

友松实验室 | 博士论文评审

论文信息：

- 标题：How School Districts Spent ESSER Funds and Learning Recovery: Evidence from Arizona School Districts
- 作者：Yixin (Crystal) Wang
- 机构：Arizona State University
- 学位：Doctor of Philosophy
- 年份：2025年12月
- 页数：117页

评审人：叶晓阳 (with Ye AI) @ 友松实验室

评审日期：2025年12月29日

总结

本博士论文研究亚利桑那州学区如何使用联邦ESSER（小学和中学紧急救济）资金以及这些支出与学习恢复的关系。基于亚利桑那州142个学区的数据，作者分析了ESSER资金在新项目、技术、设施、个人防护设备和心理健康等类别的支出模式，并探讨了这些支出与学生学习恢复的关联。

论文类型：博士论文

总体评价：本博士论文在学术诚信和方法论透明度方面表现良好。作者明确承认研究的局限性，诚实地表明“本研究不试图声称ESSER支出与学生学习结果之间存在因果关系”，并将研究定位为提供“相关性证据”而非因果推断。论文内容丰富，包含了充分的理论背景、方法论讨论和实证分析。这种诚实的学术态度值得赞赏，论文达到了博士学位的基本要求。

主要发现：

1. 小型和农村学区在几乎所有类别的人均支出都更高，特别是技术和设施方面
2. 控制人均资助总额后，学区间的支出差异基本消失
3. 大多数ESSER支出类别与学习恢复无显著关联
4. 只有技术支出显示出积极且统计显著的效果：人均技术支出每增加1%，学习恢复提高0.26%
5. 这种关联主要集中在低贫困率学区

研究贡献：

1. 地理扩展：首次将亚利桑那州纳入ESSER分析，填补了现有研究的地理空白
2. 类别细分：按支出类别细分分析，提供比总支出分析更精细的洞察
3. 长期数据：使用2016-2024年的长期面板数据，比现有研究的时间跨度更长
4. 方法论透明：明确承认研究局限性，诚实定位为相关性而非因果研究

显著特点：

1. 学术诚信：明确声明不寻求因果效应，避免了过度解释
2. 内容丰富：117页的篇幅包含了充分的理论背景和方法论讨论
3. 数据质量：使用贝叶斯方法处理分类数据，提高了分析的严谨性

4. 政策谨慎：避免了基于相关性的强烈政策建议

详细评审意见

1. 研究设计与方法论

优点：

- **研究问题重要且及时：** 利用COVID-19疫情期间的大规模联邦资金投入作为准自然实验，为教育经济学提供了独特的研究机会
- **数据质量高：** 使用亚利桑那州142个学区的完整数据，涵盖2016-2024年的长期面板数据
- **方法论合理：** 采用分层回归模型和贝叶斯估计方法，处理数据的层次结构
- **分析细致：** 将ESSER支出按类别分解，提供比总支出分析更精细的洞察
- **学术诚信：** 明确承认研究局限性，诚实定位为相关性研究

需要改进的方面：

1.1 方法论的局限性

作者诚实承认了研究的局限性，但仍需认识到相关性研究在回答政策问题时的根本限制：

内生性问题依然存在：

- **反向因果：** 预期学习恢复困难的学区可能更倾向于某些类型的支出
- **遗漏变量偏误：** 学区领导能力、教师质量、家长参与等不可观察因素同时影响支出决策和学习恢复
- **选择偏误：** 不同类型学区的支出偏好可能反映其固有特征而非资金效果

相关性的解释限制：

- 即使是统计显著的相关性也可能完全由混淆因素驱动
- 技术支出的“积极关联”可能反映学区特征而非真实效果
- 缺乏因果机制分析使得结果解释存在多种可能性

1.2 研究设计的改进空间

理论框架可以更强：

- 教育生产函数的应用可以更加精细化
- 对ESSER资金特殊性（时间限制、灵活性）的理论分析可以更深入
- 支出类别与学习恢复的理论联系可以更明确

“学习恢复”指标的验证：

- 构建方法合理，但可考虑与其他恢复指标的比较验证
- 对疫情影响测试条件、学生构成变化的考虑可以更充分
- 敏感性分析可以更全面

2. 数据分析与结果解释

2.1 支出模式分析

优点：

- 发现小型和农村学区支出更高的模式很有价值
- 控制人均资助额后差异消失的发现揭示了资金分配机制的重要性
- 按学区特征分层分析提供了丰富的政策洞察

需要深化的方面：

- 对支出差异背后机制的解释不够深入
- 缺乏对支出效率的分析
- 未充分探讨地理和规模因素如何影响支出选择

2.2 学习恢复关联分析

2.2 学习恢复关联分析

优点：

- 作者明确承认"分析识别关联模式而非因果效应"
- 将发现定位为"相关性证据"，避免了过度解释
- 为未来因果研究提供假设生成的基础

仍需注意的解释问题：

技术支出关联的谨慎解释：

- 0.26%的关联虽然统计显著，但实质意义需要谨慎评估
- 该关联可能反映学区特征、数字基础设施或其他混淆因素
- 在低贫困率学区的集中效应可能反映数字鸿沟而非支出效果本身
- 需要更多机制分析来理解这种关联的来源

其他类别无关联的合理解释：

- 可能反映实施滞后、测量误差或政策设计问题
- 也可能表明这些支出类别确实与短期学习恢复关联较弱
- 不应过度解读为这些投资"无效"

异质性分析的价值与限制：

- 按贫困率和资助水平分组提供了有价值的描述性洞察
- 但分组差异的解释仍受到因果识别限制的约束
- 可以为未来研究提供重要的假设方向

3. 学术贡献的评估

3.1 理论与实证贡献

积极贡献：

- 地理扩展：为亚利桑那州ESSE支出提供了首次系统性分析
- 方法创新：贝叶斯方法处理分类数据，技术上有一定创新
- 描述性价值：支出模式的细致分析为政策制定提供了有价值的描述性证据
- 学术诚信：明确承认局限性，为相关性研究树立了良好范例

局限性：

- **理论贡献有限：**主要是实证描述，理论创新相对较少
- **因果推断缺失：**无法回答“为什么”的问题，只能描述“是什么”
- **政策指导受限：**相关性证据对政策制定的指导作用有限

3.2 文献贡献的定位

合理定位：

- 作为描述性研究和假设生成，论文有其价值
- 为未来因果研究提供了重要的基础数据和模式识别
- 在ESSER研究的早期阶段，描述性证据具有重要意义

外部有效性考虑：

- 亚利桑那州的特殊性（教育资金不足、地理特征）需要在推广时考虑
- 单一州的分析结果推广到其他州时需要谨慎
- 样本代表性分析显示了一定的局限性

4. 政策含义与研究价值

4.1 政策洞察的价值与限制

有价值的政策洞察：

- **资金分配模式：**揭示了小型和农村学区获得更多人均资助的模式
- **支出优先级：**展示了不同类型学区的支出偏好差异
- **技术投资关联：**提供了技术支出可能重要性的初步证据
- **公平性分析：**显示ESSER在一定程度上缓解了资金不平等

政策建议的谨慎性：

- 作者避免了基于相关性的强烈政策建议，这是明智的
- 将发现定位为“为未来研究提供基础”，符合研究性质
- 承认了临时性资金的局限性和“财政悬崖”问题

4.2 对教育政策制定的启示

积极启示：

- 灵活的联邦资金可以在短期内缓解资金不平等
- 最低分配政策对小型和农村学区的重要性
- 不同支出类别可能有不同的效果模式

需要谨慎的方面：

- 技术投资的“积极关联”不应被过度解读为因果效应
- 其他支出类别的“无关联”不意味着这些投资无价值
- 政策制定仍需要更严格的因果证据支持

5. 写作质量与表述

优点：

- 结构清晰，逻辑流畅
- 数据可视化效果良好
- 统计结果报告规范

可以改进：

- 一些技术细节的解释可以更清晰
 - 对研究局限性的讨论可以更全面
 - 结论部分可以更强地连接发现与政策含义
-

具体建议

优先级1：进一步完善研究定位

1. 强化相关性研究的价值

- 更明确地阐述描述性研究在政策制定中的重要作用
- 强调为未来因果研究提供假设生成的价值
- 避免任何可能被误解为因果解释的表述

2. 深化机制讨论

- 虽然无法建立因果关系，但可以更深入讨论可能的机制
- 提供更多关于技术支出关联的理论解释
- 分析不同支出类别可能产生效果的时间差异

优先级2：增强分析的稳健性

3. 扩展敏感性分析

- 测试不同的学习恢复指标构建方法
- 考虑不同的分组方式和控制变量组合
- 分析异常值和极端情况的影响

4. 加强理论框架

- 更详细地阐述教育生产函数在ESSER背景下的应用
- 讨论紧急教育资金与常规资金的理论差异
- 提供更丰富的文献综述和理论基础

优先级3：提升政策相关性

5. 深化政策讨论

- 更详细地分析ESSER政策设计的优缺点
- 讨论临时性资金的可持续性问题
- 提供更多关于资金分配公平性的分析

6. 为未来研究提供指导

- 明确指出哪些问题需要因果研究来回答

- 建议具体的因果识别策略
 - 讨论数据收集和研究设计的改进方向
-

技术评论

统计方法

模型选择：

- 分层回归模型的使用恰当
- 贝叶斯估计方法增强了结果的稳健性
- 控制变量的选择合理

需要改进：

- 考虑使用更复杂的面板数据模型
- 探索非线性关系的可能性
- 进行更多的模型诊断

数据处理

优点：

- 数据清理过程透明
- 缺失值处理合理
- 样本代表性分析充分

建议：

- 考虑使用多重插补处理缺失数据
 - 进行更多的敏感性分析
 - 探索异常值的影响
-

贡献与意义

学术贡献

1. **数据收集：**为亚利桑那州ESSE支出提供了详细的数据整理
2. **描述性分析：**揭示了支出模式的一些有趣特征
3. **方法尝试：**尝试了支出类别的细分分析方法

严重局限性

1. **方法论缺陷：**缺乏因果识别使得主要结论不可信
 2. **理论贡献微弱：**主要是描述性统计，缺乏理论创新
 3. **政策建议无依据：**基于相关性的政策建议在学术上不可接受
 4. **外部有效性存疑：**单一州的分析难以推广
-

总体评价

本博士论文在学术诚信和方法论透明度方面表现出色。作者诚实承认了研究的局限性，明确将研究定位为相关性分析而非因果推断，这种学术态度值得赞赏。论文内容丰富，在理论背景、方法论和实证分析方面都较为充实。

主要优势：

- **学术诚信：** 明确承认无法建立因果关系，避免了过度解释
- **内容充实：** 理论背景、方法论讨论和实证分析都较为全面
- **数据质量：** 使用长期面板数据和贝叶斯方法，技术上较为严谨
- **政策价值：** 为ESSER资金使用模式提供了有价值的描述性证据

仍存在的局限：

- **因果推断缺失：** 无法回答支出"是否"以及"为什么"有效的问题
- **理论贡献有限：** 主要是实证描述，理论创新相对较少
- **政策指导受限：** 相关性证据对具体政策制定的指导作用有限
- **外部有效性：** 单一州的分析结果推广时需要谨慎

总体判断： 本论文达到了博士学位的基本要求。作者展现了良好的学术素养，诚实面对研究局限性，并在此基础上提供了有价值的描述性分析。虽然无法提供因果证据，但作为ESSER研究的早期探索，论文为该领域的后续研究奠定了重要基础。

建议： 论文可以接受，但建议作者在未来研究中考虑采用更严格的因果识别策略，以便为教育政策制定提供更强有力的证据支持。

具体修改建议

第1章 引言：

- 可以更明确地阐述研究的理论贡献
- 加强对亚利桑那州作为案例研究价值的论证
- 更清晰地界定研究边界和局限性

第2章 文献综述：

- 增加对教育技术投资效果的文献回顾
- 补充疫情期间教育不平等变化的相关研究
- 更明确地指出本研究填补的文献空白

第4章 方法论：

- 更详细地讨论因果识别的挑战和应对策略
- 加强对学习恢复指标构建的理论和实证验证
- 考虑报告更多的稳健性检验结果

第5章 结果：

- 增加对效应大小经济意义的讨论
- 深化对异质性效应机制的分析
- 更充分地讨论无效应类别的可能原因

第6章 讨论与结论：

- 结论部分很好地总结了研究发现和局限性
- 对ESSER政策效果的讨论较为平衡和谨慎
- 可以更明确地指出未来因果研究的具体方向
- 对"金钱是否重要"辩论的贡献定位合理

本论文达到了博士学位的基本要求，展现了作者良好的学术素养和诚实的研究态度。虽然在因果推断方面存在局限，但作为描述性研究和假设生成，论文具有重要价值。