

复刻Card和Krueger (1994):《最低工资与就业:新泽西州和宾夕法尼亚州快餐业的案例研究》分步指南

I. Card和Krueger (1994) 研究简介

A. 研究的意义与背景

David Card和Alan B. Krueger于1994年发表在《美国经济评论》(American Economic Review)上的论文《最低工资与就业:新泽西州和宾夕法尼亚州快餐业的案例研究》,是劳动经济学领域的一项里程碑式研究。其主要贡献在于,研究结果挑战了当时主流的观点,即提高最低工资将导致就业率下降。这项研究发现,在新泽西州快餐行业,最低工资的提高并未导致就业减少,甚至可能有所增加。此外,该研究因其严谨的方法论,对经验经济学中的“可信度革命”(credibility revolution)产生了重要影响¹。

该论文探讨了劳动经济学中的一个核心问题:最低工资对就业的影响。传统的竞争性劳动力市场模型预测,最低工资的提高会增加劳动力成本,从而导致企业减少雇佣,对就业产生负面影响²。Card和Krueger的研究结果对此提出了质疑。

一个值得注意的背景是,新泽西州于1992年4月1日将最低工资从每小时4.25美元提高到5.05美元,而此时该州正经历经济衰退³。通常情况下,经济衰退本身就会对就业产生负面压力。如果最低工资的上涨如传统理论预测那样进一步减少就业,那么在新泽西州观察到的就业变化应该会非常负面,甚至比作为对照组的宾夕法尼亚州更为显著。然而,研究并未发现这种负面效应,这使得其结论在不利的宏观经济背景下显得尤为突出和引人深思。这一背景强化了研究结果的稳健性,因为不能简单地将就业的潜在负面影响归咎于普遍向好的经济环境掩盖了最低工资的负面效应。

B. 研究问题与设计概述(自然实验)

该研究的核心问题是:州最低工资的提高对快餐行业的就业有何影响?

为了回答这个问题,研究者们巧妙地利用了一次“自然实验”:新泽西州提高了其最低工资标准,而邻近的宾夕法尼亚州的最低工资标准保持不变²。这使得研究者可以将新泽西州的快餐店(处理组)与宾夕法尼亚州的快餐店(控制组)进行比较。

数据收集通过对新泽西州和宾夕法尼亚州东部约410家快餐店进行两次问卷调查完成。第一次调查在1992年2月至3月(最低工资上涨前)进行,第二次调查在1992年11月至12月(最低工资上涨后)进行³。

选择快餐行业作为研究对象是经过深思熟虑的,原因如下³:

- 该行业雇佣了大量低薪工人,因此很可能受到最低工资变化的影响。
- 预计这些餐馆会遵守最低工资法。
- 快餐店的工作要求和产品相对同质化,便于对就业、工资和产品价格进行可靠的衡量。
- 该行业通常不涉及小费,简化了工资的衡量。
- 构建特许经营餐馆的抽样框架相对容易。

C. 主要发现及其影响概述

与传统模型的预测相反, Card和Krueger发现“没有迹象表明最低工资的上涨减少了就业”³。更具体地说, “相对于宾夕法尼亚州的餐馆, 新泽西州的快餐店平均每家店增加了2.5名员工”² (另有提及“增加了13%的就业”⁴, 这可能是指全职等效员工的百分比变化或特定子样本的计算, 复刻时需核对原文以确定精确定义)。

该研究还比较了新泽西州内那些最初支付工资已达到或超过5.00美元(因此基本不受新法影响)的餐馆与那些工资较低(不得不因新法提高工资10-15%)的餐馆的就业变化。结果显示, 未受最低工资影响的餐馆的就业增长与宾夕法尼亚州的餐馆相似, 而那些不得不提高工资的餐馆反而增加了就业²。

这些发现引发了学术界关于最低工资影响的广泛讨论和大量后续研究¹。

II. 获取与准备数据

A. 定位与下载数据集

用于本研究的数据集是公开的。主要获取来源包括:

- David Card的个人网站(加州大学伯克利分校)⁸。AveDavis在RPubs上的复刻研究明确指出:“本项目的数据集来自David Card的网站”⁹。
- 普林斯顿大学的资源, 如图书馆指南(LibGuides)⁴, 通常提供数据和配套R脚本的直接链接。⁴提供了“Card and Krueger (1994) data”和“Workshop script”的链接。

数据集通常以public.dat(文本文件) همراه با یک کدبک (codebook) 或 Stata 格式的文件(通常是card_data.dta)的形式提供¹⁰。

需要注意的是, 公开可用的数据集版本有时可能与原始出版物中使用的确切文件存在微小差异。例如, 一份资料明确指出, “数据可能与原始论文(AER 1994)中使用的数据略有不同, 因此结果也可能略有不同”¹⁰。这意味着, 在进行复刻时, 应尽可能寻找最“存档”或“官方”的版本, 这些版本通常可以从期刊(如AER)或作者的学术主页链接获得。如果在复刻过程中未能与已发表的表格完美匹配, 微小的数据差异可能是一个原因, 并应在复刻报告中予以说明。普林斯顿大学LibGuides⁴提供的文件很可能是常用的复刻文件来源。

B. 理解数据格式

- 如果使用 public.dat, 它是一个纯文本文件, 可能是空格分隔或固定宽度的。随附的编码簿¹⁰对于解析此文件至关重要。
- 如果使用 card_data.dta, 它可以直接加载到Stata中, 或由R(使用如 haven 等包)读取。
- 数据集包含410个观测值, 代表接受调查的快餐店³。

C. 将数据加载到统计软件中(推荐R或Stata)

- 对于R:
 - 如果是 public.dat:使用 read.table() 或 read.fwf()(如果是固定宽度), 并结合从编码簿中获取的列名。RPubs的示例⁹显示:df <- read.table("~/Documents/Rpubs_Projects/njmin/public.dat"), 然后使用编码簿文件

分配列名。

- 如果是 card_data.dta: 使用 `haven::read_dta()`。
- 对于 **Stata**:
 - 使用命令 `use card_data.dta, clear`。如果从 `public.dat` 开始, 则使用 `import delimited` 或 `infile` 配合基于编码簿的字典文件。
- 加载数据后的初始步骤: 验证观测值数量 (应为410个), 并检查前几行以确保数据已正确加载。

III. 理解数据集: 变量与结构

A. 编码簿详解

本节将详细利用¹⁰和¹⁰中提供的编码簿。该编码簿详细列出了变量名、列位置 (针对固定宽度文本文件)、格式和解释。

关键变量类别包括:

- 店铺标识符: SHEET (唯一的店铺ID)。
- 连锁店信息: CHAIN (1=Burger King, 2=KFC, 3=Roy Rogers, 4=Wendy's), CO_OWNED (1代表公司直营)。
- 地理和时间虚拟变量: STATE (1代表新泽西州, 0代表宾夕法尼亚州), SOUTHJ、CENTRALJ、NORTHJ、PA1、PA2、SHORE (各种地点虚拟变量)。原始论文可能使用了更简洁的NJ虚拟变量。¹⁰中提到的变量T (1代表1992年11月, 0代表1992年2月) 对于区分处理前后时期至关重要, 尽管在实践中, 通常会根据第一波和第二波变量的存在来创建此变量。
- 第一次访谈变量 (第一波 - 处理前, 1992年2月): EMPFT (全职员工数)、EMPPT (兼职员工数)、NMGRS (经理数)、WAGE_ST (起始工资)、HRSOPEN (每日营业小时数)、PSODA (中杯苏打水价格) 等。
- 第二次访谈变量 (第二波 - 处理后, 1992年11月): EMPFT2、EMPPT2、NMGRS2、WAGE_ST2、HRSOPEN2、PSODA2 等。
- 访谈过程/状态: NCALLS (回访次数)、STATUS2 (第二次访谈的状态, 表明关闭、拒绝等)。

诸如CHAIN (连锁店品牌) 和CO_OWNED (是否为公司直营) 等变量的存在具有重要意义。这些变量代表了可能独立影响就业和工资的店铺特征。为了准确分离最低工资变化的影响, 必须考虑到这些因素。例如, 不同的连锁品牌 (如肯德基与温蒂汉堡) 可能有不同的员工配置模型或薪酬策略。公司直营店与加盟店在工资设定、成本吸收能力方面也可能存在差异。如果这些特征与店铺位于新泽西州还是宾夕法尼亚州, 或者与初始工资水平相关, 那么在回归模型中忽略它们可能会导致对最低工资效应的估计产生偏差。因此, Card和Krueger很可能将这些变量 (或其中一部分) 作为控制变量纳入其回归模型, 以解释这种异质性, 从而获得对政策影响更精确的估计。复刻研究时应力求包含类似的控制变量。

B. 识别与构建用于复刻的关键变量

- 就业 (结果变量):
 - 全职等效 (Full-Time Equivalent, FTE) 就业是Card和Krueger研究中的主要结果变量。需要为两波调查分别构建。一个常用的计算公式 (应根据Card和Krueger的论文进行核实) 是: $FTE = EMPFT + 0.5 \times EMPPT + NMGRS$ 。
 - 需要创建的变量: FTE1 (使用EMPFT、EMPPT、NMGRS) 和FTE2 (使用EMPFT2、

- EMPPT2、NMGRS2)。
- 然后, 计算FTE的变化: $DFTE = FTE2 - FTE1$ 。
- 数据集中分别提供了全职员工 (EMPFT、EMPPT2) 和兼职员工 (EMPPT、EMPPT2) 以及经理 (NMGRS、NMGRS2) 的数量, 这一点至关重要¹⁰。这使得研究者能够采用比简单计算人头数更细致的就业衡量方式。简单的员工计数可能会产生误导, 特别是当劳动力构成发生变化时 (例如, 用兼职员工替代全职员工)。FTE通过将兼职员工转换为全职等效员工 (通常假设兼职员工的工作时长为全职员工的一半, 经理通常计为全职) 来标准化就业衡量。这提供了一个更准确的劳动力总投入或就业水平的度量。虽然有评论指出Card和Krueger未使用实际工作时长¹, 但使用FTE在一定程度上缓解了这个问题, 因为它是总工时的一个代理指标。复刻时必须明确说明用于计算FTE的公式, 确保与Card和Krueger的方法一致。
- 处理和 时间指示变量 (自变量):
 - NJ: 虚拟变量, 如果 $STATE == 1$ (新泽西州) 则为1, 如果 $STATE == 0$ (宾夕法尼亚州) 则为0。
 - POST: 虚拟变量, 对于第二波数据 (1992年11月) 为1, 对于第一波数据 (1992年2月) 为0。当为双重差分 (DiD) 重构数据时 (例如, 采用“长”格式或使用第二波变量时), 该变量是隐式定义的。
 - NJ_POST: 交互项 ($NJ \times POST$), 在回归框架中将捕捉DiD估计值。
- 工资变量:
 - WAGE_ST1 (之前的起始工资) 和 WAGE_ST2 (之后的起始工资)。
 - GAP: Card和Krueger使用的一个变量, 用于衡量新泽西州低工资店铺为达到新最低工资标准所需的工资增长比例。如果新泽西州的 $WAGE_ST1 < 5.05$, 则 $GAP = (5.05 - WAGE_ST1) / WAGE_ST1$, 否则为0。这对于州内部分分析至关重要。
- 控制变量: CHAIN的虚拟变量、CO_OWNED、地点虚拟变量 (CENTRALJ、SOUTHJ等, 尽管Card和Krueger可能使用了更简单的区域指标)。

C. 数据清理与准备

- 处理缺失值: 检查变量中是否存在缺失数据代码 (例如, “”, -99)。根据Card和Krueger的方法 (如果论文中有详细说明) 决定插补策略或进行列表删除。
- 处理店铺关闭: STATUS2变量¹⁰ 显示了第二波调查中店铺的状态 (例如, 1=已回答, 2=因装修关闭, 3=“永久”关闭)。Card和Krueger的论文应详细说明他们如何处理在两波调查之间关闭的店铺。如果将这些店铺从样本中剔除, 且关闭与最低工资有关, 则可能引入选择性偏差。永久关闭的店铺数量较少 (6家¹⁰)。复刻研究应报告如何处理这些店铺。
- 创建派生变量: 如上所列 (FTE, NJ, POST, NJ_POST, GAP)。
- 数据重塑: 根据某些回归命令的偏好, 可能需要将数据从“宽”格式 (每个店铺的第一波和第二波变量分别在不同列) 重塑为“长”格式 (每个店铺每个时间段占一行), 尽管DiD也可以使用“宽”格式的差分数据进行。

表1: 用于复刻的关键变量摘要

- 应包含在本节 (III.C) 讨论完变量构建之后。
- 内容:

变量名	原始变量/构建方	编码簿中的描述/	在分析中的作用	来源
-----	----------	----------	---------	----

	式	派生方式		
SHEET	SHEET	唯一店铺ID	标识符	10
CHAIN	CHAIN	连锁餐厅 (1=BK, 2=KFC, 3=Roys, 4=Wendys)	控制变量 (分类)	10
CO_OWNED	CO_OWNED	1 如果是公司直营, 否则为0	控制变量 (虚拟)	10
STATE	STATE	1 如果是新泽西州, 0 如果是宾夕法尼亚州	处理组虚拟变量的基础	10
NJ	STATE == 1	虚拟变量: 1 如果是新泽西州, 0 如果是宾夕法尼亚州	处理组指示变量	派生
FTE1	EMPFT+0.5×EMPPT+NMGRS	全职等效就业, 第一波 (1992年2月)	结果变量 (前期)	10
FTE2	EMPFT2+0.5×EMPPT2+NMGRS2	全职等效就业, 第二波 (1992年11月)	结果变量 (后期)	10
DFTE	FTE2-FTE1	全职等效就业的变化	结果变量 (用于差分模型)	派生
WAGE_ST1	WAGE_ST	起始工资 (\$/小时), 第一波	关键自变量/控制变量, GAP的基础	10
WAGE_ST2	WAGE_ST2	起始工资 (\$/小时), 第二波	结果/控制变量	10
GAP	如果 WAGE_ST1<5.05 且在新泽西州, 则为 (5.05-WAGE_ST1)/WAGE_ST1, 否则为0	新泽西州初始工资低于新最低工资标准的店铺为达标所需的工资增长比例	处理强度变量 (新泽西州内部)	派生
POST	(隐式或显式)	虚拟变量: 1 代表第二波 (1992年11月), 0 代表第一波 (1992年2月)	时间段指示变量	派生
NJ_POST	NJ×POST	交互项: 1 代表新泽西州店铺在第二波	回归模型中DiD估计量的系数	派生
其他控制变量	例如 SOUTHJ, CENTRALJ, NORTHJ	新泽西州内区域或宾夕法尼亚州区域的虚拟变量	控制变量	10

这张表格非常宝贵, 它为用户提供了一个核心参考, 将编码簿¹⁰中的原始数据字段映射到复刻所需的分析变量。它阐明了诸如“就业”或“处理组”等关键概念是如何从源数据中操作化的, 这是任何复刻工作的基本步骤。它还突出了关键变量 (如FTE和GAP) 的构建, 这些变量虽然不直接存在

于原始数据中，但对Card和Krueger的分析至关重要。

IV. 双重差分 (DiD) 方法论

A. DiD估计量的概念解释

双重差分 (Difference-in-Differences, DiD) 方法是一种准实验技术，用于通过比较处理组和控制组在干预措施前后结果变量的变化来估计特定干预措施的因果效应²。

其关键假设是“平行趋势”假设：即在没有干预的情况下，处理组和控制组的结果变量会随时间呈现相似的变化趋势。DiD估计量可以消除由组间不随时间变化的未观测差异以及由影响两组的时间趋势所带来的偏差。

数学表达式为：

$$DiD = (E - E) - (E - E)$$

其中，Y 是结果变量（例如，FTE就业），T=1 是处理组（新泽西州），T=0 是控制组（宾夕法尼亚州），P=1 是处理后时期，P=0 是处理前时期。

B. 识别处理组、控制组和时间段

- 处理组：新泽西州的快餐店（最低工资上涨的州）。变量：NJ = 1。
- 控制组：宾夕法尼亚州东部的快餐店（最低工资保持在4.25美元不变的州）。变量：NJ = 0。
- 处理前时期（第一波）：数据收集于1992年2月至3月，在1992年4月1日工资上涨之前。
- 处理后时期（第二波）：数据收集于1992年11月至12月，在工资上涨之后。

表2：描述性统计 / 平均FTE就业的2x2 DiD摘要

- 应包含在本节 (IV.B) 确定组别和时期之后，以说明核心计算。
- 内容：

	第一波：1992年2月 (平均FTE)	第二波：1992年11月 (平均FTE)	变化 (第二波 - 第一波)
新泽西州 (N=~331)	FTE_NJ1_mean	FTE_NJ2_mean	Diff_NJ = FTE_NJ2_mean - FTE_NJ1_mean
宾夕法尼亚州 (N=~79)	FTE_PA1_mean	FTE_PA2_mean	Diff_PA = FTE_PA2_mean - FTE_PA1_mean
差值 (新泽西州-宾夕法尼亚州)	FTE_NJ1_mean - FTE_PA1_mean	FTE_NJ2_mean - FTE_PA2_mean	DiD = Diff_NJ - Diff_PA

(注：新泽西州和宾夕法尼亚州的样本量N应根据处理缺失值/关闭店铺后实际使用的数据报告。~331和~79是根据Card和Krueger论文中的近似值。复刻者需要从数据集中计算这些平均值。) 这张表格提供了DiD估计量最直接和直观的展示。它将总体效应在视觉上分解为各州内部的变化，然后再计算这些变化之间的差异。这使得用户能够看到原始平均值以及它们如何组合产生核心结果，正如²的摘要中所强调的：“比较新泽西州和宾夕法尼亚州（最低工资保持不变）餐馆在工资、就业和价格方面的变化，可以简单稳健地估计最低工资上涨的影响”。这是进入回归分析之前的一个基本构建模块。

C. DiD的回归框架

估计DiD的一种常用方法是通过普通最小二乘法(OLS)回归：

$$Y_{it} = \beta_0 + \beta_1 \cdot NJ_i + \beta_2 \cdot POST_t + \beta_3 \cdot (NJ_i \cdot POST_t) + \gamma' X_{it} + \epsilon_{it}$$

其中：

- Y_{it} 是店铺 i 在时间 t 的FTE就业。
- NJ_i 是代表新泽西州店铺的虚拟变量。
- $POST_t$ 是代表处理后时期的虚拟变量。
- $(NJ_i \cdot POST_t)$ 是交互项。其系数 β_3 即为DiD估计量。
- X_{it} 代表控制变量(例如, 连锁店虚拟变量、公司所有权)。

或者, 使用一阶差分数据(就业变化)：

$$\Delta Y_i = \alpha_0 + \alpha_1 \cdot NJ_i + \delta' Z_i + u_i$$

其中 $\Delta Y_i = Y_i, POST - Y_i, PRE$ (例如, DFTE)。此处, α_1 是DiD估计量。 Z_i 是不随时间变化或基线值的控制变量。Card和Krueger通常以这种方式呈现其主要结果。

V. 核心结果的分步复刻(重点关注就业)

A. 计算平均FTE就业及其变化(填充表2)

- 计算新泽西州店铺的平均FTE1和FTE2。
- 计算宾夕法尼亚州店铺的平均FTE1和FTE2。
- 按照表2所示计算差值和DiD估计量。这应与Card和Krueger初始描述性表格(例如, 他们的表2)中报告的值一致。

B. 估计就业的基本DiD模型(复刻Card和Krueger表3或类似表格)

- 模型1: 简单DiD(平均变化之差):
 - 将DFTE对NJ进行回归(常数项被隐式包含)。NJ的系数将是DiD估计量。
 - $DFTE_i = \alpha_0 + \alpha_1 \cdot NJ_i + u_i$
- 模型2: 带控制变量的DiD:
 - 将DFTE对NJ以及相关控制变量(CHAIN虚拟变量、CO_OWNED, 可能还有初始店铺规模或区域虚拟变量的度量)进行回归。
 - $DFTE_i = \alpha_0 + \alpha_1 \cdot NJ_i + \delta_1 \cdot CHAIN1_i + \delta_2 \cdot CHAIN2_i + \delta_3 \cdot CHAIN3_i + \delta_4 \cdot CO_OWNED_i + \dots + u_i$
 - NJ的系数 α_1 仍然是关注的DiD估计量。应将其大小和统计显著性与Card和Krueger报告的结果进行比较。
- RPubs上的复刻示例⁹和普林斯顿大学提供的R脚本⁴可以作为使用R进行特定编码步骤的指南。应鼓励用户理解每行代码的作用, 而不仅仅是运行它。

C. 系数和标准误的解释

- 重点关注NJ(在差分模型中)或NJ_POST(在混合模型中)的系数。这是最低工资上涨对新泽西州相对于宾夕法尼亚州FTE就业的平均影响的估计值。
- 注意标准误和p值, 以评估统计显著性。Card和Krueger发现了一个正的、有时具有统计显著性的系数, 表明就业有所增加。
- Card和Krueger还通过比较新泽西州内高工资店铺与低工资店铺的就业变化, 进行了一项

重要的内部有效性检验²。新泽西州与宾夕法尼亚州的比较可能存在一个潜在问题，即两州可能在某些未观测到的、随时间变化的方面存在差异（违反了严格的平行趋势假设）。为了解决这个问题，Card和Krueger利用了新泽西州内部的差异。一些新泽西州的店铺（高工资店铺）已经支付等于或高于新的5.05美元最低工资，而另一些店铺（低工资店铺）支付的工资较低。最低工资法直接影响的是低工资店铺。新泽西州的高工资店铺充当了额外的控制组，它们受到相同的新泽西州特定经济条件的影响，但没有受到工资规定的直接“冲击”。Card和Krueger发现，“未受最低工资影响的店铺的就业增长与宾夕法尼亚州的店铺相同，而那些不得不提高工资的店铺则增加了就业”²。这加强了因果关系的解释，因为它显示了政策在产生实际影响的地方（即对低工资店铺）的就业效应。这涉及到创建GAP变量（或类似的、表明是新泽西州低工资店铺的指示变量），并将其与后时期虚拟变量进行交互，或者比较这些组之间的DFTE。

表3: 核心回归结果的复刻(例如, Card和Krueger 1994, 表3)

- 应包含在本节(V.D)解释完模型估计之后。
- 内容: 此表应反映Card和Krueger (1994) 中一个核心的结果表, 该表展示了他们对FTE就业变化的DiD估计。它通常包含多个列, 代表不同的模型设定。

因变量: FTE就业变化 (DFTE)	(1) 基本DiD	(2) 带店铺类型控制的DiD	(3) 带完整控制 (店铺类型、区域、初始规模) 的DiD
NJ (处理组虚拟变量)	系数_1	系数_2	系数_3
	(标准误_1)	(标准误_2)	(标准误_3)
截距	系数	系数	系数
	(标准误)	(标准误)	(标准误)
连锁店: Burger King (虚拟变量)		系数	系数
		(标准误)	(标准误)
连锁店: KFC (虚拟变量)		系数	系数
		(标准误)	(标准误)
连锁店: Roy Rogers (虚拟变量)		系数	系数
		(标准误)	(标准误)
(Wendy's 为被省略的类别)			
公司直营店 (虚拟变量)		系数	系数
		(标准误)	(标准误)
Card和Krueger表中其他控制变量			系数 (标准误)
观测值 (N)	N_1	N_2	N_3
R平方	R2_1	R2_2	R2_3

(注: 系数 (Coef) 和标准误 (SE) 应使用复刻得到的结果填充。确切的控制变量和模型设定应与Card和Krueger 1994中的目标表格相匹配。)

这张表格是计量经济学复刻的核心部分。它展示了使用回归分析重现论文核心统计发现的能力。像Card和Krueger那样展示多种模型设定, 说明了研究发现在包含各种控制变量时的稳健性。在此成功意味着用户已有效地复刻了研究的主要分析部分。

VI. 进一步分析与稳健性检验(如原文所示)

A. 按初始工资分析店铺(新泽西州内部)

- 复刻Card和Krueger根据新泽西州店铺的初始工资水平(例如, 使用GAP变量或按WAGE_ST1对店铺进行分类)进行的就业变化分析。
- 这通常涉及仅针对新泽西州店铺, 将DFTE对GAP(及控制变量)进行回归, 或者在完整样本中将GAP与NJ虚拟变量进行交互。
- 这解决了之前讨论的内部有效性问题。

B. 检验对工资的影响

- 验证最低工资的提高确实提高了新泽西州的工资, 特别是对于低工资店铺。
- 这可以通过对WAGE_ST(或WAGE_ST的变化)进行DiD分析来完成。预计新泽西州会有一个显著的正向效应。

C. 检验对价格的影响(例如, PSODA、PFRY、PENTREE)

- Card和Krueger还调查了店铺是否通过提高价格将增加的劳动力成本(如果有的话)转嫁给消费者。
- 复刻针对苏打水、薯条和主菜等商品价格变化的DiD模型。

D. Card和Krueger提及的其他稳健性检验

- 原始论文可能包括其他检验(例如, 使用替代的就业衡量标准, 关注特定类型的员工)。复刻应旨在重现主要的检验。

VII. 结论性意见与资源

A. 复刻过程与发现总结

- 简要重申所采取的步骤, 以及复刻是否成功匹配了Card和Krueger的关键结果(例如, 就业的正系数)。
- 承认任何差异及其潜在原因(例如, 如前所述的微小数据差异¹⁰, 或者如果模型设定未能完美匹配而产生的轻微变动)。

B. 常见挑战与故障排除技巧

- 数据加载和清理问题。
- 正确设定DiD模型和交互项。
- 匹配Card和Krueger使用的确切样本和控制变量。
- 解释统计软件的输出。

C. 更广泛的启示与“可信度革命”

Card和Krueger (1994) 的研究是经验微观经济学中“可信度革命”的一个典型例子¹。在此类研究之前，经济学中的许多经验工作依赖于复杂的结构模型或时间序列分析，这些方法在识别因果效应方面常常受到质疑。Card和Krueger以及Angrist和Krueger等人倡导使用自然实验、工具变量和断点回归设计来实现更可信的因果推断。该论文中使用的DiD设计虽然并非全新，但被精心应用于一个显著的政策问题，突显了一个设计良好、数据透明的研究如何能够提供有力的证据。该论文的影响超出了最低工资的争论，延伸到经济学中如何进行和评估经验研究。因此，复刻这项研究也是理解这一方法论转变的一个过程。

D. 进一步阅读指引

- Card和Krueger (2000) 的后续论文 ("Minimum Wages and Employment: A Case Study of the Fast-Food Industry in New Jersey and Pennsylvania: Reply")。
- 主要的批评性文献(例如, Neumark and Wascher, 2000, "Minimum Wages and Employment: A Case Study of the Fast-Food Industry in New Jersey and Pennsylvania: Comment")。¹和⁵提及了这些争论。
- 关于DiD方法的现代计量经济学教科书, 以获得更深入的理解。

Works cited

1. Why This Year's Nobel Prize in Economics Changes Nothing on the Minimum Wage Debate, accessed June 1, 2025, <https://thedailyeconomy.org/article/why-this-years-nobel-prize-in-economics-changes-nothing-on-the-minimum-wage-debate/>
2. Minimum Wages and Employment: A Case Study of the Fast Food Industry in New Jersey and Pennsylvania, accessed June 1, 2025, <https://irs.princeton.edu/publications/working-papers/minimum-wages-and-employment-case-study-fast-food-industry-new-jersey>
3. davidcard.berkeley.edu, accessed June 1, 2025, <https://davidcard.berkeley.edu/papers/njmin-aer.pdf>
4. Reproducing Card and Krueger (AER, 1994) - Introduction to R for ..., accessed June 1, 2025, <https://libguides.princeton.edu/c.php?g=1269845&p=9343628>
5. Taking on the great Card and Krueger (1995) study on minimum wage hikes: A re-examination - Reddit, accessed June 1, 2025, https://www.reddit.com/r/badeconomics/comments/5yvlv7/taking_on_the_great_card_and_krueger_1995_study/
6. An Evaluation of Recent Evidence on the Employment Effects of Minimum and Subminimum Wages. - David Card, accessed June 1, 2025, <https://davidcard.berkeley.edu/papers/wmin%20wage%20effects.pdf>
7. NBER WORKING PAPER SERIES REVISITING THE MINIMUM WAGE-EMPLOYMENT DEBATE: THROWING OUT THE BABY WITH THE BATHWATER? David Neumark, accessed June 1, 2025, https://www.nber.org/system/files/working_papers/w18681/w18681.pdf
8. David Card - UC Berkeley Economics, accessed June 1, 2025,

<https://econ.berkeley.edu/profile/david-card>

9. Replication of Card and Kruger 1994 study - RPubS, accessed June 1, 2025,
<https://rpubs.com/AveDavis/1225123>
10. qe4policy.ec.unipi.it, accessed June 1, 2025,
http://qe4policy.ec.unipi.it/wp-content/uploads/2015/09/Code_variables_card.pdf