



荟萃回归分析在经济学中的应用

Meta-Regression Analysis in Economics

张 莉

上海财经大学 公共经济与管理学院

2025年11月21日

内容提要



- Meta分析的发展
- Meta分析在经济学中的应用
- Meta分析在经济学中应用的分析范式
 - 理论分析
 - 数据搜集及处理
 - 异质性检验
 - 发表偏倚识别
 - 综合分析
- Meta分析示例



内容提要



- ▣ Meta分析的发展
- ▣ Meta分析在经济学中的应用
- ▣ Meta分析在经济学中应用的分析范式
 - 理论分析
 - 数据搜集及处理
 - 异质性检验
 - 发表偏倚识别
 - 综合分析
- ▣ Meta分析示例



1. Meta分析的发展



- ▣ 20世纪初，心理学、医学等领域的实验研究出现了结果不一致、甚至相矛盾的现象
- ▣ 基于实验结果的数量型文献综述方法
 - Simpson & Pearson (1904)通过计算独立实验的相关系数的平均值，对多个实验结果进行合成
 - Fisher (1932)提出了联合概率检验指标，能够将单个研究的P统计量联合起来，首次实现了对多个研究的联合显著性检验



1. Meta分析的发展



- 美国著名的心理学家、医学家、教育家Gene V. Glass (1976)
 - 首次提出Meta-Analysis的概念：以综合已有研究结果为目的，对大量单个研究结果进行综合的统计学分析方法
 - 将实证分析区分为三个层次：（1）原始数据的初步分析；（2）原始数据的再分析；（3）Meta分析
- Meta-这个前缀是“...之后”的意思，Meta-Analysis指的就是对他人分析结果的再分析
- 国内学者对Meta-Analysis的翻译多样，如：分析的分析、再分析、元分析、荟萃分析、整合分析，但很多学者倾向于直接将其称为Meta分析



1. Meta分析的发展



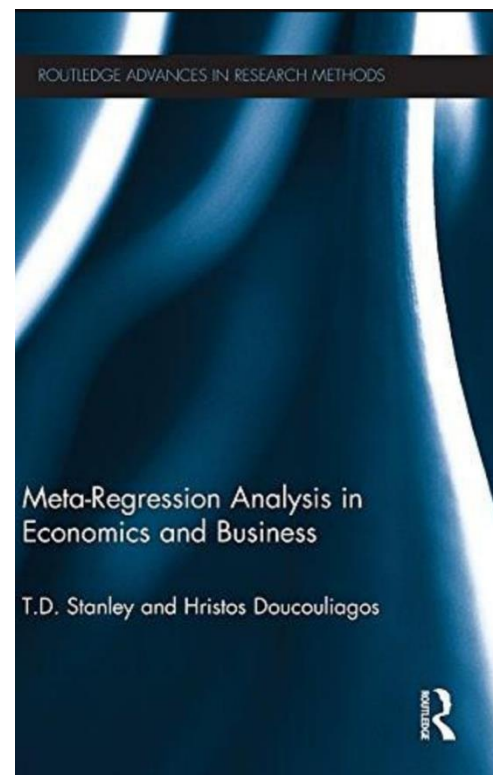
- 自然科学 → 人文社会科学
- Farly, J. U. (1982), “Patterns in parameters of buyers behavior models: Generalizing from sparse replications”, *Marketing Science* 1(2): 181 – 204.
- Assmus, G., J. U. Farley & D. R. Lehmann (1984), “How advertising affects sales: Meta-analysis of econometric results”, *Journal of Marketing Research* 21: 65 – 74.

社会科学：研究结果会受到经济、政治、文化等环境因素的影响。消费者的个体差异、研究的产品类型差异、数据的年代差异、研究的社会、政治环境差异等都会造成研究结果的不一致性。



1. Meta分析的发展

- 随着计量经济学的发展和进步，经济学领域的实证研究不断增多
- 不同实证研究在变量设定、样本获取、数据搜集、实证模型选取等方面处理方式的不同，导致实证研究结果存在不一致性，难以得到理论解释上的认同和具体问题的解决
- 1989年Stanley和Jarrell率先将Meta分析引入经济学领域，并发展出专门分析计量研究结果的Meta回归分析方法（Meta-Regression Analysis）
- 此后，Meta分析方法被应用到经济学领域的研究中，为经济学研究中解决实证结果不一致的难题提供了方法途径



1. Meta分析的发展



□ 传统文献综述：一种定性的叙述性方法

- 存在很强的主观性，难以避免“评论偏见”（Reviewer-biases）
- 对于采用统计计量方法的文献，传统综述不能对差异性计量结果予以量化评价和分析，无法回答不同实证结果为何存在差异

□ Meta分析：一种定量的更加客观的文献综述方法

- 采用定量方法对同一主题的不同实证研究的数据结果和研究结论进行综合性的统计计量分析，以解决各个独立研究结论不一致的问题
- **Meta回归分析**一方面能分析不同研究结论存在差异的原因，另一方面能识别出单独的实证研究所不能发现的变量关系和调节变量，为未来实证研究提供新的思路



内容提要



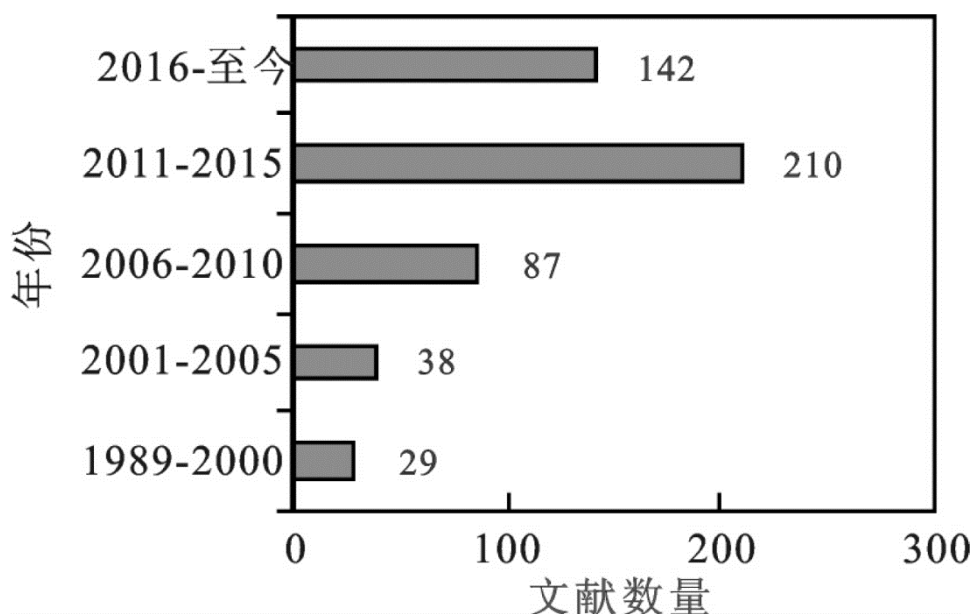
- Meta分析的发展
- Meta分析在经济学中的应用
- Meta分析在经济学中应用的分析范式
 - 理论分析
 - 数据搜集及处理
 - 异质性检验
 - 发表偏倚识别
 - 综合分析
- Meta分析示例



2. Meta分析在经济学中的应用



- 劳动经济学（最低工资问题、薪酬的性别歧视问题）、发展经济学、国际经济学、能源与环境经济学
- Elsevier, CNKI检索Meta相关的经济学研究：



内容提要



- ▣ Meta分析的发展
- ▣ Meta分析在经济学中的应用
- ▣ Meta分析在经济学中应用的分析范式
 - 理论分析
 - 数据搜集及处理
 - 异质性检验
 - 发表偏倚识别
 - 综合分析
- ▣ Meta分析示例

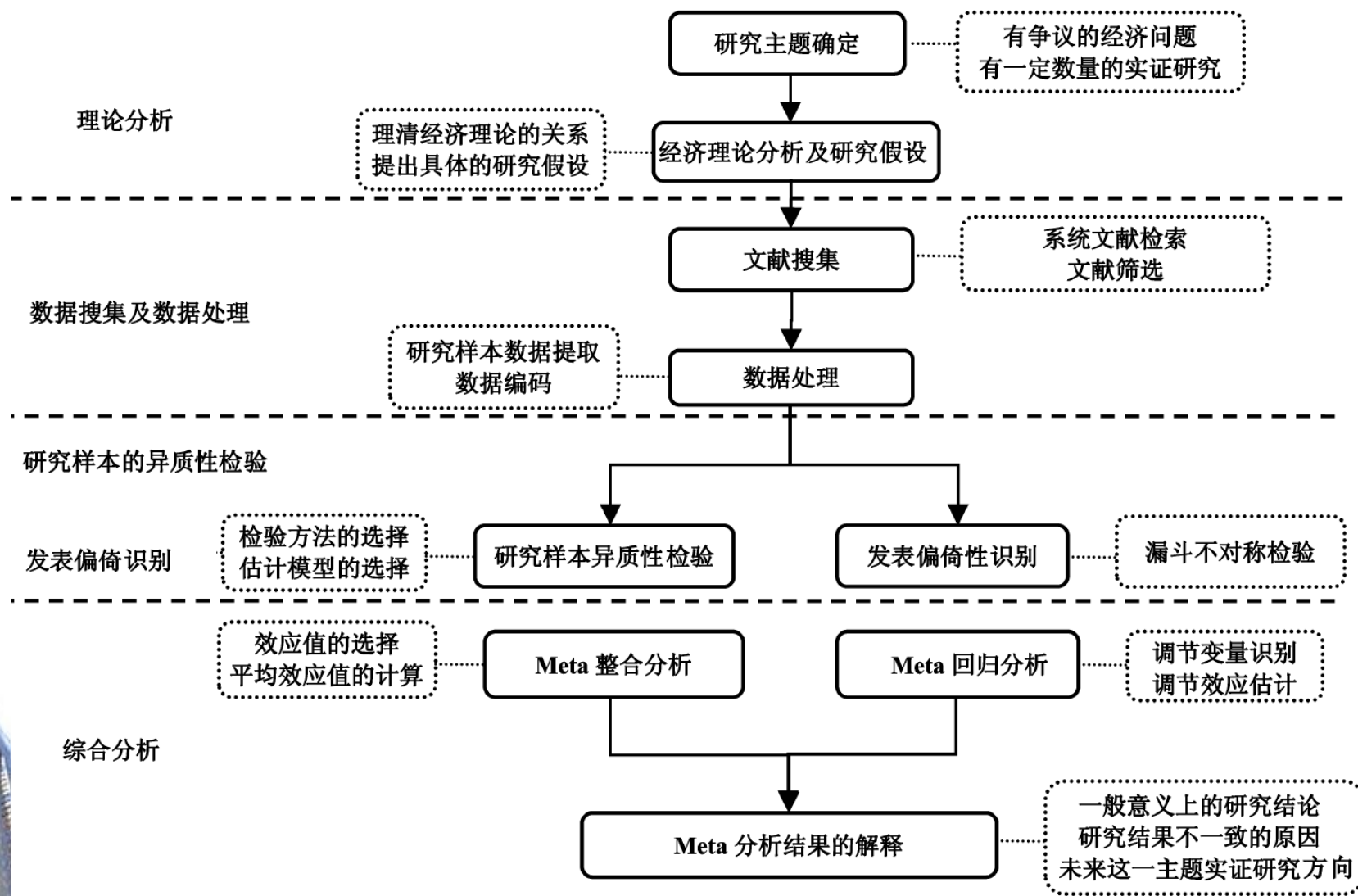


3. Meta分析在经济学中应用的分析范式

- Card, David; Krueger, Alan B. Time-series minimum-wage studies: A meta-analysis [J]. *The American Economic Review*, 1995, 85(2):238-243.
- Cao, Zhi; Lumineau, Fabrice. Revisiting the interplay between contractual and relational governance: A qualitative and meta-analytic investigation [J]. *Journal of Operations Management*, 2015, 33-34:15-42.
- 谢贞发, 张玮. 中国财政分权与经济增长——一个荟萃回归分析 [J]. *经济学(季刊)*, 2015,14(2).
- 陈立敏, 王小瑕. 国际化与绩效提升: 基于Meta整合与Meta回归两种方法的研究 [J]. *世界经济*, 2017,40(2).

粗读这些文献,
总结Meta分析的
基础研究范式

3. Meta分析在经济学中应用的分析范式



内容提要



- ▣ Meta分析的发展
- ▣ Meta分析在经济学中的应用
- ▣ Meta分析在经济学中应用的分析范式
 - 理论分析
 - 数据搜集及处理
 - 异质性检验
 - 发表偏倚识别
 - 综合分析
- ▣ Meta分析示例



3.1. 理论分析



- 具有争议性的经济理论问题作为研究主题
- 有大量针对这一经济理论问题的实证研究，同时这些实证研究的效应值可获取且具有可比性
- 效应值：文献中所报告的统计数据
 - 相关系数 (r)
 - 偏相关系数 (pcc)
 - 回归系数的标准误 (SE) 和自由度 (df)
 - 回归系数的值和自由度



3.1. 理论分析



- 根据特定经济主体的相关研究进行定性文献综述，总结出经济变量关系估计的计量模型
- $Y_i = \alpha + \beta X_i + \gamma C_i + \mu_i$
 - i 是研究个体， Y_i 是因变量， X_i 是自变量， C_i 是调节变量
- Meta回归分析关注：
 - X 对 Y 的影响：两经济变量是否相关，是正相关/负相关
 - C 对所研究的经济变量关系（ X 对 Y 的影响）的影响



内容提要



- Meta分析的发展
- Meta分析在经济学中的应用
- Meta分析在经济学中应用的分析范式
 - 理论分析
 - 数据搜集及处理
 - 异质性检验
 - 发表偏倚识别
 - 综合分析
- Meta分析示例



3.2. 数据搜集与数据处理



- 检索与所研究的经济主题高度相关的实证研究文献，保证样本文献的全面性、完整性、来源多样性
- 文献检索方法
 - 数据库检索：先确定所要检索的数据库，在根据研究主题制定检索策略在选定的数据库检索。（编写检索式：逻辑算符AND/OR/NOT；截词符*/?）
 - 滚雪球法检索：一是检索上述数据库所链接的相似文献，二是追索已有文献的参考文献中的相关文献，三是追索已有相关文献综述中提及的文献
 - 与相关作者联系：补充其提供的相关文献
- 所检索到的文献应包括已经发表的期刊文章、研究报告、会议论文，还应包括未发表的灰色文献



3.2. 数据搜集与数据处理



□ 文献筛选原则

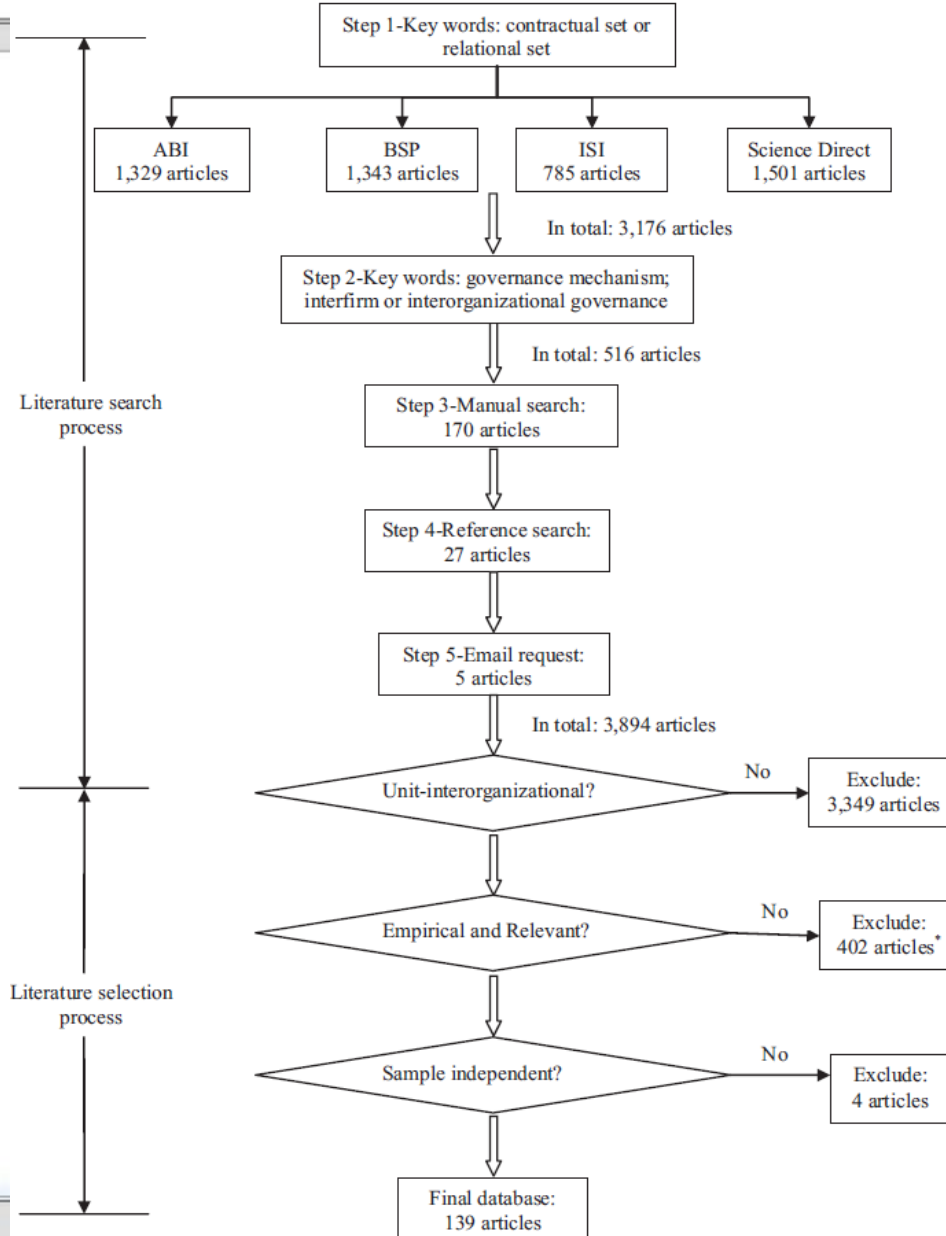
- 应是针对同一经济理论开展的定量实证研究，理论研究和综述研究等文献应予以剔除
- 文献所研究的经济关系相同，研究主题一致
- 可以直接或间接获取Meta分析所需的效应值
- 剔除重复研究：几篇文献采用相同的样本数据、相同研究方法
- 根据研究主题的实际情况选择研究的时间和空间跨度范围

□ 权衡：“宽标准-多数量” “严标准-高质量”

□ 纳入Meta分析的研究不得少于30个



3.2. 数据搜集与数据处理



3.2. 数据搜集与数据处理



▣ 提取研究描述数据

- 研究基本信息：作者、发表时间、发表期刊等文献来源信息
- 研究变量：经济关系变量的测度和代理变量的选用
- 研究样本：研究所选的样本量及样本特征、研究期间等信息
- 研究方法：文献的研究设计、所采用的数据类型、数据特点等信息

▣ 提取效应值数据

- 相关系数 r 、回归系数 β 、标准误 SE 、 t 值、 p 值、方差等
- 对于缺乏所需效应值原始数据的文献，可通过间接计算或者和作者联系获取所需数据
- 无法从上述途径获得效应值数据的文献，不能纳入到Meta分析中

▣ 不同文献的表述通常不一致，需要统一编码(coding)



3.2. 数据搜集与数据处理



Variables	Description	Mean	SD	Min	Max
Premium	Coefficient estimate for the energy label (the effect size or premium)	0.08	0.10	-0.1	0.4
Categorical Scale	= 1 if the measuring scale for EEC is categorical, = 0 for continuous	0.41	0.49	0.0	1.0
Voluntary Sales Market	= 1 if the EEC is voluntary in a country, = 0 for mandatory = 1 if the effect size is with respect to sale price, = 0 for rental value	0.67	0.47	0.0	1.0
Age of Cert Scheme	The age of the EEC since introduced in a country (= "study year" - "year introduced")	0.61	0.49	0.0	1.0
Residential Sector	The age of the EEC since introduced in a country (= "study year" - "year introduced") = 1 if the ECC effect is with respect to the residential sector, = 0 for commercial sector	15.28	5.32	5.0	24.0
Compares to non-labelled	= 1 if effect size measures premium over non labelled houses, = 0 otherwise	0.47	0.50	0.0	1.0
Published	= 1 if effect size is from published study, = 0 unpublished	0.72	0.45	0.0	1.0
No of Authors	The number of authors involved conducting the study	0.49	0.50	0.0	1.0
Year of Study	The year of study	2.55	0.69	1.0	4.0
Method of Estimation	= 1 if the study uses OLS, = 0 otherwise	2011	1.75	2008	2014
Crossectional Data	= 1 if study uses crossectional data, = 0 panel data	0.65	0.48	0.0	1.0
Sampsize	The sample size from which effect size was estimated	0.88	0.32	0.0	1.0
Asia	Singapore, Japan	41856	86399	83	323840
Europe	Austria, Belgium, Netherlands, Sweden and UK	0.20	0.40	0.0	1.0
Australia	Australia	0.23	0.42	0.0	1.0
N_America	U.S.	0.03	0.17	0.0	1.0
		0.54	0.50	0.0	1.0



内容提要



- ▣ Meta分析的发展
- ▣ Meta分析在经济学中的应用
- ▣ Meta分析在经济学中应用的分析范式
 - 理论分析
 - 数据搜集及处理
 - 异质性检验
 - 发表偏倚识别
 - 综合分析
- ▣ Meta分析示例



3.3. 异质性检验



- Meta整合分析旨在求不同研究的合并效应
- 由于纳入Meta分析的文献在研究地、研究时间、研究者、研究对象等方面不尽相同，不同研究的异质性（heterogeneity）的存在使各研究中变量关系效应的估计也不同
- 若不考虑研究间异质性，直接将不同研究进行整合分析，则会存在将“橘子”和“苹果”进行合并的错误做法
- Q检验、 H 检验、 I^2 检验



3.3. 异质性检验



□ Q检验

$$Q = \sum_{i=1}^k w_i (ES_i - \overline{ES})^2 \sim \chi^2(df)$$

- E_i (Effect Size) 为第*i*个研究的真实效应量 θ_i 的估计量
- \overline{ES} 为所有研究真实效应估计量的均值
- $w_i = 1/SE^2$ 为*i*个研究对应方差的倒数，以此作为研究的权重
- $df = k-1$ ， k 为研究总数

□ 零假设 H_0 ：各独立研究不存在异质性 $\theta_1 = \theta_2 = \dots = \theta_k = \theta$

□ 备择假设 H_1 ：各独立研究存在异质性 $\theta_1 \neq \theta_2 \neq \dots \neq \theta_k \neq \theta$

□ Q值大于临界值时，拒绝 H_0 ，即各独立研究存在异质性



3.3. 异质性检验



- Q统计值会随着研究文献数目的变化而变化，检验结果效能不高，会影响检验结果的可靠性
- I^2 检验为研究间变异占总变异的比例

$$I^2 = \begin{cases} \frac{Q - df}{Q}, & \text{若 } Q > df \\ 0, & \text{若 } Q < df \end{cases}$$

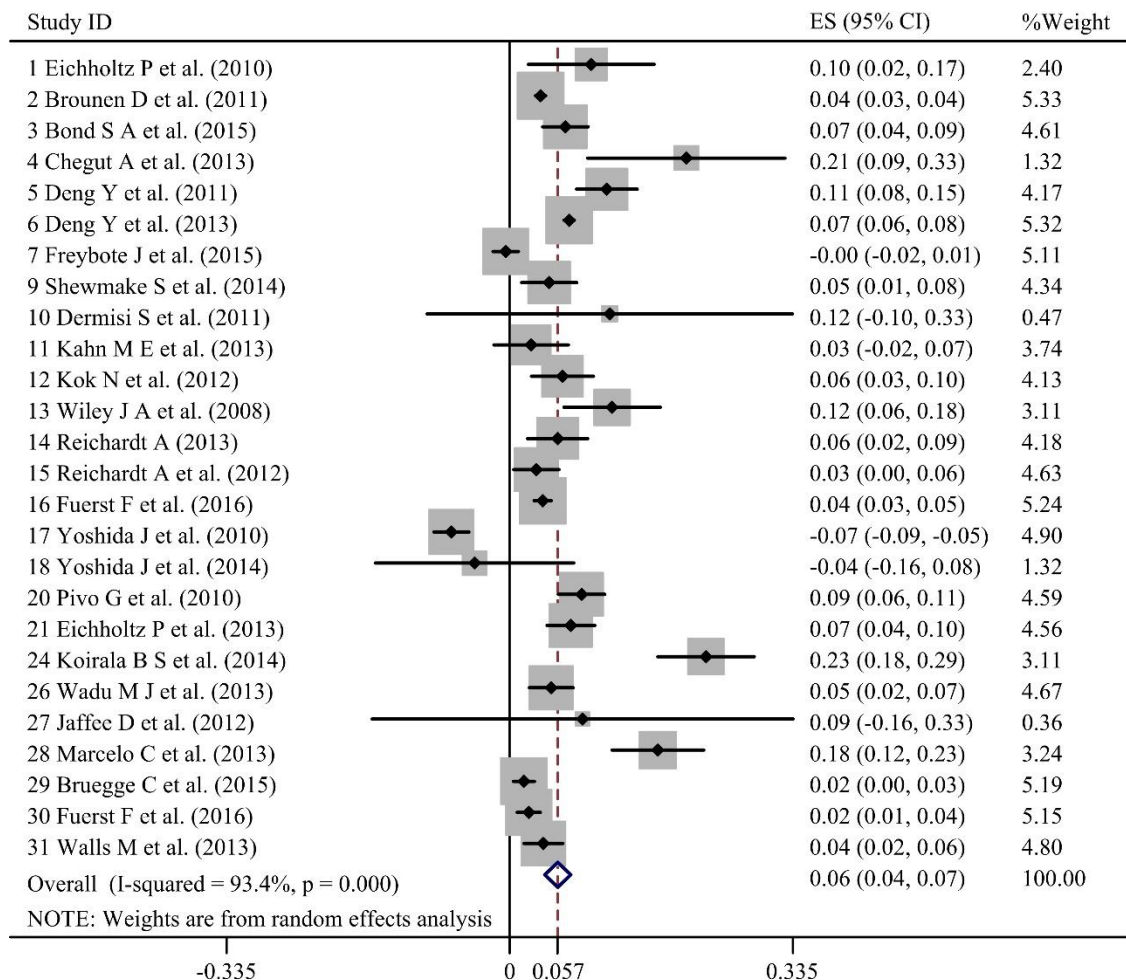
- H检验: $H=Q/df$
- 若检验结果表明各独立研究存在异质性
 - 通过亚组分析(subgroup analysis)、Meta回归分析识别异质性来源
 - 敏感性分析: 排除可能导致异质性的某个(些)研究后重做Meta分析



3.3. 异质性检验

□ 森林图

- 图形显示了每个独立研究的效应量、合并效应量、95%置信区间、各效应量加权值（方差倒数）、 I^2 统计量
- 置信区间横跨无效竖线，该估计值无统计意义
- 置信区间重叠较少，存在异质性



内容提要



- ▣ Meta分析的发展
- ▣ Meta分析在经济学中的应用
- ▣ Meta分析在经济学中应用的分析范式
 - 理论分析
 - 数据搜集及处理
 - 异质性检验
 - 发表偏倚识别
 - 综合分析
- ▣ Meta分析示例



3.4. 发表偏倚识别



- ▣ 同类研究中具有统计学意义的研究较之无统计学意义的研究更易于被接受和发表，使得Meta分析会夸大变量关系效应
- ▣ 发表偏倚（Publication Bias）源于研究人员、编辑、资金资助者等倾向于发表和报告特定类型的估计结果
 - 撰写者：论文撰写中存在主观性，会根据估计结果显著性来选择所报告的估计模型，可能会高估研究结果的显著性
 - 评审者：当评审一篇论文能否发表时，关注的是论文中解释变量对被解释变量影响的大小、方向以及显著性，具有显著性的实证研究更受论文评审者的青睐
- ▣ Meta分析的文献检索和筛选中再科学的检索策略和再严格的筛选标准，也无法避免发表偏倚问题，因此需要检验并减少影响



3.4. 发表偏倚识别



□ 检查偏倚的途径

- 先根据一个基本的入选标准收集全部的研究
- 再考虑不同的入选标准进行彻底的敏感性分析

□ 检查偏倚的程度

- 漏斗图分析 (funnel plots, funnel asymmetry tests)
- 失安全数 (fail-safe number, N_{fs})
- 森林图法
- 剪补法



3.4. 发表偏倚识别



□ 漏斗图

- 相对于样本量的效应值，是以研究的效应估计值作为横坐标，样本量作为纵坐标画出的散点图

□ 漏斗图分析

- 根据图形的不对称程度判断Meta分析中偏倚有无
- 基于效应值的精确度随样本量增大而增加这一事实

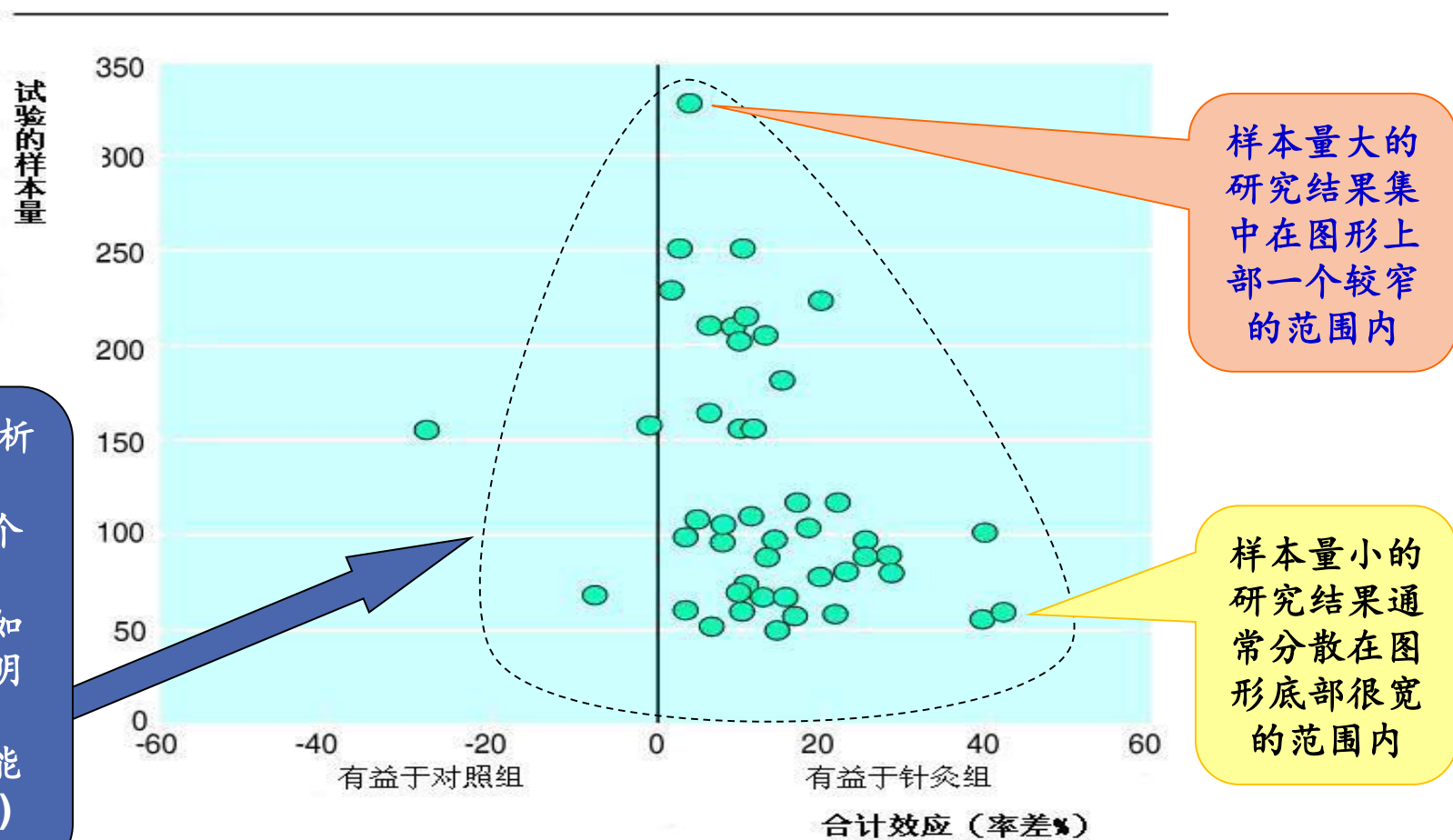
□ 漏斗图分析原理

- 对处理效应的估计，其准确性随着样本量的增加而增加
- 小样本研究的效应估计值分布于图的底部，分布范围较宽
- 大样本研究的效应估计值分布于图的顶部，分布范围较窄
- 没有发表偏倚时，图形呈对称的倒漏斗状



3.4. 发表偏倚识别

■ 针灸治疗中风的49个试验的漏斗图分析 (Tang TL, 1999)



3.4. 发表偏倚识别



□ 漏斗不对称检验是基于研究所报告标准差对效应量影响的估计

$$pcc_{ij} = \beta_0 + \beta_1 SE_{ij} + \mu_{ij}, \quad i=1, 2, \dots, N; j=1, 2, \dots, S$$

- pcc_{ij} 为第*i*个研究中第*j*个估计的效应量， SE_{ij} 为对应的标准差， N 为研究的总量， S 为所有研究的估计个数的最大值， μ_{ij} 为残差项
- 由于 μ_{ij} 的方差在各个估计之间是显著不同的，因此不满足独立同分布条件，不能直接用OLS进行估计，因此转换为：

$$t_{ij} = \beta_0/SE_{ij} + \beta_1 + v_{ij}, \quad i=1, 2, \dots, N; j=1, 2, \dots, S$$

- t_{ij} 为第*i*个研究中第*j*个估计系数的*t*统计量， $t_{ij} = pcc_{ij}/SE_{ij}$
- 原假设 H_0 : β_1 显著为0
- 备择假设 H_1 : β_1 显著不为0
- 当拒绝原假设时，存在发表偏倚问题， β_1 的正负表示偏倚方向



3.4. 发表偏倚识别

❑ 失安全数

- 定义: Meta分析中计算需多少阴性研究结果的报告才能使结论逆转
- 用途: 估计发表偏移的程度
- 意义: 失安全数大, 说明Meta分析的结果越稳定, 结论被推翻的可能性越小

■ P为0.05和0.01时计算公式如下:

$$N_{fs0.05} = (\sum Z / 1.64)^2 - S \longrightarrow \text{研究个数}$$

$$N_{fs0.01} = (\sum \underbrace{Z/2.33}_{\text{各个独立研究的效应值}})^2 - S$$

内容提要



- ▣ Meta分析的发展
- ▣ Meta分析在经济学中的应用
- ▣ Meta分析在经济学中应用的分析范式
 - 理论分析
 - 数据搜集及处理
 - 异质性检验
 - 发表偏倚识别
 - 综合分析
- ▣ Meta分析示例



3.5. 综合分析



- 一般而言，一篇经济学实证研究文献包含着多个回归估计，如果选择偏相关系数作为效应量进行分析的话，Meta回归分析的估计方程为：

$$pcc_{ij} = \alpha + \sum_{k=1}^n \beta^k Z_{ij}^k + \mu_j + \epsilon_{ij}$$

- 用两个下标*i*和*j*的好处在于区分文献层次和估计层次，以便进行多层回归分析（Multilevel Analysis）
- 考虑到异方差问题，进行如下转换：

$$t_{ij} = \gamma + \sum_{k=1}^n \frac{\beta^k Z_{ij}^k}{SE_{ij}} + u_j + e_{ij}$$



3.5. 综合分析

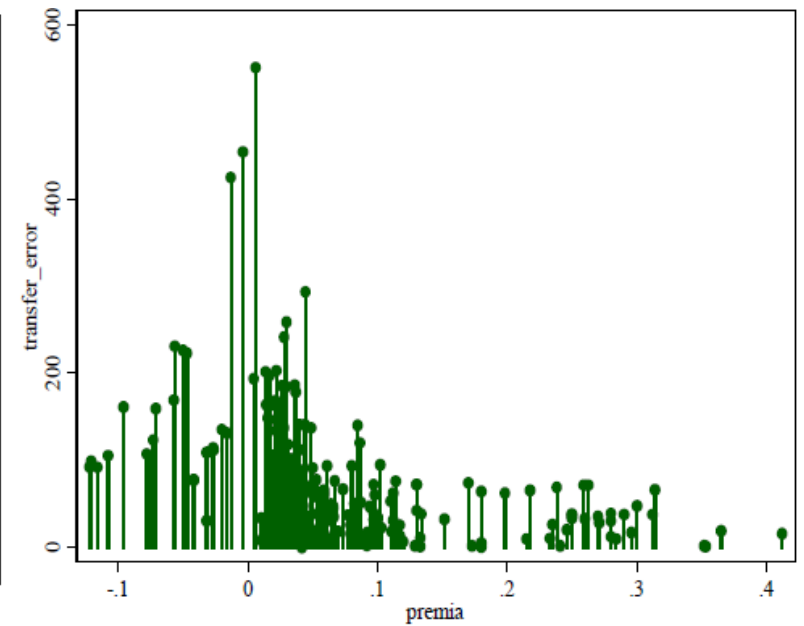
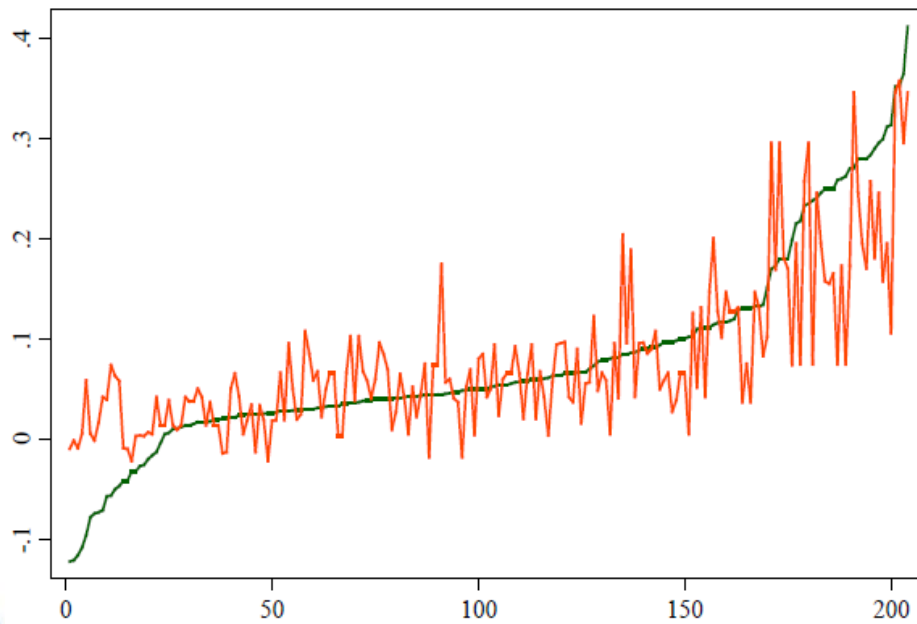


- 根据Meta整合分析结果，就已有的实证研究中存在不确定、模糊的、甚至有争议的研究结论给出更为合理且具有一般意义的结论
- 根据发表偏倚识别的分析结果分析影响某一经济理论的实证研究结论的调节因素，并分析这些调节因素的调节效应大小
- 指出这一经济主题的实证研究存在的不足，为未来该主题实证研究提供方向上的参考



3.5. 综合分析

▣ Benefit Transfer Analysis



内容提要



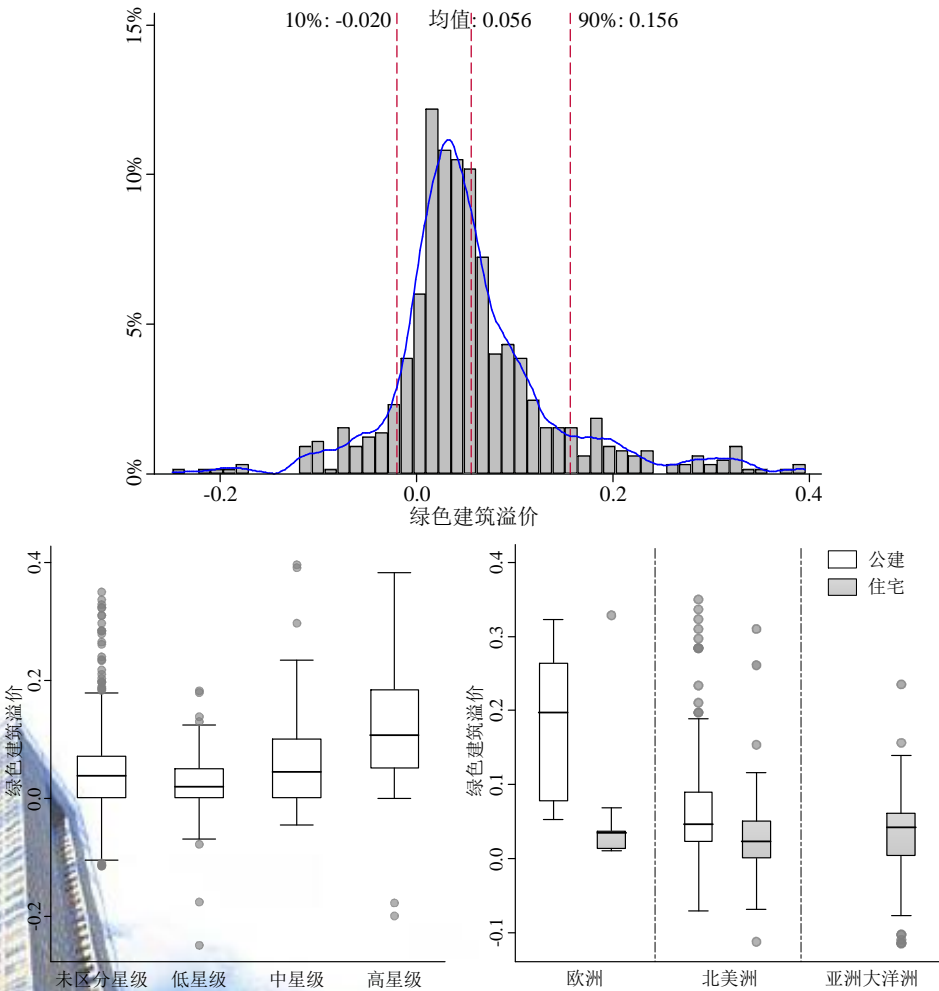
- ▣ Meta分析的发展
- ▣ Meta分析在经济学中的应用
- ▣ Meta分析在经济学中应用的分析范式
 - 理论分析
 - 数据搜集及处理
 - 异质性检验
 - 发表偏倚识别
 - 综合分析
- ▣ Meta分析示例



绿色建筑溢价的Meta分析



显示性偏好法 (38篇论文)



荟萃分析 (Meta-Regression Analysis)

控制变量

住宅/公建、销售/租赁、交易/挂牌、样本选取 (子市场/一定距离范围/倾向得分匹配)、模型选择、样本量、是否横截面、是否发表、作者数量

认证特征

被解释变量: *PREMIUM*

<i>DUAL_CERT</i>	0.0708***		
<i>CERT_AGE</i>		0.00162	0.00169
<i>ENERGY</i>		-0.0355**	-0.0426**
<i>GOV</i>			0.00688
<i>MANDATE</i>			0.0563*
论文随机效应	是	是	是

地区特征

被解释变量: *PREMIUM*

<i>GDPPC</i>	0.0381**	0.261***	0.214**
<i>GDPPC</i> ²	-0.000452***	-0.00306***	-0.00253**
<i>EDU</i>	0.00287	0.0798***	0.0866***
<i>EPRICE</i>		0.0152**	0.0146*
<i>MODERATE</i>		-0.0522	-0.0950**
认证特征	否	否	是
论文随机效应	是	是	是

本讲小结



- 经济学研究通常基于特定分析范式的重复性研究来寻求经济规律。由于每个实证研究实施条件不可能完全相同，经济学领域的重复研究是不可能的，单个研究所得出的结论不具有一般性
- 当某一课题的实证研究已经较多，并且研究结果存在不一致问题时，我们就需要运用Meta分析来研究这些文献
 - Meta分析是一种基于已有研究进行系统研究的方法，能充分利用已有的实证研究结果，将独立的研究结果进行综合分析，探究出一个具有一般性意义的结论
 - 同时还能探究实证研究差异性结果产生的原因，为进一步的实证研究提供方向性指导
- 既不能夸大Meta分析的作用，也不能否定Meta分析的用途
 - 对任何统计分析的结果，都需要结合经济学理论，对研究结果做出尽可能客观和真实的解释

