

# 自相关检验方法的检验功效研究<sup>①</sup>

王增民 李童

(北京邮电大学经济管理学院)

**研究目标:** 探究 DW 检验和 LM 检验的检验功效及其渐近性。**研究方法:** 运用蒙特卡罗模拟实验方法结合相关影响因素对两种检验方法进行分析与比较。影响因素包括样本容量、解释变量的随机性及自相关性、随机误差项的自相关程度以及分布形态。**研究发现:** DW 和 LM 检验功效与样本容量和随机误差项的自相关程度正相关, 与解释变量的自相关程度负相关; 解释变量的随机性对 DW 和 LM 检验功效无显著影响; 错误项的几种常见分布形态的变化对 DW 和 LM 检验功效的影响可以忽略; 在错误项存在一阶自相关的情况下, DW 检验效果优于 LM 检验效果。**研究创新:** 以 DW 检验和 LM 检验的假设条件为出发点, 探究比较不同条件下自相关检验方法的检验功效。**研究价值:** 在实证研究背景下为有效地选择自相关检验方法提供借鉴参考。

**关键词** DW 检验 LM 检验 检验功效 蒙特卡罗模拟

**中图分类号** F224.0 **文献标识码** A

## 引言

由于经济系统存在惯性, 经济数据往往前后期相互关联, 进而导致时间序列数据存在自相关。而回归模型的随机误差项存在自相关会导致模型的 OLS 参数估计不再有效, 假设检验不再可靠, 预测精度降低。因此, 在参数的假设检验过程中, 要对计量模型的误差项进行序列自相关检验。目前, 广泛使用的自相关检验方法是 DW 检验和 LM 检验。

DW 检验由 Durbin 和 Watson 提出, 二人证明了 DW 统计量的真实分布介于两个极限分布之间 (Durbin 和 Watson, 1950)。之后, Durbin 和 Watson 在给定显著性水平的基础上, 推导出了 DW 分布表的两个临界值, 一个下限  $d_L$  和一个上限  $d_U$ 。通过将 DW 统计量与两个临界值进行比较, 来判断误差项  $u_t$  是否存在自相关 (Durbin 和 Watson, 1951)。DW 检验过程很简单, 仅依赖于回归模型的残差估计值, 因而被频繁使用。但 DW 检验方法始终存在局限性, 即只能检验误差项存在一阶自相关的情况并且存在无结论区域。DW 统计量落入无结论区域的概率越小, DW 检验方法越有效。通过对 DW 临界值分布表的分析, 可以得出: 随着样本容量的增大, 无结论区域是逐渐变小的。

此外, DW 检验不是在任何条件下都适用, 需要以一定的条件为前提。这些重要条件有: ①解释变量不能包含被解释变量的滞后项; ②解释变量必须是确定的; ③随机误差项必须服从正态分布等 (Durbin 和 Watson, 1950)。关于对 DW 检验条件问题的讨论, 刘明和王永瑜 (2014) 在其论文中进行了探究。然而, 当有关条件不满足时, 将对 DW 检验产生怎样的后果?

<sup>①</sup> 本文获得国家自然科学基金重点项目 (71850013) 资助。

DW 检验是否还适用？刘明和王永瑜（2014）并没给出全面的解释。事实上，在经济方面的实证研究过程中，最不能满足的一个条件就是“解释变量是确定的”这个条件。因为在经济关系的研究过程中，研究者往往是数据收集的被动接受者，无法在控制解释变量不变的情况下探索其与被解释变量之间的关系，即在相同问题的研究过程中，不同研究者收集的数据是会随机变化的，不可能在抽样过程中保持解释变量不变。因而，大部分解释变量都是随机变量。当解释变量是随机变量时，DW 检验功效是否会受到影响？采用 DW 统计量进行自相关检验是否还可行？我们有必要弄清这个问题。事实上，在实证研究过程中，使用 DW 检验方法的很多学者基本上忽略这一前提。如靳涛和陶新宇（2015）在研究我国不同阶段的经济增长动力机制时，直接对随机误差项进行 DW 检验，没有考虑解释变量是否为非随机的这一假设条件。其次，因为多种原因，回归模型中的随机误差项可能不服从正态分布，特别是在小样本情况下。Gujarati 和 Porter（2009）认为若随机误差项不服从正态分布，则 DW 检验就可能不可靠；但当样本容量足够大时，仍可使用 DW 检验。当随机误差项不满足正态性假设时，DW 检验功效会受到怎样的影响？是在所有不满足正态分布的情况下 DW 检验都不可靠吗？有没有误差项不是正态分布的几种常见分布形式，也是能进行 DW 检验的？最后，由于经济活动存在惯性，解释变量  $x_t$  本身前后期可能存在联系，导致  $x_t$  自身存在自相关。Theil 和 Nager（1961）的研究告诉我们，当回归元是缓慢变化的序列时，DW 分布真正的临界值将接近于 DW 检验的上界值。那么，解释变量  $x_t$  的自相关程度对 DW 的检验有着怎样的影响？特别是当  $x_t$  是非平稳的随机游走序列时，还可以采用 DW 进行一阶自相关检验吗？针对上述问题有必要进行探究和认识。

为了克服 DW 检验的不足，Breusch（1978）和 Godfrey（1978）提出了新的自相关检验方法——LM 检验。LM 检验不仅可以检验模型中随机误差项是否存在一阶自相关的问题，还可以检验是否存在高阶自相关的情况。LM 检验的思路是利用回归模型的残差值建立辅助回归式从而构建 LM 统计量。就 LM 检验适用条件而言，无论解释变量是确定性变量还是随机变量，LM 检验都适用；同样，LM 检验也无须要求随机误差项满足正态性假设。尽管 LM 检验条件比较宽松，但 LM 统计量始终是在渐近意义上服从卡方分布（Wooldridge, 2009）。叶宗裕和王卫杰（2019）则证明了当样本容量不是很大时，LM 统计量不服从标准的卡方分布，只有在大样本情况下，LM 统计量才近似服从卡方分布。那么，当 DW 检验的相关假设条件不满足时，是否选择 LM 统计量就一定比 DW 统计量更好？针对一阶自相关，怎样在二者之间进行选择？样本容量的增大是否可以忽略二者之间的差别？针对二者的比较问题也是有必要进行探究和认识的。

通过查阅国内外有关文献，发现许多学者对随机误差项进行序列自相关检验时，都没有很好地关注自相关检验统计量的适用条件问题。这样由于忽略其适用条件，有可能导致检验结果不一定正确。本文将基于上述诸多问题，进行理论梳理和蒙特卡罗模拟实验分析，探究相关影响因素对 DW 与 LM 检验功效及渐近性的影响并进行比较分析，为实证研究过程中自相关检验方法的选择及检验效果的判断提供依据和参考。

## 一、相关理论

### 1. DW 检验

DW 检验是通过构造 DW 统计量来判断随机误差项  $u_t$  是否存在一阶自相关的检验方法。假设  $u_t$  存在一阶自相关形式，即

$$u_t = \rho \cdot u_{t-1} + v_t \quad (1)$$

DW统计量定义为：

$$DW = \frac{\sum_{t=2}^n (e_t - e_{t-1})^2}{\sum_{t=1}^n e_t^2} \quad (2)$$

其中， $e_t$  为回归模型的残差。将式(2)展开，则：

$$DW = \frac{\sum_{t=2}^n e_t^2 + \sum_{t=2}^n e_{t-1}^2 - 2 \sum_{t=2}^n e_t e_{t-1}}{\sum_{t=2}^n e_t^2} \approx \frac{2 \sum_{t=2}^n e_{t-1}^2 - 2 \sum_{t=2}^n e_t e_{t-1}}{\sum_{t=2}^n e_{t-1}^2} = 2 \left[ 1 - \frac{\sum_{t=2}^n e_t e_{t-1}}{\sum_{t=2}^n e_{t-1}^2} \right] \quad (3)$$

由于

$$\rho \approx \frac{\sum_{t=2}^n e_t e_{t-1}}{\sum_{t=2}^n e_{t-1}^2} \quad (4)$$

所以，DW统计量与 $\rho$ 值存在如下关系：

$$DW \approx 2 \cdot (1 - \rho) \quad (5)$$

在采用 DW 检验方法时，常常需要满足以下假设条件：

- (1) 回归模型中包含截距项。
- (2) 解释变量  $x_t$  是非随机变量。
- (3) 随机误差项  $u_t$  服从正态分布。
- (4) 随机误差项  $u_t$  满足一阶自回归形式。
- (5) 回归模型中不包含被解释变量  $y_t$  的滞后项。
- (6) 样本数据无缺失项。

### 2. LM 检验

LM 检验是通过对回归模型的残差项  $e_t$  建立辅助回归式来构造 LM 统计量，从而判断随机误差项  $u_t$  是否存在自相关的检验方法。其辅助回归式的一般形式为：

$$e_t = \hat{\rho}_1 e_{t-1} + \cdots + \hat{\rho}_q e_{t-q} + \beta_0 + \beta_1 x_{1t} + \cdots + \beta_k x_{kt} + \varepsilon_t \quad (6)$$

则 LM 统计量定义为：

$$LM = (n - q) \cdot R^2 \quad (7)$$

其中， $n - q$  为式(6)的样本容量， $R^2$  为式(6)的可决系数。在大样本条件下，LM 统计量渐进服从  $\chi^2(q)$  分布，其中  $q$  为  $u_t$  的自回归阶数。若 LM 统计量所对应的  $p$  值小于显著性水平  $\alpha$ ，则拒绝原假设  $H_0$ ，说明该回归模型的随机误差项  $u_t$  存在  $q$  阶自相关，否则不拒绝原假设  $H_0$ ，即不能推断  $u_t$  存在自相关。

### 3. 检验功效

假设检验是基于样本数据对提出的假设做出推断的过程。但提出的假设并非总是正确的，

有可能做出错误的推断。因此，进行假设检验时所犯错误一般分为两类：第Ⅰ类错误为当原假设为真时，拒绝原假设的概率，即“弃真”的概率，可表示为：

$$P\{\text{拒绝 } H_0 \mid H_0 \text{ 为真}\} = \alpha \quad (8)$$

第Ⅱ类错误为当原假设为假时，接受原假设的概率，即“取伪”的概率，可表示为：

$$P\{\text{接受 } H_0 \mid H_0 \text{ 为假}\} = \beta \quad (9)$$

检验功效则是衡量假设检验方法检验效果的重要指标，一般是指当原假设为假时，检验出原假设为假的概率。也就是说，检验功效是当原假设错误时拒绝原假设的概率，即

$$P\{\text{拒绝 } H_0 \mid H_0 \text{ 为假}\} = 1 - \beta \quad (10)$$

就本文研究而言，针对模型  $u_t = \rho * u_{t-1} + v_t$  提出原假设  $H_0: \rho = 0$ ，备择假设  $H_1: \rho \neq 0$ 。

## 二、蒙特卡罗模拟实验设计

下面通过蒙特卡罗模拟实验的办法来探究误差项的自相关强度、样本容量、解释变量的确定性与随机性、解释变量的自相关程度、误差项的几种分布形态等诸多因素是如何影响 DW 及 LM 统计量的检验功效及其渐进性的，并将它们进行对比分析。

蒙特卡罗模拟实验以简单的双变量模型进行。事实上，本文也尝试了多元回归模型、对数模型及多项式模型等常见的数学模型形式，实验结果与简单的双变量模型是一致的，不改变研究结论。从第二部分给出的理论描述也可以看出，DW 统计量与模型中包含变量的个数及模型形式是没有关系的；而对 LM 统计量来说，模型中包含变量的个数仅影响 LM 统计量的自由度，模型形式对 LM 统计量也没有影响。

假定数据生成过程为：

$$y_t = \beta_0 + \beta_1 x_t + u_t \quad (11)$$

其中，

$$u_t = \rho * u_{t-1} + v_t \quad (12)$$

从式 (11) 中可以看出，为了在计算机上能模拟生成数据，首先必须对截距  $\beta_0$  和  $\beta_1$  进行赋值。本文在探究过程中，使用 0、0.5 及 1 这三个数分别对  $\beta_0$  和  $\beta_1$  进行不同组合的赋值实验，发现对  $\beta_0$  和  $\beta_1$  赋什么样的值，对结论没有影响。因为本文探究的是自相关检验统计量 DW 及 LM 的检验功效问题，而在理论上 DW 统计量与 LM 统计量均与  $\beta_0$  和  $\beta_1$  取什么值是无关的，所以对  $\beta_0$  和  $\beta_1$  赋什么值，在原理上不影响结论。故本文后面在介绍数据生成及实验设计过程中，将式 (11) 具体化为： $y_t = 1 + x_t + u_t$ ，也就是截距和斜率分别设定为 1。下面分三种情况进行蒙特卡罗模拟实验设计。

第一种情况：探究解释变量  $x_t$  是确定性变量的情况下，误差项  $u_t$  的自相关强度  $\rho$ 、 $u_t$  的分布形态及样本容量  $n$  对 DW 及 LM 检验功效的影响。

实验设计方案是：第一步，样本容量  $n$  分别取值为 20、40、60、80 和 100， $\rho$  取值为 0.2、0.5 和 0.8（仅探究正自相关）。第二步，利用计算机随机数发生器生成服从标准正态分布的随机序列  $x_t$ 。事实上，无论是理论上还是实验上，独立同分布的  $x_t$  序列分布形态的变化对结果没有影响。第三步，利用计算机随机数发生器生成几种服从不同分布的  $v_t$ 。 $v_t$  的几种不同分布分别取自正态分布、均匀分布、卡方分布、指数分布、二项分布和 F 分布，且所有分布的均值

都设为0, 方差都设为1。将 $u_t$ 的初始值 $u_0$ 赋值为0, 按公式 $u_t=\rho \cdot u_{t-1}+v_t$ 生成序列 $u_t$ 。根据 $v_t$ 的几种不同分布就可以得到几种不同的分布序列 $u_t$ 。第四步, 根据方程 $y_t=1+x_t+u_t$ 生成序列 $y_t$ 。第五步, 将 $y_t$ 对 $x_t$ 进行回归, 计算得到DW统计量值, 同时构造辅助回归计算出LM统计量值及其所对应的 $p$ 值。第六步, 将DW统计量值与DW分布表的下临界值 $d_L$ 进行比较(因为 $\rho>0$ ), 以及将LM统计量所对应的 $p$ 值与显著性水平 $\alpha$ 进行比较, 从而判断回归模型 $y_t=1+x_t+u_t$ 的随机误差项 $u_t$ 是否存在正的一阶自相关。依次重复第三步至第六步, 重复循环2000次。最后计算出DW统计量值小于 $d_L$ 的比率以及LM统计量对应的 $p$ 值小于显著性水平 $\alpha$ 的比率, 相应的比率作为两种统计量的检验功效。因为在此循环过程中, 没有重复前两步, 这样保证了在循环实验过程中, 解释变量 $x_t$ 是固定不变的, 即保证了 $x_t$ 的确定性。

第二种情况: 探究 $x_t$ 是随机变量的情况下, 误差项 $u_t$ 的自相关强度 $\rho$ 、 $u_t$ 的分布形态及样本容量 $n$ 对DW及LM检验功效的影响。

实验设计方案是: 修改第一种情况下的循环步骤, 将第一种情况下第三步至第六步循环改为第二步至第六步循环, 重复循环2000次。因为循环过程是从第二步开始的, 所以在每次循环过程中,  $x_t$ 都重新生成了一遍, 这样保证了 $x_t$ 的随机性。同样计算出2000次循环中DW统计量值小于 $d_L$ 的比率以及LM统计量对应的 $p$ 值小于显著性水平 $\alpha$ 的比率。

第三种情况: 探究 $x_t$ 的自相关程度 $\theta$ 对DW及LM检验功效的影响。

实验设计方案是: 第一步, 样本容量 $n$ 分别取值为20、40、60、80和100,  $\rho$ 仅取0.5。第二步,  $x_t$ 按如下方法生成: 首先对 $x_t$ 的初始值 $x_0$ 赋值为0, 然后按公式 $x_t=\theta \cdot x_{t-1}+w_t$ 生成 $x_t$ 序列值。其中,  $w_t$ 序列按标准正态分布生成,  $\theta$ 的取值分别为0、0.5和1。这样得到三种形式的 $x_t$ 序列数据: 一是在时间上不存在序列相关的平稳序列, 二是存在序列相关的平稳时间序列, 三是非平稳的随机游走序列。第三步, 对 $u_t$ 的初始值 $u_0$ 赋值为0, 按公式 $u_t=\rho \cdot u_{t-1}+v_t$ 生成序列 $u_t$ 。其中 $v_t$ 序列按标准正态分布生成。第四步, 根据三种形式的 $x_t$ 序列数据及方程 $y_t=1+x_t+u_t$ 生成序列 $y_t$ 。第五步, 将相应的 $y_t$ 对 $x_t$ 进行回归, 计算得到DW统计量值, 同时构造辅助回归计算出LM统计量值及其所对应的 $p$ 值。第六步, 将DW统计量值与DW分布表的下临界值 $d_L$ 进行比较(因为 $\rho>0$ ), 以及将LM统计量所对应的 $p$ 值与显著性水平 $\alpha$ 进行比较, 从而判断出回归模型 $y_t=1+x_t+u_t$ 的随机误差项 $u_t$ 是否存在正的一阶自相关。依次重复第二步至第六步, 重复循环2000次。同样计算出DW统计量值小于 $d_L$ 的比率以及LM统计量对应的 $p$ 值小于显著性水平 $\alpha$ 的比率。

### 三、蒙特卡罗模拟实验结果分析

#### 1. $x_t$ 是确定性变量的情况下, $u_t$ 的分布形态、 $n$ 和 $\rho$ 对DW及LM检验功效的影响分析与比较

在研究DW和LM检验功效的过程中, 本文提出的自相关假设是 $H_0: \rho=0$ , 即误差项不存在自相关。由于在实验过程中,  $\rho$ 分别取值为0.2、0.5和0.8, 导致原假设 $H_0: \rho=0$ 始终不成立。因此, DW及LM检验功效可用不接受原假设 $H_0$ 的概率度量。Durbin和Waston(1951)研究表明, DW统计量值会落入DW分布的五个区域, 若 $0<DW<d_L$ , 则表明回归模型中随机误差项 $u_t$ 存在正的自相关关系。在本文的探究过程中, 用2000次循环过程中DW值小于下界 $d_L$ 值的比率, 作为DW统计量的检验功效的度量, 以及LM统计量所对应的 $p$ 值小于显著性水平 $\alpha$ 的比率作为LM统计量的检验功效的度量。

根据第一种情况的蒙特卡罗模拟实验设计, 通过改变样本容量 $n$ 、 $u_t$ 的自回归系数 $\rho$ 及 $u_t$

的分布形态来进行检验功效探究。在显著性水平  $\alpha=0.05$  的条件下, 模拟结果如表 1、表 2 所示。

表 1 DW 统计量检验功效

$\rho$	$n$	$u_1$	$u_2$	$u_3$	$u_4$	$u_5$	$u_6$
0.2	20	0.096	0.096	0.081	0.085	0.095	0.076
	40	0.223	0.236	0.212	0.228	0.246	0.218
	60	0.351	0.378	0.307	0.323	0.367	0.301
	80	0.467	0.458	0.408	0.441	0.450	0.410
	100	0.546	0.566	0.508	0.538	0.570	0.513
0.5	20	0.459	0.441	0.451	0.466	0.447	0.467
	40	0.861	0.852	0.899	0.883	0.856	0.916
	60	0.965	0.959	0.982	0.984	0.972	0.990
	80	0.990	0.996	1	0.998	0.994	0.999
	100	0.999	0.998	1	0.999	0.999	1
0.8	20	0.819	0.829	0.847	0.839	0.828	0.878
	40	0.995	0.998	0.999	0.997	0.995	0.999
	60	1	1	1	1	1	1
	80	1	1	1	1	1	1
	100	1	1	1	1	1	1

注: 表中数据是通过在循环过程中 DW 统计量落入  $[0, d_L]$  区域的次数除以总循环次数 (2000) 的方法求得。

表 2 LM 统计量检验功效

$\rho$	$n$	$u_1$	$u_2$	$u_3$	$u_4$	$u_5$	$u_6$
0.2	20	0.072	0.073	0.057	0.061	0.071	0.047
	40	0.155	0.169	0.141	0.160	0.177	0.153
	60	0.263	0.276	0.224	0.240	0.271	0.214
	80	0.358	0.359	0.317	0.344	0.348	0.302
	100	0.442	0.452	0.397	0.427	0.457	0.390
0.5	20	0.386	0.359	0.363	0.368	0.364	0.361
	40	0.796	0.799	0.833	0.828	0.797	0.854
	60	0.942	0.935	0.957	0.967	0.954	0.973
	80	0.985	0.987	0.998	0.993	0.987	0.997
	100	0.998	0.996	0.999	0.998	0.997	0.999
0.8	20	0.748	0.759	0.782	0.775	0.763	0.799
	40	0.992	0.994	0.997	0.995	0.990	0.996
	60	1	1	1	1	1	0.998
	80	1	1	1	1	1	1
	100	1	1	1	1	1	1

注: 表中数据是通过在循环过程中 LM 统计量所对应的  $p$  值小于显著性水平  $\alpha$  的次数除以总循环次数 (2000) 的方法求得。

通过对表1和表2中的数据进行分析,可以发现:①当 $u_t$ 在相同分布的条件下, $n$ 和 $\rho$ 越大,DW和LM统计量的检验功效就越大;②当 $n$ 和 $\rho$ 大到一定程度,DW和LM的检验功效都趋近于1。如两表中的数据显示,当 $\rho=0.8$ 且 $n\geqslant 60$ 时,DW和LM检验功效都为1,此时,DW统计量和LM统计量检验效果都达到最优。

为了更清楚地探究 $u_t$ 的不同分布形态是如何影响DW及LM统计量检验功效的,根据表1中的数据绘制出图1,根据表2中的数据绘制出图2。

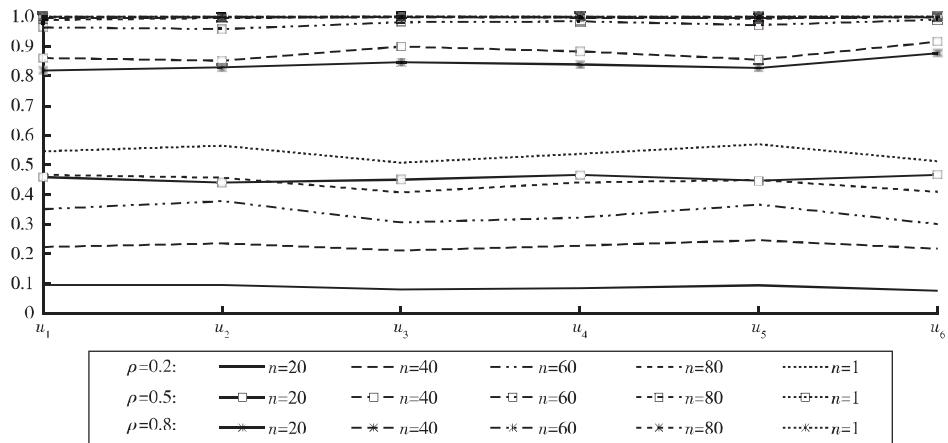


图1  $x$  非随机条件下, DW 检验功效趋势

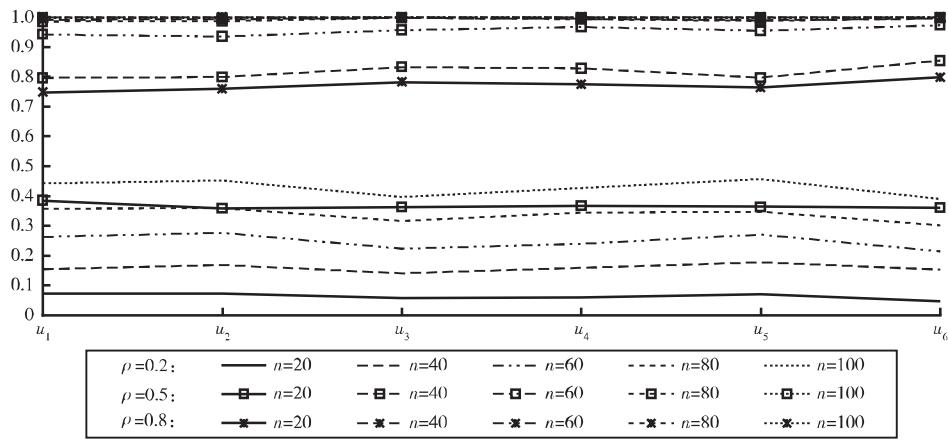


图2  $x$  非随机条件下, LM 检验功效趋势

从图1可以看出:当 $\rho$ 与 $n$ 都比较大或 $\rho$ 与 $n$ 都比较小的情况下,DW检验功效趋势图对应的曲线都比较平坦。当 $\rho$ 与 $n$ 都比较大时,DW检验功效趋近于1;当 $\rho$ 与 $n$ 都比较小时,DW检验功效非常小,接近于0,说明在这两种情况下, $u_t$ 分布形态的变化不能对DW检验功效产生较大的波动。仔细观察图1还可以看到:有几条曲线随着 $u_t$ 分布形态的不同出现一定的波动,有波动的几条曲线主要出现在 $\rho$ 比较小而 $n$ 比较大的情况、 $\rho$ 比较大而 $n$ 比较小的情况或 $\rho$ 取0.5而 $n$ 取40、60几个中间值情况,但其波动幅度都不是很大。事实上,根据表1中的数据可以推断出,随着 $u_t$ 分布形态的变化,无论 $\rho$ 与 $n$ 取何值,DW检验功效波动幅度都介

于 0 到 0.06 的范围内。

因此,根据本文的模拟实验研究,可以得出结论:尽管相关文献和教科书中常常提到 DW 检验的前提条件之一是要求  $u_t$  服从正态分布,但  $u_t$  常见的几种分布形态的变化对 DW 统计量检验功效的影响是有限的。特别是  $\rho$  与  $n$  都比较大时,更是可以忽略  $u_t$  的分布形态对检验功效的影响。当然,或许存在  $\rho$  与  $n$  比较小的情况下本文研究以外的  $u_t$  的某种分布形式可能会导致 DW 检验存在一定的误差,但通过本文对  $u_t$  的几种常见分布形式的探究,还是为我们大胆采用 DW 统计量进行一阶自相关检验迈进了一步。

图 2 中 LM 检验功效趋势图与图 1 中 DW 检验功效趋势图走向基本一致。同样可以得出结论:当  $\rho$  与  $n$  都较大或  $\rho$  与  $n$  都较小时,LM 检验功效随着  $u_t$  分布形态的变化基本不变。尽管  $\rho$  和  $n$  取某范围内的值,LM 检验功效会随着  $u_t$  分布的变化有一定的波动,但波动幅度同样介于 0 到 0.06 的范围内。所以,当  $x_t$  是非随机变量时,  $u_t$  常见的几种不同分布形态的变化不会显著影响 LM 统计量的检验功效。

为了对 DW 和 LM 检验方法的检验效果进行比较,以  $u_t$  服从正态分布为例,结合表 1 数据和表 2 数据绘成图 3。本文也在  $u_t$  服从其他分布的条件下对 DW 和 LM 检验功效进行了比较研究,实验结果与图 3 的结论一致。

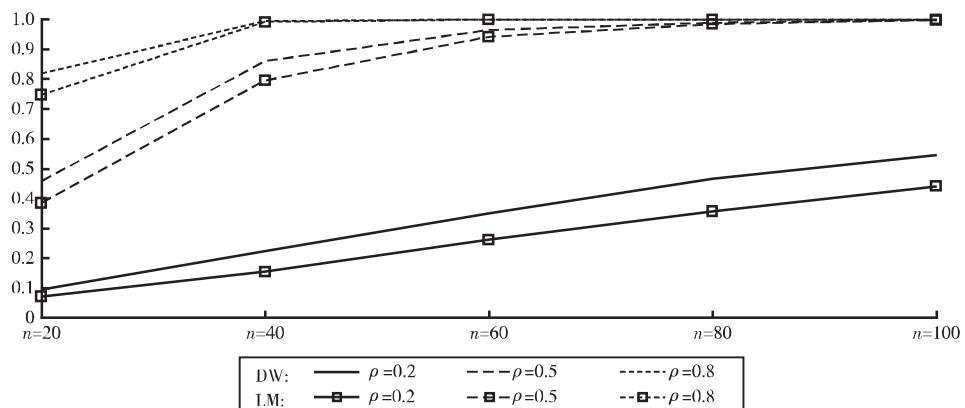


图 3  $x_t$  非随机与  $u_t$  正态条件下, 检验功效趋势对比

图 3 为误差项服从正态分布的条件下 DW 统计量和 LM 统计量检验效果对比。从图 3 可以看出:①误差项的自回归系数  $\rho$  越大,DW 和 LM 检验功效越趋近于 1;②无论  $n$  和  $\rho$  取何值,DW 检验功效都高于或不低于 LM 检验功效;③当  $\rho$  比较小时,如  $\rho=0.2$  时,随着  $n$  的增大,DW 检验功效高于 LM 检验功效的差距越来越大;但可以推测出随着  $n$  的进一步增大,二者的检验功效差距又会逐渐缩小,进而检验功效都趋近于 1;④当  $\rho$  比较大时,如  $\rho=0.8$  时,随着  $n$  的增大,LM 检验功效增长速度快于 DW 检验功效,进而 DW 检验功效优于 LM 检验功效的差距逐渐缩小,二者的检验功效都趋近于 1;⑤当  $\rho$  取中间值,如  $\rho=0.5$  时,起初随着  $n$  的增大,DW 和 LM 检验功效的增长速度大体相同。但随着  $n$  的进一步增大,LM 检验功效改进程度更快些,进而二者的差距逐渐缩小并趋近于 1。

2. 在  $x_t$  是随机变量且不存在自相关的情况下,  $u_t$  的分布形态、 $n$  和  $\rho$  对 DW 及 LM 检验功效的影响分析与比较

DW 检验假设条件之一为解释变量是确定的,该假设条件是 Durbin 和 Waston 确定 DW 分

布表临界值  $d_L$  和  $d_U$  的基础。然而在实证研究过程中，解释变量  $x_t$  几乎都是随机变量。当  $x_t$  是随机变量时，仍以  $d_L$  和  $d_U$  作为临界值计算，会对 DW 检验的有效性产生怎样的影响？LM 检验并不受限于  $x_t$  是确定性变量这一条件，也不限制  $u_t$  的分布形态。那么，当  $x_t$  是随机变量时，DW 检验效果还会比 LM 检验效果好吗？在  $x_t$  是随机变量的情况下， $u_t$  分布形态的变化对 DW 及 LM 统计量的检验功效会产生怎样的影响呢？

为了探究上述问题，根据第二种情况的蒙特卡罗模拟实验设计，在显著性水平  $\alpha=0.05$  的情况下，其模拟结果如表 3、表 4 所示。

表 3 DW 统计量检验功效

$\rho$	$n$	$u_1$	$u_2$	$u_3$	$u_4$	$u_5$	$u_6$
0.2	20	0.102	0.100	0.077	0.085	0.102	0.085
	40	0.232	0.229	0.217	0.215	0.229	0.192
	60	0.337	0.358	0.324	0.346	0.359	0.291
	80	0.451	0.431	0.425	0.430	0.464	0.407
	100	0.559	0.551	0.534	0.558	0.556	0.532
0.5	20	0.436	0.437	0.431	0.437	0.429	0.440
	40	0.837	0.824	0.885	0.866	0.841	0.890
	60	0.962	0.968	0.985	0.979	0.972	0.986
	80	0.992	0.993	0.998	0.997	0.994	0.998
	100	0.999	1	1	1	0.998	1
0.8	20	0.814	0.797	0.838	0.829	0.803	0.853
	40	0.995	0.995	0.999	0.997	0.994	0.999
	60	1	1	1	1	1	1
	80	1	1	1	1	1	1
	100	1	1	1	1	1	1

表 4 LM 统计量检验功效

$\rho$	$n$	$u_1$	$u_2$	$u_3$	$u_4$	$u_5$	$u_6$
0.2	20	0.069	0.072	0.057	0.062	0.076	0.057
	40	0.173	0.168	0.151	0.150	0.171	0.132
	60	0.240	0.276	0.241	0.249	0.266	0.211
	80	0.347	0.327	0.332	0.325	0.371	0.289
	100	0.455	0.436	0.418	0.454	0.448	0.414
0.5	20	0.355	0.353	0.327	0.355	0.347	0.337
	40	0.775	0.766	0.815	0.789	0.786	0.823
	60	0.935	0.941	0.967	0.961	0.952	0.975
	80	0.988	0.986	0.993	0.991	0.987	0.995
	100	0.998	0.998	1	0.999	0.995	0.999
0.8	20	0.752	0.725	0.768	0.756	0.730	0.780
	40	0.991	0.988	0.998	0.993	0.989	0.998
	60	1	1	1	1	1	1
	80	1	1	1	1	1	1
	100	1	1	1	1	1	1

通过分析表3和表4的数据可以发现：①当 $x_t$ 是随机变量时，无论 $u_t$ 服从何种分布，随着 $n$ 和 $\rho$ 的增大，DW及LM统计量的检验功效逐渐增大并逐渐趋近于1。这与 $x_t$ 是确定性变量时DW及LM检验功效的特征是一致的。②当 $\rho$ 较小时， $u_t$ 的不同分布形态会导致DW和LM检验功效出现细微变化。为了更直观地呈现 $u_t$ 的分布形态对检验功效的影响，这里只绘出DW检验功效的变化趋势，如图4所示。LM检验功效的变化趋势也呈现类似的走向，故而省略。

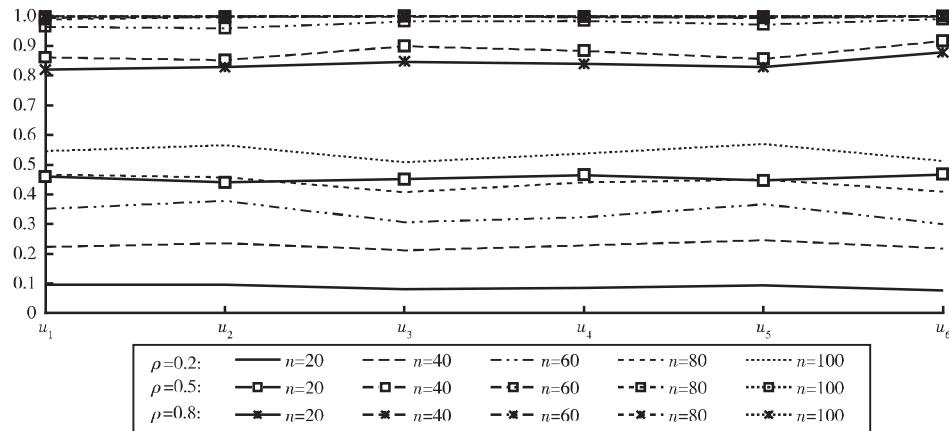


图4  $x_t$ 随机条件下，DW检验功效趋势对比

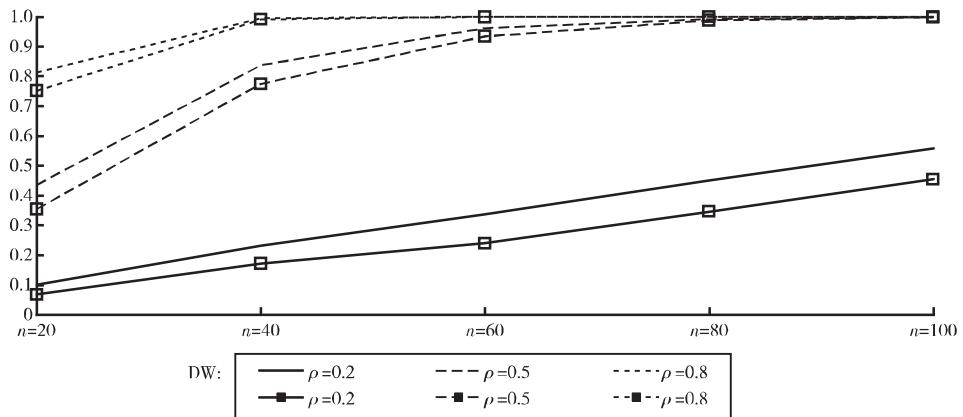
图4中DW检验功效趋势的变化规律基本与图1相同。因此，可以得出结论：①无论 $x_t$ 是随机的还是确定的，针对 $u_t$ 的常见几种分布形式，随着 $n$ 和 $\rho$ 的增大，DW检验功效是逐渐增大的，且检验功效逐渐趋近于1；②在 $x_t$ 是随机变量的情况下，尽管 $n$ 和 $\rho$ 在某些范围内取值时，DW检验功效会随着 $u_t$ 分布形态的不同而有所波动，但波动幅度介于很小的范围内，变化差异仍然控制在0到0.06的范围内，且当 $\rho$ 较大时，DW检验功效更是不再随 $u_t$ 分布形态的变化而变化，其趋势图呈一条平整的直线。因此，可以认为在 $x_t$ 是随机变量的情况下，DW检验功效在 $u_t$ 的常见几种分布形态之间没有显著差异。

为了探究在 $x_t$ 是随机变量的情况下，DW检验功效和LM检验功效之间的差异，以 $u_t$ 服从正态分布为例，结合表3和表4数据绘成图5。当 $u_t$ 服从其他分布时，DW和LM检验功效的对比图也表现出相同的结论。

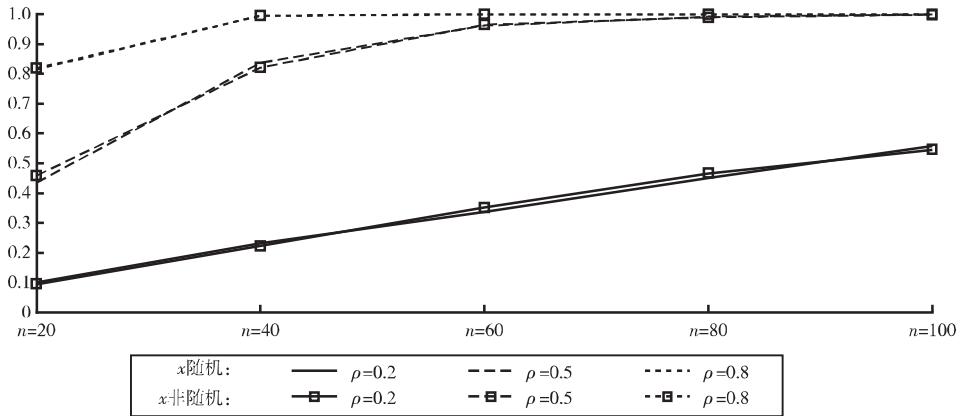
图5给出的DW与LM检验功效对比图的变化趋势与图3基本一致。可以得出结论：①当误差项服从正态分布时，无论 $x_t$ 是确定性的还是随机的，相同条件下，DW统计量的检验效果优于或不差于LM统计量的检验效果；②当 $\rho$ 比较小时，如 $\rho=0.2$ 时，随着样本容量 $n$ 的增大，起初DW检验功效优于LM检验功效的差距不断增大，当 $n$ 进一步增大时，二者的差距又逐渐减小，并且检验功效都趋近于1；③当 $\rho$ 较大时，如 $\rho=0.8$ 时，随着样本容量 $n$ 的增大，LM检验功效的增长速度快于DW检验功效的增长速度，二者检验功效的差距逐渐减小，最后都趋近于1。

### 3. 在 $x_t$ 是确定性变量与时间上独立同分布随机变量的两种情况下，DW检验功效的分析与比较

当 $x_t$ 从确定性变量转变为随机变量时，DW检验功效是否会发生变化呢？为了对比DW检验功效在 $x_t$ 是确定性变量与独立同分布随机变量两种情况下的差异，将 $u_t$ 的分布形态按正

图 5  $x_t$  随机与  $u_t$  正态条件下, 检验功效趋势对比

态分布生成, 结合表 1 和表 3 数据, 绘成图 6。事实上, 通过模拟研究发现, 当  $u_t$  来自其他分布时,  $x_t$  两种情况下的 DW 检验功效对比效果与图 6 的结论相同, 因而省略。

图 6  $u_t$  正态条件下, DW 检验功效趋势对比

根据图 6 可以看出, 无论  $n$  和  $\rho$  取何值,  $x_t$  是独立同分布随机变量时的 DW 检验功效趋势图与  $x_t$  是确定性变量时的 DW 检验功效趋势图基本重合。因此, 可以得出结论: 在  $x_t$  是确定性变量和随机变量两种情况下 DW 检验功效及其渐近性没有差异。因此, 在采取 DW 检验方法对误差项进行一阶自相关检验的过程中, 当  $x_t$  是独立同分布的随机变量时仍可以大胆采用 DW 检验。

#### 4. $x_t$ 的自相关程度 $\theta$ 对 DW 及 LM 检验功效的影响分析与比较

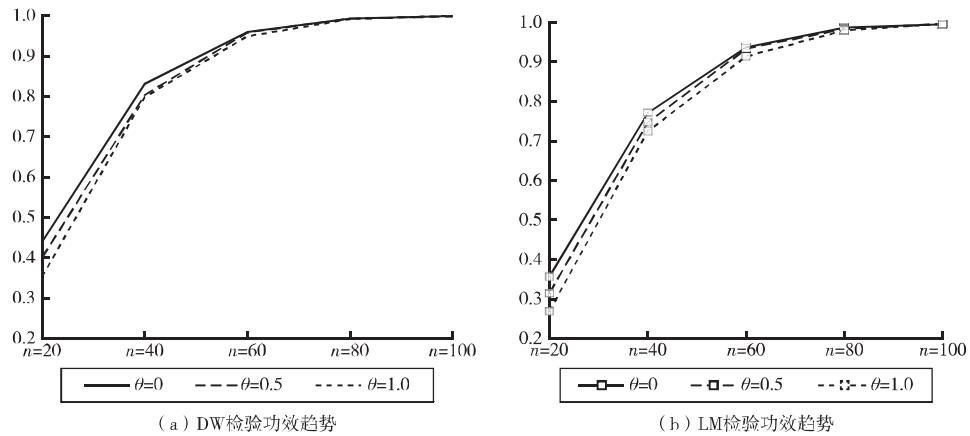
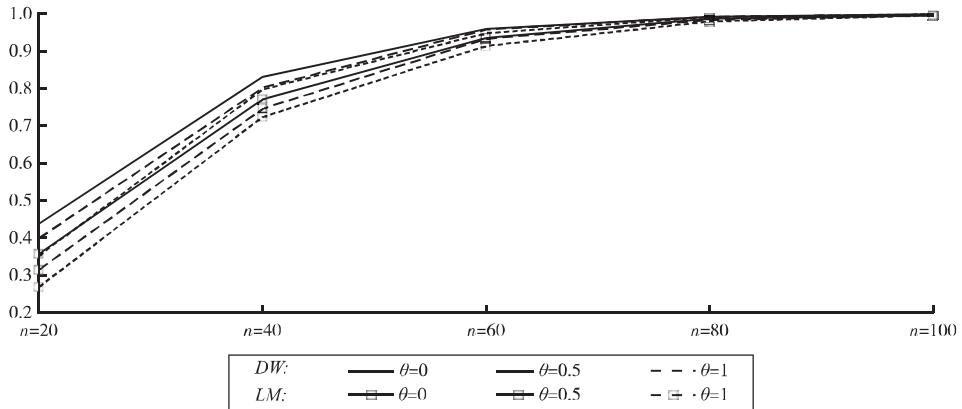
在现实的经济生活中, 时间序列本身往往前后期存在相关性, 特别是在协整关系研究中, 解释变量和被解释变量都是非平稳的随机游走序列或单位根序列。当  $x_t$  是自相关序列甚至是单位根序列时,  $x_t$  的自相关强度  $\theta$  将会对 DW 和 LM 检验功效及其渐近性产生怎样的影响? 为了探究这一问题, 根据蒙特卡罗模拟实验设计的第三种情况, 通过改变样本容量  $n$  及  $x_t$  的自相关强度  $\theta$  来进行蒙特卡罗模拟实验, 在显著性水平  $\alpha=0.05$  的情况下, 其模拟结果如表 5 所示。

表 5

DW 和 LM 统计量检验功效

检验方法	$\theta$	$n$				
		20	40	60	80	100
DW	0	0.439	0.831	0.960	0.993	0.999
	0.5	0.400	0.803	0.958	0.993	0.998
	1	0.353	0.798	0.949	0.992	0.999
LM	0	0.358	0.772	0.937	0.987	0.996
	0.5	0.316	0.747	0.934	0.984	0.996
	1	0.270	0.725	0.914	0.980	0.996

从表 5 可以看出：①在样本容量相同且比较小的条件下，随着  $\theta$  的增大，DW 和 LM 检验功效逐渐减小；②无论  $\theta$  取何值，即无论  $x_t$  是平稳序列还是随机游走序列，随着样本容量  $n$  的增大，DW 和 LM 检验功效逐渐趋于一致并趋近于 1。为了更容易地看清  $\theta$  的变化对 DW 和 LM 检验功效的影响，根据表 5 数据绘成图 7。同时，为了对比分析 DW 和 LM 统计量的检验效果，绘制出图 8。

图 7  $x$  随机与  $u$  正态条件下，自相关检验功效趋势图 8  $x$  随机与  $u$  正态条件下，检验功效趋势对比

从图 7 中可以看出：DW 检验功效趋势图与 LM 检验功效趋势图都呈上凸趋势。在样本容量相同的情况下， $\theta$  越大，即  $x_t$  自身存在的自相关性越强，DW 和 LM 检验功效就越差；无论  $\theta$  取

何值，随着样本容量  $n$  的增大，DW 和 LM 检验功效逐渐增大并逐渐趋近于 1，但增长速度由快变慢，即随着样本容量  $n$  的增大， $x_t$  的自相关强度  $\rho$  的变化对两种检验功效的影响逐渐忽略。

从图 8 中可以看出：无论  $x_t$  是平稳序列还是随机游走序列，DW 检验功效都比 LM 检验功效大，且随着样本容量的增大，DW 和 LM 检验功效差距逐渐缩小并趋于一致。

### 5. 关于小样本下的进一步探讨

在上述的研究过程中，为了从趋势上把握规律，模拟实验设计过程中，样本容量  $n$  取值从 20 跨度到 100，分别为 20、40、60、80 和 100。事实上，特别是针对时间序列数据而言，在实证研究过程中，我们可获取的样容量通常比较小。针对小样本情况，本文做了进一步的探究和分析。我们将表 1 中的样本容量  $n$  的取值改为 10、20、40、60 和 80，即去掉样本容量为 100 对应的数据添加样本容为 10 对应的数据。重新运行模拟程序并生成数据（为了篇幅所限，省略数据展示），根据生成的数据绘制出图 9。

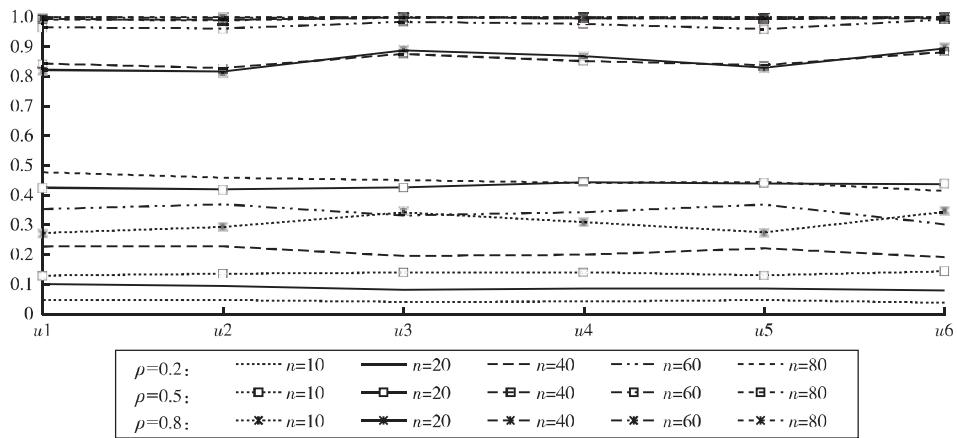


图 9  $x$  非随机条件下，含有样本容量  $n=10$  的 DW 检验功效趋势

根据图 9，不难发现，当  $n$  为 20、40、60 和 80 所对应的趋势线的形态与位置基本与图 1 中相应趋势线是一致的。观察  $n$  为 10 所对应的趋势线可以发现，相比其他样本容量，对应的 DW 检验功效更加小，并且随着误差项  $u_t$  分布形态的变化，波动幅度很小。

将表 2 中的样本容量 100 替换成 10，其他样本容量不变可以得到类似的结论。

为了进一步在小样本下比较 DW 与 LM 检验功效，我们将图 3 给出的功效趋势对比图在小样本容量下绘制出，如图 10 所示。

根据图 10，我们可以得出结论：①随着样本容量  $n$  的增加，DW 与 LM 检验功效仍然逐渐增强，且任何情况下 DW 检验功效都不低于 LM。②当自相关系数  $\rho$  介于 0.5 左右时，即使样本容量达到 30，二者的检验功效分别是 73% 与 64%，即仍然会有约 27%~36% 的概率无法检验出存在一阶自相关；在  $\rho$  等于 0.5 的情况，样本容量  $n$  从 10 逐渐增大到 30，DW 检验功效始终明显优于 LM。③当  $\rho$  很小如等于 0.2 时，无论  $n$  在 10 至 30 取何值，二者的检验功效都很低，低于 16%，且随着  $n$  的增大，二者的检验功效增长也都很缓慢，即基本上无法判断出存在一阶自相关。④当自相关系数  $\rho$  等于 0.8 时， $n$  从 10 增大到 30，二者的检验功效增长速度都较快，且 LM 的增长快于 DW。然而，当样本容量  $n$  等于 10 时，DW 检验功效仅仅约是 34%，LM 检验功效仅仅约是 22%；当  $n$  等于 20 时，DW 检验功效约是 78%，LM 检验功效约是 70%。只有  $n$  等于 30 时，检验功效达到 90% 以上。

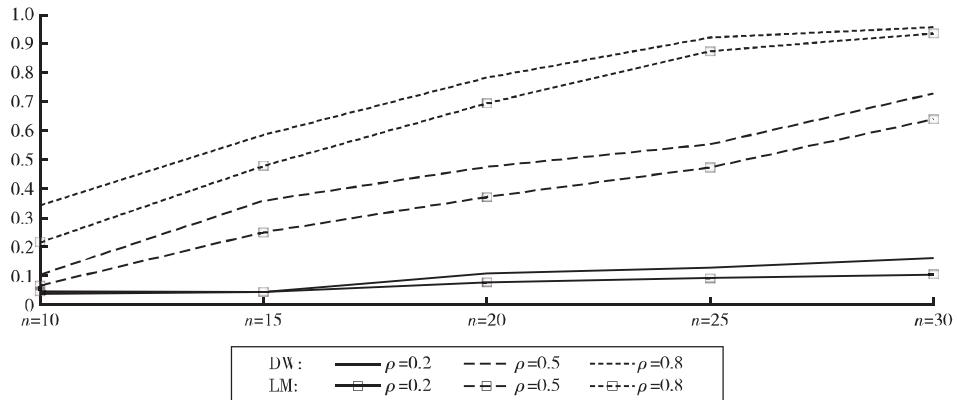


图 10  $x$  非随机、 $u$  正态条件及小样本情形下，检验功效趋势对比

将图 5 在小样本情形下绘制仍然会得到类似的结论。

#### 四、结论

本文通过蒙特卡罗模拟实验的方法，在误差项存在一阶自相关的条件下，探究了样本容量  $n$ 、误差项  $u_t$  的自回归系数  $\rho$ 、 $u_t$  的分布形态、解释变量  $x_t$  的确定性和随机性、 $x_t$  的自相关强度  $\theta$  等因素对 DW 和 LM 检验功效及其渐近性的影响，并对 DW 和 LM 检验方法进行了对比分析。可以归纳如下结论：

第一，无论是 DW 检验还是 LM 检验，在误差项存在一阶自相关的前提下，随着  $n$  和  $\rho$  的逐渐增大，二者的检验功效都逐渐增大，并且趋近于 1。 $x_t$  的自相关强度  $\theta$  越大，二者的检验功效越小。

第二，当误差项  $u_t$  存在一阶自相关时，在任何情况下，DW 优于或不差于 LM 检验功效；随着  $n$  和  $\rho$  的逐渐增大，二者检验功效的差异逐渐减小，并且二者检验功效逐渐趋于一致。

第三，尽管有关文献在阐述 DW 检验时，常把  $u_t$  服从正态分布作为 DW 检验条件之一，但当  $u_t$  来自几种常见的分布形式时， $u_t$  分布形态的变化对 DW 和 LM 检验功效的影响基本可以忽略。因此，在实证研究过程中，对回归模型的误差项进行一阶自相关检验时，可以适当放松  $u_t$  的分布形态变化。

第四，有关文献在阐述 DW 检验时，把  $x_t$  是确定性变量作为前提条件，但是，这一条件在现实研究过程中很难得到满足。本文研究发现，当  $x_t$  是时间上独立同分布的随机变量时，DW 及 LM 统计量的检验功效及其渐进性与  $x_t$  是确定性变量条件下的结果是一致的。当  $x_t$  是存在自相关的随机变量时， $x_t$  的自相关强度越大，DW 及 LM 统计量的检验功效越差，但通过扩大样本容量  $n$ ，可以逐渐克服  $x_t$  的自相关强度对 DW 及 LM 检验功效的影响。

第五，尽管当误差项仅存在一阶自相关的情况下，DW 始终优于 LM 检验，但当误差项存在高阶自相关的时候，DW 检验就不适用。所以当采用两种方法进行一阶自相关检验后，还怀疑误差项存在高阶自相关的时候，最好用 LM 检验方法进行进一步的高阶自相关检验。若是模型中解释变量不是严格外生的，如模型中包括滞后的解释变量，那么 DW 检验也不适用，此时最好采用 LM 统计量或其他方法进行检验。

自相关检验是在实证研究过程中不可回避的问题。通过在知网（CNKI）上的北大核心及 CSSCI 期刊库中进行搜索，搜索全文中既包含“一阶自相关”同时又包含“DW 检验”词语的论文，目前大约有 135 篇。通过查阅一些论文，可以发现一些研究现象并总结出一些

问题，进而有利于认识和把握本文的研究意义。通过文献梳理，我们发现：①很多的实证研究论文在采用 DW 检验的时候，基本上都没有考虑解释变量是随机性还是确定的，也没有验证误差项是不是服从正态分布的。这种忽略解释变量可能是随机的以及误差项可能不服从正态分布的情况下，是否会造成检验结果的不正确？应该说本文的研究结论对其以往的研究结果起到了一定的支撑作用。②当采用 DW 统计量进行一阶自相关检验的时候，若不能拒绝原假设，那么打算再次利用 LM 统计量进行进一步的检验，以此判断是否会得到一个不一样的结果是不明智的。靳庭良和张宝青（2009）的论文中却采用了这样一种机制。他们在原模型的自由度为 21 的情形下，首先利用 DW 统计量进行了一阶自相关检验，发现无法确定存在一阶自相关，他们进而采用 LM 检验进行了进一步的探索，结果仍然是无法拒绝原假设。基于 LM 检验功效低于 DW 检验，想利用 LM 检验进行进一步的判断应该是意义不大的。明智的办法应该是扩大样本容量进行进一步的检验或计算估计系数稳健标准误。③根据本文的研究结论，当对模型中的误差项进行一阶自相关检验的时候，若是 DW 检验条件得到满足，最好采用 DW 检验，而不是先采用 LM 检验进而接受原假设。孙秋柏等（2014）的论文中，在没有说明 DW 检验条件是否满足的情况下，却采用了 LM 来检验一阶自相关，且因 LM 统计对应的 P 值是 0.072，所以接受了不存在自相关的原假设，这种接受原假设的推断是值得怀疑的，最好判断一下 DW 检验条件进而采用 DW 统计量来对一阶自相关进行检验。④当实证研究过程中，若是怀疑模型中误差项存在二阶及以上的高阶自相关，那么首先应该采用 LM 检验方法进行高阶自相关检验，若是拒绝存在二阶以上自相关，就有必要在 DW 检验条件成立的情况下，采用 DW 检验方法来检验一阶自相关。崔海燕（2007）的论文采用了这个策略。

近十几年来，计量经济学理论应用在我国得到了长足而广泛的发展，然而轻理论重应用的现象一直存在，不少作品不能很好地把握计量经济学的建模理论和建模思想，单纯地将计量经济模型作为一种工具照搬套用，不能灵活把握这一工具的内涵，出现不少照猫画虎的“赝品”。早在 2010 年，李子奈与齐良书（2010）就撰文写道：“计量经济学应用研究中，问题和错误大量存在，究其原因，对计量经济学模型的方法基础，特别是它的哲学基础缺乏深入研究和正确理解。”2017 年，李子奈（2017）再一次强调：“提高应用研究水平已经刻不容缓。……否则，有可能影响人们对计量经济学的热情和信心，甚至葬送计量经济学发展的大好局面。”所以，加强计量经济学理论的深入研究仍然十分重要。掌握自相关检验是在实证研究过程中不可回避的理论问题，自 DW 检验方法成立以来，一直有学者对 DW 检验的缺陷性及功效性问题进行探讨。通过本文的基础性研究，进一步验证、扩展和明确了 DW 检验的适用条件以及 DW 与 LM 检验功效的差异，这必将对大量实证研究的科学性起到支撑作用。

#### 参 考 文 献

- [1] Brcusch T. S. , 1978, *Testing for Autocorrelation in Dynamic Linear Models* [J], Australian Economic Papers, 17 (31), 334~355.
- [2] Durbin J. , Watson G. S. , 1950, *Testing for Serial Correlation in Least Squares Regression I* [J], Biometrika, 37 (3~4), 409~428.
- [3] Durbin J. , Watson G. S. , 1951, *Testing for Serial Correlation in Least Squares Regression II* [J], Biometrika, 38 (1~2), 159~178.
- [4] Godfrey L. G. , 1978, *Testing against General Autoregressive and Moving Average Error Models When the Regressors Includes Lagged Dependent Variables* [J], Econometrica, 46 (6), 1293~1301.

- [5] Gujarati D. N., Porter D. C., 2009, *Basic Econometrics* [M], 5th ed, McGraw-Hill.
- [6] Theil H., Nager A. I., 1961, *Testing the Independence of Regression Disturbances* [J], Journal of the American Statistical Association, 56, 793~806.
- [7] Wooldridge J. M., 2009, *Introductory Econometrics: A Modern Approach* [M], 4th ed, South-Western Cengage Learning.
- [8] 崔海燕:《我国农村居民消费行为的实证研究》[J],《山西财经大学学报》2007年第Z1期。
- [9] 靳涛、陶新宇:《中国持续经济增长的阶段性动力解析与比较》[J],《数量经济技术经济研究》2015年第11期。
- [10] 靳庭良、张宝青:《回归分析中的t检验与F检验关系的进一步探讨》[J],《统计与决策》2009年第21期。
- [11] 刘明、王永瑜:《Durbin-Watson 自相关检验应用问题探讨》[J],《数量经济技术经济研究》2014年第6期。
- [12] 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年第2期。
- [13] 李子奈:《我国计量经济学发展的三个阶段与现阶段的三项任务》[N],《企业家日报》2017年12月22日,第W3版。
- [14] 孙秋柏、于思敏、李华:《影响国有控股工业企业利润总额的因素及其实证分析》[J],《经济问题探索》2014年第10期。
- [15] 叶宗裕、王卫杰:《基于蒙特卡罗模拟的误差序列自相关检验研究》[J],《统计与信息论坛》2019年第9期。

## Research on the Test Power of Autocorrelation Test Methods

Wang Zengmin Li Tong

(School of Economics and Management, Beijing University of Posts and Telecommunications)

**Research Objectives:** To explore the power and asymptotic behavior of DW and LM test.

**Research Methods:** Monte Carlo simulation experiment is used to compare and analyze the influencing factors including sample size, the randomness and autocorrelation of explanatory variables, the autocorrelation degree and the distribution pattern of random error terms.

**Research Findings:** The power of DW and LM test is positively correlated with the autocorrelation degree of random error term and sample size, and negatively correlated with the autocorrelation degree of explanatory variables. The randomness of explanatory variables has no significant influence on the power of DW and LM test. The influence of the variation of error term distribution on the efficiency of DW and LM test can be ignored. When the error term has first-order autocorrelation, the power of DW test is better than that of LM test.

**Research Innovations:** Starting from the hypothesis conditions of DW and LM test, the test power of autocorrelation test methods under different conditions is explored and compared.

**Research Value:** Under the background of empirical research, it can provide reference for the effective selection of autocorrelation test methods.

**Key Words:** DW Test; LM Test; Test Power; Monte Carlo Simulation

**JEL Classification:** C15

(责任编辑:白延涛)