

对研究方法的一些体悟

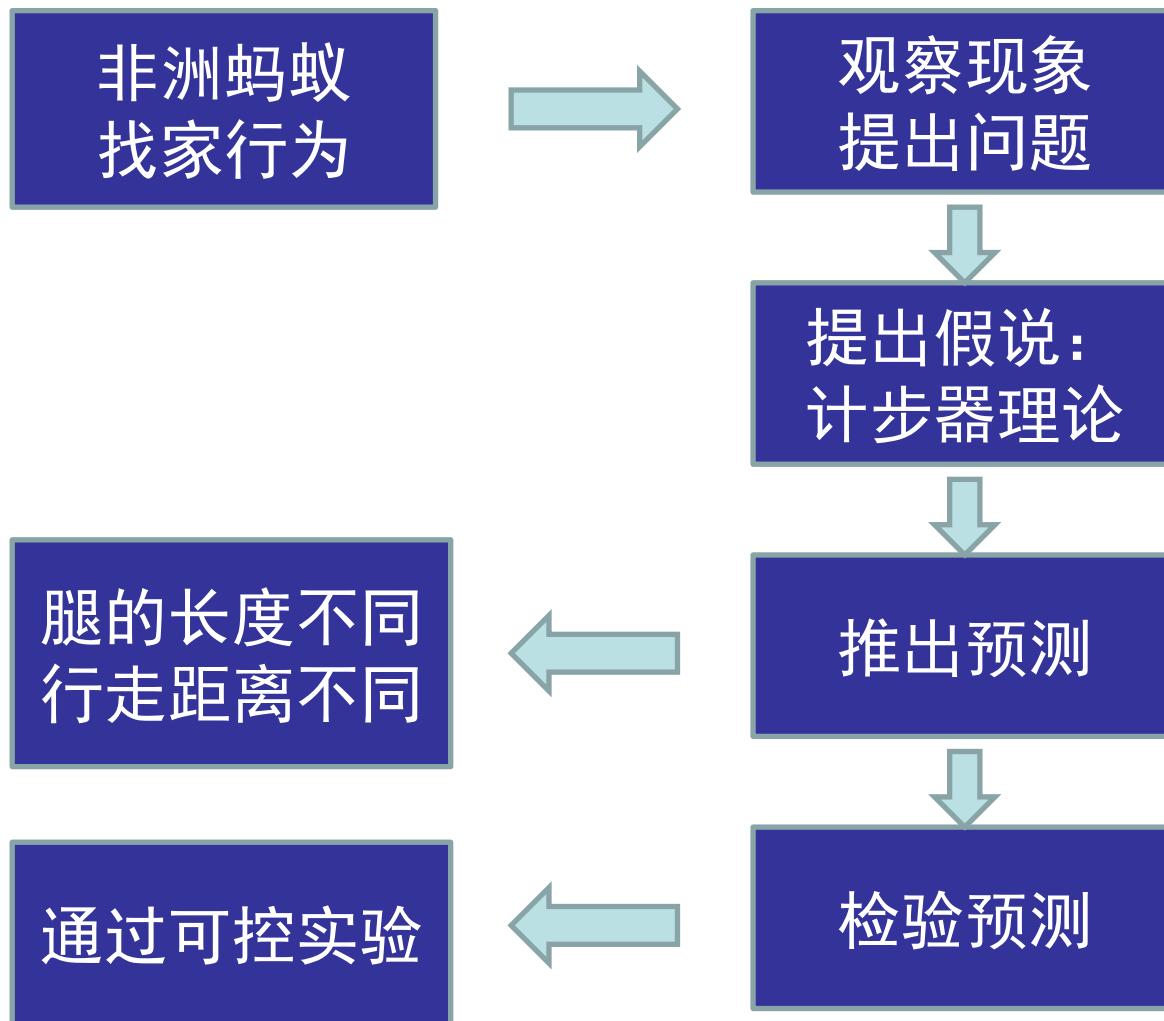
安信证券宏观研究



非洲蚂蚁找家

- 非洲蚂蚁不依靠气味回家
- 科学家提出“计步器”理论
- 利用改变腿长短的实验来验证理论

非洲蚂蚁找家



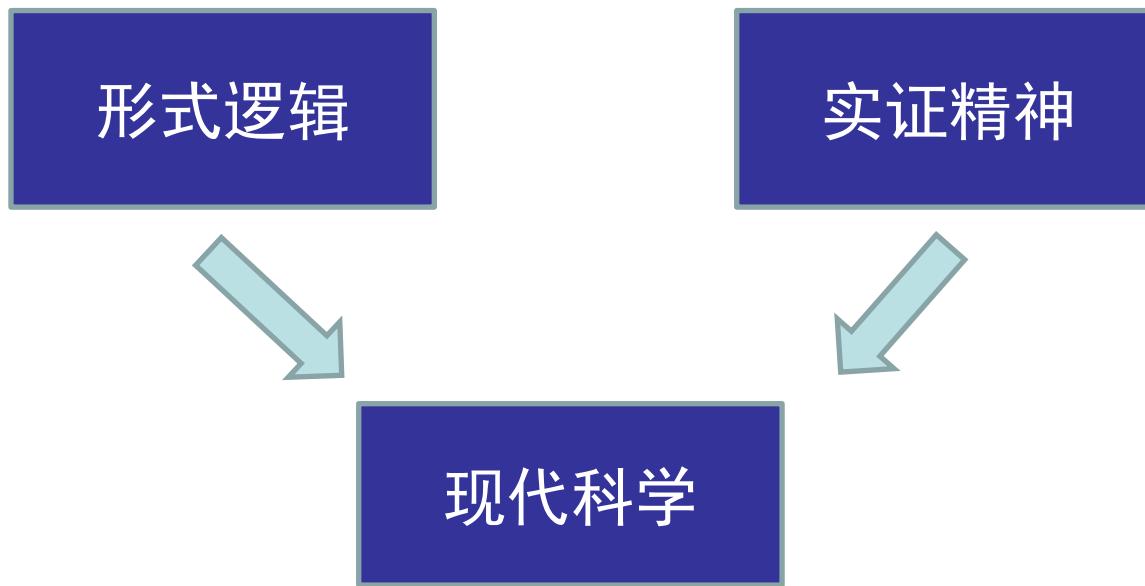
科学研究的4个步骤

1. 观察现象，提出问题
2. 提出假说，解释现象
3. 从理论的内在逻辑出发，提出预测
4. 系统性的搜集经验事实，验证理论预测

重的物体下落更快

- 古希腊人：发达的形式逻辑—欧式几何
- 亚里士多德：重的物体下落更快
- 伽利略： 可控实验检验

形式逻辑+实证精神=现代科学



光线是可以弯曲的

- 爱因斯坦的广义相对论：思想和形式异常复杂
 - 做出的预测：简单且可验证
-
- 1918年，英国科学家通过观察日食验证光线弯曲

民可使由之，不可使然之

- 历史上，中国人缺乏形式逻辑和实证精神
- 大量史料事无巨细记载事件
- 知识分子醉心于解读古代典籍
- 中医理论

下边，没有了

- 中国人缺乏形式逻辑和实证精神延续至今
- 找不到密实的逻辑和基于逻辑提出的预测
- 看不到通过对证据系统性地梳理来验证预测
- 看到：
 - (1) 材料的堆积
 - (2) 孔子似的灵感
 - (3) 中医一样突然的论断

科学研究的步骤不可或缺

- 所有学科都遵循的方法
- 霍金：“这注定是我们的知识能够取得大的进步的方法。除了这样的方法，我们不知道还有其他的方法可以帮助我们进步。”

研究本质上是讨论因果关系

因果关系：一项输入变化带来的输出的变化

保持其他不变
改变一项

变化



研究因果关系需要隔离

- 万有引力：只与质量和温度相关，其他无关
- 大量物理学研究原则上是可以隔离的和控制
- 这使得相关的研究变得非常容易
- 经济学和生物学隔离和可控实验难以做到
- 一般采用统计学的方法

因果关系与相关关系

雄鸡一唱天下白

鸡叫稳定地领先于天亮

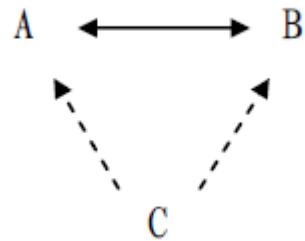
但鸡叫并不是天亮的原因

类似的结论有很多：

哈佛大学中途退学的收入水平高于完成学业

中国大学高消费的学生收入较高

因果关系与相关关系



A与B之间存在相关关系，可能存在三种情况：

- 1. A cause B, A是B的原因
- 2. B cause A, B是A的原因
- 3. C同时cause了A和B

C既是A的原因，也是B的原因

红酒和心脏病

适量饮用葡萄酒的人，心血管发病率较低

得出：适度饮红酒，软化血管，降低心脏病发病率

因果关系：喝红酒=>降低心脏病

其他C同时决定了两者：收入，生活习惯

因果关系还可能是反着的

考虑伴随因素，饮酒的保健作用很低

上帝是存在的

- 科学研究：可证伪

例. 命题 “树叶是绿的”
是可证伪的

简单的原则

两种方法：

- 十种理论解释一百个现象----每个都很完美
- 一种理论解释一百个现象----不那么完美

正确的应该是后一种

简单理论 战胜 复杂理论

“日心说” 战胜 “地心说”

世界是简单的、可以理解的

“上帝难以琢磨，但他并不邪恶。”

“上帝隐藏他的秘密，是因为他本性高傲，而不是凭借什么狡猾的手段。”

——爱因斯坦

方法论思考题

1. 过去几年，农产品价格出现此起彼伏地上涨。农产品价格上涨的原因是什么？如何检验？
2. 去年蔬菜价格的涨幅惊人。有人认为摊位费、租金的上涨，是重要原因。如何在总量数据层面检验这种看法？
3. 去年四季度,大家认为节能减排对经济的影响较大。如何在总量数据层面进一步验证这个看法？

方法论思考题续一

4. 最近两年，消费的增速一直维持在历史上较高的水平。大家认为低端劳动力收入快速地上升是支撑消费的重要原因。那么如何检验这个看法？
5. 目前，我们一直在讨论新一轮经济周期的启动。那么，新经济周期的启动如何确认，有什么广谱的数据可以观察到？

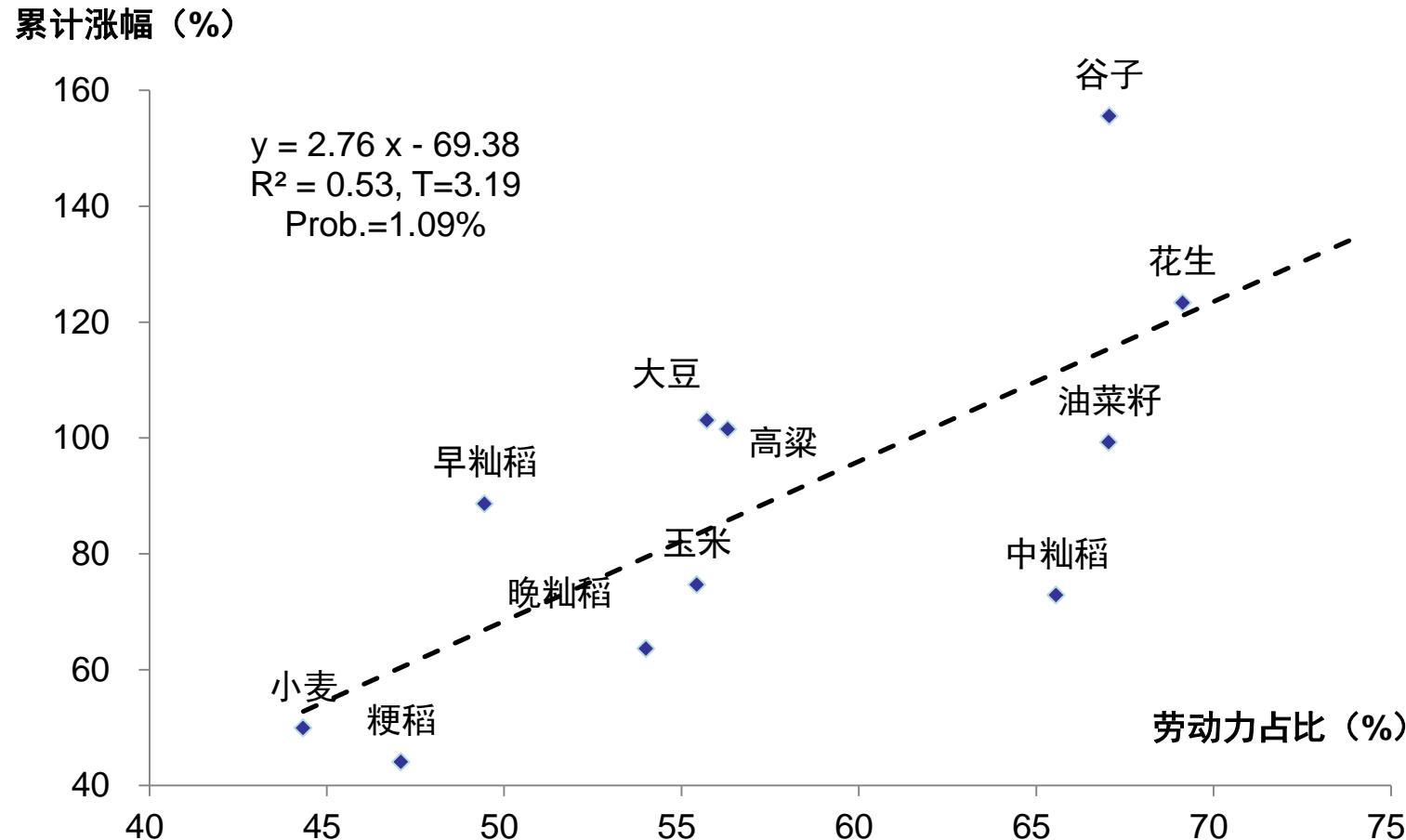
方法论思考题续二

6. 智商、环境和教育都会对个人能力和日后的收入造成影响。那么如何检验教育对个人收入的影响？
7. 经济学中，劳动力价格取决于劳动力的边际产出，而边际产出取决于技术进步和劳动力数量。如何简单有效地说明劳动力数量影响劳动力价格？
8. 人的心理暗示往往会影响人体的反应和主观判断。那么如何能很好地验证一种药物是有效的？

方法论思考题续三

9. 为什么千百年来，对绝大多数文明而言，男人总是比女人更花心？如何验证？
10. 现在人们普遍比几十年，几百年前要胖，而且无论是发达国家还是发展中国家，近几十年，肥胖症的比例有显著地抬升。肥胖的成因是什么？如何验证？
11. 流行文化中，星座一般认为和性格、职业甚至命运高度相关。如何设计多种实验检验这一点？

中国不同劳动力密集农产品的累计涨幅（2000年-2007年）



数据来源：CEIC，全国农产品成本收益统计年鉴，安信证券研究中心

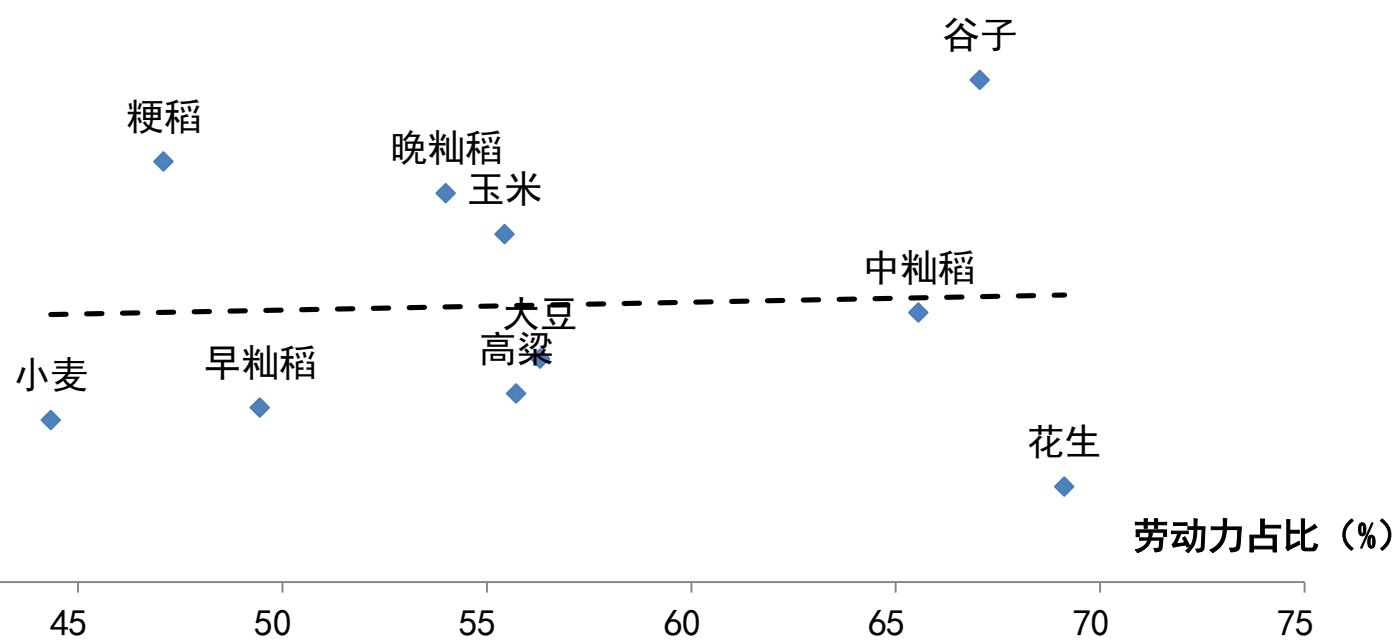
数据说明：劳动力占比的计算方法是在产值中抛去物质费用和土地成本等占比

中国不同劳动力密集农产品的累计涨幅（1990年-2000年）

累计涨幅 (%)

160
140
120
100
80
60
40

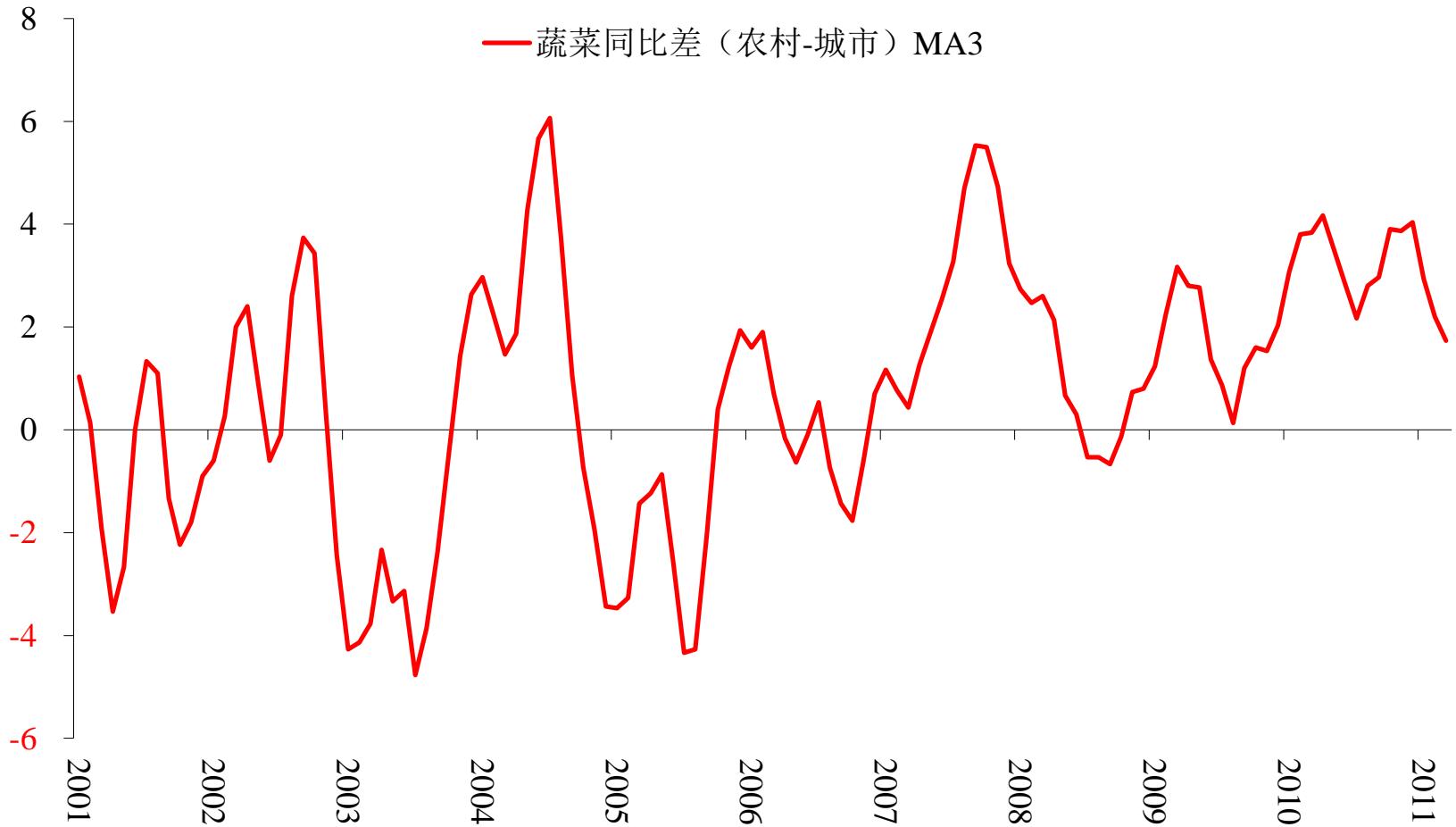
$$y = 0.1255x + 77.003$$
$$R^2 = 0.0026$$



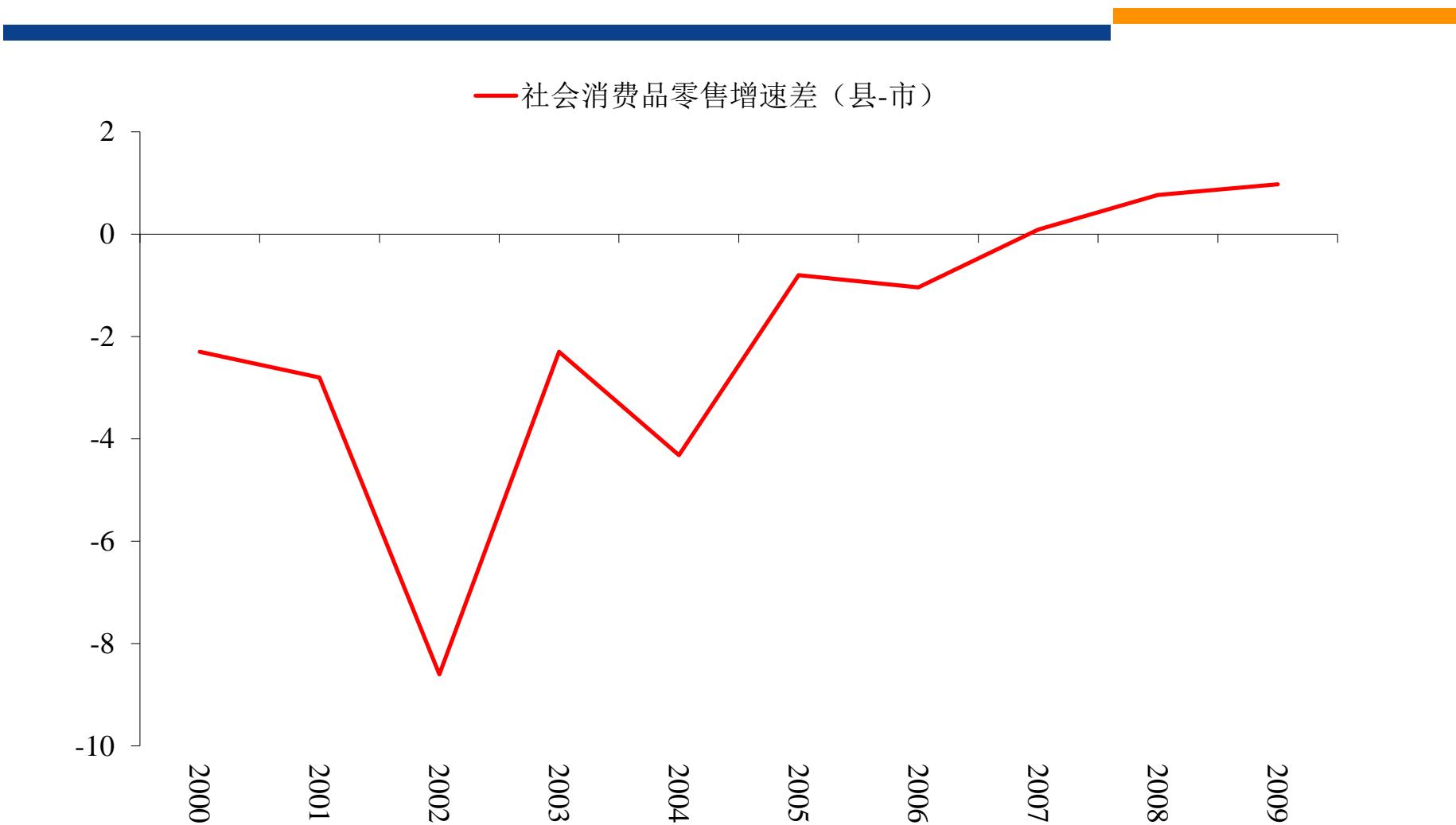
数据来源：CEIC，全国农产品成本收益统计年鉴，安信证券研究中心

数据说明：劳动力占比的计算方法是在产值中抛去物质费用和土地成本等占比

城乡月度CPI蔬菜价格增速差（农村-城市，%）



城乡社会消费品零售增速差（县-城市，%）



思考题总结

实证精神和验证技巧

- 广谱数据
- 横断面数据
- 自然实验
- 精巧选取



致谢！