

doi: 10.12012/CJoE2020-0001

理解现代计量经济学

洪永淼

(1. 中国科学院数学与系统科学研究院, 北京 100190; 2. 中国科学院大学经济与管理学院, 北京 100190)

摘要 本文基于现代计量经济学的发展历程, 介绍了现代计量经济学的思想、理论、主要内容体系、模型、方法与工具. 文中首先回顾经典计量经济学中经典线性回归模型的基本假设, 并考察通过扬弃这些假设, 发展而来的现代计量经济学的历史背景, 进而阐述现代计量经济学的理论体系与主要内容; 同时讨论了在大数据时代, 大数据为计量经济学带来的挑战与机遇, 以及计量经济学今后发展的若干重要方向与趋势.

关键词 非实验性; 线性回归模型; 非线性模型; 模型设定; 正态分布; 条件异方差; 内生性; 工具变量; 广义矩估计; 平稳性; 结构变化; 模型不确定性; 大数据; 高维数据; 机器学习; 预测; 因果关系; 政策评估

Understanding Modern Econometrics

HONG Yongmiao

(1. Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China;
2. School of Economics and Management, University of Chinese Academy of Sciences, Beijing 100190, China)

Abstract This paper aims to introduce the philosophy, theories, fundamental content systems, models, methods and tools of modern econometrics based on its development history. We first review the classical assumptions of the linear regression model and discuss the historical development of modern econometrics by various relaxations of the classical assumptions to further illustrate the modern theoretical system and fundamental contents. We also discuss the challenges and opportunities for econometrics in the Big Data era and point out some important directions for the future development of econometrics.

Keywords non-experimental; linear regression model; non-linear model; model speci-

收稿日期: 2020-03-22

基金项目: 国家自然科学基金“计量建模与经济政策研究”基础科学中心项目(71988101)

Supported by “Econometric Modelling and Economic Policy Studies” Fundamental Scientific Center Project of National Natural Science Foundation of China (NSFC) (71988101)

作者简介: 洪永淼, 世界计量经济学会会士、发展中国家科学院院士、中国科学院数学与系统科学研究院特聘研究员、中国科学院大学经济与管理学院特聘教授, 研究方向: 计量经济学、时间序列分析、金融计量学、统计学, E-mail: ymhong@amss.ac.cn.

fication; normal distribution; conditional heteroskedasticity; endogeneity; instrumental variable; generalized method of moments; stationarity; structure changes; model uncertainty; Big Data; high-dimensional data; machine learning; forecast; causal inference; program evaluation

1 引言

经济学特别是现代经济学的研究方法中, 主要有逻辑分析方法、历史分析方法与计量分析方法, 以及这些研究方法的交叉。例如, 计量经济史就是基于经济历史数据应用计量经济学方法研究经济史的重要问题。计量经济学是以经济观测数据为基础的实证研究最重要的研究方法论, 在推动经济学科学化过程中发挥了极其重要的作用。经济学研究的主要目的是透过大量复杂的经济现象, 识别经济变量之间的因果关系, 揭示经济系统的运行规律。由于经济观测数据具有非实验性的特点, 人们通常无法像自然科学那样通过可控实验识别经济变量之间的因果关系。因此, 计量经济学在经济分析中具有十分重要的方法论作用。正如使用数学是一门科学成熟的标志那样, 以数据为基础的计量经济学实证分析是经济学研究精确化的重要标志, 尤其在大数据时代更是如此。关于计量经济学在经济研究中的地位、作用及局限性, 可参见洪永森 (2007)、李子奈和齐良书 (2010)。

计量经济学作为一门学科已有近百年的发展历史。在中国, 计量经济学的发展也有近 40 年, 极大推动了中国经济学教育与研究的学术化、规范化、国际化, 成为经济学研究理论联系实际的主要方法与工具。很长一段时间, 中国经济学研究以定性分析为主, 缺少对现实经济的定量分析和实证研究, 计量经济学的引进与广泛应用, 使中国经济学研究水平得到很大的提升, 并且在国际经济学术界初显其学术影响力。

但是, 不少人对计量经济学还存在一些认识误区, 例如照搬照抄国外实证研究所用的计量经济学模型、方法与工具, 不注意其适用的前提与条件, 不注意原始经济理论或经济假说与基于计量经济学模型的统计假说之间的差别, 不注意数据证据与模型证据之间的异同, 缺乏对计量经济学统计推断结果的经济学解释。此外, 一些人认为中国特色社会主义政治经济学研究, 不宜采用量化分析或数量分析方法, 等等。

本文的主要目的, 是通过 40 年来现代计量经济学发展的历史背景与历史逻辑, 阐述现代计量经济学的理论体系、思想、主要内容, 包括各种计量经济学模型与方法适用的前提与条件, 计量经济学模型与经济理论之间的关系, 并探讨在大数据时代计量经济学发展的若干方向与趋势。

2 计量经济学的发展历程

2.1 经典计量经济学

作为一门方法论学科, 计量经济学已有近百年的发展历史, 而现代计量经济学是在传统计量经济学的基础上, 在近 40 年发展起来的。因此, 要了解现代计量经济学的理论体系与主要内容, 经典计量经济学可以作为一个出发点。经典计量经济学的一个核心版块是经典线性回归模型, 其基本假设是:

1) 线性回归模型, 即 $Y_t = X_t' \beta^0 + \varepsilon_t, t = 1, 2, \dots, n$, 其中 Y_t 是因变量, X_t 是由经济解释变量及其非线性变换所构成的 K -维自变量 (regressors) 向量, β^0 是 K -维未知参数向量, ε_t 是不可观测的随机扰动项, 代表除了自变量 X_t 之外所有其他因素对 Y_t 的总效应, 而 n 是样本容量.

2) 严外生性, 即 $E(\varepsilon_t|X) = 0$, 其中 $X = (X_1, X_2, \dots, X_n)'$ 是 $n \times K$ 矩阵. 这个条件意味着随机扰动项 ε_t 对 Y_t 的平均效应不受 X 的影响. 一个充分条件是随机扰动项序列 $\{\varepsilon_t\}$ 和 X 互相独立.

3) 条件同方差与零自相关, 即 $E(\varepsilon\varepsilon'|X) = \sigma^2 I$, 其中 $\varepsilon = (\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n)'$ 是 n -维向量, 而 I 是 $n \times n$ 单位矩阵. 这个条件意味着 ε_t 的条件方差或波动不受 X 的影响. 此外, $\{\varepsilon_t\}$ 序列不存在自相关. 类似地, 一个充分条件是 $\{\varepsilon_t\}$ 与 X 是互相独立的.

4) 随机扰动项服从独立同正态分布, 即 $\varepsilon|X \sim N(0, \sigma^2 I)$. 条件正态分布假设隐含着 ε 和 X 是互相独立的. 正态分布假设为有限样本 (即样本容量 n 为有限整数) 条件下的统计推断提供了很大便利, 因为很多重要的参数估计量, 如最小二乘法 (OLS) 估计量, 是随机扰动项 $\{\varepsilon_t\}$ 的加权平均, 其抽样分布亦为正态分布, 这样, 有限样本条件下的经典统计推断理论便可适用.

5) 自变量样本矩阵 $X'X$ 为非奇异矩阵, 这个条件是对样本数据的一个约束, 即任何一个自变量不能是其他自变量的线性组合, 这排除了共线性的存在.

在 $X'X$ 为非奇异矩阵条件下, OLS 估计量存在. 在严外生性条件下 OLS 估计量是未知参数 β^0 的无偏差估计. 在条件同方差与零自相关条件下, OLS 估计量为最佳线性无偏估计量 (BLUE). 而当随机扰动项服从独立同正态分布时, 对任意有限样本容量 $n > K$, OLS 估计量的抽样分布为正态分布, 这个有限样本抽样分布可用于构建未知参数置信区间估计量, 也可用于构建参数假说的检验统计量, 包括著名的学生 t -检验和 F -检验统计量, 等等. 参见洪永淼 (2011, 第 3 章).

经典计量经济学是理解现代计量经济学的一把钥匙. 经典计量经济学是现代计量经济学的基础与出发点, 现代计量经济学正是通过扬弃线性回归模型的经典假设而建立起来的计量经济学理论与方法, 其假设更贴近经济现实, 更一般化, 涵盖了更多的计量经济学模型, 发展了更多的计量经济学理论与方法, 因而大大扩展了其应用范围与空间, 整个理论体系也更为严谨, 更加科学化. 计量经济学通过以经济观测数据为基础的实证推断研究, 推动了现代经济学的创新与发展. 在以下各节, 我们通过逐个放松经典回归模型的各个假设, 来介绍现代计量经济学发展的历史背景与主要内容.

2.2 从正态分布到非正态分布假设

经典线性回归模型的随机扰动项正态分布假设, 主要是为了得到有限样本条件下 OLS 估计量以及相关统计量的抽样分布, 从而方便统计推断. 正是在正态分布假设下, 经典 t -检验与 F -检验统计量在有限样本条件下才分别服从学生 t -分布与 F -分布.

但是, 大多数经济金融数据的一个典型特征经验事实是非正态分布, 具有厚尾特点, 其主要标志是大多数经济金融数据的峰度均大于 3. 这样, 建立在随机样本正态分布假设基础上的有限样本经典抽样理论不再适用. 计量经济学家因此提出了不少检验线性回归模型残差是否服从正态分布的检验方法, 以便判断经典线性回归模型理论是否适用, 其中一个著名的残差正态性检验是 Jarque and Bera (1980) 检验. 现代计量经济学的一个主要发展就是放弃随机扰动项的正态分布

假设, 运用渐近理论与方法 (特别是大数定律与中心极限定理), 推导出大样本 (即 $n \rightarrow \infty$) 条件下 OLS 估计量及其他估计量仍然是未知参数的一致估计, 并且推导出这些估计量的渐近正态分布. 从大样本分析中可以看到, 即使随机扰动项不服从正态分布, 在独立同分布的随机样本条件下, 只要随机扰动项存在条件同方差, 则当样本容量很大时, 经典 t -检验与 F -检验还是适用的, 而 OLS 估计量也是 BLUE. 换言之, 在存在条件同方差的条件下, 经典 OLS 理论在样本容量足够大时, 照样是适用的. 这个结论, 当随机扰动项是鞅差分过程且满足条件同方差时, 对平稳时间序列的随机样本也是适用的. 见文献洪永森 (2011, 第 4 章和第 5 章)

在计量经济学大样本分析方面, Halbert White 发挥了重要作用, 他 1984 年出版、2001 年再版的 *Asymptotic Theory for Econometricians* 一书也成了计量经济学大样本分析的经典参考书.

大样本分布在实证应用中简单方便, 但是在有限样本特别是小样本条件下, 参数估计量和检验统计量的渐近分布与真实的未知的有限样本分布可能相差甚远, 这在统计推断时将产生很大的 Type I 和 Type II 误差, 导致推断结论不可靠. 为了改进渐近分布在有限样本条件下的近似程度, 计量经济学家和统计学家 (如 Klein and Spady (1993), Phillips (1977a, 1997b, 1977c), Ullah (1990)) 曾经致力于发展有限样本特别是小样本条件下的渐近理论, 其中包括所谓的 Edgeworth 展开和 Saddle Point 近似. 但是, 这些方法相当复杂, 在实际中没有得到广泛的应用. 近年来, 随着计算机技术的快速进步, 计量经济学的一个重要发展是 Bootstrap 方法的发展及其广泛应用. Bootstrap 方法的理论基础也是 Edgeworth 展开, 但它巧妙地利用计算机对观测数据进行多次重复抽样而产生的随机样本, 大大改进了对参数估计量和检验统计量的真实有限样本分布的近似程度, 在实证研究中可提供更为可靠的统计推断结论, 因此得到了广泛的应用. 更多讨论可参考 Hall (1992) 和 Horowitz (2001).

2.3 从条件同方差与零自相关到条件异方差与自相关

经典线性回归模型的另一个重要假设是随机扰动项服从条件同方差 (其条件方差不随自变量取值的变化而变化) 与零自相关, 即随机扰动项 ε_t 的条件二阶矩不随 X 取值的变化而变化. 在此假设条件下, OLS 估计量是 BLUE. 当条件同方差或零自相关不成立时, 不仅 OLS 估计量不再是 BLUE, 经典 t -检验与 F -检验统计量也不再分别服从学生 t -分布与 F -分布, 甚至在大样本条件下经典 t -检验与 F -检验也不再适用. 经典计量经济学很早就认识到条件同方差与零自相关这两个假设的局限性, 因此提出了广义最小二乘法 (GLS) 理论. GLS 理论假设随机扰动项存在条件异方差与自相关, 但条件异方差与自相关的形式是已知的 (只存在一个未知常数), 因此可以通过分解已知形式的条件方差来消除条件异方差与自相关, 将原始的线性回归模型转变为一个满足条件同方差与零自相关的线性回归模型, 从而对变换后的线性回归模型进行 OLS 估计, 这时经典线性回归理论可以适用. 例如, 在静态时间序列线性回归模型中, 如果随机扰动项服从一个固定阶数的自回归过程, 则可通过 Cochrane-Orcutt 方法消除随机扰动项的自相关, 得到至少是渐近最优线性无偏估计的 GLS 估计量.

但是, “条件异方差与自相关的形式已知” 这个假设显然不适合大多数经济观测数据. 在实际应用中, 条件异方差与自相关的形式是未知的. 在零自相关条件下 (通常是横截面数据), 可通过非参数方法一致地估计 OLS 残差的条件异方差, 再代入 GLS 估计量公式中, 从而得到适应

性可行 GLS 估计量, 这种方法在大样本条件下具有 BLUE 性质 (参见 Robinson (1988), White and Stinchcombe (1991)).

但是, 在实际应用中, 大部分研究还是偏好采用比较简单的 OLS 估计及相应的统计推断方法. White (1980) 推导了条件异方差条件下 OLS 估计量的渐近方差公式并提出了其估计方法, 这被称为 White's heteroskedasticity-consistent variance-covariance matrix estimator. 基于这个正确的方差估计量, 可修正经典 t -检验量, 使之在存在条件异方差但样本容量足够大时, 依然可以适用. 所得到的检验通常称为稳健性 (robust) t -检验. 另一方面, 经典 F -检验量因为无法修正, 因此在存在条件异方差时, 不再适用, 即使是大样本时也是如此. 但是, 可构造使用正确方差估计量的稳健性 Wald 检验与拉格朗日乘子检验.

对时间序列线性回归模型, 当随机扰动项不但存在未知形式的条件异方差而且还存在未知形式的自相关时, OLS 估计量的渐近方差估计量不仅要考虑条件异方差的影响, 也要考虑未知形式的自相关的影响. 这需要估计所谓的长期方差 - 协方差矩阵. Newey and West (1987, 1994) 和 Andrews (1991) 等提出了用非参数核方法估计长期方差 - 协方差矩阵, 这些方法广泛应用于实证研究中. 但是, 经验研究与计算机模拟实验发现, 当存在较强的自相关时, 基于核估计的长期方差 - 协方差估计量, 经常导致相关的统计检验量在有限样本条件下会过度拒绝正确的参数原假设, 即存在很大的 Type I 误差, 这个问题至今没有得到彻底的解决, 虽然已有各种改良方法.

2.4 从线性模型到非线性模型

在计量经济学中, 线性回归模型 ($Y_t = X_t'\beta^0 + \varepsilon_t$) 是指因变量 Y_t 与自变量 X_t 及未知参数 β^0 之间的线性关系, 其中自变量 X_t 是由某个或某些经济解释变量及其非线性变换 (如平方项或对数) 所构成, 因此不是指因变量和原始解释变量之间的线性关系. 例如, 当因变量的条件均值 (或回归函数) 是某个经济解释变量的一个多项式时, 这仍被视为线性回归模型. 但是, 在不少计量经济学模型中, 当因变量的条件均值不是未知参数的线性函数时, 因变量与自变量之间的关系一般是非线性的. 而且, 当模型不是刻画因变量的条件均值, 而是刻画因变量的条件方差, 条件分位数, 条件矩甚至整个条件分布时, 更是一种非线性关系, 不管是对参数或者原始经济解释变量而言, 都是如此.

在时间序列计量经济学中, 线性时间序列模型通常是指因变量与解释或预测变量及随机扰动项之间的线性关系, 例如所谓的 ARMA 模型. 目前比较流行的非线性时间序列模型包括自回归门槛模型 (TAR), 马可夫链机制转移模型 (MCRS), 平稳转换自回归模型 (STAR) 等, 这些模型均是对时间序列因变量的条件均值进行非线性建模. 这些非线性均值模型可解释为由不同状态下的线性模型混合组成的, 在某一个状态下, 因变量是一个线性时间序列过程, 在另一个状态下, 因变量是另一个线性过程, 而非线性特征主要由这些不同状态如何互相转换的机制假设而定. 非线性均值模型可用于刻画经济金融数据中的非线性关系, 例如经济周期的非对称性和金融市场之间联动的非对称性.

在 20 世纪 70 年代, 石油危机, 浮动汇率制度以及美联储高利率政策, 导致世界经济充满不确定性. 如何测度经济不确定性及其对经济金融市场的影响便成为当时一个迫切的重要问题. 由于波动量化分析的重要性, 计量经济学家便提出各种类型的条件方差模型, 包括 Engle (1982) 的

ARCH 模型, Bollerslev (1986) 的 GARCH 模型, Nelson (1991) 的 EGARCH, 以及 Glosten et al. (1993) 的门槛 GARCH 模型, 等等. 这些模型并不是对整个条件概率分布建模, 它们只是对某个时间序列因变量的前二阶条件矩建模, 因此不能用最大似然法 (MLE) 估计未知模型参数. 在实证研究中, 为了估计波动模型的未知参数, 一般需要假设额外的辅助条件, 从而推导出时间序列因变量的整个条件概率分布, 因此可用 MLE 方法估计模型参数. 由于辅助假设可能不正确 (研究者事先也知道这一点), 波动模型的似然函数因此可能误设, 这种方法被称为 Quasi-MLE, 或 QMLE. 只要前二阶条件矩模型设定正确, QMLE 仍可一致地估计模型参数, 但代价是其估计量的渐近方差将大于基于正确条件概率分布的 MLE 的渐近方差, 因此 QMLE 估计量比较不精确. QMLE 渐近方差的结构, 与 MLE 的渐近方差结构有显著的不同, 它类似于线性回归模型中 OLS 估计量在存在条件异方差与自相关时的渐近方差结构, 因此必须使用似然函数误设时仍然适用的稳健方差公式及其一致估计量. 而 MLE 的渐近方差, 其结构则类似于线性回归模型 OLS 估计量在存在条件同方差与零自相关时的渐近方差. 与线性回归模型的 F -检验类似, 当似然函数误设时, 著名的似然比检验量不再适用, 因为它相当于使用了 MLE 的渐近方差公式. 人们可以构建基于 QMLE 渐近方差公式的稳健统计检验量, 如稳健 Wald 检验和稳健拉格朗日因子检验. 参见洪永森 (2011, 第 9 章).

另一类计量经济学模型 (包括很多线性与非线性模型) 由一个或一组总体矩条件来刻画. 总体矩条件一般是从经济理论推出, 例如, 宏观经济学理性预期理论意味着随机资产定价误差相对于经济主体的历史信息条件均值为零, 即每个时期均不存在系统定价偏差. 根据这个性质, 可选择一些合适的工具变量, 构造一组总体矩, 当其在真实参数值处取值时, 总体矩为零. 由于不知道经济变量的整体条件分布, 无法使用 MLE. 广义矩方法 (GMM) 估计的基本思想, 就是构建一组样本矩, 其中样本矩的维度不小于未知参数的维度, 然后选择让样本矩尽量接近总体矩的参数值, 作为未知参数值的估计量. 在数学上, GMM 估计量使样本矩的一个加权二次项最小化, 其中的权重一般会影响 GMM 估计量的精确度. 一个渐近最优权重是样本矩的方差 - 协方差估计量, 其作用与 GLS 方法类似, 可消除样本矩的异方差以及样本矩之间的相关性, 从而获得渐近最优 GMM 估计. 这个方法由 Hansen (1982) 提出. GMM 可视为是对统计学经典矩估计方法的拓展, 但它是为了估计与检验经济学模型, 特别是理性预期模型而提出的. GMM 应用非常广泛, 大多数计量经济学估计量均可视为其特例, 包括 OLS 和二阶段最小二乘法 (2SLS). QMLE 和 GMM 是估计非线性计量经济学模型的两个最常用方法.

2.5 从外生性到内生性

经典线性回归模型假设随机扰动项对于自变量在各个时期不管取什么值时其条件均值为零, 即自变量在当期、过去或将来的任何取值均不影响随机扰动项的平均效应, 这个条件称为严外生性. 在计量经济学中, 有多个关于外生性的定义. 如果随机扰动项序列 $\{\varepsilon_t\}$ 和自变量序列 $\{X_t\}$ 互相独立, 或自变量为非随机变量, 则称为存在强外生性. 如果随机扰动项 ε_t 相对于当期自变量 X_t , 不管取什么值, 其条件均值为零, 则称为弱外生性. 弱外生性意味着线性回归模型是条件均值 $E(Y_t|X_t)$ 的正确设定, 即条件均值 $E(Y_t|X_t) = X_t'\beta^0$ 是自变量 X_t 的线性函数.

经典线性回归模型的严外生性假设介于强外生性与弱外生性两个条件之间, 强外生性意味着

严外生性成立, 而严外生性意味着弱外生性成立, 但反之不然. 经典线性回归模型之所以假设严外生性条件, 是为了方便推导有限样本条件下 OLS 估计量及相关检验统计量的抽样分布. 对大样本渐近理论, 弱外生性条件就足够了.

当弱外生性不成立时, 随机扰动项相对于当期自变量的条件均值不为零, 这种情形通常被称为存在内生性 (endogeneity), 此时自变量 X_t 称为内生变量. 产生内生性的原因很多, 包括自变量存在测量误差, 存在遗漏变量, 存在联立方程偏差, 等等. 所谓联立方程偏差是指除了所考虑的线性回归方程外, 还有一个或几个遗漏方程, 这个或这些遗漏方程描述自变量如何由因变量及其他变量决定. 在这种情况下, 自变量与因变量一般具有双向因果关系, 即自变量影响因变量, 同时因变量也影响自变量. 计量经济学家常将这种双向因果关系导致的自变量和随机扰动项之间存在的相关性, 称为内生性, 因为这意味着自变量 X_t 也是一个内生变量, 由联合方程组共同决定. 严格地说, 内生性导致正交条件 $E(\varepsilon_t|X_t) = 0$ 不成立, 从而 OLS 估计量不是未知参数 β^0 的一致估计量. 应该指出, 其他原因, 如自变量存在测量误差, 回归模型存在遗漏变量, 或函数形式误设等, 也会导致 $E(\varepsilon_t|X_t) = 0$ 不成立, 但严格上说这些原因与内生性无关. 简单起见, 一般统称 $E(\varepsilon_t|X_t) \neq 0$ 为存在内生性.

当存在内生性时, 如果只对自变量决定因变量的线性回归模型进行估计, OLS 估计量将不是真实参数值的一致估计. 这时, 可采用 2SLS 估计法, 借助一组与随机扰动项不相关但与自变量密切相关的工具变量, 先将自变量“影射”到工具变量, 然后将因变量对“影射变量”进行回归, 这样可获得单向因果关系 (从自变量到因变量) 的参数值的一致估计. 2SLS 估计已有近百年的历史, 最早是 20 世纪 20 年代在病理学实证研究中发展起来的, 其关键与难点在于如何寻找有效的工具变量 (参见 Stock and Trebbi (2003)).

在实证研究中, 经常发现所使用的工具变量与自变量相关性很低, 导致 2SLS 估计量不稳定甚至不能一致估计未知真实参数值. 这种情况称为“弱工具变量” (参见 Staiger and Stock (1997)). 弱工具变量的研究成为 20 年来计量经济学的一个重要研究方向.

内生性问题不仅在线性回归模型存在, 在其他类型的模型 (如条件方差模型、条件分位数模型、条件分布模型等) 中也都可能存在. 另外, 内生性非参数与半参数回归模型也得到很大关注 (参见 Blundell and Powell (2004)). 内生性问题在计量经济学中占有核心地位, 其主要原因, 乃是经济学家的最主要任务是识别、估计经济变量之间的因果关系, 从而揭示经济运行规律. 有人说, 在大数据时代, 只需要相关性, 不需要因果关系. 这一点不适合于经济学.

由于经济现象与经济观测数据的非实验性特点, 人们不能通过控制其他变量取值不变, 利用实验手段研究某个或某些经济变量的变化是否引起因变量的变化. 因此, 如何识别因果关系是经济实证研究的一个重要难点, 也是现代计量经济学研究的一个热点问题. 微观计量经济学 20 年来的一个研究热点是所谓的“处理效应 (treatment effect)”, 这个领域的研究, 借鉴了统计学特别是生物/医学统计学随机试验 (randomized experiments) 的思想与方法, 发展出了一套识别、估计经济因果关系的计量经济学理论与方法, 诞生了政策评估计量经济学 (econometrics of program evaluation) 这个新兴学科, 可用于在非实验性条件下通过估计虚拟事实来量化评估各种经济社会公共政策. 定量评估一个经济政策的效应, 其基本思想是在同等条件下, 比较实施该政策的结果与假设没有实施这项政策的虚拟结果, 两者之差就是该项政策的效应. 这里的关键与难点是

在政策已实施的条件下, 如何准确估计假设政策没有实施时的虚拟结果. 不少方法, 包括双差分 (difference-in-difference, DID), 回归断点设计 (regression discontinuity design, RDD), 倾向评分匹配 (propensity score matching, PSM), 面板数据评估方法 (panel data approach to program evaluation) 等, 已广泛应用于各种政策和项目评估 (参见 Imbens and Wooldridge (2009) 的综述文章). 例如, Hsiao et al. (2011) 提出一个基于面板数据的政策评估方法, 并应用于评估香港与内地 2002 年签定的“关于建立更紧密经贸关系的安排 (CEPA)”对香港经济增长的影响. 另一方面, 近 30 年兴起的实验经济学, 通过控制实验室条件研究经济主体经济、心理行为, 也是发现经济因果关系的一种新方法与新学科. 此外, 还有“田野研究 (field study)”, 这是一种准实验方法, 在现实社会经济环境下, 人为引入一些实验控制手段与方法, 跟踪研究经济变量之间的因果关系.

2.6 从模型正确设定到模型误设

当线性回归模型存在遗漏变量, 或函数形式误设时, 我们称线性回归模型是条件均值的误设. 具体地说, 如果存在一个未知参数值 β^0 , 使得线性回归模型与条件均值相等, 即 $E(Y_t|X_t) = X_t'\beta^0$, 则称线性回归模型是条件均值的正确设定, 而这个未知参数值称为真实参数值. 如果不存在任何一个参数值, 可使线性回归模型等于条件均值, 即对所有参数值 β , 有 $E(Y_t|X_t) \neq X_t'\beta$, 则称线性回归模型是条件均值的误设.

当线性回归模型误设时, 模型参数不能被解释为真实参数值, 更不能被赋予经济涵义. 例如, 线性消费函数若存在误设, 则参数值不能解释为边际消费倾向, 因为边际消费倾向是消费对收入的偏导数, 当消费函数不是线性时, 边际消费倾向不等于一个常数. 总的说来, 不管是对条件均值或条件概率分布的其他方面 (如条件方差、条件分位数、乃至整个条件分布) 建模, 若存在模型误设, 参数便不能解释为真实模型参数值, 更不能被赋予经济解释. 经济解释的有效性取决于模型设定的正确与否. 除了经济解释的有效性之外, 模型误设在实际应用中也可能导致严重后果. 例如, 在金融市场, 使用误设模型会导致所谓的“模型风险 (model risk)”. 不少人将 2008 年席卷全球的美国次贷危机归咎于华尔街当时普遍使用的用于金融衍生产品定价的高斯连结模型 (Gaussian copula model), 该模型不能正确刻画金融市场之间的非对称联动关系.

但是, 模型误设并不意味着这个模型一无是处或不能在实际中使用. 例如, 假设某个经济变量的动态演变由大约 80% 的线性因素和大约 20% 的非线性因素共同决定. 显然, 线性回归模型是错误的, 但它仍有不错的预测能力, 虽然其参数不能解释为自变量的边际效应. 另外, 有些误设模型的部分参数仍可赋予经济涵义. 例如, 在研究教育回报率的线性回归模型中, 因变量是工资或收入, 如果自变量只包括教育程度和工作经验, 这个模型可能存在遗漏变量, 因为观测不到的个人能力也会影响工资或收入, 而个人能力与教育和工作经验相关. 但是, 如果研究者的兴趣只是教育回报率, 可采用工具变量方法将教育程度这个解释变量的参数值一致估计出来, 这样, 遗漏个人能力这一重要变量的线性回归模型照样值得研究.

模型误设还可能出现在以下情形. 一个线性回归模型, 如果观测数据存在若干缺陷, 如存在删失数据 (censored data) 或截断数据 (truncated data), 那么 OLS 估计量便不是真实参数值的一致估计. 在这个情况下, 可增加辅助假设, 例如假设随机扰动项服从正态分布, 然后用 MLE 方法估计未知模型参数. 由于数据存在缺陷, 因变量与自变量之间的关系事实上已变为非线性关

系, 但是 MLE 可以将原始线性回归模型的参数一致估计出来. 在这里, 辅助假设是否正确, 会影响线性回归模型参数估计的一致性. 另一方面, 波动模型是对因变量的前二阶条件矩建模. 在波动模型正确设定的情况下, 我们并不知道因变量的条件概率分布, 此时通常也需要增加辅助假设 (如假设标准化随机扰动项服从标准正态分布), 以便推导出因变量的条件概率分布, 从而使用 MLE 方法估计未知模型参数. 由于辅助假设可能是错误的, 条件分布因此可能误设, 这种方法称为 QMLE. 但是, 如果条件均值和条件方差仍然正确设定, 则即使条件分布 (即条件高阶矩) 误设, QMLE 仍可一致估计条件均值和条件方差模型中的未知参数值, 这与上述线性回归模型因数据缺陷而需要辅助假设的情形不同. 但是, 模型误设将影响 QMLE 估计量的方差或精确度, 因此需要使用在概率分布模型误设时也适用的稳健性渐近方差公式. 参见 Bollerslev and Wooldridge (1992), Lee and Hansen (1994), Lumsdaine (1996).

由于回归模型误设 (即 $E(\varepsilon_t|X_t) \neq 0$) 影响到参数的经济解释的有效性以及参数估计的一致性, Hausman (1978) 提出了一个检验回归模型误设的方法, 被称为 Hausman 检验. White (1981), Newey (1985), Tauchen (1985) 和 White (1990) 将此方法拓展到以矩条件为基础的更一般的模型设定检验, 称为矩检验或 m - 检验. 这些检验均不是一致检验 (consistent tests), 即无法检测出一些误设模型, 甚至在大样本时也是如此. Bierens (1982, 1990), Hong and White (1995), Fan and Li (1996), Hong and Lee (2013) 等采用非参数等方法提出了线性或非线性回归模型设定的检验方法. 当样本容量足够大时, 这些方法可以检测任何形式的模型误设, 因此是一致检验.

在统计学与计量经济学, 模型检验还有另一种形式, 称为模型验证 (model validation). 这种方法是将数据分为训练数据 (training data set) 和测试数据 (test data set), 其中训练数据用于估计模型参数, 测试数据则用于评估已估计的模型对另一个数据的预测能力. 这种模型验证方法主要是验证模型的样本外预测能力, 在时间序列预测和机器学习中广泛使用. 样本外模型验证可以避免样本内模型过度拟合 (overfitting). 此外, 如果产生训练数据的生成过程与产生测试数据的生成过程有显著不同 (可因结构变化或样本主体差异性所致), 那么即便模型对训练数据生成过程的设定是正确的, 它在验证数据的预测很可能不会太好.

2.7 从平稳性到非平稳性

在假设遍历平稳时间序列随机样本的条件下, 经典线性回归的 OLS 理论, 当随机扰动项满足零差分与条件同方差时, 在大样本情形下, 还是适用的. 如果存在条件异方差或自相关时, OLS 估计量即使在大样本条件下也不是 BLUE, 经典的 t - 检验和 F - 检验也不再适用. 但是, 可以修正经典统计检验量, 特别是 t - 检验, 修正后的稳健统计检验量在大样本时是可用的.

但是, 上述 OLS 理论在因变量和自变量时间序列为非平稳过程时, 一般不再适用. Granger and Newbold (1974) 通过计算机模拟研究发现所谓伪回归 (spurious regression) 现象, 即两个互相独立的非平稳单位根时间序列, 如果将其中一个变量对另一个变量进行回归, 则回归系数估计量的 t - 检验量在统计学上是显著的. Phillips (1986) 从理论上给了严谨的解释. Nelson and Plosser (1982) 通过应用 Dicky-Fuller 检验, 发现大部分宏观经济金融时间序列均是非平稳单位根过程. 因此, 计量经济学家自 20 世纪 80 年代起便致力于发展非平稳时间序列计量经济学理论与方法, 包括 Engle and Granger (1987) 的协整理论, Phillips (1987a, 1987b) 的单位根理论, 等

等. 对线性回归模型, 单位根时间序列大样本理论与平稳时间序列计量经济学理论完全不同. 详细介绍可参考 Hamilton (1994).

除单位根过程之外, 非平稳时间序列也可表现为趋势平稳过程. 所谓趋势平稳序列, 是指其均值是时间的函数, 因此不是一个平稳时间序列. 当均值是时间的线性函数时, 这个时间序列便会显示一个长期线性变化趋势. 若将时间趋势消除, 时间序列将变为平稳过程. 如果样本容量不大, 趋势平稳过程与单位根过程产生的数据非常类似, 不容易区分. 在宏观经济实证研究中, 判断宏观经济变量是单位根过程还是趋势平稳过程, 是非常重要的, 因为它们具有不同的政策含义. 例如, 这关系到宏观财政货币政策是否具有长期效应 (当相关经济变量是单位根过程时) 或只有短期效应 (当相关经济变量为趋势平稳过程时).

非平稳时间序列还可表现为结构变化 (structure changes). 结构变化通常分为两种, 一种是结构突变 (structure breaks), 另一种是缓慢结构变化 (smooth structural changes). 存在结构突变的时间序列所产生的数据, 与单位根过程产生的数据有时看上去非常类似. 对于结构突变, 自 Chow (1960) 提出针对单个已知变点 (change point) 的结构突变检验方法以来, 计量经济学家在这个方面取得了长足的发展, 特别是允许多个变点, 以及不知道变点的情形下如何检验是否存在结构突变, 并且建立了比较完善的大样本分布理论. 可参考 Andrews and Ploberger (1994, 1995), Bai and Perron (1998) 等文献.

除了结构突变之外, 另一种结构变化是缓慢结构变化, 即模型参数随时间缓慢变化. 一个例子是所谓的时变性 STAR 模型. 趋势平稳过程也是一个例子, 如果其均值是时间的平滑函数的话. 现实经济由于人口偏好、技术进步、政策变化、体制改革、外部条件变动等原因, 经常导致经济行为与经济结构发生变化. 此外, 由于经济主体可以理性预测或感知政策等变化, 并适应性地改变其行为, 这也会导致经济结构发生变化, 这就是著名的卢卡斯批判 (Lucas (1976)). 但是, 即使上述因素突然发生, 由于习惯与调整成本等原因, 经济主体的行为及经济结构也可能会缓慢地变化. 甚至即使经济个体的行为发生了突变, 如果个体行为突变的时间点不同, 则加总的宏观经济序列也会出现缓慢变化态势. 总之, 很多情形下, 缓慢结构变化可能更接近经济现实. 一个典型案例是美国宏观经济波动的“大缓和 (Great Moderation)”现象. 美国自 20 世纪 80 年代中期以来, 宏观经济波动, 特别是 GDP 增长率和通货膨胀率的波动, 呈现出一个逐渐缩小的趋势, 这称为“大缓和”现象 (参见 Bernanke (2004)). 其实, 中国宏观经济也存在类似的“大缓和”现象, 中国的 GDP 增长率和通货膨胀率, 分别从 1992 年和 1999 年起出现了波动逐渐变小的趋势 (参见 Sun, Hong and Wang (2019)). 但是, 对中国宏观经济的“大缓和”现象, 似乎至今还没有人深入研究其原因.

Chen and Hong (2012, 2016) 提出了检验时间序列线性回归模型和 GARCH 模型是否存在缓慢结构变化的方法, 其中一个实证研究发现, 结构变化可能是资产收益率预测不准的一个重要原因. 缓慢结构变化也可能产生伪长记忆现象 (spurious long memory), 即经济金融时间序列数据的长记忆现象有可能系缓慢结构变化导致, 而非真正具有长记忆性质. Hong, Wang and Wang (2017) 提出检验严平稳性和弱平稳性的方法, 并应用于检验宏观经济金融时间序列的平稳性. 他们的研究表明, 差分后的宏观经济金融时间序列并不满足平稳性假设, 特别是其均值或方差随时间缓慢变化. 差分或去趋势后的时间序列满足平稳性条件, 是非平稳时间序列计量经济学的一个

基本假设.

在时间序列分析中, 20 年来出现了一个新的研究领域, 即局部平稳时间序列分析 Dahlhaus (1996). 所谓局部平稳时间序列, 是一种非平稳过程, 其参数值随时间缓慢变化, 因此, 在一小段时间内, 可以用一个平稳时间序列模型来近似, 但不同时期需要用不同的平稳时间序列模型来近似, 而这些不同模型的参数值是随时间连续缓慢改变的. 高频数据可获得性的提高, 有助于渐近性结构变化的研究与局部平稳时间序列模型的应用.

2.8 从计量经济学模型到经济理论

上述介绍中, 似乎看不到经济理论在计量经济学建模中所起的作用. 我们看到, 现代计量经济学有各种各式的模型, 包括条件均值模型, 条件方差模型, 条件分位数模型, 条件矩模型, 条件概率分布模型, 还有大家熟悉的广义线性回归模型, 如 Probit 模型, Logit 模型, Cox 比例危险率模型, 泊松回归模型等. 在实证研究中, 使用哪一类模型需要由所要研究的经济问题的本质决定, 即“问题导向”. 并不是一个计量经济学模型或方法可用于研究任何经济问题. 例如, 研究金融市场有效性或可预测性时, 需要对资产回报率的条件均值建模; 研究实体经济和金融市场之间的波动溢出效应时, 需要对条件方差建模; 而研究风险价值 (value at risk, VaR) 时, 则需要对条件分位数或条件分布建模.

经济理论的作用还体现在建模时解释变量的选择上. 选择哪些重要经济解释变量, 除了研究经验及应用计量经济学、统计学与人工智能方法之外, 还需要依靠经济理论, 因为经济理论可帮助判断经济变量之间的因果关系. 在模型设定正确时, 经济理论还可为模型参数赋予经济涵义, 增强模型的可解释性. 对计量经济学模型本身以及实证分析所得到的统计结果, 从经济学视角进行解释, 是至关重要的. 在识别因果关系时, 单凭计量经济学、统计学与人工智能方法是不够的, 需要借助经济理论. 那么, 如何检验经济理论是否能够解释所观测的经济现象? 换言之, 如何检验经济理论的有效性或正确性? 一个基本思想是将经济理论或经济假说转变为对计量经济学模型的一种约束, 然后通过数据和计量经济学方法检验这种约束是否正确. 值得强调的是, 在这个分析过程中, 一般需要增加一些辅助假设 (例如设定模型的函数形式). 因此, 当拒绝原假设时, 必须区分究竟是因为经济理论不成立, 还是因为辅助假设不成立. 一般说来, 大多数经济理论或经济假说均为 model-free, 即不依赖于某个具体形式的计量经济学模型, 当我们将经济理论或经济假说转变为对所选择的计量经济学模型的统计约束时, 经济假说与统计假说两者之间会存在差别. 换言之, 两个假说并不等价, 此时在解释统计假说的检验结果时需要非常谨慎. 以检验市场有效性为例, 如果资产收益率历史信息对将来资产收益的条件均值没有任何预测能力, 我们称弱式有效市场假说成立 (Malkiel and Fama (1970)). 这个经济假说是一个 model-free 的假说. 现在假设选择一个 p 阶线性自回归模型检验这个经济假说. 当弱式有效性假说成立时, 自回归模型没有预测能力, 所有自回归系数当为零. 如果发现至少有一个自回归系数不为零, 那么弱式有效性假说不成立. 但是, 如果“所有自回归系数为零”这个统计假说没有被拒绝, 我们能否说弱式有效性假说成立? 不能. 这是因为线性自回归模型只是预测资产收益率的众多模型中的一个而已. 所有自回归系数为零, 只能说明线性自回归模型没有预测能力, 不能说明资产收益率历史数据对未来资产收益率没有预测能力. 因此, 如果所有自回归系数为零, 我们只能说没有发现拒绝弱式有

效性假说的证据。总之, 需要注意经济假说和统计假说的差别, 并且区分数据证据和模型证据, 两者并不等价。

计量经济学在推动经济理论创新与发展方面, 发挥着关键作用。随着数据的不断积累, 以及新的计量经济学方法的不断产生, 我们会比较容易找到拒绝现有经济理论的证据, 即现有经济理论不能解释观测数据, 此时便需要新的经济理论。经济理论的发展就是这样一个过程, 在这个过程中, 计量经济学本身也获得了发展。

由于经济数据样本容量有限和所使用的计量经济学方法的局限性等各种原因, 有时候可能存在两个或两个以上的经济理论或经济假说, 同时通过数据的检验, 换言之, 存在多个经济理论可以解释同一经济现象的可能性。这时, 到底哪个经济理论是正确的? 这种情形在经济学称为模型不确定性 (model uncertainty) 或模型模糊性 (model ambiguity)。在大数据时代, 通常存在众多潜在的解变量或预测变量, 因此有很大可能性出现共线性现象。在实际应用中, 可能存在几个不同的计量经济学模型 (如包含不同解释变量的线性回归模型, 或不同函数形式的模型), 根据一定的统计学准则, 其表现非常相近或几乎一样。此外, 有可能对于某个数据, 一个计量经济学模型表现最好, 但如果对数据施加“微扰”, 如增加或改变几个数据点, 则表现最佳的模型可能会变成另一个。这种现象, 在统计学也称为模型不确定性, 类似于经济理论的不确定性。更多讨论参见 Breiman (2001)。在存在模型不确定性条件下, 如何解释模型不确定性以及研究其对统计推断的影响, 是一个十分重要的课题。这不仅仅关系到模型的解释, 也关系到现代计量经济学的一些基本假设, 特别是模型唯一性假设, 即存在一个唯一的计量经济学模型正确刻画数据生成过程。在预测领域, 常常使用模型组合或模型平均来减少模型不确定性对预测的影响。

2.9 从传统数据到大数据

随着互联网与移动互联网等信息技术的发展与广泛应用, 经济观测数据正以指数增长的速度快速增加, 这是所谓的大数据 (Big Data)。除了传统数字型结构化数据之外, 大数据还包括大量的非结构化和半结构化数据, 如文本、图形、音频、视频等数据。即使是数字型数据, 也包含一些新型数据, 例如区间数据、函数数据、符号数据等。大数据来源包括互联网电商公司的交易数据、企业与政府部门的网站, 各种社交平台与感应器, 等等。这些大数据大多是实时或几乎实时收集记录, 其数量、体量大, 但很多情况下信息密度低。

大数据提供了传统数据所没有的很多信息, 这些信息可以用来构建以前不能精确测度的一些重要变量。例如, 通过社交平台和新闻媒体文本数据, 可构建投资者或消费者的情绪指数 (如 Baker and Wurgler (2006, 2007)) 和经济政策不确定性指标 (如 Baker et al. 2016), 然后研究这些变量对经济运行和金融市场的影

响。大数据也使得构建高频化的重要经济指数或经济变量成为可能。例如, 宏观经济物价指数, CPI 或 PPI, 目前抽样最高频率是月度数据。通过互联网消费品和生产产品价格信息, 完全有可能借助人工智能方法构建每周甚至每天的消费者物价指数和生产者物价指数。这些高频宏观经济变量对研究宏观经济的当期或近期的运行状况, 具有重要的现实意义。可以预料, 在不远的将来, 高频宏观经济学将有可能应运而生, 这样的新兴学科能够更好地研究宏观经济和金融市场之间的互动关系。事实上, 大数据已催生了实时预测 (nowcasting) 计量经济学。“实时预测”这个术语源

于气象学,通常是指 2 小时之内的短期天气预测. 计量经济学的“实时预测”,是指利用实时大数据预测当季 GDP 及其它宏观经济指标. 由于季度 GDP 的公布时间有较长的滞后期,实时预测当季 GDP 等重要宏观经济指标,对监控宏观经济运行具有重要意义. 不少国家的中央银行已开始对各自国家的 GDP 进行实时预测. 有关实时预测的更多讨论参见 Giannone et al. (2008). (Wang et al. (2005)) 提出了一个 TEI@I 预测方法论,将文本挖掘、计量建模与人工智能等方法综合集成,解决了将统计历史数据与异质数据、定性与定量分析相结合的难题,显著地提升了样本外预测精度.

大数据产生了不少有别于传统数据的新型数据. 例如,每笔交易数据、日内资产价格数据、区间数据、函数数据等. 这些新型数据呼唤新的计量经济学模型与方法. 过去 20 年,高频和超高频金融数据已经推动了高频金融计量学的快速发展. 例如,Engle and Russell (1998) 基于超高频每笔交易数据对价格变动之间的时间间隔进行建模,提出一个 autoregressive conditional duration (ACD) 的新模型,这个模型可利用超高频金融交易数据预测下一个价格变动的的时间,以及未来每个时间点价格变动的概率,这些预测可用于量化投资交易策略设计及衍生品定价. 另一方面,利用日内资产收益率数据,金融计量经济学家提出了新的估计每日波动率及每日资产收益率之间的协方差的方法,这些高频波动率与协方差可用来改善投资组合管理和金融风险管理. 参见 Aït-Sahalia et al. (2010), Andersen et al. (2001, 2003, 2004), Barndorff-Nielsen and Shephard (2002, 2004), Noureldin, Shephard and Sheppard (2011), Shephard and Sheppard (2010).

大数据产生了一类被称为函数数据的新型数据,如每天温度变化曲线,上市公司每天股票价格曲线,等等. 面板数据是函数数据的一个特例. 这类数据在统计学上催生了一组新模型,叫函数数据模型 (functional data models) (参见 Müller (2005), Müller and Stadtmüller (2005), Ramsay and Silverman (2002, 2005)). 另一类新型数据是区间数据 (interval-valued data),即由某个变量的最高值和最低值所组成的数的集合. 传统数据大多为点数据 (point-valued data). 相对点数据来说,区间数据包含着更多信息,但长期以来没有得到有效利用. 区间数据并不少见,例如每天的最高温最低温,高血压低血压,每天股票最高价最低价,金融资产的买卖差价 (bid-ask spread),等等. 区间数据是所谓符号数据 (symbolic data) 的一个特例. Han, Hong and Wang (2017) 率先提出了基于区间数据的 autoregressive conditional interval (ACI) 模型,利用历史区间数据预测将来区间大小. 这个模型可视为时间序列分析中 ARMA 模型的区间版本. 也参见 Han et al. (2016) 和 Sun et al. (2018).

大数据在很多情况下样本容量相当大,特别是高频数据. 如果数据样本容量远大于大数据中潜在解释或预测变量的维数,这类数据称为“高大数据 (Tall Big Data)”. 与传统数据相比,高大数据有助于挖掘大数据中变量之间的复杂关系,特别是非线性关系. 因此非线性计量经济学模型在大数据时代可望有更广的发展与应用空间. 用于大数据分析的很多机器学习方法,例如决策树,随机森林,人工神经网络,深度学习等,都是非线性计算机算法,很大程度上类似于统计学的非参数分析. 以 White 为代表的计量经济学家,构建了人工神经网络非参数建模的理论与方法 (参见 White (1992)). 人工神经网络模型曾被广泛应用于经济金融数据的实证研究中. 事实上,大数据最重要的特征并不在于样本容量大,而更在于拥有很多潜在的解释或预测变量,甚至有可能出现潜在解释或预测变量的维数超过样本容量,这称为“胖大数据 (Fat Big Data)”或高维大数据. 高

维解释变量提供了变量选择或模型选择的很多可能性。对高维大数据,很多情形下可能只有少数未知的解释变量具有重要的解释或预测能力。机器学习已发明不少计算机算法(如决策树、随机森林等),来选择重要解释或预测变量(参见 Varian (2014) 的介绍与实例)。如何识别、选择重要解释变量并估计其影响,需要借助机器学习的方法与工具,并与计量经济学理论结合起来。这方面的一个典型案例是从高维数据中选择重要变量的 LASSO 方法,这是结合机器学习的一个“统计学习”方法(参见 Hastie, Tibshirani and Wainwright (2015))。基于高维大数据,可结合机器学习方法进行研究的计量经济学理论问题中,包括:如何在大量潜在工具变量中寻找最显著的工具变量,以达到最佳工具变量估计效果,从而解决弱工具变量带来的困扰;如何在大量经济领先指数中(其个数可能超过样本容量),识别、选择其中最具有预测能力的领先指数,达到最佳样本外预测效果;如何在构建资本资产定价模型时,从大量潜在的风险因子中,分别识别每个资产的少数几个重要风险因子并估计其影响,以及如何在高维非参数函数估计中(特别是存在多个解释变量时),运用机器学习方法降低非参数函数估计量的维度,以得到更好的估计效果,等等。

大数据和机器学习也可望在研究经济因果关系方面提供很大的帮助。如前文所述,基于非实验性经济观测数据的定量政策评估的一个基本思想是,在同等条件下,比较实施某一政策后的结果与假设没有实施该政策时的虚拟结果。由于政策已经实施,假设该政策没有实施的虚拟结果是无法观测的,因此需要用计量经济学方法来估计或预测。显然,政策效应评估的准确性取决于估计或预测虚拟结果的计量经济学方法的科学性与有效性,其本质在某种意义上可视为样本外预测。由于机器学习的主要应用是基于大数据提供的海量信息,通过计算机算法提供比较准确的样本外预测,可以预计,基于大数据和机器学习的预测方法可为精准政策评估提供新的方法与工具。机器学习的算法预测并不一定是基于经济变量之间的因果关系,事实上算法预测常常看上去像个“黑箱”,不具备很清楚的可解释性。但是,正如 Varian (2014) 指出,大数据和机器学习对虚拟事实的精确预测可为经济因果关系的识别与精准量化评估政策效应提供巨大帮助。

应该强调,大数据并不代表全样本。例如,由于经济系统的复杂性、异质性与时变性,具有不确定性,因此哪怕充分利用当前所有数据,也不能精确进行样本外预测。如果大数据是“胖大数据”,即潜在解释或预测变量的维数大于样本容量,这样的高维大数据从计量经济学建模的角度看,其实是“小样本”。对“大数据”,即样本容量远大于大数据中的解释或预测变量的维数,我们在建模与统计推断时,可以不必像传统数据分析那样,主要关注由随机抽样产生的参数估计不确定性。例如,当样本容量为几万甚至几十万时,假设参数估计值的 t -检验量刚好在 1% 或 5% 显著水平上是显著的,这样的解释变量究竟能够解释因变量数据的多大比例? 对这么大样本容量,一个具有统计显著性的解释变量是否在经济效应上也是显著的? 在大样本条件下,模型参数估计与检验的重要性可能不及模型选择的重要性,特别是当存在模型不确定性时,模型不确定性的影响可能远比模型参数估计不确定性的影响要重要得多。因此,在大数据建模时,应该更关注模型选择与模型不确定性。总之,大数据不论多“大”,总是复杂经济系统的一种样本信息,建立在抽样理论基础上的统计分析 with 计量建模,其基本统计思想并不过时,但是需要创新与发展。

3 结论

本文从经典计量经济学特别是经典线性回归模型出发,通过放松线性回归模型的经典假设,

介绍了 40 年来现代计量经济学的发展历史与理论体系,并讨论了在大数据时代背景下计量经济学若干重要的发展方向与趋势.可以看出,现代计量经济学是对经典计量经济学的扬弃,尤其通过对正态分布,条件同方差与零自相关,严外生性,模型正确设定,结构平稳性等经典假设的扬弃,大大拓展了计量经济学的适用、应用范围,丰富了计量经济学的模型、方法与工具,提升了计量经济学理论的严谨性与科学性,使计量经济学成为现代经济学实证研究的最主要的科学方法论,在促进经济学科学化,推进经济理论的创新过程中,发挥着方法论的关键作用.在大数据时代,现代计量经济学还将迎来更大的发展.

本文是根据作者在对外经济贸易大学国际经济贸易学院、中央财经大学经济学院、华东师范大学经济管理学院、南开大学金融发展研究院、厦门大学邹至庄经济研究中心 2018 计量经济学师资培训班、东北财经大学经济学院和厦门大学王亚南经济研究院 2018 全国研究生计量经济学暑期学校、上海社会科学院研究生院 2018 全国研究生计量经济学暑期学校的演讲整理而成.

作者感谢国家自然科学基金委员会“计量建模与经济政策研究”基础科学中心项目(71988101)的资助.

参 考 文 献

- 洪永淼, (2007). 计量经济学的地位、作用和局限 [J]. 经济研究, (5): 139–153.
- Hong Y M, (2007). The Status, Roles and Limitations of Econometrics[J]. Economic Research Journal, (5): 139–153.
- 洪永淼, (2011). 高级计量经济学 [M]. 北京: 高等教育出版社.
- Hong Y M, (2011). Advanced Econometrics[M]. Beijing: Higher Education Press.
- 李子奈, 齐良书, (2010). 计量经济学模型的功能与局限 [J]. 数量经济技术经济研究, (9): 133–146.
- Li Z N, Qi L S, (2010). The Strength and Disadvantage of Econometric Models[J]. The Journal of Quantitative & Technical Economics, (9): 133–146.
- Aït-Sahalia Y, Fan J, Xiu D, (2010). High-frequency Covariance Estimates with Noisy and Asynchronous Financial Data[J]. Journal of American Statistical Association, 105: 1504–1517.
- Andersen T G, Bollerslev T, Diebold F X, Ebens H, (2001). The Distribution of Realized Stock Return Volatility[J]. Journal of Financial Economics, 61: 43–76.
- Andersen T G, Bollerslev T, Diebold F X, Labys P, (2001). The Distribution of Realized Exchange Rate Volatility[J]. Journal of American Statistical Association, 96: 42–55.
- Andersen T G, Bollerslev T, Diebold F X, Labys P. (2003). Modeling and Forecasting Realized Volatility[J]. Econometrica, 71: 579–625.
- Andersen T G, Bollerslev T, Meddahi N, (2004). Correcting the Errors: Volatility Forecast Evaluation Using High-frequency Data and Realized Volatilities[J]. Econometrica, 73: 279–296.
- Andrews D W K, (1991). Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation[J]. Econometrica, 59: 817–858.
- Andrews D W K, Ploberger W, (1994). Optimal Tests when a Nuisance Parameter is Present Only under the Alternative[J]. Econometrica, 62: 1383–1414.

- Andrews D W K, Ploberger W, (1995). Admissibility of the Likelihood Ratio Test when a Nuisance Parameter is Present Only under the Alternative[J]. *Annals of Statistics*, 23: 1609–1629.
- Bai J, Perron P, (1998). Estimating and Testing Linear Models with Multiple Structural Changes[J]. *Econometrica*, 66: 47–78.
- Baker M, Wurgler J, (2006). Investor Sentiment and the Cross-section of Stock Returns[J]. *Journal of Finance*, 61: 1645–1680.
- Baker M, Wurgler J, (2007). Investor Sentiment in the Stock Market[J]. *Journal of Economic Perspectives*, 21: 129–152.
- Baker S R, Bloom N, Davis S J, (2016). Measuring Economic Policy Uncertainty[J]. *Quarterly Journal of Economics*, 131: 1593–1636.
- Barndorff-Nielsen O E, Shephard N, (2002). Estimating Quadratic Variation Using Realized Variance[J]. *Journal of Applied Econometrics*, 17: 457–477.
- Barndorff-Nielsen O E, Shephard N, (2004). Econometric Analysis of Realized Covariation: High Frequency Based Covariance, Regression, and Correlation in Financial Economics[J]. *Econometrica*, 72: 885–925.
- Bierens H J, (1982). Consistent Model Specification Tests[J]. *Journal of Econometrics*, 20: 105–134.
- Bierens H J, (1990). A Consistent Conditional Moment Test of Functional Form[J]. *Econometrica*, 58: 1443–1458.
- Blundell R W, Powell J L, (2004). Endogeneity in Semiparametric Binary Response Models[J]. *Review of Economic Studies*, 71: 655–679.
- Bollerslev T, (1986). Generalized Autoregressive Conditional Heteroskedasticity[J]. *Journal of Econometrics*, 31: 307–327.
- Bollerslev T, Wooldridge J M, (1992). Quasi-maximum Likelihood Estimation and Inference in Dynamic Models with Time-varying Covariances[J]. *Econometric Reviews*, 11: 143–172.
- Breiman L, (2001). Statistical Modeling: The Two Cultures (with Comments and a Rejoinder by the Author)[J]. *Statistical Science*, 16: 199–231.
- Chen B, Hong Y, (2012). Testing for Smooth Structural Changes in Time Series Models via Nonparametric Regression[J]. *Econometrica*, 80: 1157–1183.
- Chen B, Hong Y, (2016). Detecting for Smooth Structural Changes in GARCH Models[J]. *Econometric Theory*, 32: 740–791.
- Chow G C, (1960). Tests of Equality between Sets of Coefficients in Two Linear Regressions[J]. *Econometrica*, 28: 591–605.
- Dahlhaus R, (1996). Maximum Likelihood Estimation and Model Selection for Locally Stationary Processes[J]. *Journal of Nonparametric Statistics*, 6: 171–191.
- Engle R F, (1982). Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation[J]. *Econometrica*, 50: 987–1007.
- Engle R F, Granger C W J, (1987). Co-integration and Error Correction: Representation, Estimation, and Testing[J]. *Econometrica*, 55: 251–276.
- Engle R F, Russell J R, (1998). Autoregressive Conditional Duration: A New Model for Irregularly Spaced Transaction Data[J]. *Econometrica*, 66: 1127–1162.
- Fan Y, Li Q, (1996). Consistent Model Specification Tests: Omitted Variables and Semiparametric Functional Forms[J]. *Econometrica*, 64: 865–890.
- Giannone D, Reichlin L, Small D, (2008). Nowcasting: The Real-time Informational Content of Macroeconomic Data[J]. *Journal of Monetary Economics*, 55: 665–676.

- Glosten L R, Jagannathan R, Runkle D E, (1993). On the Relation between the Expected Value and the Volatility of the Nominal Excess Return on Stocks[J]. *Journal of Finance*, 48: 1779–1801.
- Granger C W J, (1990). *Modelling Economic Series: Readings in Econometric Methodology*[M]. Oxford: Oxford University Press, 369–383.
- Granger C W, Newbold P, (1974). Spurious Regressions in Econometrics[J]. *Journal of Econometrics*, 2: 111–120.
- Hall P, (1992). *The Bootstrap and Edgeworth Expansion*[M]. New York: Springer Science & Business Media.
- Hamilton J D, (1994). *Time Series Analysis*[M]. Princeton: Princeton University Press.
- Han A, Hong Y, Wang S, (2017). Autoregressive Conditional Models for Interval-Valued Time Series Data[R]. Working Paper, Department of Economics, Cornell University.
- Han A, Hong Y, Wang S, Yun X, (2016). A Vector Autoregressive Moving Average Model for Interval-valued Time Series Data[C]// *Essays in Honor of Aman Ullah Advances in Econometrics*, 36: 417–460.
- Hansen L P, (1982). Large Sample Properties of Generalized Method of Moments Estimators[J]. *Econometrica*, 50: 1029–1054.
- Hastie T, Tibshirani R, Wainwright M, (2015). *Statistical Learning with Sparsity: The Lasso and Generalizations*[M]. Boca Raton: Taylor & Francis.
- Hausman J A, (1978). Specification Tests in Econometrics[J]. *Econometrica*, 46: 1251–1271.
- Heckman J J, Leamer E, (2001). *Handbook of Econometrics*[M]. Holland: Elsevier, 5: 3159–3228.
- Hong Y, Lee Y J, (2013). A Loss Function Approach to Model Specification Testing and Its Relative Efficiency[J]. *Annals of Statistics*, 41: 1166–1203.
- Hong Y, Wang X, Wang S, (2017). Testing Strict Stationarity with Applications to Macroeconomic Time Series[J]. *International Economic Review*, 58: 1227–1277.
- Hong Y, White H, (1995). Consistent Specification Testing via Nonparametric Series Regression[J]. *Econometrica*, 63: 1133–1159.
- Hsiao C, Ching S H, Wan S K, (2011). A Panel Data Approach for Program Evaluation: Measuring the Benefits of Political and Economic Integration of Hong Kong with Mainland China[J]. *Journal of Applied Econometrics*, 27: 705–740.
- Imbens G W, Wooldridge J M, (2009). Recent Developments in the Econometrics of Program Evaluation[J]. *Journal of Economic Literature*, 47: 5–86.
- Jarque C M, Bera A K, (1980). Efficient Tests for Normality, Homoscedasticity and Serial Independence of Regression Residuals[J]. *Economics Letters*, 6: 255–259.
- Klein R W, Spady R H, (1993). An Efficient Semiparametric Estimator for Binary Response Models[J]. *Econometrica*, 61: 387–421.
- Koenig E F, (2004). *The Taylor Rule and the Transformation of Monetary Policy*[M]. Stanford: Hoover Press.
- Lee S W, Hansen B E, (1994). Asymptotic Theory for the GARCH(1,1) Quasi-maximum Likelihood Estimator[J]. *Econometric Theory*, 10: 29–52.
- Lucas R E, (1976). *Econometric Policy Evaluation: A Critique*[C]// *Carnegie-Rochester Conference Series on Public Policy*, 19–46.
- Lumsdaine R L, (1996). Consistency and Asymptotic Normality of the Quasi-maximum Likelihood Estimator in IGARCH(1,1) and Covariance Stationary GARCH(1,1) Models[J]. *Econometrica*, 64: 575–596.

- Müller H G, (2005). Functional Modelling and Classification of Longitudinal Data[J]. *Scandinavian Journal of Statistics*, 32: 223–240.
- Müller H G, Stadtmüller U, (2005). Generalized Functional Linear Models[J]. *Annals of Statistics*, 33: 774–805.
- Malkiel B G, Fama E F, (1970). Efficient Capital Markets: A Review of Theory and Empirical Work[J]. *Journal of Finance*, 25: 383–417.
- Nelson C R, Plosser C R, (1982). Trends and Random Walks in Macroeconomic Time Series: Some Evidence and Implications[J]. *Journal of Monetary Economics*, 10: 139–162.
- Nelson D B, (1991). Conditional Heteroskedasticity in Asset Returns: A New Approach[J]. *Econometrica*, 59: 347–370.
- Newey W K, (1985). Generalized Method of Moments Specification Testing[J]. *Journal of Econometrics*, 29: 229–256.
- Newey W K, West K D, (1987). A Simple, Positive Semi-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix[J]. *Econometrica*, 55: 703–708.
- Newey W K, West K D, (1994). Automatic Lag Selection in Covariance Matrix Estimation[J]. *Review of Economic Studies*, 61: 631–653.
- Noureldin D, Shephard N, Sheppard K, (2011). Multivariate High-frequency-based Volatility (HEAVY) Models[J]. *Journal of Applied Econometrics*, 27: 907–933.
- Phillips P C B, (1977a). An Approximation to the Finite Sample Distribution of Zellner's Seemingly Unrelated Regression Estimator[J]. *Journal of Econometrics*, 6: 147–164.
- Phillips P C B, (1977b). Approximations to Some Finite Sample Distributions Associated with a First-order Stochastic Difference Equation[J]. *Econometrica*, 45: 463–485.
- Phillips P C B, (1977c). A General Theorem in the Theory of Asymptotic Expansions as Approximations to the Finite Sample Distributions of Econometric Estimators[J]. *Econometrica*, 45: 1517–1534.
- Phillips P C B, (1986). Understanding Spurious Regressions in Econometrics[J]. *Journal of Econometrics*, 33: 311–340.
- Phillips P C B, (1987a). Time Series Regression with a Unit Root[J]. *Econometrica*, 55: 277–301.
- Phillips P C B, (1987b). Towards a Unified Asymptotic Theory for Autoregression[J]. *Biometrika*, 74: 535–547.
- Ramsay J, Silverman B W, (2002). *Applied Functional Data Analysis: Methods and Case Studies*[M]. New York: Springer.
- Ramsay J, Silverman B W, (2005). *Functional Data Analysis*[M]. New York: Springer.
- Robinson P M, (1988). Root- N -Consistent Semiparametric Regression[J]. *Econometrica*, 56: 931–954.
- Shephard N, Sheppard K, (2010). Realising the Future: Forecasting with High-frequency-based Volatility (HEAVY) Models[J]. *Journal of Applied Econometrics*, 25: 197–231.
- Staiger D, Stock J H, (1997). Instrumental Variables Regression with Weak Instruments[J]. *Econometrica*, 65: 557–586.
- Stock J H, Trebbi F, (2003). Retrospectives: Who Invented Instrumental Variable Regression?[J]. *Journal of Economic Perspectives*, 17: 177–194.
- Sun Y, Han A, Hong Y, Wang S, (2018). Threshold Autoregressive Models for Interval-valued Time Series Data[J]. *Journal of Econometrics*, 206: 414–446.
- Sun Y, Hong Y, Wang S, (2019). Out-of-Sample Forecasts for China's Economic Growth and Inflation Using Rolling Weighted Least Squares[J]. *Journal of Management and Engineering*, 4(1): 1–11.

- Tauchen G, (1985). Diagnostic Testing and Evaluation of Maximum Likelihood Models[J]. *Journal of Econometrics*, 30: 415–443.
- Ullah A, (1990). *Finite Sample Econometrics: A Unified Approach*[M]. New York: Springer.
- Varian H R, (2014). Big Data: New Tricks for Econometrics[J]. *Journal of Economic Perspectives*, 28: 3–28.
- Wang S, Yu L, Lai K K, (2005). Crude Oil Price Forecasting with TEI@I Methodology[J]. *Journal of Systems Science and Complexity*, 18: 145–166.
- White H, (1980). A Heteroskedasticity-consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity[J]. *Econometrica*, 48: 817–838.
- White H, (1981). Consequences and Detection of Misspecified Nonlinear Regression Models[J]. *Journal of the American Statistical Association*, 76: 419–433.
- White H, (1984). *Asymptotic Theory for Econometricians*[M]. San Diego: Academic Press.
- White H, (1992). *Artificial Neural Networks: Approximation and Learning Theory*[M]. Oxford: Blackwell Publishers, Inc.
- White H, Stinchcombe M, (1991). Adaptive Efficient Weighted Least Squares with Dependent Observations[M]. New York: Springer.