the precision and clarity with which structural errors can be defined in the interventional or counterfactual formalism (e.g. Pearl 1998, 2000). When applied to structural errors, the standard error-based criterion of exogeneity coincides formally with that of (3), as can be verified using the back-door test (Pearl 1995) (with $Z = \emptyset$). Consequently, the standard definition conveys the same information as that embodied in more complicated and less communicable definitions of exogeneity. I am therefore convinced that the standard definition will eventually regain the acceptance and respectability that it has always deserved.

References

[Aldrich, 1993] J. Aldrich. Cowles' exogeneity and core exogeneity. Technical Report Discussion Paper 9308, Department of Economics, University of Southampton, England, 1993.

[Angrist et al., 1996] J.D. Angrist, G.W. Imbens, and Rubin D.B. Identification of causal effects using instrumental variables (with comments). *Journal of the American Statistical Association*, 91(434):444–472, June 1996.

[Balke and Pearl, 1997] A. Balke and J. Pearl. Bounds on treatment effects from studies with imperfect compliance. *Journal of the American Statistical Association*, 92(439):1172–1176, September 1997.

[10] Imbens prefers definitions in terms of experimental metaphors such as "random assignment assumption," fearing, perhaps, that "typically the researcher does not have a firm idea what these disturbances really represent" (Angrist et al. p. 446).

[Bowden and Turkington, 1984] R.J. Bowden and D.A. Turkington. *Instrumental Variables.* Cambridge University Press, Cambridge, England, 1984.

[Christ, 1966] C. Christ. *Econometric Models and Methods.* John Wiley and Sons, Inc., New York, 1966.

[Darnell, 1994] A.C. Darnell. *A Dictionary of Econometrics.* Edward Elgar Publishing Limited, Brookfield, VT, 1994.

[Dhrymes, 1970] P.J. Dhrymes. *Econometrics.* Springer-Verlag, New York, 1970.

[Engle et al., 1983] R.F. Engle, D.F. Hendry, and J.F. Richard. Exogeneity. *Econometrica*, 51:277–304, 1983.

[Greene, 1997] W.H. Greene. *Econometric Analysis.* Prentice Hall, Upper Saddle River, NJ, 1997.

[Greenland et al., 1999] S. Greenland, J.M. Robins, and J. Pearl. Confounding and collapsibility in causal inference. *Statistical Science*, 14(1):29–46, February 1999.

[Hendry and Morgan, 1995] D.F. Hendry and M.S. Morgan. *The Foundations of Econometric Analysis.* Cambridge University Press, Cambridge, 1995.

[Hendry, 1995] David F. Hendry. *Dynamic Econometrics.* Oxford University Press, New York, 1995.

[Imbens, 1997] G.W. Imbens. Book reviews. *Journal of Applied Econometrics*, 12, 1997.

[Leamer, 1985] E.E. Leamer. Vector autoregressions for causal inference? *Carnegie-Rochester Conference Series on Public Policy*, 22:255–304, 1985.

[Maddala, 1992] G.S. Maddala. *Introduction to Econometrics.* McMillan, New York, NY, 1992.

[Orcutt, 1952] G.H. Orcutt. Toward a partial redirection of econometrics. *Review of Economics and Statistics*, 34:195–213, 1952.

[Pearl, 1995a] J. Pearl. Causal diagrams for empirical research. *Biometrika*, 82(4):669–710, December 1995.

[Pearl, 1995b] J. Pearl. On the testability of causal models with latent and instrumental variables. In P. Besnard and S. Hanks, editors, *Uncertainty in Artificial Intelligence 11*, pages 435–443. Morgan Kaufmann, 1995.

[Pearl, 1998] J. Pearl. Graphs, causality, and structural equation models. *Sociological Methods and Research*, 27(2):226–284, 1998.

[Pearl, 2000] J. Pearl. *Causality: Models, Reasoning, and Inference.* Cambridge University Press, New York, 2000.

[Richard, 1980] J.F. Richard. Models with several regimes and changes in exogeneity. *Review of Economic Studies*, 47:1–20, 1980.