

连享会 · 2025 暑期班

Update: 2025.6.23 , 网页版

连享会 · 2025 暑期班

课程概览

A. 课程概要

B. 授课教师简介

C. 课程特色

初级班

高级班

论文班

D. 观念要转变: AI+ 实证分析的研究模式

1. 初级班

1.1 课程导引

1.2 专题介绍 (初级班)

A0. 课前准备

A1. Stata 简介和数据处理

A2. 线性回归分析

A3. 模型设定和假设检验

A4. 实证分析可视化

A5. 面板数据模型

A6. 一篇 Top 期刊论文重现

A7. 自行研读-提供复现数据和代码

2. 高级班

2.1 课程导引 (高级班)

2.2 专题介绍 (高班)

B1. 动态面板模型

B2. 面板 ARDL 模型

B3. 面板交互固定效应

B4. 面板变系数模型

B5. 聚类分析 (bunching)

B6. 随机边界分析前沿方法

3. 论文班

主讲嘉宾

3.1 课程导引

3.2 专题介绍 (论文班)

C1. 首篇论文发表: 如何通过复制AER经典文献开启学术发表之路?

C2. 首次独作: 遇到研究瓶颈时如何破局脱困、化危为机?

C3. 最新发表与在研论文: 如何结合前沿方法开拓研究新径?

4. 报名和缴费信息

缴费方式

5. 听课指南

5.1 软件和课件

课程概览

A. 课程概要

班次

- 初级班：2025 年 8 月 4-6 日 (3 天)，连玉君
- 高级班：2025 年 8 月 8-10 日 (3 天)，连玉君
- 论文班：2025 年 8 月 12-14 日 (3 天)，张宏亮

安排

- **时间**：2025 年 8 月 4-14 日
- **方式**：网络直播 + 回放
- **授课教师**：连玉君 (初级+高级) || 张宏亮 (论文班)
- **PDF 课纲**：https://file.lianxh.cn/KC/lianxh_PX.pdf，[参考文献和预读资料](#)
- **课程主页**：<https://www.lianxh.cn/PX.html> ([往期答疑和板书](#))
 - **Note**: 预习资料、常见问题解答等都将通过该主页发布。
- **报名链接**：<https://www.wjx.top/vm/tC1UWC.aspx#>
- **助教招聘**：<https://www.wjx.top/vm/eCkXO8U.aspx#>

回放安排：每班 30 天回放

- **初级班**：8 月 7 日 - 9 月 5 日 (30 天)
- **高级班**：8 月 11 日 - 9 月 9 日 (30 天)
- **论文班**：8 月 15 日 - 9 月 13 日 (30 天)
- **多班组合**:
 - **初级+高级**：8 月 11 日 - 10 月 9 日，共计 60 天；
 - **初级+论文**：8 月 7 日 - 8 月 11 日 (5 天)，8 月 15 日 - 10 月 8 日 (55 天)，共计 60 天；
 - **高级+论文**：8 月 15 日 - 10 月 13 日，共计 60 天；
 - **全程班**：8 月 15 日 - 11 月 12 日，共计 90 天。

B. 授课教师简介



连玉君，西安交通大学经济学博士，中山大学岭南学院副教授，博士生导师。已在 China Economic Review、Energy Economics、Stata Journal、经济研究、管理世界、经济学(季刊)、金融研究、统计研究等期刊发表论文 70 余篇。目前已完成 Panel VAR、Panel Threshold、Two-tier Stochastic Frontier 等计量模型的 Stata 实现程序，并编写过几十个小程序，如 `lianxh`，`ihelp`，`sftt`，`winsor2`，`xtbalance`，`bdiff`，`getiref` 等。连玉君老师团队一直积极分享统计和计量方法，开设了 [连享会-博客](#)，[连享会-知乎](#) 等专栏，已在微信公众号 ([连享会](#)，ID: `lianxh_cn`) 分享推文 1300 余篇，各平台阅读量逾 3000 万人次。



张宏亮，美国麻省理工学院 (MIT) 博士，浙江大学经济学院新百人计划研究员，博士生导师。主要从事经济学微观实证研究，尤其擅长因果推断与机器学习方法在劳动经济学、公共经济学、发展经济学、城市经济学等领域的应用。研究成果见诸 International Economic Review (IER), Journal of European Economic Association (JEEA), Journal of Public Economics (JPubE), Journal of Development Economics (JDE, 2 篇), Journal of Urban Economics (JUE) 等专业领域顶尖期刊。

C. 课程特色

本期课程的主要特色包括：

初级班

- 五次公开课：讲解 Stata 和 Python 的基本应用。
- 两次公开课：开课前，学员可以通过 [连享会-Books](#) 以及公开课讲解，预先完成 Stata 和 Python 运行环境的配置，建立起使用 AI 辅助进行实证研究的思维方式和工作模式。
- 课件进行了全新改版，更注重基本原理的讲解 + 论文复现。前者是合理应用各种计量方法的基础，后者则有助于大家掌握规范的实证研究套路和研究设计方法。至于具体的命令细节，学员可以通过 AI 辅助工具 (如 ChatGPT, Claude 等) 来完成。
- 提供了 4 篇完整的论文复现文档，辅以「中文精要」对论文中的难点进行解读，以帮助学员将所学内容迁移到自己的论文中。

- 目标：帮助学员实现「Stata + Python + AI助手」的工作模式，学会编写提示词，逐步实现使用「自然语言」进行编程的实证分析新思路。

高级班

- B1 - B4 覆盖了一些最主流的面板数据模型，包括动态面板模型、动态面板门槛模型、面板交互固定效应模型、面板 ARDL 模型、面板变系数模型等。为研究中分析长期和短期动态关系、异质性边际效应、结构变化和非线性关系提供了强有力的工具。
- B5 系统介绍了聚类分析 (bunching) 方法，适用于研究税收政策、补贴政策等对经济行为影响的计量方法。
- B6 包括两个新近发展的随机边界模型：异质性双边随机边界模型 (Het-TTSFA) 和稳健非参随机边界模型 (Robust Nonparametric SFA)。前者可以应用面板数据模型，并允许干扰项存在截面相关；后者采用非参数方法估计边界函数，无需预先设定函数形式，能有效克服现有方法普遍存在的模型误设和离群值导致的估计偏差。二者在 TFP 估算、效率分析、议价行为的分析方面具有重要应用。
- 每个专题都提供 1-2 篇完整的论文复现资料包，包括复现数据、代码、中文精要 (论文核心公式和关键代码解读)。

论文班

- **创新的研究思维**：通过精读和复现经典文献，学员能够发现新的研究切入点，提升论文的边际贡献。课程鼓励从多个维度借鉴其他文献，提升文章的学术深度和创新性。
- **因果推断与机器学习结合**：课程特别强调运用机器学习技术（如 Causal Forest 和 Gradient Boosting）增强因果推断能力，结合前沿工具（如随机对照试验、Causal Forest、Fuzzy DiD 等），为因果分析提供强有力的支持。
- **完整的论文复现资料**：每个专题提供完整的论文复现资料包，包括数据、代码和中文精要解读，帮助学员理解论文中的核心公式和关键代码，轻松掌握实证分析技巧。
- **写作与投稿指导**：课程注重提升学员的写作能力，帮助学员完成第一篇独立作者文章，面对审稿意见时学会审慎应对，不断提升论文质量，精益求精。
- **全面的因果分析方法**：深入讲解微观面板数据的因果识别方法，如 FE 模型、FE-IV 模型、FD-FE-IV 模型等，同时介绍局部平均处理效应 (LATE)、分布处理效应 (DTE) 等方法，并结合经典文献复现，帮助学员掌握因果推断的核心技能。

D. 观念要转变：AI+ 实证分析的研究模式

by 连玉君 2025.6.21

2023 年 1 月，我首次使用 ChatGPT 时便被其强大的自然语言处理能力所震撼。在过去的两年里，我尝试了市面上主流的各种 AI 工具，也在焦虑和磨合中不断寻求新的平衡。时至今日，我开始积极拥抱 AI，并对实证分析的研究模式有了新的认识。

最核心的感悟是：AI 也好，Stata、R、Python 也罢，终归都是工具，所有的工具最终都是为了实现某种目标而服务的。当我发现很多机器学习方法都是使用 R 语言来实现的，便在半年前开始学习 R 语言。借助 AI 工具的帮助，学习效率提高了很多。两个月前，当应邀给研究生开设「数据分析」课程时，我强迫自己用 Python 作为授课语言。在花了两天时间配置好 Python 的运行环境后，我的备课效率提高了 5 倍以上。在我目前完成的 [数据分析与 Python 应用讲义](#) 中，我基本上没有自己动手写过 Python 代码，所有代码都是通过 Github Copilot 和 ChatGPT 生成的。我的班上有 60 多名在职研究生，来自各个领域，几乎都没有编程基础，但他们在第二周的课程结束后，就已经可以借助 AI 工具，完成 [上市公司财务资料的合并、清洗和统计分析](#) 了，还有些学生已经可以 [爬取并统计股市基本信息](#)，还有学生详细分析了 [当当网 Python 类图书的销售数据](#)。这让我既感叹又欣慰。

在此过程中，我的授课模式和实证分析的理念都发生了很大的改变：

1. **从单一工具到多工具结合**：不再局限于 Stata 或 Python，而是基于目标导向，根据任务需求灵活选择和组合工具。
2. **从敲代码到编写提示词**：更多地使用自然语言提示词来生成代码，降低编程门槛，提高效率。
3. **流程和框架比掌握具体的命令细节重要得多**：由于不必花大量时间学习各种命令和函数，反而可以将精力集中在研究问题本身，思考如何凝练研究问题、设定识别策略、设计研究方案。在此过程中，可以有效利用 AI 辅助进行选题、研究设计和文献梳理。
4. **统计和计量基础比任何时候都重要**。写提示词的关键是准确表达我们的需求，这要求我们有扎实的理论基础，对基本概念和模型的基本原理、适用条件有清晰的认知。例如，在多元线性回归分析中，最重要的不是具体的估计方法，而是外生性假设和 FWL 定理。前者是模型设定的理论基础，扩展后就演变出了 IV 估计、GMM 估计，以及基于实验设计思想的 DID, RDD, SCM 等方法。而 FWL 定理则是理解多元回归系数的经济含义的基础，也是应对高维数据的重要武器。理解了 FWL 定理，就很容易理解高维固定效应模型、Lasso 估计，以及机器学习中的正则化方法、双重稳健性估计方法 (DDML)。又如，实证分析中最棘手的问题是因果识别，那么你必须了解因果识别的基本原理和方法，了解政策的发生过程、手头的数据是否存在样本选择和自选择问题，实验组和控制组样本的设定。以此为基础，才能确定该用 DID, RDD 还是 PSM, Heckman, DDML, 亦或是合成 DID 或者因果森林等基于机器学习的方法。
5. **见识要广**。以实证分析中最常用的固定效应模型为例，如果不理解「混淆变量」、「遗漏变量」等概念，就无法理解「辛普森悖论」，也就无法理解一阶差分、组内去心这些简单易行的处理方法的统计原理，致使在学习双向固定效应、高维固定效应和交互固定效应模型时，会面临很大的心理障碍。这种情况下，你向 AI 提问的质量会非常低。

本期课程的内容设计和讲义都将围绕上述理念进行调整和优化。我们会将更多的时间和精力放在帮助学员建立起实证分析的整体框架和思维方式上，而不是单纯地教授某个工具或方法的具体细节。为此，每个专题中，在介绍完原理和核心方法后，都会讲解 1-2 篇相关的论文，以便帮助大家建立以「框架意识」和「研究设计意识」。

1. 初级班

- **时间**：2025 年 8 月 4-6 日 (三天)
- **方式**：网络直播 + 30 天回放
- **授课嘉宾**：连玉君 (中山大学)
- **授课安排**
 - **授课方式**：幻灯片+Stata17 实操演示，全程电子板书+Stata 演示截图，课后以 PDF 形式分享给学员
 - **授课时间**：上午 9:00-12:00，下午 14:30-17:30 (17:30-18:00 答疑)。
 - **全程答疑**：由 10 位经验丰富的同学组成的助教团队会在课程群中全程答疑，并对答疑接龙文档进行详细的记录和分类，公布于 [课程主页](#)。
- **课程详情**：<https://www.lianxh.cn/PX.html>
 - **板书和答疑**：<https://gitee.com/arlionn/PX/wikis>
 - **PDF 课纲**：https://file.lianxh.cn/KC/lianxh_PX.pdf
 - **预读资料**：[点击下载参考文献](#)
- **报名链接**：<https://www.wjx.top/vm/tC1lUWC.aspx#>

1.1 课程导引

实证分析中，最伤神和耗时的事情莫过于研究设计和数据处理。在以往的授课中，有不少学员在听完了高级班后，又返回来参加初级班。大家的感触是：若没有扎实的基础，以及对实证分析整体架构的认识，后续的学习进度总感觉力不从心，进展缓慢。

在初级班中，我们力求将三天的课程设置成一个比较完整的体系，希望大家经过几天的学习（尚需另外花费 1-2 个月演练吸收），能掌握基本的统计和计量分析方法，能理解多数期刊论文中使用的分析方法。

翻阅 Top 期刊上的论文，文中的方法我们似乎都会。细细想来，原因在于这些论文的想法或视角通常都比较独特，并使用了恰当的方法来论证。关键在于研究设计，而这一点在当前的计量经济学教科书中却较少涉及。为此，本次课程突出两个特点：一方面，我会努力把基础知识讲解透彻，进度上不求快；另一方面，我会通过论文复现来讲解这些方法的合理应用。

在**内容安排**上，遵循了由浅入深，循序渐进的原则。

A1 讲 介绍 Stata 的基本用法、数据处理、编程和绘图基础。学习这些内容无需太多的计量经济学基础，但对于提高实证分析能力和分析效率，大有裨益。本讲中，我会以一篇文章为实例，说明 Stata 的基本语法结构，并对数据处理过程中的关键问题进行介绍，如离群值的处理、文字变量的处理等。就我个人的经验而言，数据处理能力的高低直接决定实证分析的效率，而对于离群值等问题的处理是否妥善会直接影响全文结果的稳健性，是多数人不够重视但却至关重要的问题。此前有不少学完了高级班的同学又回炉初级班，便是感悟到了这一点。

但凡提及写程序，很多人都会产生恐惧心理，其实，一旦掌握了最基本的原理和语法格式，Stata 中的程序设定并没有想象的那么困难。更为重要的是，对于多数人而言，由于并不需要写完整的 ado 文档，因此只需要学会最基本的条件语句和循环语句即可，难度又会进一步降低。一旦掌握了基本的编程知识和理念，你的实证分析便开始进入「快车道」了。

A2-A3 介绍文献中使用频率最高的线性回归模型，包括 OLS 的原理、结果的解释，以及虚拟变量和交乘项的使用等。对于这些内容的深刻理解和熟练掌握，构成了后续，多种主流实证模型的基础，例如，目前文献中广泛使用的固定效应模型 (FE)，倍分法 (DID)，断点回归设计 (RDD) 等方法，本质上就是在传统的线性模型基础上，增加一些虚拟变量或交乘项，配合巧妙的研究设计，来实现对不可观测的个体效应的控制，以及对政策效应的估计。

很多人会觉得 OLS 很简单，但 Top 期刊中使用最多的仍然是 OLS，如何合理的构建模型、解释结果便成为实证分析中必须掌握的。我精选了大家经常面临的几个专题并结合论文进行讲解，包括：虚拟变量的使用、交乘项的使用和解释、分组回归的合理设定和假设检验。我会重点强调对条件期望函数和 FWL 定理的解读，这构成了理解因果模型、面板模型以及机器学习中多种方法的基础。首经贸的一个博士生发信息给我：「连老师，我发现只要把你初级里面的虚拟变量相关的知识完全掌握，很多复杂的方法就都好理解了，甚至可以自己解决问题。」，我的回复是：「那看来你是把相关的东西基本搞明白了，我每次上初级班的时候会花很多时间讲虚拟变量和交乘项，这构成了双重差分、断点回归、时间中断分析、面板数据模型等一系列模型的重要基础。」

A4 讲介绍实证分析中的各种可视化呈现方法。学习本讲的目的有二：其一，在实证研究过程中，借助各种密度函数图、散点图、柱状图，可以让我们对变量的分布、关系有更为直观地理解，有助于加深对隐藏在数据背后的规律，长期积累下来会逐渐形成直觉；其二，目前主流期刊越来越强调结果呈现的可读性和直观性，各种可视化工具也应运而生，最为典型的是，多期 DID、RDD、Bunching、合成控制法等因果推断方法，以及稳健性检验、敏感性分析、安慰剂检验等都会借助图形来直观呈现高密度信息。

A5 介绍固定效应模型 (FE)，涉及基本的 FE，TWFE 模型和进阶的高维固定效应模型、交互固定效应，以及长差分模型。在讲解这些模型的基本思想和估计方法的过程中，我会将重点放在模型含义、使用场景和条件上来。例如，对于同一笔数据而言，何时采用 OLS 进行估计，何时采用 FE 估计？不同的方法之间有何差异和关联？结果背后的经济含义如何解读？掌握这些方法有助于大家合理控制内生性问题，以便得到更为可信的结论。更为重要的是，本讲介绍的**高维固定效应**、**交互固定效应**是理解各类面板因果推断模型 (如 DID，回归控制法等) 的基础，也是进一步学习动态面板、面板 VAR 等模型的基础。

A6 会拆解一篇发表于 QJE 的论文。该文基本上涵盖了前几讲中的主要内容。我们可以尝试用新的方法来研究文中涉及的问题。在此过程中，既能巩固对现有模型和方法的理解，也能够对比后续文献来确定新的研究主题。

1.2 专题介绍（初级班）

A0. 课前准备

Note: 这部分内容会在开课前通过 [连享会-Books](#) 发布，并由连玉君老师在开课前进行直播讲解，以便大家了解 Stata 和 Python 的基本运行环境，掌握基本的数据处理和可视化技能。最重要的是，帮助大家建立起使用 AI 辅助进行实证研究的思维方式和工作模式。

- 搭建 AI 辅助代码环境
 - Anaconda + VScode 环境配置
 - 基于提示词的编程和数据分析
 - Stata+Python: Jupyter Notebook
- 论文复现与实证研究流程
 - 如何借助 AI 复现论文
 - 在哪里找论文
 - 如何复现论文
 - 如何写复现报告
 - 新选题和研究设计
 - 提供 3 篇论文的复现包

A1. Stata 简介和数据处理

- 实证分析的基本流程
- AI 辅助：环境搭建、提示词编程
- 数据的导入/导出、转换及合并
- 离群值、文字变量的处理
- 基本统计分析
- do 文档、帮助文件和外部命令
- 局域暂元和全局暂元 (local, global)
- 控制语句（条件语句、循环语句）
- 返回值
- Stata 中的各类函数

A2. 线性回归分析

- 条件期望函数：OLS, MLE 与 MM 的关系
- 线性概率模型
- OLS 估计和系数含义
- FWL 定理 (Frisch-Waugh-Lovell)
- 假设检验和统计推断
- 稳健性标准误：Bootstrap、Jackknife、聚类调整
- 结果输出与呈现

A3. 模型设定和假设检验

- 控制变量：选取、含义、可视化
- 变量缩放和转换
- 取对数：弹性与半弹性
- 虚拟变量与固定效应

- 交乘项、平方项、高阶项与调节效应
- 因子变量与边际效应分析
- RDD 与 RKD 简介
- 系数组合检验、组间差异检验

A4. 实证分析可视化

- 可视化的基本原则
- 直方图与密度函数图: `histogram` , `kdensity` , `biplot`
- 分仓散点图: `binscatter`
- 系数及系数差异的可视化呈现: `coefplot`
- 调节效应、倒 U 型关系及边际效应的可视化
- 面板数据、多个控制变量、高维固定效应模型的可视化

A5. 面板数据模型

- 何谓个体效应? FE v.s. RE
- 高维固定效应模型
- 长差分 (long difference)
- 异方差和序列相关 (Bootstrap、Cluster 调整标准误)
- 面板模型中的非时变变量和宏观变量如何分析?
- 双重差分 (DID) 简介
- 参考文献
 - Cameron, C. A., D. L. Miller, 2015, A practitioner's guide to cluster-robust inference, *Journal of Human Resources*, 50 (2): 317-372. [-Link-](#) , [-PDF-](#)
 - Correia, S. 2016. `reghdfe` , Linear Models with High-Dimensional Fixed Effects: An Efficient and Feasible Estimator. Working Paper. [-PDF-](#) , [Examples](#)
 - Akcigit, U., J. Grigsby, T. Nicholas, S. Stantcheva, 2022, Taxation and innovation in the twentieth century, *The Quarterly Journal of Economics*, 137 (1): 329-385. [-Link-](#) , [-PDF-](#) Table 3, 长差分

A6. 一篇 Top 期刊论文重现

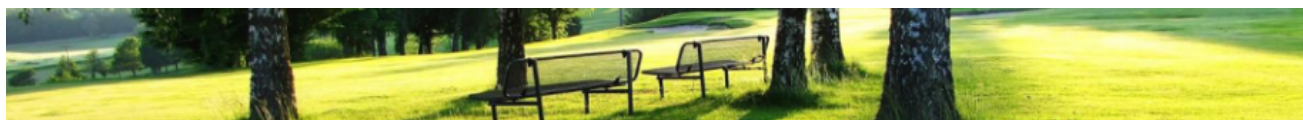
- Akcigit, U., J. Grigsby, T. Nicholas, S. Stantcheva, 2022, Taxation and innovation in the twentieth century, *The Quarterly Journal of Economics*, 137 (1): 329-385. [-Link-](#) , [-PDF-](#) , [-Appendix-](#) , [-cited-](#) , [-Replication-](#)
 - **简介:** 文章研究了美国公司税和个人税对创新的影响。文中采用了多种识别策略, 综合使用了高维固定效应模型、长差分模型和交互固定效应模型等方法, 辅以多种方式的可视化手段。通过复现和精读这篇论文, 初级班中所学到的多数内容都能得到巩固和应用。最重要的是, 文中的研究设计和实证研究思路可以顺畅地移植到大家自己的研究论文中。
 - 方法:
 - 高维固定效应、长差分、交互固定效应
 - 实证结果可视化: 分仓散点图、长期效应
 - 交乘项
- Note: (1) 我只挑选一些与 A1-A5 相关的内容来讲解; (2) 学员课前务必精读原文, 撰写读书笔记, 学有余力者, 可以根据作者提供的复现资料进行实操。

A7. 自行研读-提供复现数据和代码

这里提供三篇论文的复现包，供大家自行研读和练习。每篇论文的代码都经过整理，并附有中文精要，帮助大家理解论文的研究设计和实证分析方法。

- Sherman M G, Tookes H E. Female representation in the academic finance profession. **Journal of Finance**, 2022, 77(1): 317-365. [-Link-](#), [-cited-](#), [-PDF-](#), [-Replication-](#)
 - 该文研究了金融学术圈中的性别失衡现象。2009-2017 年，美国排名前 100 的商学院中的金融教师中，女性仅占 16.0%。性别失衡表现在几个方面：其一，在控制了研究能力后，更多的女性在排名较低的机构中任职，晋升为正教授的可行性相对较低，伴以薪酬较低。其二，女性发表的论文数量较少，但质量上不存在差异。其三，女性多与同性合作，表明她们的社会关系网较小。时间序列数据表明，上述性别差距正在缩小。
 - 该文没有使用任何复杂的回归方法，仅使用了固定效应模型。但在统计分析和结果可视化方面做了很多工作，是 Stata 入门学习的绝佳范本。
 - 方法：
 - 各种统计分析，列表和图形呈现： `egen` , `foreach`
 - OLS, 高维固定效应模型, 交乘项, 因子变量： `reghdfe`
 - 结果可视化： `coefplot`
 - 结果输出： `estadd` , `estout` , `esttab` ,
- Feng, L., Qi, J., & Zheng, Y. (2025). How can AI reduce carbon emissions? Insights from a quasi-natural experiment using generalized random forest. **Energy Economics**, 141, 108040. [Link](#) , [PDF](#) , [Google](#) .
 - 该文采用多期双重差分模型(DID)分析2013-2022年中国A股上市公司数据，评估AI试点政策(AIPZ)对企业碳绩效的影响及异质性机制。通过将AIPZ政策视为准自然实验，研究了企业数字化水平、内部控制等因素对政策效果的调节作用。此外，运用广义随机森林(GRF)模型识别企业财务健康状况（如ROA、托宾Q）对减排效果的非线性影响。研究发现，AIPZ政策在高数字化和强治理能力的企业中效果显著，而技术基础薄弱的企业则未能充分受益。本研究为AI政策的精准设计和目标群体识别提供了实证依据。
 - 方法：FE-DID, PSM-DID, 广义随机森林 (GRF), 可视化
 - 文中使用的主要 Stata 命令：
 - `reghdfe` : 高维固定效应回归，处理个体和时间固定效应。
 - `ivreg2` : 工具变量回归，用于解决内生性问题。
 - `outreg2` : 导出回归结果，支持多种文件格式。
 - `eventdd` : 事件研究分析，检验政策或干预的动态效应。
 - `honestdid` : 基于随机化的DID估计，控制时间和空间异质性。
 - `did_multipligt` : 多期DID分析，探索不同时间点的效应。
 - `psmatch2` : 倾向得分匹配 (PSM) , 估计处理效应的异质性。
 - `pwcorr_a` : 计算相关系数矩阵，检验变量间相关性。
- Tian, X., Tu, G., & Wang, Y. (2024). The Real Effects of Shadow Banking: Evidence from China. **Management Science**, 70(12), 8556–8582. [Link](#) , [PDF](#) , [Google](#) .
 - 该文通过多种实证方法探讨影子银行（特别是委托贷款）对企业技术创新的影响。首先，采用双重差分法 (DID) 估计委托贷款对创新产出的影响，并通过熵平衡匹配 (Entropy Balancing Matching) 方法解决样本选择偏差问题。为进一步验证结果的稳健性，研究进行了伪随机测试 (placebo test) 并使用工具变量 (IV) 方法解决内生性问题。研究假设委托贷款通过资本重新分配机制促进资金匮乏企业的创新产出。研究结果表明，委托贷款显著促进了企业的创新，尤其在金融资源配置不均的环境中，影子银行通过灵活的市场化机制支持了中小企业的技术创新。
 - 文中使用的主要 Stata 命令：
 - `tabstat` : 计算并展示分组统计量。

- `ttest` : 进行独立样本 t 检验, 比较不同组间均值差异。
- `reghdfe` : 高维固定效应回归, 处理个体和时间固定效应。
- `pctile` : 计算给定百分位数的分位值。
- `ivreg2` : 进行工具变量回归分析, 解决内生性问题。



2. 高级班

- **时间**: 2025 年 8 月 8-10 日 (三天)
- **方式**: 网络直播 + 30 天回放
- **授课嘉宾**: 连玉君 (中山大学)
- **授课安排**
 - **授课方式**: 幻灯片+软件实操演示, 全程电子板书+演示截图, 课后以 PDF 形式分享给学员
 - **软件**: Stata17 + Python (Jupyter Notebook)
 - **授课时间**: 上午 9:00-12:00, 下午 14:30-17:30 (17:30-18:00 答疑)。
 - **全程答疑**: 由 10 位经验丰富的同学组成的助教团队会在课程群中全程答疑, 并对答疑接龙文档进行详细的记录和分类, 公布于 [课程主页](#)。
- **课程详情**: <https://www.lianxh.cn/PX.html>
 - **板书和答疑**: <https://gitee.com/arlionn/PX/wikis>
 - **PDF 课纲**: https://file.lianxh.cn/KC/lianxh_PX.pdf
 - **预读资料**: [-点击下载参考文献-](#)
- **报名链接**: <https://www.wjx.top/vm/tC1lUWC.aspx#>

2.1 课程导引 (高级班)

高级班包括六个专题, 详情如下:

B1-B4 集中于实证分析中使用最频繁的面板数据模型, 涵盖了动态面板模型、面板 ARDL 模型、面板交互固定效应模型和面板变系数模型。这些模型在经济学、金融学和社会科学研究中具有广泛的应用, 能够有效地处理面板数据中的异质性、动态关系、结构变化和非线性关系。

B5 聚束分析 (bunching) 是一种用于研究税收政策、补贴政策等对经济行为影响的计量方法。它通过分析个体或企业在某一特定点附近的行为集中现象, 来推断政策对经济决策的影响。该方法特别适用于研究税率变化、补贴门槛等政策对个体行为的激励效应。

B6 介绍两个新近发展的随机边界模型: 异质性双边随机边界模型 (Het-TTSFA) 和稳健非参随机边界模型 (Robust Nonparametric SFA)。这两种模型在处理生产率和效率分析中具有重要应用, 能够更好地捕捉数据中的异质性和非参数特征。在研究 TFP、新质生产力、投资效率、议价和定价行为方面有独特的优势。

在讲授过程中，我会穿插介绍如何借助 AI 工具来辅助实证分析，包括数据清洗、模型设定、结果解释等环节。例如，在使用双边随机边界模型时，最棘手的并不是模型的估计和检验，而是如何设定模型的经济含义和理论基础。我会介绍如何撰写提示词，将我们的研究背景、已经想到的经济机制和理论分析视角传递给 AI 工具，让它辅助我们进行更为系统、严谨的论述，从而提升理论分析的深度，并为模型设定提供更为清晰的指导。

更为重要的是，我们可以借助 AI 工具来辅助选题和研究设计。这意味着，高级班的每个专题中介绍的方法都可以支撑起一篇新的论文。我期望在课程结束后一个月内，班上至少有 50% 的同学已经基本完成了这篇论文的初稿。

2.2 专题介绍（高班）

B1. 动态面板模型

多数经济行为的调整都会因为「调整成本」的存在而表现出粘性，即只能实现部分调整；同时，即使基于理性预期模型进行决策，但由于信息存在滞后，变量在时序上就会表现出较强的相关性。这些都构成了动态关系的理论基础，也使得动态面板模型广泛出现于经济增长、公司金融、国际贸易、劳动经济学等领域。

本专题首先介绍动态面板模型的理论基础、模型设定和估计方法，进而扩展至允许结构变化的情形，此时门槛值和异质性调整速度是我们关注的焦点。这两套模型既可以在论文中单独使用，也可以作为「稳健性分析」或「进一步分析」部分的扩展工具，以便从动态视角和结构变化角度提供更丰富的经验证据。最后，我会通过两篇论文的复现，展示如何在实证研究中应用动态面板模型和动态面板门限模型。

- 一阶差分 GMM 估计 (FD-GMM)
- 系统 GMM 估计 (SYS-GMM)
- 序列相关和过度识别检验 (Sargan 检验)
- 模型设定常见问题 (弱工具变量问题)
- 动态面板门限模型
- 参考文献：
 - Roodman, David. 2009, How to do Xtabond2: An Introduction to Difference and System GMM in Stata, *Stata Journal*, 9(1): 86–136. [-PDF-](#)
- 复现论文：
 - Deseau, A., Levai, A., & Schmiegelow, M. (2025). Access to justice and economic development: Evidence from an international panel dataset. *European Economic Review*, 172, 104947. [Link](#), [PDF](#), [Google](#).
 - Lee, C.-C., Li, M., Li, X., & Song, H. (2025). More green digital finance with less energy poverty? The key role of climate risk. *Energy Economics*, 141, 108144. [Link](#), [PDF](#), [Google](#). [-Replication-](#)

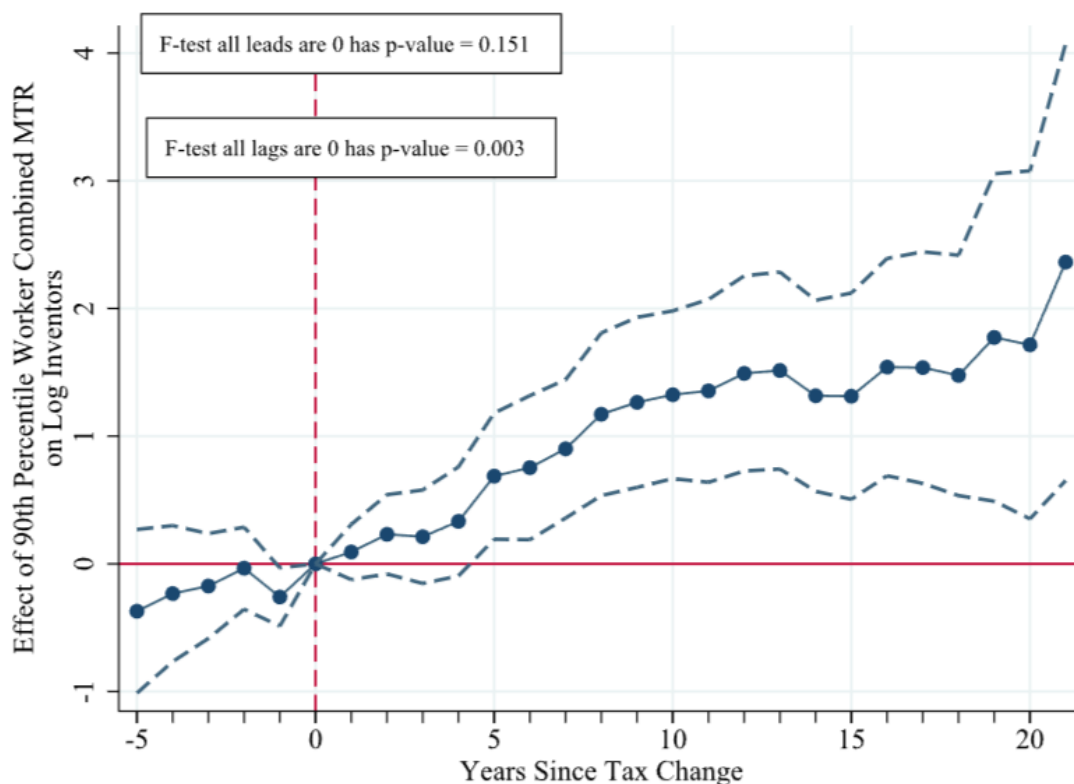
B2. 面板 ARDL 模型

面板 ARDL 模型全称「面板自回归分布滞后模型」，主要用于估计变量之间的长期关系。在时间序列分析中，该模型主要用于分析具有协整关系的非平稳序列之间的长期和短期关系。在面板数据中，我们可以更好地控制各类固定效应、考虑空间相关以及异质性特征，以便分析一项政策或某个变化缓慢的变量 (如气候、制度) 对经济增长、创新、贸易等结果变量的影响。本专题介绍面板 ARDL 模型的主要设定形式，以及该模型在经济和金融领域的应用场景。

我会重点讲解设定此类长期动态模型的理论基础和机制分析。以部分调整模型和理性预期模型为基础，介绍如何基于特定的故事背景设定长期目标函数和调整成本函数，以增强模型的可解释性，也为机制分析和异质性分析提供了理论依据。

- 时间序列简介：AR 与 MA 过程
- ARDL 模型：AutoRegression Distributed Lag

- 理论基础：部分调整模型与理性预期模型
- 面板 ARDL 模型
- 长期乘数与短期效应
- 模型扩展与经济含义



Source: Akcigit et al., 2022, Taxation and innovation in the twentieth century, *QJE*, 137 (1): 329-385. [-Link-](#), [-PDF-](#), [-Replication-](#)

• 复现论文：

- Dell, M., B. F. Jones, B. A. Olken, **2012**, Temperature shocks and economic growth: Evidence from the last half century, *American Economic Journal: Macroeconomics*, 4 (3): 66-95. [-Link-](#), [-PDF-](#), [-Replication-](#)
- Ditzgen, J. 2021. "Estimating long-run effects and the exponent of cross-sectional dependence: An update to xtdcce2". *The Stata Journal*, 21 (3): 687-707. [Link](#), [-PDF-](#), [PDF2](#).
- Kahn, M. E., K. Mohaddes, R. N. C. Ng, M. H. Pesaran, M. Raissi, J.-C. Yang, 2021, Long-term macroeconomic effects of climate change: A cross-country analysis, *Energy Economics*, 104: 105624. [-Link-](#), [-PDF1-](#), [-PDF2-](#), [-Replication-](#), [Cited](#). [lincom](#), [xtmg](#)

B3. 面板交互固定效应

交互固定效应模型和因子模型在高维数据分析和因果推断中扮演着尤为重要的角色。以面板数据（二维数据）为例，归属于同一个省份或行业的公司会受到一些共同因素（common factors）的影响，致使它们的行为表现出一定的相关性。在模型设定和统计推断中，若不考虑这些共同因素的影响，会导致遗漏变量偏误，或有偏的标准误。在因果推断中，「反事实」的构造和估计本质上是一个预测问题。除了随机森林、人工神经网络等方法外，因子模型和交互固定效应模型（IFE）虽然简单，但却具有很好的预测和外推能力，使其广泛应用于合成控制法、DID、RDD、回归控制法等模型。提出的「合成 DID, SDID」，以及 [Arkhangelsky](#) and Samkov (2024) 新近提出的「次序 SDID, Seq-SDID」都是以 IFE 为基础的。

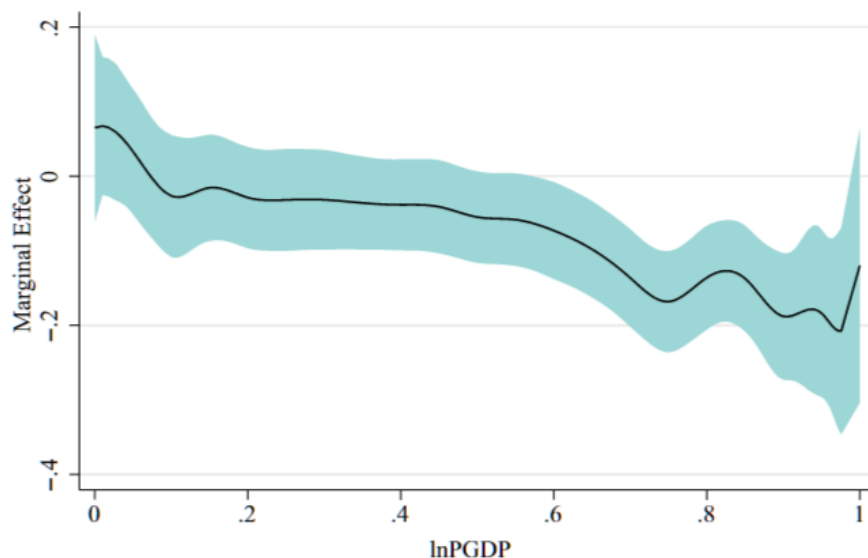
本讲从因子模型入手，介绍交互效应模型的设定、估计和检验方法，以及他们在因果推断中的典型应用。

- 共同因子和因子载荷
- 面板交互固定效应模型 (IFE) 的设定和估计方法
- 包含宏观变量的高维固定效应模型 (混频数据问题)
- 应用 1: 回归控制法
- 应用 2: 广义合成控制法
- 应用 3: 合成 DID (SDID)
- 参考文献:
 - Bai, J. 2009. Panel Data Models With Interactive Fixed Effects. (2009). *Econometrica*, 77(4), 1229–1279. [Link](#), [PDF](#), [Google](#).
 - Hsiao, C., H. S. Ching, S. K. Wan, **2012**, A panel data approach for program evaluation: Measuring the benefits of political and economic integration of hong kong with mainland china, *Journal of Applied Econometrics*, 27 (5): 705-740. [-Link-](#), [-PDF-](#)
 - Arkhangelsky D, Athey S, Hirshberg D A, et al. Synthetic difference-in-differences[J]. *American Economic Review*, 2021, 111(12): 4088-4118. [Link](#), [-PDF-](#) [-Replicate-](#) [-Github-](#)
 - Clarke, D., Pailañoir, D., Athey, S., & Imbens, G. (2024). On synthetic difference-in-differences and related estimation methods in Stata. *The Stata Journal*, 24(4), 557–598. [Link](#), [PDF](#), [-PDF2-](#), [Google](#).
 - Wager, Stefan. 2024, Causal Inference: A Statistical Learning Approach. [-PDF-](#), [Chap 13](#)
 - Sul, D. Panel data econometrics: Common factor analysis for empirical researchers[M]. 2019. [-Link-](#), [-PDF-](#), [Book-review](#), [Codes-Stata/Gauss/Matlab](#), [R-codes-readme](#)
 - Yan, G., & Chen, Q. (2022). rcm: A command for the regression control method. *The Stata Journal*, 22(4), 842–883. [Link](#), [PDF](#), [Google](#).
- 复现论文:
 - Ko, D. G. (2025). Did the American Rescue Plan cause inflation? A synthetic control approach. *Economic Modelling*, 143, 106935. [Link](#), [PDF](#), [Google](#). [-Replication-](#)
 - Stata codes + Jupyter Notebook

B4. 面板变系数模型

结构变化现象广泛存在于经济研究中，主要体现为变量之间的非线性关系，如变量 x 对变量 y 的影响随 z (经济环境或政策力度等变量) 的变化而不同。当这种边际影响具有非线性或结构跳跃特征时，如「J 型，S 型或倒 U 型」关系时，传统方法 (如交乘项或分组回归) 不再适用或存在严重偏误。

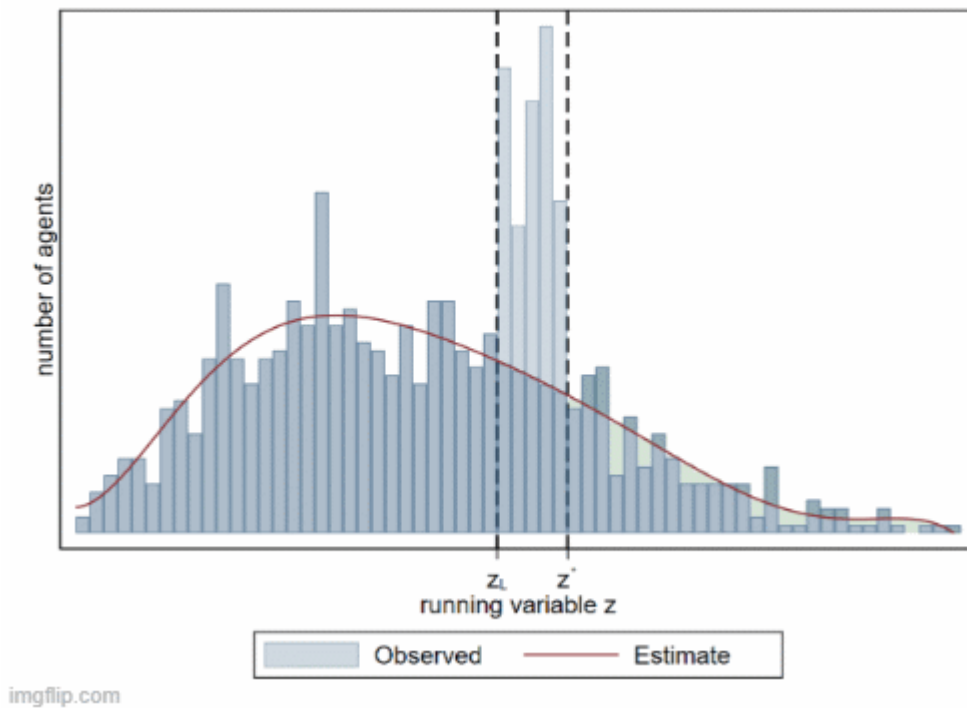
本专题介绍基于非参数估计的**函数系数模型 (FCM)**，它能把异质性分析和机制分析结合起来，有助于挖掘非线性关系背后的经济机制。相对于面板平滑转换模型和动态门槛模型，该模型在处理非线性关系时更为灵活，且不需要预先设定门槛值或转换函数。通过对核心参数的非参数估计，我们可以更好地捕捉变量之间的非线性关系，并进行异质性分析。该方法的早期应用文献 Du et al. (2021) 在发表以来的引用次数已经高达 936 次 便充分说明了该方法的有效性和广泛应用。



- 非参数估计简介
- 样条函数和基函数
- 结构变化的理论基础
- 函数系数模型 (FCM) 的设定和估计方法
- **复现论文:**
 - Du, K., Cheng, Y., & Yao, X. (2021). Environmental regulation, green technology innovation, and industrial structure upgrading: The road to the green transformation of Chinese cities. **Energy Economics**, 98, 105247. [Link](#), [PDF](#), [-Replication-](#), [Google](#).
 - Du, C., Cao, Y., Ling, Y., Jin, Z., Wang, S., & Wang, D. (2024). Does manufacturing agglomeration promote green productivity growth in China? Fresh evidence from partially linear functional-coefficient models. **Energy Economics**, 131, 107352. [Link](#), [PDF](#), [Google](#). [-Replication-](#)

B5. 聚束分析 (bunching)

聚束分析 (bunching) 是一种用于研究税收政策、补贴政策等对经济行为影响的计量方法。它通过分析个体或企业在某一特定点附近的行为集中现象，来推断政策对经济决策的影响。该方法特别适用于研究税率变化、补贴门槛、价格调整等政策对个体行为的激励效应。



- 聚束分析的理论基础
- 聚束分析的设定和估计方法
- 应用场景：税收政策、补贴政策、价格调整等
- 反事实检验和稳健性分析
- 复现论文：
 - Ewens, M., Xiao, K., & Xu, T. (2024). Regulatory costs of being public: Evidence from bunching estimation. *Journal of Financial Economics*, 153, 103775. [Link](#), [PDF](#), [Google](#). -[Appendix](#)-, [github-part-of-Data](#), -[Replication](#)-

B6. 随机边界分析前沿方法

在经济学和计量经济学中，随机边界分析 (SFA) 是一种用于测量生产率和效率的计量方法。传统的 SFA 模型假设生产函数已知，多设定为线性形式，但在实际应用中，这种假设往往过于严格。随着数据规模的扩大和复杂性的增加，传统 SFA 方法面临着模型设定刚性、观测误差异质、离群点干扰等问题。

本专题介绍两种新近发展的随机边界模型：异质性双边随机边界模型 (Het-TTSFA) 和稳健非参随机边界模型 (Robust Nonparametric SFA)。这两种模型在处理生产率和效率分析中具有重要应用，能够更好地捕捉数据中的异质性和非参数特征。

本次课程介绍的双板随机边界模型可谓「有理有据」。在理论层面，我将基于 Papadopoulos & Parmeter (2025) 的新作，从议价理论、信息不对称、遗漏变量 (如生产能力、管理能力等)、潜变量等角度提供设定双边随机边界模型的多种可能的理论框架。实证层面，我将基于 Lian et al. (2023) Papadopoulos & Parmeter (2025)，对传统的 TT-SFA 模型进行如下拓展：

- 非效率项的分布不再局限于指数分布，允许半正态分布等更为灵活的设定，使我们可以更好地捕捉数据中的异质性特征。
- 支持面板数据：可以设定双向固定效应
- 允许干扰项截面相关，这是审稿人经常质疑的一个问题。

Robust Nonparametric SFA (SFMA) 方法克服了传统 DEA/SFA 方法在大样本、观测误差和离群点较多情况下的失真和异常值干扰问题。SFMA 通过非参数样条灵活拟合效率前沿，利用修剪机制自动识别和排除异常观测，更准确地估计随机前沿。在能源效率评估、碳排放强度和绿色生产效率分析中，SFMA 能有效应对测量误差和离群值，自动剔除极端观测点。同时，在农业、医疗、银行和制造业等生产率和效率分析场景中，SFMA 准确识别高效单位，显著提升结果的经济学解释力，为政策制定和产业评价提供更可靠的依据。

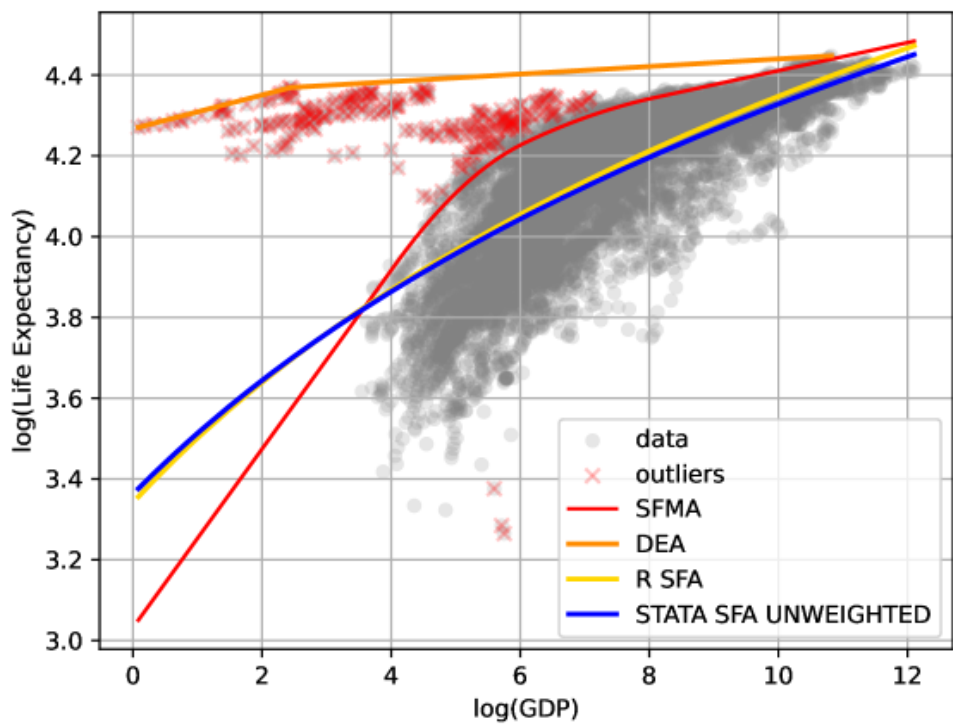


Table 1: Model comparison

Model	Stochastic	Non-parametric	Specified Obs. SE	Outlier-Robust
DEA		✓		
SFA	✓			
StoNED	✓	✓		
SFMA	✓	✓	✓	✓

• 参考文献:

- Lian Y., Chang L., Parmeter, C. F., Two-tier stochastic frontier analysis using Stata. **Stata Journal**, 2023, 23(1): 197-229. -Link-, -PDF-, Codes & Data, github. 安装: `net install st0705.pkg, all replace` (Stata)
- Papadopoulos, A., & Parmeter, C. F. (2025). Two-Tier Stochastic Frontier Analysis for the Social Sciences. Springer Nature Switzerland. Link, PDF, Google, github
- Zheng P., Worku, N., Bannick, M., Dielemann, J., Weaver, M., Murray, C., & Aravkin, A. (2024). Robust Nonparametric Stochastic Frontier Analysis. arXiv (Preprint submitted to **Journal of Econometrics**). Link, PDF, Google. github

3. 论文班

- **时间**：2025 年 8 月 12-14 日 (三天)
- **方式**：网络直播 + 30 天回放
- **授课嘉宾**：张宏亮 (浙江大学)
- **授课安排**
 - **授课方式**：幻灯片，Stata 实操演示，全程电子板书，课后以 PDF 形式分享给学员
 - **授课时间**：上午 9:00-12:00，下午 14:30-17:30 (17:30-18:00 答疑)。
 - **全程答疑**：由 10 位经验丰富的同学组成的助教团队会在课程群中全程答疑，并对答疑接龙文档进行详细的记录和分类，公布于 [课程主页](#)。
- **课程详情**：<https://www.lianxh.cn/PX.html>
 - **往期板书和答疑**：<https://gitee.com/arlionn/PX/wikis>
 - **PDF 课纲**：https://file.lianxh.cn/KC/lianxh_PX.pdf
 - **预读资料**：[点击下载参考文献](#)
- **报名链接**：<https://www.wjx.top/vm/tC1lUWC.aspx#>

主讲嘉宾



张宏亮，美国麻省理工学院 (MIT) 博士，浙江大学经济学院新百人计划研究员，博士生导师。主要从事经济学微观实证研究，尤其擅长因果推断与机器学习方法在劳动经济学、公共经济学、发展经济学、城市经济学等领域的应用。研究成果见诸 International Economic Review (IER), Journal of European Economic Association (JEEA), Journal of Public Economics (JPubE), Journal of Development Economics (JDE, 2 篇), Journal of Urban Economics (JUE) 等专业领域顶尖期刊。

3.1 课程导引

张宏亮老师将结合自己论文选题、写作和发表的经历，与大家分享研究选题、数据收集、实证分析、文章写作、投稿修改等历程中的经验、教训和心得。

撰写一篇论文是一个系统工程，需要协调掌控「开局」、「技法」和「章法」三大要素，并在写作过程中不断优化迭代已臻完善。把控这项系统工程，尤其是从「章法」角度统领全文，需要具备“全才型舵手”的素养，这对研究新手（甚至有经验的学者）来说都颇具挑战性。本次论文班将着重传授「章法」——如何将文章的各个模块巧妙串联成一篇逻辑完整、富有说服力的学术论文。张老师对「章法」的讲解不是简单罗列要点泛泛而谈，而是通过深入剖析自己已发表和在研的高水平论文，结合具体研究主题和文献背景，从作者视角探讨研究过程中遇到的各类挑战及其解决方案。

在分享个人经验和精讲相关研究之余，张宏亮老师也将在课程中贯穿介绍这些论文中用到的微观实证研究方法，并带领大家应用面板数据分析、工具变量、机器学习等方法复制精选论文的实证结果。

课程亮点：

- 从 Top 期刊论文作者角度讲解论文选题、研究设计、英文写作等方面的经验；
- 研究贡献的挖掘、梳理和论证；
- 结合论文的外审意见，探讨如何应对挑战，如何与审稿人和编辑进行有效沟通；
- 精讲 3-4 篇论文，并详细讲解多种前沿方法在论文中的综合应用。

精讲的实证研究方法

- 面板数据分析 (Panel Data Analysis)
- 移位-份额工具变量 (Shift-Share Instrument)
- 基于随机实验的因果分析 (RCT-based Causal Analysis)
- 双样本工具变量估计 (Two-Sample IV Estimation)
- 可信的双重差分法 (Creditable DID)
- 梯度提升模型 (Gradient Boosting)
- 因果森林模型 (Causal Forest)

3.2 专题介绍 (论文班)

C1. 首篇论文发表：如何通过复制AER经典文献开启学术发表之路？

如何写好第一篇学术论文是很多初学者面临的困惑。张宏亮老师将通过自己最早的两篇论文发表经历分享如何从“复制经典、师法范文”中扬帆起航、开启学术发表之旅。

2011 年，Duranton & Turner 在 *American Economic Review* (AER) 上发表了一篇实证论文，验证城市规划领域的“交通拥堵基本定律”。他们发现美国都市圈内州际高速公路的车辆行驶里程数对道路容量的弹性为 1，由此得出结论：拓展道路容量无法从根本上解决城市主干道的拥堵问题。

在复制该文章的实证结果后，张宏亮老师对其将历史交通基础设施存量作为当前州际高速公路道路容量的工具变量产生了一些质疑，认为影响城市历史交通基础设施发展水平的因素也可能影响当今的交通需求。基于这一思考，他与合作者另辟蹊径，利用日本第四次国家综合开发计划中的高速公路远景规划，构建了各都市圈不同时期国家高速公路道路容量增量的工具变量，重新检验了“交通拥堵的基本定律”。

论文投稿到城市经济学领域的顶尖国际期刊 *Journal of Urban Economics* (JUE) 后，编辑和审稿人对文章的内容、定位和创新提出了非常严格的要求，经过认真修改和仔细打磨后，论文质量得到了大幅提升，在理论框架、因果识别和实证发现上都较 AER 的文章做出了重大创新和拓展，发表后在城市交通拥堵研究领域产生了重要的影响。

张宏亮老师将与大家分享这篇源于复制经典文献的 JUE 论文诞生的全部历程，指导大家对这篇文章的实证结果进行复现，并应用近年“移动-份额工具变量” (Shift-Share IV) 文献中的两个重要分析框架 (Goldsmith-Pinkham, Sorkin & Swift, AER, 2020; Borusyak, Hull & Jaravel, RES, 2022) 来重新审视和改进当年文章中使用的实证模型。

第一天的课程中，张宏亮老师还将介绍自己2014年发表于发展经济学顶尖期刊 *Journal of Development Economics* (JDE) 的另一项研究——探讨父母离乡打工对留守儿童认知发展的影响。在分享这篇论文的完成和发表经历时，张老师将重点阐述该研究如何全面借鉴 Andrabi et al. (AEJ:Applied, 2011) 的研究框架和实证方法，将动态面板方法 (Dynamic Panel Method) 与双重差分法 (Difference-in-Differences, DiD) 相结合，用于评估父母离乡打工状态变化对留守儿童认知发展的影响。

专题亮点：

- 从精读/复现经典文献发现新的研究切入点；
- 寻找自己文章对现有文献的边际贡献，提炼文章亮点；
- 诚恳、谨慎、积极对待审稿意见，提升文章质量；
- 微观面板数据的因果识别方法（FE模型、FE-IV模型、FD-FE-IV模型）；
- 移动-份额工具变量（Shift-Share Instrument）；
- 动态面板方法（Dynamic Panel Method）；
- 复现 JUE 文章的实证结果。

核心论文：

- Duranton, G., & Turner, M. A. (2011). The Fundamental Law of Road Congestion: Evidence from US Cities. *American Economic Review*, 101(6), 2616–2652. [Link](#) (rep), [PDF](#), [Appendix](#), [Google](#).
- Hsu, W.-T., & Zhang, H. (2014). The fundamental law of highway congestion revisited: Evidence from national expressways in Japan. *Journal of Urban Economics*, 81, 65–76. [Link](#), [PDF](#), [Google](#).

其他相关论文：

- Andrabi, T., Das, J., Ijaz Khwaja, A., & Zojonc, T. (2011). Do Value-Added Estimates Add Value? Accounting for Learning Dynamics. *American Economic Journal: Applied Economics*, 3(3), 29–54. [Link](#) (rep), [PDF](#), [Appendix](#), [Google](#).
- Borusyak, K., Hull, P., & Jaravel, X. (2021). Quasi-Experimental Shift-Share Research Designs. *The Review of Economic Studies*, 89(1), 181–213. [Link](#) (rep), [PDF](#), [Google](#).
- Goldsmith-Pinkham, P., Sorkin, I., & Swift, H. (2020). Bartik Instruments: What, When, Why, and How. *American Economic Review*, 110(8), 2586–2624. [Link](#) (rep), [PDF](#), [Appendix](#), [Google](#).
- Zhang, H., Behrman, J. R., Fan, C. S., Wei, X., & Zhang, J. (2014). Does parental absence reduce cognitive achievements? Evidence from rural China. *Journal of Development Economics*, 111, 181–195. [Link](#), [PDF](#), [Google](#).
-cited-

C2. 首次独作：遇到研究瓶颈时如何破局脱困、化危为机？

一篇论文的发表需要克服研究选题、数据收集、实证分析、写作投稿等阶段出现的各种挑战，对第一次作为独立作者的新手殊为不易。

张宏亮老师将分享自己第一篇独作论文发表的艰辛历程，特别是面对研究瓶颈如何沉淀自我、突破困境、实现升华。这篇论文最初的定位是，利用初中录取电脑派位这一随机实验作为工具变量，研究重点中学的教育回报率。然而，虽然他克服层层障碍完成了数据收集、实证分析和初稿的撰写，但由于写作技巧不成熟等原因，文章在投稿时屡屡碰壁。同时，以 Pop-Eleches & Urquiola (AER, 2013) 和 Abdulkadiroglu et al. (Econometrica, 2014) 为代表的——一批相似主题的论文在此期间陆续发表，降低了张老师这篇论文仅凭实证贡献在顶刊发表的可能性。

面对这一困局，他将文章重新定位，聚焦于解决电脑派位数据库与中考数据库之间无法通过姓名实现完全匹配的方法挑战。为此，他提出了一套全新的「非完全匹配条件下」的双样本工具变量估计方法。这一改进将 Angrist & Imbens 获得诺奖的“局部平均处置效应”理论框架 (Econometrica, 1994) 的适用范围拓展到包含工具变量 (Z) 的数据集与包含处置变量 (D) 和结果变量 (Y) 的数据集之间仅能通过非唯一标识符 (non-unique identifier) 进行不完全匹配的情形。

这一计量方法创新使文章有了更清晰的定位，论文内容也更加扎实。该文最终发表在公共经济学领域的顶尖国际期刊 *Journal of Public Economics* (JPubE)。

在第二天的课程中，张宏亮老师将详尽介绍他在独立完成这篇 JPubE 论文面临困境时，和如何从破局脱困，转“危”为“机”，并指导大家对这篇文章的实证结果进行复现。此外，他还将结合这篇论文详细讲解基于随机实验的因果识别方法及其在实证研究中的应用和拓展。

专题亮点：

- 如何完成第一篇独立作者文章；
- 提升写作能力，精益求精；
- 面对拒稿败而不馁，锲而不舍；
- 基于随机实验的因果分析方法；
- 局部平均处置效应（Local Average Treatment Effect）
- 分布处置效应（Distributional Treatment Effects）
- 双样本工具变量估计方法（Two-Sample IV Estimation）
- 复现 JPubE 文章的实证结果。

核心论文：

- Zhang, H. (2016). Identification of treatment effects under imperfect matching with an application to Chinese elite schools. *Journal of Public Economics*, 142, 56–82. [Link](#), [PDF](#), [Google](#). -cited-,
- Imbens, G. W., & Angrist, J. D. (1994). Identification and Estimation of Local Average Treatment Effects. *Econometrica*, 62(2), 467. [Link](#), [PDF](#), [-PDF2-](#), -cited-, [Google](#).

其他相关论文：

- Abadie, A. (2002). Bootstrap Tests for Distributional Treatment Effects in Instrumental Variable Models. *Journal of the American Statistical Association*, 97(457), 284–292. [Link](#), [PDF](#), [Google](#). -cited-
- Abdulkadiroglu, A., Angrist, J. D. and Parag, P. (2014). The Elite Illusion: Achievement Effects At Boston And New York Exam Schools. (2014). *Econometrica*, 82(1), 137–196. [Link](#) (rep), [PDF](#), [Google](#). -cited-
- Angrist, J. D., & Krueger, A. B. (1992). The Effect of Age at School Entry on Educational Attainment: An Application of Instrumental Variables with Moments from Two Samples. *Journal of the American Statistical Association*, 87(418), 328. [Link](#), [PDF](#), [Google](#).
- Arellano, M., & Meghir, C. (1992). Female Labour Supply and On-the-Job Search: An Empirical Model Estimated Using Complementary Data Sets. *The Review of Economic Studies*, 59(3), 537. [Link](#), [PDF](#), [Google](#).
- Pop-Eleches, C., & Urquiola, M. (2013). Going to a Better School: Effects and Behavioral Responses. *American Economic Review*, 103(4), 1289–1324. [Link](#) (rep), [PDF](#), [Google](#). -cited-, [-Replication-](#), [-Appendix-](#)

C3. 最新发表与在研论文：如何结合前沿方法开拓研究新径？

随着大数据、机器学习、人工智能和新计量方法的迅速发展，学者们需要掌握并融会贯通多种前沿工具来解决复杂的研究问题。在第三天的课程中，张宏亮老师将结合自己最新发表和在研的论文，分享如何整合前沿思路与方法，博采众长，开拓研究新径。

张宏亮老师将首先介绍其最新发表在经济学权威综合性期刊 *International Economic Review* (IER) 上的论文。该论文研究了一项在中国农村开展的课后辅导实验，探究其对成绩落后小学生的家庭教育投入和学业成绩的影响。张老师将深入剖析其写作过程中如何从「开局」、「章法」和「技法」三个维度分别借鉴 Banerjee et al. (QJE, 2007)、Das et al. (AEJ:Applied, 2013) 和 Davis & Heller (ReStat, 2020) 这三篇经典实证论文。在文章发表后，张老师注意到 Carlana et al. (Econometrica, 2022) 采用了类似方法，评估了一项针对成绩优异的意大利初中移民学生的学业辅导和职业咨询干预实验。结合这两篇最新随机实验论文中使用机器学习方法进行的异质性处置效应（Heterogeneous Treatment Effects, HTEs）估计，张老师将详细讲授 Wagner & Athey (JASA, 2018) 提出的因果森林 (Causal Forest) 算法的原理和应用，并指导大家复现 Carlana et al. (Econometrica, 2022) 的实证结果。

当政策变化（如最低工资上调、禁止性别歧视性招聘广告）引发个人或企业调整行为时，政策评估面临一个关键挑战：难以识别在反事实条件下——即政策未调整的情况下——哪些群体本应受到新规则影响。张宏亮老师将通过 Card et al. (RES, 2024) 和 Cengiz et al. (JoLE, 2022) 介绍如何结合人口统计学和机器学习预测方法来应对这一挑战，从而估计政策对那些外生识别的、最可能受影响的特定群体的效应。

在此基础上，张老师还将结合自己使用荷兰全国人口的行政数据评估2012年托育补贴政策改革影响的工作论文，介绍如何进一步整合 Fuzzy DiD (De Chaisemartin and d'Haultfoeuille, RES, 2017) 和 DiD with Unobserved Treatment (Botosaru & Gutierrez, JAE, 2018) 等前沿计量方法，以评估政策变化对真正受影响群体的带来的因果效应。

专题亮点：

- 如何在既有文献的基础上探索新的研究聚焦点，提升文章的边际贡献；
- 如何在论文写作过程中博采众长，从多个维度借鉴其他文献；
- 运用机器学习 (Causal Forest) 增强因果推断的分析能力；
- 应用机器学习 (Gradient Boosting) 辅助因果推断；
- 如何整合新的前沿工具 (Randomized Controlled Trial, Causal Forest, Gradient Boosting, Fuzzy DiD, DiD with Latent Treatment) 进行因果推断研究；
- 复现 Carlana et al. (Econometrica, 2022) 文章的实证结果。

核心论文：

- Behrman, J. R., Fan, C. S., Guo, N., Wei, X., Zhang, H., & Zhang, J. (2023). TUTORING EFFICACY, HOUSEHOLD SUBSTITUTION, AND STUDENT ACHIEVEMENT: EXPERIMENTAL EVIDENCE FROM AN AFTER-SCHOOL TUTORING PROGRAM IN RURAL CHINA. *International Economic Review*, 65(1), 149–189. Portico. [Link](#), [PDF](#), [Google](#).
- Carlana, M., La Ferrara, E., & Pinotti, P. (2022). Goals and Gaps: Educational Careers of Immigrant Children. *Econometrica*, 90(1), 1–29. [Link](#) (rep), [PDF](#), [Google](#).
- Cengiz, D., Dube, A., Lindner, A., & Zentler-Munro, D. (2022). Seeing beyond the Trees: Using Machine Learning to Estimate the Impact of Minimum Wages on Labor Market Outcomes. *Journal of Labor Economics*, 40(S1), S203–S247. [Link](#) (rep), [-PDF-](#), [-PDF2-](#), [PDF+Appendix](#), [Google](#). [-cited-](#), [-Replication-](#), [Data](#)

其他相关论文：

- Banerjee, A. V., Cole, S., Duflo, E., & Linden, L. (2007). Remedying Education: Evidence from Two Randomized Experiments in India. *The Quarterly Journal of Economics*, 122(3), 1235–1264. [Link](#), [PDF](#), [Google](#).
- Botosaru, I., & Gutierrez, F. H. (2017). Difference-in-differences when the treatment status is observed in only one period. *Journal of Applied Econometrics*, 33(1), 73–90. Portico. [Link](#) (rep), [PDF](#), [Google](#).
- Card, D., Colella, F., & Lalive, R. (2024). Gender Preferences in Job Vacancies and Workplace Gender Diversity. *Review of Economic Studies*. [Link](#) (rep), [PDF](#), [Google](#).
- Das, J., Dercon, S., Habyarimana, J., Krishnan, P., Muralidharan, K., & Sundararaman, V. (2013). School Inputs, Household Substitution, and Test Scores. *American Economic Journal: Applied Economics*, 5(2), 29–57. [Link](#) (rep), [PDF](#), [Appendix](#), [Google](#).
- Davis, J. M. V., & Heller, S. B. (2020). Rethinking the Benefits of Youth Employment Programs: The Heterogeneous Effects of Summer Jobs. *The Review of Economics and Statistics*, 102(4), 664–677. [Link](#), [PDF](#), [Google](#).
- De Chaisemartin, C., & D'Haultfoeuille, X. (2017). Fuzzy Differences-in-Differences. *The Review of Economic Studies*, 85(2), 999–1028. [Link](#) (rep), [PDF](#), [Google](#).
- Wager, S., & Athey, S. (2018). Estimation and Inference of Heterogeneous Treatment Effects using Random Forests. *Journal of the American Statistical Association*, 113(523), 1228–1242. [Link](#), [PDF](#), [Google](#).

最后的话

需要特别强调的是，自我提升从来都不是件轻松的事情。因此，在开课之前，大家务必认真研读每一篇论文，了解其研究背景、研究思路、计量方法和主要结论，带着问题听课。同时，也建议大家在开课务前务必掌握文献的检索方法，学会使用 [文献管理](#) 和 [笔记](#) 工具，以便追踪每篇论文的后续进展，发掘新的研究主题。

4. 报名和缴费信息

- **主办方：**太原君泉教育咨询有限公司
- **标准费用**
 - 初级班：3300 元/人
 - 高级班：3600 元/人
 - 论文班：3700 元/人
- **单班报名优惠方案：**
 - **专题课老学员单班报名：**9 折
 - **学生（需提供学生证/卡照片）：**9 折
 - **会员单班报名：**85 折
- **组合报名优惠价：**
 - **初级+高级：**5700 元/人
 - **初级+论文：**5800 元/人
 - **高级+论文：**6100 元/人
 - **全程班报名：**8800 元/人
- **温馨提示：**以上各项优惠不能叠加使用。
- **联系方式：**
 - 邮箱：wjx004@sina.com
 - 王老师：18903405450 (微信同号)
 - 李老师：18636102467 (微信同号)

报名链接： <https://www.wjx.top/vm/tC1UWC.aspx#>

▶▶ 长按/扫描二维码报名：



缴费方式

方式 1：对公转账

- 户名：太原君泉教育咨询有限公司
- 账号：35117530000023891 (晋商银行股份有限公司太原南中环支行)
- **温馨提示：**对公转账时，请务必提供「汇款人姓名-单位」信息，以便确认。

方式 2：扫码支付



温馨提示：

- 可以使用已经绑定公务卡的微信/支付宝/云闪付等扫码付款
- 微信转账时，请务必在「添加备注」栏填写「汇款人姓名-单位」信息。
- 扫码支付后，请将「付款记录」截屏发给王老师-18903405450（微信同号）

5. 听课指南

5.1 软件和课件

听课软件：支持手机，ipad，平板以及 windows/Mac 系统的笔记本，**但不支持台式机以及 Surface 平板式电脑**

特别提示：

- 为保护讲师的知识产权和您的账户安全，系统会自动在您观看的视频中嵌入您的「用户名」信息。
- **一个账号绑定一个设备**，且听课电脑**不能外接显示屏**，请大家提前准备好自己的听课设备。
- 本课程为虚拟产品，**一经报名，不得退换**。
- 为保护知识产权，课程不允许以任何形式录屏及传播。

5.2 实名制报名

本次课程实行实名参与，具体要求如下：

- 高校老师/同学报名时需向连享会课程负责人 **提供真实姓名，并附教师证/学生证图片**；
- 研究所及其他单位报名需提供 **能够证明姓名以及工作单位的证明**；
- 报名即默认同意「**连享会版权保护协议条款**」。

6. 助教招聘

说明和要求

名额：30 名 (初级、高级和论文班各 10 名)

- **A. 课前准备：**协助完成 2 篇介绍 Stata 或 Python 或 R 语言 和计量经济学基础知识的文档，可以在 [这里](#) 查看推荐选题，风格类似于 [lianxh.cn](#)；
- **B. 开课前答疑：**协助学员安装课件和软件，在微信群中回答一些常见问题；
- **C. 上课期间答疑：**针对前一天学习的内容，在微信群中答疑 (8:00-9:00, 19:00-22:00)；
 - Note: 下午 5:30-6:00 的课后答疑由主讲教师负责。
- **要求：**热心、尽职，熟悉 Stata/R/Python 中至少一种语言的基本语法和常用命令，能对常见问题进行解答和记录。
- **特别说明：**往期按期完成任务的助教可联系连老师直录。
- **截止时间：**2025 年 7 月 16 日 (将于 7 月 18 日公布遴选结果于 连享会主页 [lianxh.cn](#))

申请链接： <https://www.wjx.top/vm/eCkXO8U.aspx#>

扫码填写助教申请资料：



课程详情： <https://www.lianxh.cn/PX.html>