

◆ 王今朝¹
Jin-chao Wang

计量经济方程识别因果关系的理论潜力和限度

Theoretical Potentials and Limits of Econometric Equation to Identify Causal Relationship

摘要：计量经济学研究无疑取得了不少的成果，但其可信性一直在遭受质疑。计量研究在识别因果上的能力受到了来自数据和方程设定两个方面的影响。从计量所约定的“因果”概念来说，数据概念和衡量的不确定性、数据解释的主体依赖性、数据产生过程对外生变量的依赖性是导致计量难以识别因果的原因之一。而从计量方程角度看，其逻辑结构设定正确与否，以及它与事物决定的开放性本质的一致性程度的高低，决定了计量方程在识别因果上的效力。而从已有研究看，计量经济学在识别哲学因果上的作用是微弱的。从这三个角度看，计量经济学识别因果关系上的可信性是一个极为复杂的事情。

关键词：计量，因果，数据，方程

中图分类号：F064.1 文献标识码：A

一、引言

因果关系构成包括计量经济学在内的所有科学的核心。从1930年计量经济学会成立至今，计量经济学产生了大量理论研究成果。²既有通过引入概率论使计量经济学严谨化，通过明确限定其应用的条件而提高其科学化程度的奠基之作，也有在此基础上的基本方法的创新：50年代出现的工具变量方法、80年代出现的向量自回归方法和协整理论、90年代出现的实验方法，等等。在政策实证和

¹作者简介：王今朝，武汉大学经济发展研究中心教授。

基金项目：国家自然科学基金（71071118）、国家社会科学基金（12BJL009）和教育部人文社会科学规划基金（11YJAZH089）的支持。

²关于计量经济学发展史的轨迹，参见 David A. Bessler, Jeffrey H. Dorfman, et al., *Econometric Developments in Agricultural and Resource Economics: The First 100 Years*, *American J. of Agricultural Economics*, Vol. 92, Issue 2, 2010, pp. 571-589. 关于《计量经济学杂志》（*Journal of Econometrics*）1973年创刊后计量经济学发展的一些情况，参见 Takeshi Amemiya, *Thirty-Five Years of Journal of Econometrics*, *Far Eastern Meetings of Econometric Society in Singapore* in July 16-18, 2008.

理论建设领域，计量研究也取得相当进展。比如，针对货币主义学派主张的货币在解释经济周期上的显著效力，计量经济学研究得出了如下一般性结论：货币因素只能解释 0-25% 的经济波动，而历史上的所有经济衰退都不能归因为货币政策。这可以看作是为宏观经济学的一个重大问题画上了一个句号。又如，计量的工作确实已经让宏观经济学改变了许多（比如，关于 IS-LM 曲线、供给函数、需求函数、Philips 曲线），也正是在这个过程中，计量展现出了它对宏观经济学的工具价值。

尽管理论和实证研究在取得进展，但计量是在肯定-否定-否定之否定的螺旋中上升的。这些否定有的来自计量领域自身的学者，有的来自经济学领域内而计量领域外的学者。比如，对计量经济学用于识别因果的某种具体方法的肯定引起了对这种肯定的方法论否定，而方法论否定自身又引起了某种否定之否定。20 世纪 30 年代，凯恩斯称丁伯根(J. Tinbergen)的宏观计量为炼金术，而弗里希(R. Frisch)认为当时的大量回归研究所确认的因果是虚假的；在 70 年代，卢卡斯(R. Lucas)指出克莱因(R.L. Klein)等人的宏观计量方程研究无法有效地用于它所宣称的政策分析，而在 80 年代，西姆斯(C.A. Sims)指出，克莱因等人的宏观计量把因果关系建立在不可信的识别假设基础之上。在 20 世纪最后 10 年，在计量经济学内部，半参数和非参数计量领域的学者批评参数计量领域的学者依赖太多假设，在 21 世纪第一个十年结束的时候，后者反过来指出前者的假设一点儿也不少。

中国在引进西方计量近 30 年后，有几位重要学者基于因果识别的角度对计量的模型设定、数据依赖、研究结构等进行了方法论和哲学高度的研究。如洪永淼指出，计量经济学建立在两个基本公理基础之上：任何经济系统都可以看作是服从一定概率分布的随机过程；任何经济现象都可以看作是随机数据产生过程的实现；尽管计量拥有区别于数学、统计的特征，因而具有较为独特的功能实现机制，但它无法被单独用于识别因果。李子奈概括出，计量总体模型设计必须遵循“唯一性”、“一般性”、“现实性”、“统计检验必要性”、“经济主体动力学关系导向”五大原则。李子奈还总结出，计量既可检验又可发现，既可归纳又可演绎，既可证伪又可证实，但在构建模型的方法上要从一般到特殊，而在结论上只能认为它具有相对性。王美今等指出，计量经济学在 21 世纪的西方开始经历第三次可信性革命，而在中国，计量经济学在识别因果上的可信性也亟待加强。

上述西方计量发展轨迹和中国当前研究焦点提出了这样一个元方法论问题，即计量经济学在识别因果上究竟受到什么限制？西方目前尽管提出了可信性革命的问题，但它依然试图依靠在具体方法上的突破实现这种革命。这种可信性革命是否可能成功呢？本文则在上述研究基础上，利用现代经济学方法论特别是开放系统方法论的成果，探讨计量经济学在识别因果上受到的本体论、方法论和认识论限制。本文结构如下：第二部分研究计量经济学中的因果关系概念，第三和第四部分分别研究数据性质和方程设定对计量识别它所定义的因果上的限制，第五部分研究哲学意义上的因果以及计量对识别这种因果的限制，第六部分对如何改进计量识别因果提出方案。

二、计量经济学中的因果定义：作为一事物与其它事物之间外在联系的因果关系

从事物的自身规定性的角度看，一个事物有其自身发展规律，同时又与其它事物发生相互作用。因此，因果关系有两种：反映事物自身发展规律的因果关系和表示一个事物与其它事物外在联系的因果关系。本文延续哲学的传统，把前者中的原因称为内因，把后者中的原因称为外因（或条件）。

在20世纪50年代，有几位重要学者从逻辑的角度探讨了因果关系的定义，却没有区分内因与外因。比如，后来获诺贝尔经济学奖的西蒙（H. Simon）用集合论的语言对因果给出了一个高度形式化的定义。他明确指出他的定义是逻辑的，而非本体论的，这就为他把雨量充沛看成小麦丰收的原因并把小麦丰收看成是小麦价格下降的原因提供了前提假设。而从物质自身规定性角度看，雨水与小麦是两种在本体上完全不同的东西，前者不是后者的内因，尽管在逻辑上二者之间可以建立函数关系。又如，著名数学家、控制论之父维纳（N. Wiener）提出，如果关于P的信息可以用于提高对Q的预测，P就定义为Q的原因。¹这里，P和Q可以是变量或向量。维纳的这个因果关系定义中的原因也是既可以是内因，也可以是外因。几乎与他们同时，一位经济学家专们著文，探讨了计量与因果的关系，表达了类似的观点。很显然，他们的定义都没有区分内因和外因。尽管如此，但从表述上看，他们的观点都是建立在唯物主义基础上的，即都没有把主观（内因和外因）与客观（内因和外因）混淆起来。

在他们成果出现约50年后，却有重要的学者不仅混淆了内因和外因，而且混淆了主观与客观。2005年，诺贝尔经济学奖2000年得主海克曼（J.J. Heckman）提出了一个关于因果的“科学模型”。他认为，在计量经济学中，因果关系是一个建立在假说基础上的模型的性质（property of a model of hypotheticals）。他进一步认为，模型是在脑海中存在的，因而，因果也是在脑海中存在的。在海克曼对计量因果的这个定义中，可以看到弗里德曼方法论的影子，也可以看到黑格尔哲学的影子。如果脑海中的模型是对现实世界的真实的反映，那么，这种因果就具有客观性；而如果脑海中的模型只是对现实世界的扭曲的，或者是过度简化的反映，那么，这种因果关系就是带有主观偏向性的了。2008年，他又提出“计量因果”（econometric causality）的概念和框架，用以处理政策效应这种短期性问题。有学者在考察海克曼的因果推断何时有效。同时也有学者认为，海克曼等人的因果推断需要被一般化。但无论怎样一般化，如果一个学者对因果所做的定义既容忍对世界的扭曲的认识，又容纳对世界的客观的认识，就不具有应有的严谨性。皮尔等考察了计量研究的外在合理性（external validity）问题，即一项研究得出的因果关系能否用于另一个问题。运用皮尔等所考察的问题可以向海克曼提出这样的问题：一个脑海中存在的主观因果关系在什么条件下才是与客观一致的因果关系？

无论是主观还是客观意义上的内因和外因的区分，它们与结果所形成的关系都可用如下逻辑公式来表示： $P \rightarrow Q$ ，读作如果P则Q。 $P \rightarrow Q$ 表示P在主观或客观上成为Q的内因或外因。这里，P、Q表示一个集合，即 $P = \{P_1, \dots, P_n\}$ ， $Q = \{Q_1, \dots, Q_m\}$ 。但是，在其中，我们允许 P_1, \dots, P_n 可能对 Q_i 起不同的作用，而且允许每个 P_i 对 Q_j 的作用发生变化，以反映有的 P_i 可以作为内因，而

¹这种定义也被美国法学界所广泛使用。

有的 P_i 可作为外因。如果不对各种变量进行这种区分,就相当于从一开始就假设众多变量是某个变量的同等意义上的原因了。¹在有的问题情境中,可能无需区分内因和外因,这时所有的 P_i 可能都被视为外因。这个逻辑模型足够一般,能够容纳不同的经济理论体系。比如,按照马歇尔理论,劳动、投资、土地、组织是经济增长率和收入分配的原因,而按照马克思理论,劳动是经济增长率的原因。这个逻辑模型也足够基本,它能够容纳各种计量方程的形式和数据的性质。比如,当 $m=1$ 和 $n=1$ 时, $P \rightarrow Q$ 表示一元计量方程;当 $m=1$ 而 $n>1$ 时,它表示多元计量方程;当 $m>1$ 而 $n \geq m$ 时,它表示联立计量方程。如果 P 是截面数据,它表示截面方程;如果 P 是时间序列数据,它表示时间序列方程;如果 P 是面板数据,它表示面板方程。因而以下我们就可以用它来研究计量方程在识别因果关系上的潜力和限度。为了引用的方便,我们写成如下的方程形式:

方程(1)(1)几乎被所有计量经济学教材都列为基本模型。尽管公式(1)(1)容纳内因-结果关系,但计量经济学的实践大都回避了这种关系,而集中于探讨外因-结果关系。由于不同属性的条件可能很多,条件之间可以是互补关系(资本和技术对于经济增长是互补关系),也可以是替代关系(如为了达到一定的经济增长率,技术可以部分替代资本),再加上事物本身(宏观的、微观的、金融的)的多元性,仅仅研究这些外因-结果关系就可以产生许多成果。那么,计量在利用方程(1)(1)识别外因-结果关系上的能力如何呢?按照丁伯根等,构成计量确认因果关系的困难来自于两个方面:一是数据的性质,二是由对这些数据进行抽象所得到的变量所构成的方程(不管是微观还是宏观)的性质。下面第三部分和第四部分就从这两个方面展开分析。

三、数据性质对计量方程识别外因-结果关系的影响

“让数据说话”被一些学者看作是计量工作的一个信条。然而,数据能够说出怎样的话取决于数据的性质。为了探讨数据性质对计量方程识别因果能力的影响,我们在这部分假设计量方法不存在任何瑕疵。在这一假设下,下面的分析表明,计量研究中的数据会因存在无法被计量模型所吸纳的不确定性而对外因-结果识别产生重大影响。在这种情况下,即使计量所用数据通过各种必要检验而具有“计量充分性”,关于因果的计量结论也不具有足够的可信性。

(一) 概念内涵的不确定性

在方程(1)(1)中,每个 P_i 仿佛都具有良好的定义。实际不然。即使著名的经济学家在一些重要概念上也存在严重分歧甚至对立。在宏观领域,经济学家之间关于资本、GDP、劳动、人力资本、技术水平、价格等诸多概念都存在许多争论,而他们与社会学家、政治学家在这些概念上的争论会更大。在微观领域,经济学家关于个人特征(包括偏好、收入、储蓄等)的定义存在许多分歧;并且由于不同个体的差异性,同一收入的经济含义是不一样的,这就如同水对城市人群的意义和水对在沙漠中的人的意义不同一样。在截面数据中,不同对象所对应的同一变量的意义是不同的。比如,工人的工资与同时期企业家的收入虽然都可以称为消费者收入,它们的产生则是服从不同规律的。

¹比如,经济学中就有这样的例子。西方经济理论把投资、土地、组织看作是经济增长率的同等意义上的原因,它们都服从边际生产力分配规律,最多只是因为生产函数不同以及各自数量不同而得到不同的收入。

将工资与利润混同在 GDP 中建立消费与 GDP 的关系与其说带来科学的因果关系的发现,毋宁说是用相关性代替了因果性。在时间序列数据中,同一对象在不同时间所产生的数据存在差别,更不用说去用过去的拟合出的计量方程去预测未来了。

在这种情况下进行以得到定量估计为目的的计量分析就只能是建立在某些明显或隐含做出的假定(用 A 来表示)基础上了。如果换作其他假定(用 B 来表示),就需要对计量估计所依赖的数据的背后概念进行修订。如在宏观经济学统计中,修订数据已经是一种常态。而且,统计史上不乏统计体系发生转换的案例。当概念变化后,对于实际的经济状况,我们至少会有两组数据:

$\{P_1^A, \dots, P_n^A\}$ 和 $\{P_1^B, \dots, P_n^B\}$ 。这两组数据所对应的因变量将会是 Q^A 、 Q^B 。鉴于这些假定之间的

区别不应该被忽视,理性的预期是 $Q^A \neq Q^B$, 并且 $Q^A \neq Q^B + \varepsilon$, 其中, ε 为 0 均值的随机变量;

换句话说,我们总可以找到两类不同的合理假定,建立在它们基础上所得到的 Q^A 、 Q^B , 满足

$\|Q^A - Q^B\| \square 0$, 而且它们各自在自己所依赖的假定中可能都是正确的。那么, A 和 B 之间的差异

又是由于什么原因呢? 做出这些假定的学者的学术观点、政治观点、个人能力、个人目的等上的差异都是可能的原因。这些与认识论有关的原因就不让计量经济学的公理之一失效了。因此,基于任何概念基础上的计量研究就只能算作一家之言了。难怪经济学方法论领域的学者已经发现,证伪一个命题不是如原来所想象的那么容易。

基于不同概念必然产生不同的数据组,这会对计量研究产生重大影响。如果计量研究的目的仅仅是对过去进行一个估计,则同一客观事物对应两组变量,如 $\{Q^A; P_1^A, \dots, P_n^A\}$ 和 $\{Q^B; P_1^B, \dots, P_n^B\}$, 意味着基于这样两个数据组中的任何一组所进行的任何计量方法的估计都会受到另一组数据所产生的计量估计的挑战。在两种不同概念下存在两种不同数据不是一个概率事件,而是一个必然事件。计量经济学无论怎样发展都无法消除这样产生出来的两种估计。于是,计量所确定的因果关系就是不确定的了。

如果计量研究的目的是对政策效应进行估计,同一客观事物对应两组变量又意味着什么呢? 政策效应的出现在时间上总是要晚于政策制定,而政策制定总是需要设定一个经验模型。¹不妨假设所有人都会对政策实施后的效果达成一致意见,我们不妨把这个人们具有一致意见的结果设为

$\{P_1^C, \dots, P_n^C\}$, 并认为它将对应 Q^C 。于是,从不同的评估主体来看,无论采用怎样的计量工具,

在所有评估主体对计量工具也具有一致意见的情况下,政策的效应性也有两个: $\|Q^C - Q^A\|$ 和

$\|Q^C - Q^B\|$ 。根据三角不等式, $\|Q^C - Q^A\| + \|Q^C - Q^B\| \geq \|Q^A - Q^B\|$, 既然已有 $\|Q^A - Q^B\| \square 0$,

¹比如, 参见 CristophRothe, Nonparametric estimation of distributional policy effects, Journal of Econometrics, 155 (2010), 56-70.

那么，要么 $\|Q^c - Q^A\| \neq 0$ ，要么 $\|Q^c - Q^B\| \neq 0$ 。即基于不同的评估主体，政策效应性将很可能会有完全不同的评价。而这两个不同的评价中哪个是衡量政策效用性的合适指标呢？如果再考虑到不同的评估主体对计量工具合理性的评估也会不同，¹这两个指标的作为合适指标的不确定性就更为明显了。这时，科学研究就不得不变成一个政治过程，即需要通过某种权威来强制采取某种概念基础上的数据。美国的国民经济研究局、穆迪公司为美国社会提供权威数据，但2008年的经济危机表明，这种数据的权威只是来自于约定，而不是来自于数据的科学性。

社会科学领域概念内涵的上述不确定性使得它的数据性质与自然科学领域的数据性质具有根本的差异性。这样，原来基于自然科学领域内的数据性质所产生的计量工具（最小二乘法）的逻辑基础就变动了。如果说自然科学领域内最小二乘法的理论和应用是建立在坚实的基础之上，社会科学领域内最小二乘法的理论和应用就缺乏那么坚实的基础了。

（二）数据衡量的不确定性

即使所有研究者都对所有概念的内涵形成了相同意见，对这些概念的数量衡量通常也并不唯一，而是对研究主体形成依赖。首先，在客观上，可能缺乏对数据衡量的不变标准。比如，就资本衡量而言，假设英国剑桥学者同意美国剑桥的意见，承认资本同质对于某些目的的合理性，但还有一个问题，即到底是用历史成本还是重置成本来计算它们的总量呢？随着技术进步，资本的重置成本可能远远低于历史成本。即使固定了重置成本法或历史成本法，资本的衡量是依赖于利率的，但利率是可变的。因此，资本总量在理论上可以存在无数个数值。这意味着将会有无数个计量结果。其次，数据衡量还受到数据提供者和采集者主观意志的系统性影响。大量的计量研究所依赖的数据不是如物理数据那样出于经受过高级训练的学者之手，而是出于仅仅受过初等训练，面对经费、时间等限制的人之手，并且这些数据来自于他们对他人的回忆的记录。这些回忆是建立在使用所有信息的基础上的吗？这些回忆是否存在选择性记忆呢？回忆出来的信息是否准确地表达出来呢？调查员如何保证这些信息的真实性呢？这种调查过程使得最高级的调查员所得出的数据也与物理实验室里专业工作者所得到的数据在产生机理上具有根本的性质差异。第三，数据衡量还受到数据使用者的主观意志的影响。在计量应用中，建模者为了证实某种先验理论而有意掩盖或摒弃与理论相悖的数据信息可谓普遍现象。比如，如果商品房价格不显著，就换成住宅价格。如果房子价格作为解释变量不合适，就挑选其他变量。在这三种因素作用下产生的数据衡量不确定性对计量结果的影响与概念内涵不确定性对后者的影响类似，它与自然科学领域的测量误差具有根本不同的意义。

任何计量数据都无法避免上述概念不确定性和衡量不确定性问题。而计量对大量数据的要求使上述问题更为严重。数据越多，概念内涵和数据衡量导致的问题就越会积累。由于存在这两种不确定性以及相关累计效应，即使数据通过了各种异方差检验、内生性检验、单位根检验、序列相关检

¹大量教育收益率的实证研究把受教育年限作为工具变量，而如果不考虑出身差异对个人收入的影响，这种以教育年限为工具变量的方法就会导致极大偏差。在美国，那些处于收入最高等级的人通常都是出身最好的人，这让那些犹太裔等少数民族种裔的精英都愤愤不平。参见：莫尼卡·兰利：《花旗帝国：金融奇才桑迪·韦尔传奇》，刘晶晶译，中信出版社，2005年。

验等，也只是说明在形式上数据满足了相应统计工具的要求，并不能保证这些数据不存在概念确定和衡量上的问题，更不能保证基于它们所得出的因果推断的合理性。也就是说，它们只是保证计量因果识别的初步必要条件。研究者即使把这些工作做得再好，这些检验本身不存在这样那样的问题，¹也只能保证作者进一步语言严谨化，无法保证计量工具能够真正识别出外因-结果关系。

（三）数据解释的主体依赖性

数据解释也是影响计量识别外因-效果关系效应性高低的一个变量。首先，对原始数据的解释中可能存在主体依赖性。比如，一个 4.3% 增长率和 2.0% 标准差的经济（美国 20 世纪 60 年代），与 3.0% 的增长率和 1.5% 标准差的经济相比（美国 90 年代），哪个更具有波动性呢？即使由经济学领域的专家来回答，答案也取决于衡量标准的选择。其次，由于社会人群的异质性（西方社会学理论的利益集团理论和马克思的阶级理论都是基于社会人群的异质性而产生），同一经济数据对于不同人群的意义不同，不同人群对同一经济数据的感知也不可能一样。更一般地看，在社会科学中许多种类的数据都存在解释的主体依赖性。比如，成本收益分析是标准的经济学方法，却在经济衡量上存在严重的不合理性。在一个贫富分化的世界里，低收入人群和高收入人群对于同一商品的评价不可能一样。用市场价格作为信号指导资源配置，必然会偏向高收入人群，除非把高收入人群的经济评价缩小到合适的程度。可是，这个合适程度又怎样来决定呢？如果成本收益法存在衡量的不合理性，那么，在这个方法支配下所产生的任何数据（比如具有高度确定性的股票市场交易数据）就都会产生合理性问题。第三，对于结论数据的解释也会有主体依赖性。大量应用计量研究对于结论的解释过于简单化。比如，当一个计量研究考察了人力资本与经济增长率之间的关系之后，就不考虑本国实际情况而把人力资本作为本国最主要的经济增长政策。类似这样的结论是认识论和方法论简单化的表现。根据(1)(1)式，我们知道，经济增长率取决于一系列因素，比如，如果没有大规模的技术创新，人力资本投资再多，最后的结果也只能是造成更多知识失业，而较少经济增长。又如，宏观协整数据所表现出的计量稳定性可能意味着其分量的高度不稳定性，而分量的不稳定性如果发生在低收入人群，社会动荡就会产生，这样就会导致宏观计量研究忽视对社会稳定发展的最显著影响变量。

（四）数据产生过程对其它变量的依赖性

除了上述四种认识论方面的原因外，数据的本体属性对于计量能否识别外因-结果关系也非常重要，相对上述四种认识论因素也许更为重要，尽管这种影响的数学结构与概念内涵影响的数学结构类似。首先，任何数据都是在一种观念下产生的，而观念是可变的。当观念变化后，数据就会发生变化。其次，任何一组相关数据都是在一定环境下产生的，环境是可变的。当环境成为变量后，数据就会发生变化。比如，一个国家的宏观经济数据是与其他国家密切相关的，而一个企业的数据是与其他企业、国家政策相关的。而任何一个计量模型都无法预测其它国家、其它企业的变动，而后

¹比如，关于单位根检验存在的问题，参见 Sims, Christopher A. [1988]. "Bayesian Skepticism on Unit Root Econometrics," *Journal of Economic Dynamics and Control*, 12, 463-74.

者一定是在变动的，这就使得任何使用经验数据的计量研究所得出的因果关系从长期看一定是经验主义性质的研究，也就是它所识别出来的因果在长期内是无效的。比如，美国经济的繁荣会由于它国经济的长足发展而受到巨大影响，因而基于目前西方增长理论所得出的美国经济增长源泉分析在长期内都将失去效力。用政治哲学的术语来说，任何基于经验数据而不考虑数据性质的动态性本质所得到的计量结果都属于经验主义。这种经验主义与经济学、管理学等所主张的创新是格格不入的，也是与体制转变（regime switching）作为经常性事实不吻合的。可变系数的计量方程在一定程度上或许可以抓住历史数据所具有的结构变化，但有什么方法能够用来检验未来可能发生的联系呢？

以上探讨表明，概率论框架并没有让计量经济学在因果识别上摆脱数据性质的限制。使用概率工具确实为计量经济学提供了坚实的数学基础，但它并没有解决弗里希等人对数据问题的担忧，特别是没有解决由社会科学领域自身性质导致的数据性质与自然科学领域数据性质的差异。如果说概率论在自然科学领域的适用性是很高的，它在社会科学领域内的作用必然较低。早已有统计学权威指出，在没有足够证据时，假设一个总体具有某种分布形式并进行参数估计和检验是不负责的，结果是不可靠的，甚至是灾难性的。在现实中，因数据过多、过少、数据质量差、数据质量与目的不一致、数据过滤、动态性、异质性、观察的非同质性（non-homogeneous）、数据对其他变量的依赖性而导致计量估计的偏差无法避免。已有的计量经济学研究对于上述数据问题的重视程度是不够的。

四、经济学的复杂系统本质对计量方程识别外因-结果关系的影响

计量经济学的理论研究没有想出什么有效的办法根本解决数据问题，他们不得不满足于一定的数据质量，而寄希望于找到合适的方法来提高计量识别外因-结果的能力。计量无疑可以通过方法的科学性来提高自身识别外因-结果的能力，但提高的程度受到如下两大重要因素的制约而极为有限。

（一）逻辑结构设定对计量方程识别外因-结果关系的影响

方程(1)(1)在形式上把所有自变量之间的关系视为一种平行结构，而实际上，自变量之间存在着其他形式的逻辑结构。图 1 表示， P_1 是决定 Q 的根本原因，同时也是直接原因之一。此外， P_1 还决定 P_2 和 P_3 ，并通过它们间接影响 Q 。 P_2 和 P_3 之间又存在相互作用。比如， P_1 表示市场经济发育程度， P_2 和 P_3 分别表示财政和金融制度，而 Q 表示一个国家的经济增长率。对于这样一种关系，在数学上，它可以表示为如下系统：

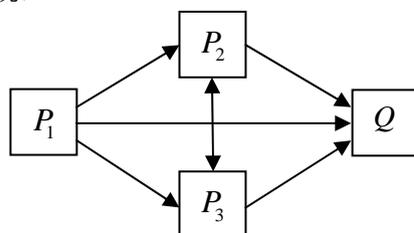


图 1 简单分层的逻辑因果关系

形式上，(2)(2)是一个联立方程组。不管是用联立方程组的传统解法求解，还是运用自回归方程对它进行估计，不仅面临着技术上的困难，如上述函数关系究竟对应于怎样的实际经济数据（如当期数据还是滞后数据，滞后几期等），而且面临一个根本的局限：即便能够给出方程(2)(2)的定量估计，也不能认为是真正的因果关系。这是因为由于世界普遍联系的本质， P_1 并不是一个外生变量，而是会受到其他变量的影响，也就是说，图 1 可能变为图 2。图 2 中虚线表示一种可能的决定关系。如果图 2 代表真正的逻辑关系，那么， P_0 就成为最根本的决定因素，而 P_1 就不是根本的决定因素了。这样，根据图 1 所做出的任何计量结果都必然产生估计偏差。因为基于图 1 的参数估计失去了关于 P_0 的信息。这时， P_1 、 P_2 和 P_3 对于 Q 连恩格尔等人所定义的弱外生性(weak exogeneity)都没有，更不用说强外生性了。这违背了计量检验的基本假设。

(二) 开放系统本质对计量方程识别条件-结果关系的影响

如果我们把上面的 P_0 视为一个新变量，就可以发现导致计量识别因果的另一个困难，即经济学问题的开放系统、普遍联系本质对它的影响。不管是在平行的逻辑结构还是在分层的逻辑结构中，

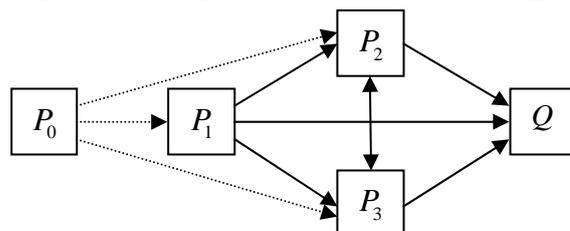


图 2 复杂分层的逻辑因果关系

新显著变量的加入都会对原有系统产生影响。计量经济学文献已经对遗漏显著变量会产生怎样的估计结果进行过研究，也已经提出过对模型设定是否遗漏显著变量的统计检验，也已经提出了一些解决方法，如使用工具变量。可是，按照现代开放系统方法论，由于变量之间的相互依赖作用，对于任何给定的方程，都会存在遗漏变量问题。

开放系统方法论认为，任何事物都应该视为具有自身结构并包含构成该结构的各种因素，而这些内部因素与外部环境、事物相互影响、作用，具有输入(input)和输出(output)共性本质，具有建立新的、重组旧的关系、现象及规律的共性本质。比如，任何生产、交换和消费行为都受到政治、制度、文化等变量的影响，因而处于一个开放系统之中，无所谓同方差异方差、稳定过程不稳定过程、非结构计量模型结构计量模型。而任何一个计量方程一经设定，就具有封闭系统的特性。这个封闭系统是否总是能够逼近那个作为本体的开放系统呢？答案是否定的。任何一个封闭系统下得出的结论只能是在这个封闭系统中才能严格成立。而任何一个现实问题，不管是试图对它进行解释，还是试图进行政策分析，都具有开放系统的本质，甚至对物质、能源、信息等非经济学变量开放。

世界的普遍联系本质说明，对于任何一种研究，不大可能只有一种外生变量，而是在已有的由

$\{P_1, \dots, P_n\}$ 所描述的外生变量之外，还有众多的其他外生变量，以至于形成了一个外生变量的无穷序列： $\{P_0, P_{-1}, P_{-2}, \dots, P_{-k}, \dots\}$ 。因此，计量研究可以形成如下的方程：

可以想见，每加入一个外生变量，就会导致原来 P_1 、 P_2 和 P_3 的系数发生一次变化。每个估计都会或多或少地否定另外的估计。对于企业发展而言，格兰杰提出了这样一种可能性：一位生意不大的企业家，提出一项使他的企业成功的特殊计划。而其他从事同样生意的企业家普遍认为这项计划不可能成功。可是，这位企业家确信会发生一个特殊的情形，使得他的企业取得特别的成功。很显然，这种特殊情形构成其他企业家用以得出他们观点的封闭系统以外的变量。

著名的卢卡斯批评指出政策对计量方程有效性的影响，也是说明经济的开放系统本质的一个经典案例。与格兰杰的例子不同的是，卢卡斯批评所依赖的机理是原来未曾想到的关系而非变量。不妨用 P_0 来表示政策， P_1 、 P_2 和 P_3 分别表示三种不同的行为。卢卡斯批评认为， P_0 的出现会对 P_1 、 P_2 和 P_3 产生影响，结果，在 P_0 出现后， Q 对 P_1 、 P_2 和 P_3 的回归将产生不同的参数，这使得 P_0 不再是最优的政策，因而否定了原来的计量方程作为因果关系的合理性。

在计量研究中，未被观察到的潜在变量（被称为过度的异质性）、调整过程或未被注意到的变量和调整过程对计量方程识别因果会产生类似卢卡斯批评和格兰杰案例所包含的效应。比如，我们没有见到任何一个计量研究能够考虑一个国家的基本制度对该国经济中的所有变量产生的最根本影响。由于技术能力、人们的认知能力的变化导致了一个国家的基本制度的可变性（如资本主义自身已经与200年前有较大的不同），很显然，计量研究还没有与这类基础性研究相结合。又如，在计量研究中通常被视为外生变量的偏好实际是内生变量，而衡量偏好内生的计量研究似乎还较少。经济行为作为一种或多或少受到成本收益比较方法影响的行为，当那些主客观条件发生变化后，经济行为的成本结构和收益结构都会发生变化，因而经济行为也一定会发生变化，而相应的因果关系也就必然发生变化了。到目前为止，计量研究的动态化并没有能够刻画出这些动态。

一国增长源泉分解研究和“跨国增长回归”构成方程(3)(3)-(6)(6)模式的两个典型研究领域。¹任何增长理论都可以被模型化为如下形式的生产函数，

其中， f 表示产出，而 x_i 表示要素投入， n 是“要素”种类的数量。从已有的经济理论看，对经济增长能够产生影响的变量有：土地、劳动、资本、技术、组织、制度、人力资本、信息、高有形资本投资率、生育率、距赤道远近、发病率和条件适当的开放政策等等。理论上，随着经济复杂性的增加， n 可以无限地扩大。毫不奇怪，随着考虑的变量的复杂性，利用这种跨国增长回归证明的大量统计学关系中有不少是相互矛盾的：如有的证明外国援助能够促进增长，有的则说外国援助不能促进增长，还有的说外国援助可以在特定条件下促进增长。

正是由于这种机制，一种方法主张计量方程的设定要从一般到特殊。一些计量经济学实践同时列出了多种形式的回归方程设定。这种努力和想法是好的，但开放系统性质使得所有模型都可能面

¹这种开放系统结构在微观计量中也开始被强调，如母亲对孩子的照顾时间对孩子智力发展的影响，参见 Michael Keane, structural vs. atheoretic approaches to econometrics

临误设模型的问题，面临无法确定所有显著变量的问题。在审慎模型选择基础上都要进行模型设定检验，其目的也只是尽可能地排除模型设定的错误，而并不能排除这种错误。即使经过“模型设定-模型估计-模型设定检验-模型再设定”的过程，设定检验内在的一致性也只是保证模型正确分析的必要条件，而非充分条件。“检验，检验，再检验”这一计量经济学的金科玉律最多只能是计量经济分析过程中对部分“杂草”的清除机制。作为开放系统的一个自然推论，在新的时间、新的地点、新的环境下产生的用同一名称命名的数据是否还能具有原来时间、地点和环境产生的数据所具有的结构特性，是并不确定的。或用最新的方法论术语来说，原来的时间、地点和环境下所产生的计量模型是否具有外部合理性，那是需要再检验的。而从同一国家来看，时间差别越大，环境就差别越大；从不同国家来看，地点差别越大，环境差别就越大。根据这些考虑，凯恩斯20世纪30年代指责为炼金术的计量也可以用于包括工具变量显著性的F检验、方程结构的邹至庄检验、嵌套与否的戴维森-麦金农检验，适用在回归模型的敏感性分析、稳健性分析，残差项的0均值正态分布检验，毕竟它们最多是提高了计量识别因果上的概率，并不能保证识别出因果。就连通过了Granger因果关系检验的估计也只是提供了因果关系可能存在的初步证据。

五、本体性质对计量工具识别内因-结果关系的影响

在识别条件-结果上存在的困难也或多或少地存在于识别内因-结果关系上，但内因-结果的关系与条件-结果的关系相比，具有自身的特点。内因-结果关系的数量要比外因-结果的数量要少得多，最重要的内因-结果关系更是稀少。比如，哲学把一切问题归结为物质和意识的关系问题，从而就把物质和意识的因果关系提了出来，并提高到唯一重要的地位。在经济学中，劳动与增长的关系以及所有制/产权与分配的关系在所有经济因果关系中所处的地位类似于物质和意识的关系在哲学中所处的地位。而这两个基本因果关系所衍生出的次级因果关系比计量经济学所探讨的条件-因果关系的数量依然要少。仅从这一点来看，计量工具用于识别内因-结果关系时，具有产生更好结果的条件，但实际并非如此。这种内因-结果关系的计量研究同样会遭受识别困难。这可以用最简单的一元方程来说明。

首先，由于内因对外因的依赖以及内因自身的异质性所产生的识别内因-结果关系的困难。由于只有一个内因时的计量方程才表现为一元方程，不妨把这个内因记为 P_1 ，把这个内因对外因的依赖记为 $P_1 = P_1(P_2, \dots, P_n)$ 。于是，方程(1)(1)写为：

比如，按照李嘉图、马克思的劳动价值论，劳动是价值的唯一的源泉，而土地、资本、技术都是劳动创造价值的条件，也就是说使用不同质量和数量的土地、资本、技术的劳动者可以认为是不同的。¹但从计量经济学的实践看，一元回归方程不会考虑劳动这个变量的这种内生性，而是采用简单统计劳动的办法，把GDP对劳动进行回归。这种回归所得出的参数不可能通过计量经济学所要求的那些检验。对这种回归进行改进，一种可能的办法是把劳动分成若干种，但这也会产生诸多统计

¹同一劳动者支配不同数量的土地、资本和技术会表现出不同的劳动生产率。

问题，其中一个问题就是，当不同的劳动力之间存在重大的外部性时，怎样从统计上把这些劳动力区分开来，又怎样把他们的贡献区分开来呢？比如，如果不是城市的环卫工人把垃圾及时运走，城市里就没有人能够创造价值。从这个角度看，由于一个经济中各种劳动者之间的关系既有分工又有合作，并且这种合作的形式是多种多样的，无论按照什么分工标准来区分他们的贡献，都会在合作领域找到其他标准来质疑这种区分（如工人与农民之间的互相依赖、企业家与工人之间的互相依赖、政府官员与普通百姓之间的互相依赖）。因此，那种细分变量（这里是劳动者）的方法并不合适。

其次，由于结果可能并不表现出来所产生的识别内因-结果关系的困难。这种情形可以用 $Q \in \{0,1\}$ 来说明。 $P_1 = P_1(P_2, \dots, P_n)$ 本身可以看作是一个数据产生过程，由于 $P_1 = P_1(P_2, \dots, P_n)$ 的函数形式可能不同，在不同函数的作用下，即使 (P_2, \dots, P_n) 保持不变， Q 也会有时取值 0，有时取值 1。可以认为，当 $Q = 0$ 时，因果关系不表现出来；当 $Q = 1$ 时，因果表现出来。比如，一些国家拥有大量的劳动力，却长期没有经济增长；有时在劳动力数量和质量都没有显著变化的情况下，经济增长率却可能出现巨变。在这种情况下，用计量经济学检验出来的关系都有什么呢？1840-1949 年，中国拥有世界最庞大的人口，却是世界最贫困的国家之一，把这个时期的财富对劳动做回归得出的结果可以自然地解释为劳动是财富源泉这个命题被证伪。但这个命题真的被这个回归证伪了吗？

第三，反事实情况的存在所产生的识别内因-结果关系的困难。西方国家在过去 200 多年中实现了长期的增长。如果利用诺斯等人的新制度经济学理论，在经济增长与资本主义产权之间进行回归得到的结果，将会证明资本主义产权最有利于经济增长吗？答案是否定的。如果要证明资本主义产权制度最有利于经济增长，那就要证明，在这个世界上不可能存在一种制度比资本主义让经济增长更快。一方面，从理论上讲，这种反事实的分析无法完全穷尽，另一方面，中国等国家的经验证明，社会主义确实曾经实现比资本主义更高的经济增长率。因此，西方发达国家的高富裕程度并不能证明他们的制度就是最优的。

六、建立提高计量识别因果效能的理论基础

至今为止的计量自身发展，不管是从一元回归到多元回归，再到联立方程组、时间序列回归，面板回归、半参数和非参数计量，都试图通过具体方法的改进来提高其识别因果的效能。从本体论的角度看，如果不首先确认一个事物的内在的原因，就谈不上这个原因与这个事物有关的任何数量关系，就谈不上对与这种因果有关的关系的反反复复的认识过程（从实践到认识再到实践）。而把内因和外因概念混合起来，特别是用外因替代内因的分析会导致逻辑上的极大混乱，用一种表面的因果关系去替代真正的因果关系。因此，内因-结果这种哲学上反映自身规律的因果关系具有最为基本的重要性，它是研究外部事物与该事物数量关系的基础。计量研究如果可以辨识因果，首先的、首要的任务就是辨识这种因果。

令人遗憾的是，在计量经济学界，维纳、西蒙、海克曼的因果关系定义都没有尊重哲学中重视本体的传统，混淆了条件与原因。西蒙在讨论因果时所举出的雨水与小麦，是两种完全不同的东西，二者之间的因果不是物质意义上的，即前者不是后者的物质原因。政策与福利也是完全不同的东西。无论是统计还是计量的学者，都是在实用主义地探究原因的效应（effects of causes），而不是在真正地使用反事实方法探究效应的原因（causes of effects），即他们只是在关心政策的效应性，而不是关心事物、现象的根本性的原因，不是在对根本性原因诊断的基础上有针对性地采取政策。对于某些领域的问题来说，他们的框架体系不无道理，也不无继续探讨的必要，但对于经济学的根本问题而言，他们的框架会导致他们以及接收他们观点的人重蹈凯恩斯的覆辙。早就有计量领域的学者指出，哲学家通过把“因果”从他们的辞典里清除出去而代之以“函数关系”或“可预测性”的做法不可能解决问题，如果所有包含因果内容的概念都被抛弃，那么，在学术研究中将会只有描述和正式主义留下来。

为了使得计量工具发挥识别因果的作用，需要建立一种科学的哲学。在哲学界，有些哲学家试图否认物质原因的概念，他们认为，在不同本质的事物之间定义因果关系具有合理性。比如，休谟就否定物质的存在性，因而否定存在物质原因的概念。但他无法回避不同事物之间存在的联系，因而就用任何事物之间的“恒常联系（constant conjunction）”这个看起来重要的特征形式化地定义因果关系。这种形式化的定义抽象掉了不同物质之间的本质差别，为在不同本质事物之间建立种种因果概念提供了“前提基础”。表面上看，休谟的定义似乎具有合理性。在这种定义下所得到的关系或许能够具有他所宣称的“恒常”性，但一个事物自身内在的规律性这种更为重要、最为重要的、内在的、稳定的、单一的恒常性就可能因被那种外在的、表面的、暂时的、多样化的恒常性掩盖、取代而不被揭示出来。

为了使得计量工具发挥识别因果的作用，需要建立一种科学的经济学方法论。从服务于现实的目的来看，函数分析方法比计量方法更具有可靠性，因为数据衡量的上述不确定性不会对函数分析产生影响。函数分析不对数据数量和统计质量产生要求，而只要求数据的典型性。凯恩斯的《货币、利息与就业通论》是函数关系分析优于数量分析的典范，它表明，重大的科学研究可以在极少数据下完成。如果所有的计量研究都可以进行必要的函数分析，则计量在识别因果上的效能将可能会有本质性的特征。

为了获得更为准确的计量模型，需要交叉使用不同的理论，甚至需要发展新的理论。比如，对于给定偏好的模型就不得不需要考虑制度经济学和信息经济学等科学的发展。每个学科都有它自己的弱点，但这些学科交叉起来就可能出现科学的结果。越来越多的经济学家（例如 Granger, 2001; Reiss and Wolak, 2007; Acemoglu, 2010; Keane, 2010）认为必须在宏观经济学、发展经济学、产业组织、劳动经济学等各个领域的经验研究中引入更多的经济理论（结构），才能从本质意义上解释因果效应背后的逻辑链条与作用机制。对于某些问题而言，恐怕只有把包括物质、能源、信息在内的相关非经济变量和经济变量进行纵横交错的交叉科学检验，发现交互作用和关系，在此基础上设立

的计量方程才能对那个具体问题具有对称性。当考虑不同时间、不同空间的类似问题时，这种交叉检验就需要重做一遍。

即使进行了上述工作，也不能把识别因果看作是一个一劳永逸的过程，而应该把任何计量结果都看作是“具有初步证据的因果性”（*prima facie causality*），这可以帮助人们在科学上获得一种科学相对性和开放性的心态，不把自己的结论作为绝对的、唯一的、权威的、定论的，而是时刻准备根据新的情况接受研究结论的必要的修正。这样，当在实际情况中发现了与计量理论不一致的情况时，应该去修正理论，而不是用理论去修正实际。

参考文献

- [1] James J. Heckman. *Econometric Causality*[J]. *International Statistical Review*, vol. 76, no. 1, 2008. pp. 1-27.
- [2] Sargan, J.D. The estimation of economic relationships using instrumental variables[J]. *Econometrica*. 1958, 26, 393-415.
- [3] Granger, C.W.J. Some properties of time series data and their use in econometric model specification[J]. *Journal of Econometrics*, 1981,16, 121-130.
- [4] Angrist, J. Lifetime earnings and the Vietnam era draft lottery: Evidence from Social Security administrative records[J]. *American Economic Review* , 1990,80, 313-336.
- [5] An Interview with Chris Sims[R], 2011 Nobel Laureate, April 2, 2012, http://www.frbatlanta.org/news/conferences/12zero_degrees_sims_transcript.cfm
- [6] David F. Hendry. *Econometrics-Alchemy or Science?*[J]. *Economica*, vol. 47, No. 188, 1980, 387-406.
- [7] Frisch, R. *Statistical Confluence Analysis by Means of Complete Regression Systems*[M]. Oslo: University Institute of Economics, 1934.
- [8] Lucas, Robert E. Jr. *Econometric Policy Evaluation: A Critique*[R]. *Carnegie-Rochester Conference Series on Public Policy* 1, 1976, pp. 19-46.
- [9] Sims, Christopher A. *Macroeconomics and Reality*[J]. *Econometrica*. 1980 ,Vol. 48, No. 1, pp. 1-48.
- [10] Michael Keane. *structural vs. atheoretic approaches to econometrics*[J]. *Journal of Econometrics*, 2010, 156, 3-20.
- [11] 洪永淼. 计量经济学的地位、作用和局限[J]. *经济研究*, 2007, (5)
- [12] 李子奈. 计量经济学应用研究的总体回归模型设定[J]. *经济研究*, 2008, (8).
- [13] 李子奈、齐良书. 关于计量经济学模型方法的思考[J]. *中国社会科学*, 2010, (2).
- [14] 王美今、林建浩. 计量经济学应用研究的可信性革命[J]. *经济研究*, 2012, (2).
- [15] Imbens, Guido W., and Jeffrey M. Wooldridge. *Recent Developments in the Econometrics of Program Evaluation*[J]. *Journal of Economic Literature*, 2009, 47(1): 5-86;
- [16] Aviv Nevo and Michael D. Whinston. *Taking the Gogma out of Econometrics: Structural Modeling and Credible Inference*[J]. *Journal of Economic Perspectives*, vol. 24, No. 2, 2010, pp. 69-82.
- [17] 金守庚. 内因和外因[J]. *哲学研究*, 1981, (5).
- [18] Hebert Simon. *On the definition of the causal relation*[J]. *Journal of Philosophy* 49 (16): 517-528 (1952);

- [19] Wiener, N. The theory of prediction. In: Beckenbach[M].E. (Ed.), Modern Mathematics for Engineers. McGraw-Hill, New York. 1956.
- [20] Herman O. A. Wold, Causality and Econometrics[J]. *Econometrica*, Vol. 22, No. 2 (Apr., 1954), pp. 162-177.
- [21] James J. Heckman, the scientific model of causality[J]. *Sociological Methodology*, Vol. 35, No. 1, 2005, pp. 1-97.
- [22] Damien Fennell. Why and When Should We Trust our Methods of Causal Inference? Lessons from James Heckman on RCTs and Structural Models[R]. <http://www.lse.ac.uk/CPNSS/projects/CoreResearchProjects/ContingencyDissentInScience/DP/DP0607FennellHeckman.pdf>.
- [23] Fernando Martel García. A Unified Approach to Generalized Causal Inference[R]. <http://polmeth.wustl.edu/media/Paper/FMARTELM002v01D.pdf>.
- [24] Judea Pearl and Elias Bareinboim. Transportability across studies: A formal approach[R]. http://ftp.cs.ucla.edu/pub/stat_ser/r372.pdf, 2013.
- [25] Tinbergen, J. An Econometric Approach to Business Cycle Problems[M]. Paris: Herman & Cie Editeurs, 1937.
- [26] Stefan Hoderlein, Joachim Winter, Structural measurement errors in nonseparable models[J]. *Journal of Econometrics*, 157, 2010, 32-440.
- [27] Woodridge, J. M. Introductory Econometrics: A Modern Approach[M]. Cincinnati, OH: South-Western College Publishing, 2003.
- [28] 约翰·奎金. 僵尸经济学[M]. 格致出版社, 上海人民出版社, 2012:27.
- [29] Mishan, E. J. The New Controversy about the Rationale of Economic Evaluation[J]. *Journal of Economic Issues*, Vol. 16, No. 1, Mar., 1982, pp. 29-47..
- [30] Haavelmo, T. The probability approach in economics[J]. *Econometrica* 12 (Suppl.), iii-vi and 1-115, 1944.
- [31] 陈希孺、柴根象. 非参数统计教程[M]. 上海: 华东师范大学出版社, 1993:1.
- [32] Jean-Marie Dufou, Abderrahim Taamouti, Short and long run causality measures: Theory and inference. 关于这种复杂性, 参阅卡尔·E·瓦什. 货币理论与政策[M]. 中国人民大学出版社, 2001:1-23.
- [33] Christian Gourieroux, Peter C.B. Phillips, Jun Yu. Indirect inference for dynamic[J]. *Journal of Econometrics* 157, 2010, pp. 68-77.
- [34] Bertalanffy, L. von. General system theory[M]. 1988, New York: George Braziller.
- [35] Milakovich, M. E., G. J. Gordon. Public Administration in America[M], New York: Bedford/St. Martin's, 2001.
- [36] 克莱夫·W·J·格兰杰. 经济学中的经验建模: 设定与评价[M]. 中国人民大学出版社, 2005:44.
- [37] Andrew Chesher, Excess heterogeneity, endogeneity and index restrictions[J]. *Journal of Econometrics* 152 2009, pp37-45.
- [38] João Rodrigues, Endogenous preferences and embeddedness: A Reappraisal of Karl Polanyi[J]. *Journal of Economic Issues*, vol. 38, No. 1, 2004, pp. 189-200.
- [39] Hendry, D. F., 1995, Dynamic Econometrics[M], Oxford: Oxford University Press.
- [40] T. Haavelmo. Econometrics and the Welfare State (1989 Nobel Lecture)[J]. *American Economic Review*, 87, 1997, 13-17.
- [41] Axelos, K. The World: Being Becoming Totality, from Systematique ouverte (Trans. Gerald Moore, Les Editions de

Minuit: Paris[J]. *Environment and Planning D: Society and Space*, 2006, vol. 24, pp. 643-651.