

金融发展与农村经济转型——来自中国县域的证据*

王雪融

内容摘要：过去四十年中国农业农村经历了迅速的发展并取得了巨大成就。2020 年底，中国打赢了脱贫攻坚战，实现了全面小康社会的阶段性目标，取得了举世瞩目的成就。乡村振兴和农业农村现代化的实现，离不开金融活水的深度滋润。然而农村金融一直以来是发展中国家农村改革所面临的重点和难点问题，如何建立和发展健全、有效的农村金融体系来支持农业现代化和农村转型，既是保障未来农业发展的核心，也是实现农业农村朝向现代化方向转型的必要条件。本文利用了 1993-2016 年分县域数据，探究金融机构发展对于农业劳动生产率的影响和作用机制，在解决了样本选择偏误和渐进差分法中的负权重问题后，实证结果发现，金融机构的设立对提升县域农业劳动生产率并无显著影响，然而开展信贷业务的金融机构可以提升县域农业劳动生产率，和上个世纪 90 年代金融机构大多只“吸储不贷”造成农村资本大量外流的局面相比，开展贷款业务通过促进资本投入提升农业劳动生产率 9.5 个百分点。此外，金融机构发展对于农业劳动生产率的影响在那些经济发展水平处于中等水平、财政自足状况相对良好，金融业务开展有一定基础的县域的促进作用更加显著。

关键词：金融机构发展 农业劳动生产率 农村经济转型

一、引言

过去四十年中国农业农村经历了迅速的发展，在保障国家粮食供应安全的基础上，为经济总体发展和全球贫困问题的解决出巨大的贡献。从 1978 年到 2022 年，农业实际总产值以年均 5.4% 的速度增长，远远高于人口年均增长 1.0% 的平均速度，成功实现人均粮食供应的大幅提升。在仅拥有世界 5% 的淡水资源下，利用不到 9% 的世界耕地，中国为接近 20% 的世界人口不仅解决了温饱问题，而且显著改善了全体国民的食物营养结构。迅速发展的农业和农村也带来农民收入的大幅提升和农村贫困的发生率的下降。到 2020 年底随着脱贫攻坚战取得全面胜利，中国成功实现现行标准下的农村贫困人口的全部脱贫，提前 10 年完成了联合国提出的 2030 可持续发展目标，为在全球范围内实现零饥饿和零贫困做出表率。

尽管中国农业和农村的发展在过去取得了巨大成就，但是未来进一步加快农业农村发展仍面临着诸多挑战。一是以劳动密集型为主的小农生产方式抑制了农业生产力的增长，导致农产品国际比较优势的迅速下降。二是农业产出的提升很大程度上是以牺牲资源环境为代价，水土资源约束和环境污染等问题与日俱增。三是农民从事农业生产的收入增长明

* 王雪融，北京大学现代农学院，邮政编码：100871，电子信箱：wangxuerong@pku.edu.cn

显放缓，带来城乡收入差异的扩大，成为制约全民共同富裕的主要短板。为解决上述问题，党的二十大提出了建设农业强国新目标，要求在加快乡村振兴的基础上逐步实现我国由农业大国向农业强国的转型，这不仅是推进农业农村现代化发展的必然趋势，也是走中国式现代化道路的必由之路。然而，要实现这一目标，首先需要加快农村生产要素市场的改革，通过推动资本和劳动的优化配置，改变农业生产方式、提高农业生产力，以推动农业生产的工业化和商业化进程。

一系列文献关注农业技术进步、资本积累、农业机械化和农村劳动力迁移在提高中国农业劳动生产率方面的作用 (Colin A Carter et al., 2003; Junichi Ito, 2010)。中国粮食和农业生产力增长的主要驱动力来自于，农村制度创新、农业技术进步、农产品市场改革和农业基础设施投入增长四大因素 (黄季焜, 2018)。其中，金融机构改革作为一项关键的制度，在推动农业农村转型中发挥着极其重要的作用。广泛认为，建立有效的金融市场是现代经济发展的基础，金融市场可以通过提供信贷的方式分散生产者投资和经营的风险，优化行业内资源的配置，鼓励市场竞争和技术进步，从而推动经济增长 (Yair Mundlak, 2005; Philip G Pardey & Julian M Alston, 2019)。鉴于此，发展农村金融体系不仅是保障未来农业发展的核心，也是实现农业农村朝向现代化方向转型的必要条件。

中国作为世界上最大并且发展最快的发展中国家，农村金融发展经历了 70 多年的从创立、改革到创新、发展的曲折过程。特别是自 1990 年代中期以来，农村金融体系逐步实现了从国有银行主导到国有私有共存、从计划经济到市场经济的转变，并在进入 21 世纪后转向开展以优化金融服务体系，创新金融服务产品和服务为主的金融创新。2020 年，全国银行网点乡镇覆盖率达 97.13%，行政村基础金融服务覆盖率达 99.97%。2021 年末，全国涉农贷款余额 43.21 万亿，普惠型涉农贷款余额 8.88 万亿。近期走向通过发展数字金融产品与服务创新来实现农村普惠金融跨越性发展道路。

农村金融在中国的迅速发展不仅为农业农村发展和转型做出了重要贡献，也积累了大量的经验和教训。一方面，农村金融机构在设立和商业化改革的过程中，在逐利属性和涉农群体高风险特征的共同驱动下，加速了农村地区资金外流 (温铁军, 2013; 周振等, 2015)，导致资本错配，不利于当地经济的发展 (Claudio EV Borio et al., 2016; Daniela Marconi & Christian Uppner, 2017)，对于贷款规模小、成本高、风险大的涉农群体来说，增加了他们从正规金融部门获得贷款的难度 (汪昌云等, 2014)。另一方面，在放松信贷和增加市场竞争力的驱动下，金融机构发挥重新分配资金的作用 (Ross Levine & Sara Zervos, 1998; James R Brown et al., 2012; Petra Valickova et al., 2015)，提高服务质量和效率，扩大服务范围 (Abdul Abiad & Ashoka Mody, 2005)，增加金融机构对农村居民的信贷供给，加快农业农村发展 (Roberta Gatti & Inessa Love, 2008; Catherine Guirkingner &

Stephen R Boucher, 2008)。学者将金融机构在农村地区发挥的正向作用分为直接效应和“涓滴效应”。前者在农村地区资本稀缺时，金融机构面对农村资本边际收益率更高的情况下，会增加对农村居民的信贷供给（姚耀军和李明珠, 2014；吕勇斌和赵培培, 2014），后者通过带动区域整体的经济增长，使得农村地区在这一过程中获益（David Dollar & Aart Kraay, 2002）。

自 20 世纪 80 年代中国农村金融体系改革以来，银行主导的县域金融机构体系发生了深刻的变化（王雪和何广文, 2019）。打破了国有银行控制下的行政垄断和农村信用社的市场垄断，各金融机构走向商业化的转型，形成了一个具有商业金融、政策性金融和合作金融机构并存的多元化、竞争性的银行体系，农信社从农行中独立出来,不再受农行领导和管理，农业银行退出县域市场，农业发展银行和邮政储蓄银行相继设立。邮政储蓄银行作为农信社、农村商业银行和农村合作银行在农村地区的最大竞争者，在县域拥有最多的分支机构。县域内的金融服务也从最初的“只吸储不放贷”转向了深化县域金融服务，提高县域信贷供给的局面。

二、文献综述

本文致力于探究农村金融机构发展对于农业劳动生产率的影响及作用机制，农业劳动生产率是农业经济发展的重要指标。根据国际农业发展基金会（IFAD）的定义，农村经济转型是指“农业劳动生产率不断提高、农产品商品化和可销售量不断提升，生产结构不断多样化，非农就业和创业机会不断增加以及基础设施和可获性服务不断改善的过程”。总体说来，农村经济转型包含两个内容：一是实现农业生产力的提升和农业生产朝向专业化和市场化方向转变；二是实现农业产业结构和就业结构的转变（既农业非农就业的创造）。虽然现有文献表明：农村经济转型背后有许多决定因素（包括自然资源禀赋、制度安排、农村政策和农村投资等），但是金融市场体系的完善和发展发挥着不可替代的作用。相关文献可以从以下几个方面进行综述。

首先是金融发展和生产力之间的关系，大部分理论和实证研究表明金融发展具有提升生产力的作用。例如，在早期研究中，Ross Levine & Sara Zervos (1998)用 47 个国家 1976-1993 年的股票市场的流动性和银行业发展来衡量金融发展程度，发现金融市场发展和生产力之间存在正向的相关关系。随后一系列的研究利用公司层面的数据，探讨金融发展与生产力的因果关系。Dan Andrews & Federico Cingano (2014)的研究表明更高的金融发展水平和更宽松的银行管制，会为企业进入市场提供更好的环境，形成有活力的竞争格局（Albert Bravo-Biosca et al., 2016），利于生产力的提升。金融改革会通过增加投资加快农村转型和发展。标准的信贷模型预测放松信贷约束往往会增加投资，从而提高收入水平

(Abhijit V Banerjee & Andrew F Newman, 1993; Philippe Aghion & Patrick Bolton, 1997; David S Evans & Boyan Jovanovic, 1989), 该影响在富裕家庭中 (Abhijit Banerjee et al., 2019), 以及拥有创业经历的家庭中更为明显 (Tamara Broderick et al., 2020), 信贷供应的增加可能会对投资产生积极的影响, 从而通过较富裕的借款人直接影响来影响家庭的收入。对于那些贫困的家庭, 即使他们不使用贷款来进行投资, 但是农村地区注入正规信贷, 可能会扩大当地的经济活动 (Huw Lloyd-Ellis & Dan Bernhardt, 2000; Anita Ghatak & Ferda Halicioglu, 2007), 相对富裕的家庭将资金从农业部门转出并用到企业投资中, 企业通过雇佣相对贫困家庭的人, 提高他们的劳动参与率以及劳动收入。

其次是金融市场对劳动力市场的影响, 一部分文献是通过 RCT 实验或者调研数据, 分析信贷供给增加对于个人的劳动收入和劳动参与率的影响 (Abhijit Banerjee et al., 2021; Dean Karlan & Jonathan Zinman, 2010; Joseph P Kaboski & Robert M Townsend, 2012), 克服劳动力市场准入的空间限制可能需要大量的前期投资, 提供信贷便利有助于个人获得在地理上相对较远的劳动力市场从事工作的机会, 进而产生高而持久的回报 (Julia Fonseca & Bernardus Van Doornik, 2022)。另一部分文献是通过企业数据, 探究当企业的信贷约束发生变化时, 对劳动力雇佣的影响 (Joe Peek & Eric S Rosengren, 2000; Gabriel Chodorow-Reich, 2014; Burcu Duygan-Bump et al., 2015)。比如当金融危机发生的时候, 当地银行的倒闭抑制企业获得融资的机会, 对企业雇佣产生巨大的负面影响 (Efraim Benmelech et al., 2019), 非农部门的工人在金融危机之后的就业岗位更是大幅减少 (Huang et al., 2011)。与之相反的是, 当企业面临的信贷约束减少时, 工人的劳动收入会增加, 对于高技能工人的效果更为明显 (Julia Fonseca & Bernardus Van Doornik, 2022)。

然而, 农业部门不同于其他部门, 其具有独特的特征, 例如大量的气候和环境不确定性以及生产中涉及的风险。因此在该部门的金融市场面临效率低、风险大、发展不完善等问题。其中信贷约束是发展中国家农民面临的最大挑战, 严重影响农业投资 (Michael R Carter & Pedro Olinto, 2003) 和农业收益 (Jeremy D Foltz, 2004)。实证结果显示, 信贷约束的缓解对农业生产力提升、农业产出提高都有显著的正向影响 (Roberta Gatti and Inessa Love, 2008; Catherine Guirkingier & Stephen R Boucher, 2008)。而在中国农村地区长期面临着农户信贷需求高度分散、经营规模狭小、缺乏担保和抵押品、融资风险大等特点。相比城市, 中国县域和农村地区的金融服务覆盖率、可得性和满意度依然较低 (王雪和何广文, 2019), 农村金融的需求主体, 如农户、新型农业经营主体和小微企业受到金融排斥的现象仍然较为普遍 (王定祥等, 2009; 何广文和 刘甜, 2018)。农村地区长期存在正规金融体系和非正规金融体系两个部分。其中, 正规金融体系整体主要是以银行机构为主体的间接融资主导的。中国县域金融市场长期由农村信用社主导, 其“小规模、本土化”的特征在

与本土信贷需求主体方面建立了紧密的长期关系，这种“软信息”的存在使得金融机构将资源集中在了具有比较优势的技术上（Andrea F Presbitero & Alberto Zazzaro, 2011）。

综上所述，国内外文献在金融发展与经济增长、金融发展与农村结构转型以及农村金融发展及其在发展中国家（包括中国）农村转型中的影响已有一定的研究，对我们理解我国新时期农村金融对农村经济转型和发展的影响具有一定借鉴意义，但大多数文献研究的重点放在金融对总体经济增长或对制造业生产力的影响上，一些文献研究农村金融对农业农村经济的影响也主要以发达国家为对象，只有少数研究以发展中国家（包括中国）为研究对象。尽管从理论上，金融的发展可以提高产出，促进劳动参与率，然而由于农业的独特性，其中生产更不确定，并且严重依赖土地作为固定投入，经济理论和实证文献尚未完全解开农业领域中金融发展和农业生产的作用机制。金融机构的发展对农业的影响可能与非农业行业有所不同，一方面它既可能放松农场的金融限制，导致新技术的采用增加，又可能对小型、生产力较低的农场进行风险对冲，从而减缓行业整合的速度。这些相互冲突的影响可能相互抵消，使金融发展对农业劳动生产率的整体影响成为一个实证问题。

本文力图在以下几个方面有所创新：第一，基于生产函数模型揭示技术进步和资本投资对于农业劳动生产率的影响，特别是金融机构发展在这其中的作用；其次，利用商业银行体系形成到现在二十多年的数据，剖析农村金融发展对于农业劳动生产率的影响和作用机理；第三，进一步挖掘金融发展对农业劳动生产率影响的异质性，丰富金融发展与农村经济转型的相关文献。

三、背景介绍

（一）中国农业劳动生产率

改革开放 40 多年以来，经济结构转型使得农业与非农业部门的劳动生产率绝对差异不断缩小，数据显示农业与非农部门的劳动生产率之比从 1:4.8 下降到了 1:3.7，但相对差异仍然显著。为在 2035 年基本实现和 2050 年全面实现农业现代化和全民共同富裕，未来仍需要转移 1.47 亿劳动力到非农部门就业，实现劳动生产率在农业和工业服务业趋同，从而提升从事农业生产的农民收入，消除城乡收入和工农收入差异。农业 GDP 占比从 1978 年的 25.82% 下降到 2022 年的 7.69%，农业就业占比从 1978 年的 70.5 下降到 2022 年 24.1%，非农部门和农业部门的劳动生产率在呈现下降趋势，2022 年该比值为 3.81。从要素投入角度来看，提高农业劳动生产率，需要提升物质资本、人力资本（薛国琴, 2018）以及技术进步（Philip G Pardey & Julian M Alston, 2019）等。物质资本投入包括农业机械化的投入、农业基础设施的投入等（薛超等, 2020），人力资本投入包括提升农民的受教育水平，加强农业职业技术教育等。提升技术水平加快创新能力，既包括提高农业服务，

如金融服务水平（张爱英和孟维福, 2021），也包括制度改革，如土地流转、户籍制度等（谭词 et al., 2022）。

（二）农村地区代表性金融机构

邮政储蓄银行的前身为邮政储汇局，早在 1986 年开始，邮政储蓄作为回笼货币的一个手段予以恢复，定位于“积聚资金、增加货币回笼和稳定金融”的功能。作为回笼货币的媒介，邮政储蓄近三分之二的网点和 42% 的储蓄余额来源于农村地区（周颖辉和彭建刚, 2010）。相比于其他商业银行，邮储银行的核心竞争力是几十年积累下来的庞大的网络优势和客户资源优势。自 1986 年开办以来，邮政储蓄的网点遍布全国大小城乡，邮递员兼任储蓄代办员，为农村居民提供便捷服务的同时，也完成了贷款服务，网点和储户优势使得邮政储蓄在社会经济生活中的地位极为稳固。早期的邮储吸收存款之后，只存不贷，吸收存款全部转存中央银行，并拥有更高的利率，享受巨大的政策红利，在存款利息和转存利率存在息差的背景下，邮储的存款规模迅速上升。与此同时，给人民银行带来了沉重的负担，1998 年后邮储转存利率甚至高于央行的再贷款利率，邮储吸收的资金只转存中央银行，造成当地存款流失，但又不产生贷款，大量的农村资金外流。在该背景下，邮储银行逐渐进行了改革，先是完善邮储转存利率机制，建立农村资金回流机制，其次将邮储从邮政中划分出来，2004 年银监会下发的《邮政储蓄机构业务管理暂行方法》明确要求邮政储蓄与邮政业务实行“财务分开和分账经营”。2006 年，银监会下发《关于加强邮政储蓄机构小额贷款业务试点管理的意见》，解决了邮政“只存不贷”的尴尬局面，批准邮政存款可以开展银团贷款业务。邮政基本具备了商业银行的雏形，在 2007 年，正式注册成立，定位服务“三农”、城乡居民和中小企业，建立中国银行业唯一的“自营+代理”运营模式，并于 2012 年完成股份制改革，2016 年正式在香港上市。

回顾邮政储蓄银行几十年的发展历程可以发现，邮储具有“高覆盖、低成本”的优势，利用几十年积累下来的网点和客户资源优势，邮储银行深扎农村市场，拥有强大的农村市场根基。相比于农业银行网点偏重于城市，以及农村涉网点集中在乡镇，邮储银行网点横贯城乡，在县域金融发挥重要的作用。在一些偏远的农村地区，邮储是当地居民唯一可以获得金融服务的机构（周颖辉和彭建刚, 2010）。数据显示，邮储有 70% 以上的网点分布于县及县以下。在 20 世纪末，各家商业银行重组收缩网点之时，邮储银行不但没有收缩网点，还继续增设至农村各个乡镇，为邮储银行拓展金融业务奠定了坚实的基础。因此，本文在分析金融机构发展时，主要以邮储银行在各个县域的进入情况为研究对象，探究金融机构设立和发展对于县域农业劳动生产率的影响。这样不仅可以避免不同金融在县域布点规则和战略不一致带来的偏差，也由于邮储银行在县域的“高覆盖率”使得结果有一定的代表性。

本文以邮储银行在县域开设网点与贷款服务为冲击，探索了金融机构设立与开展贷款业务对于县域农业劳动生产率的影响。第四部分和第五部分从理论模型出发设立实证模型，并分析其中的影响机制，特别是开始贷款业务后对于县域的资本投入与农业劳动生产率的影响。在解决实证模型中可能存在的样本选择偏误、负权重等问题之外，本文讨论了该影响的异质性，除了将样本简单的分为贫困县与非贫困县、粮食主产县和非粮食主产县外，以广义随机森林的方法（Generalized Random Forest, GRF）的方法分析该种异质性效果的来源。实证的结果表明，金融机构的设立对县域劳动生产率的提升主要源于其贷款业务的开放，相比于 2007 年前“只贷不存”造成的农村资金外流外，贷款业务的开设会增加县域地区的资本投入，促进县域的农业劳动生产率，这一效果在有一定经济和金融基础的县域中发挥了更大的服务作用。

四、理论模型

为了支持本论文的实证分析，首先通过生产增加值的模型来推导传统的金融市场的发展与劳动生产率的理论关系。在本文假设总的农业增加值为 Y ， Y 是关于资本投入 K 、土地投入 Z 、以及劳动投入 L 的方程，设置具体形式如下：

$$Y = AK^b Z^c L^\theta \quad (1)$$

其中 $\theta = 1 - b - c$ 。通常 A 表示的每单位总投入的增加值。在宏观经济增长和发展的文献，关注于总投入的回报，因此这部分也经常被称为是全要素生产率或者效率。

两边同时除上劳动投入 L ，并且取对数关系，即可以得到：

$$\ln y = a + b \ln k + c \ln z \quad (2)$$

其中 $a = \ln A$, $y = \frac{Y}{L}$, $k = \frac{K}{L}$, $z = \frac{Z}{L}$ 。通过上式可以看出，增加值农业劳动生产率有三个驱动因素：技术进步、资本密集度和土地投入。

本文假设有金融机构进入的地区和没有金融机构进入的地区代表两种的生产过程。为了区分这两种生产过程的差异，分别用 $\ln y(0)$ 和 $\ln y(1)$ 来表示如下：

$$\ln y(0) = a(0) + b(0) \ln k + c(0) \ln z$$

$$\ln y(1) = a(1) + b(1) \ln k + c(1) \ln z \quad (3)$$

用 R 来表示金融机构进入县域的概率，如果 $R=1$ 表示金融机构进入了县域，反之 $R=0$ 表示金融机构未进入到该县域。用 $\ln y$ 来表示观测到的结果，那么观测到的劳动生产率和潜在的结果之间的关系如下：

$$\ln y = R \ln y(1) + (1 - R) \ln y(0)$$

$$= \alpha(0) + \beta(0)lnk + \gamma(0)lnz + \alpha R + \beta Rlnk + \gamma lnz \quad (4)$$

其中， $\alpha \equiv a(1) - a(0)$ 衡量了有金融机构进入县域前后，该地区的农业生产力的区别， $\beta \equiv b(1) - b(0)$ 衡量的是金融机构进入前后，该地区的资本产出弹性的差异。 α 和 β 都是本文在研究中关注的重点。上述理论模型把影响农业劳动生产率的两大因素，即资本密集度以及金融改革，以一个简单透明的方式勾勒出来，为深入分析金融机构如何影响农村经济转型的机制提供了理论基础。

五、实证策略和数据

(一) 实证模型设定

将理论模型转化到可观测到的设置中，用实证的模型来表示如下：

$$lnyl_{ct} = \alpha R_{ct} + b_0 lnk_{ct} + c_0 lnz_{ct} + \beta R_{ct} \times lnk_{ct} + \gamma R_{ct} \times lnz_{ct} + c_0 + u_c + v_t + \epsilon_{ct} \quad (0)$$

其中角标 ct 代表的是第 c 个县域在 t 年， R_{ct} 是一个哑变量表示的是县域 c 在时间 t 是否有金融机构进入，如果在 t 期第一次有金融机构进入，那么从 t 期及之后，该变量均为 1。 yl_{ct} 代表的是县域 c 在 t 年的劳动生产率， k_{ct} 代表的是县域 c 在 t 年的资本密集程度，类似地， z_{ct} 代表的是县域 c 在 t 年的每单位劳动力所投入的土地， u_c 是县域的固定效应， v_t 是时间的固定效应， ϵ_{ct} 是误差项。

本文的实证策略分为三个步骤。首先第一步是估计一个双向固定效应模型（Two-way fixed effect, TWFE）。以分析金融机构设立和发展对于县域农业劳动生产率的平均影响。计量模型如下所示。

$$\Delta lnyl_{ct} = \alpha \Delta R_{ct} + b \Delta lnk_{ct} + c \Delta lnz_{ct} + \beta \Delta (R_{ct} \times lnk_{ct}) + \gamma \Delta (R_{ct} \times lnz_{ct}) + \Delta v_t + \Delta \epsilon_{ct} \quad (6)$$

其中角标 ct 代表的是第 c 个县域在 t 年， $\Delta lnyl_{ct} = lnyl_{ct} - \frac{1}{T} \sum_t lnyl_{ct}$ ， $\Delta lnk_{ct} = lnk_{ct} - \frac{1}{T} \sum_t lnk_{ct}$ ， $\Delta R_{ct} = R_{ct} - \frac{1}{T} \sum_t R_{ct}$ ， $\Delta (R_{ct} \times lnk_{ct}) = R_{ct} \times lnk_{ct} - \frac{1}{T} \sum_t R_{ct} \times lnk_{ct}$ ， $\Delta (R_{ct} \times lnz_{ct}) = R_{ct} \times lnz_{ct} - \frac{1}{T} \sum_t R_{ct} \times lnz_{ct}$ ， $\Delta v_t = v_t - \bar{v}$ ， $\Delta \epsilon_{ct} = \epsilon_{ct} - \frac{1}{T} \sum_t \epsilon_{ct}$ 。本文中最关注的两个系数是 α 和 β ，其中 α 反映的是金融机构设立对农业劳动生产率的直接影响，即金融机构设立对当地的全要素生产率的影响， β 反映的是金融机构设立对于资本加深的边际影响。

上式提供了金融机构设立对农业劳动生产率影响的基准检验，但在模型估计中面临几

个识别方面的难题。其中第一个挑战是，金融机构的设点在各个县域可能并不是随机的，进而引出样本自选择的问题，举例来说，当邮储银行在各个县域准备开展相关业务的时候可能会根据当地特定的宏观经济表现，如经济增长、产业结构来抉择设点的时间。如果较早设点的县域内部的农业劳动生产率较高，而较晚设点的县域的农业劳动生产率较低，那么这两组之间的比较可能会收到样本选择问题的干扰。尽管在前文的地图中可以看出，邮储的布局策略是相对在全国比较均匀的，到那时为了避免以上问题的干扰，使得计量估计结果更加准确，本文使用了 Imbens (2015) 的近邻匹配的方法。假设样本的数据满足独立性条件，即 $R \perp (lny|X)$ ，“ \perp ”表示的是两个随机变量之间的独立关系， X 是协变量的一组向量，在给定 X 的情况下，分配给设点组和非设点组的概率在 $(0, 1)$ 范围内，本文对 $Pr(R = 1|X) \in (0,1)$ 的表示为 logit 形式，其中 X 包含了四个协变量，即县域 GDP、财政支出、财政收入以及第一产业产值。这四个变量是县域经济发展状况的关键指标，银行前期为了提升利润往往会将网点设立在经济较为发达的地区。不同于横截面数据或者传统双重差分法的匹配过程，本文以 1993 年为基准，将改革时间在 1993 年之后的与 1993 年已经拥有邮储网点的县域进行匹配，匹配的过程中产生了一个 994 个县域的子样本。匹配的有效性在平行趋势中显现出来。除此之外，本文考虑每次设立网点时，每万人所拥有的网点数量，构造金融机构设立连续变量作为冲击，来探究每万人所拥有金融机构数量增加是否会对农业劳动生产率产生影响。

第二个挑战是，在用双向固定效应进行随时间变化的处理效应分析时候，可能会产生“负权重”的问题，即现在文献中大多数讨论的渐进双重差分法的局限性 (Kirill Borusyak et al., 2021; Clément De Chaisemartin & Xavier d'Haultfoeuille, 2020; Andrew Goodman-Bacon, 2021)。举例来看，加入 A 县和 B 县，他们在不同的时间设立了邮储银行的网点，且 A 县设点的年份要比 B 县早，因此金融机构设立对农业劳动生产率的影响会存在一定的差异。如果使用双向固定效应模型对这两个县进行差异分析，则估计出来的平均影响是这两个效应的加权平均。A 县的处理效应相对来说是准确的，因为是和那些没有邮储银行网点的县域进行了对比，但是 B 县的处理效应估计可能是不准确的，因为双向固定效应模型将 A 县的“已处理 (Already treated group)”视为了“控制组 (Control group)”，引入了不正确的信息进行了比较 (Jonathan Roth et al., 2023)。为了解决这一问题，本文使用了最新对于渐进双重的稳健估计量来检验金融机构设立对于农业劳动生产率的影响和机制检验 (Brantly Callaway & Pedro HC Sant'Anna, 2021; Clément De Chaisemartin & Xavier d'Haultfoeuille, 2020; John Gardner, 2022; Liyang Sun & Sarah Abraham, 2021)。具体来看，

假设用 l 和 k 来表示金融机构设立时间不同的两个群体， k 为“先处理的组（Early-treated group）”， l 为“后处理的组（Late-treated group）”， U 代表的群体是从未受到处理效应的群体，则金融机构设立带来的平均影响可以分解为：

$$\hat{\alpha} = \sum_{k \neq U} s_{kU} \hat{\alpha}_{kU} + \sum_{k \neq U} \sum_{l > k} [s_{kl}^k \hat{\alpha}_{kl}^k + s_{kl}^l \hat{\alpha}_{kl}^l] \quad (7)$$

其中 $\hat{\alpha}$ 代表的是金融机构设立带来的平均处理效应，表达形式为：

$$\begin{aligned} \hat{\alpha}_{kU} &= \left(\overline{\ln y}_k^{Post(k)} - \overline{\ln y}_k^{Pre(k)} \right) - \left(\overline{\ln y}_U^{Post(k)} - \overline{\ln y}_U^{Pre(k)} \right) \\ \hat{\alpha}_{kl}^k &= \left(\overline{\ln y}_k^{Mid(k,l)} - \overline{\ln y}_k^{Pre(k)} \right) - \left(\overline{\ln y}_l^{Mid(k,l)} - \overline{\ln y}_l^{Pre(k)} \right) \\ \hat{\alpha}_{kl}^l &= \left(\overline{\ln y}_k^{Post(k)} - \overline{\ln y}_k^{Pre(k)} \right) - \left(\overline{\ln y}_U^{Post(k)} - \overline{\ln y}_U^{Pre(k)} \right) \end{aligned} \quad (8)$$

含义分别是“先处理组（Early-treated group）”与“从未处理组（Never-treated group）”的差、“先处理组（Early-treated group）”与“后处理组（Late treated group）”的差、“后处理组（Late-treated group）”与“从未处理组（Never-treated group）”的差。上角标“Pre”、“Mid”以及“Post”代表的是金融机构设立的阶段，权重 $s(\cdot)$ 与实践组的大小以及每对处理变量的方差成比例， $\sum_{k \neq U} s_{kU} + \sum_{k \neq U} \sum_{l > k} [s_{kl}^k + s_{kl}^l] = 1$ 。同样的方法，本文对交乘项 $\hat{\beta}$ 也进行了这样的估计，作为本文所关注变量的稳健检验。

本文的第二步是探究金融机构设立和发展对于县域农业劳动生产率的动态影响。考虑到跨时间和跨县域变化的影响，可能会在估计中出现两个问题。一个是跨期加总问题，即金融机构设立对农业劳动生产率的影响在首次实施后可能会随着时间变化而变化，如果金融机构设立的影响在短期内是正向（或负向的），但随着时间的推移而逐渐减弱，最终的影响将会大过短期的影响，并且对金融机构设立对于农业劳动生产率的平均影响估计可能无法反映其实际的影响。第二个问题是交叉滞后的加总，即不同县域在不同时间设立金融机构。在这种情况下，金融机构设立对于农业劳动生产率的影响会因县域特征不同而不同。平均处理效应就无法很好地捕捉时间滞后和异质组之间复杂的交互效应，因此可能会缺乏适当的经济学解释意义。为了解决上述所提到的平均效应结果所产生的潜在问题，本文采用了事件研究方法（Event-study approach）来研究金融机构设立与县域农业劳动生产率的动态影响。本文首先在实证研究方法中增加金融机构设立的年份虚拟变量来探究金融机构设立对于农业劳动生产率的直接影响，其次估计金融机构设立对于资本投资边际回报的间接影响。同样地，与第一步中估计平均效应的方法类似，采取了近邻匹配的方法来解决样本选偏问题，和第一步略微不同之处在于，在估计平均效应的时候，本文以 1993 年作为基准年份，对已经设点的县域与尚未设点的县域进行匹配，此方法在动态效应估计过程中

可能会低估年份变化带来的影响，因此在每年，将已经设点的县域和尚未设点的县域进行了匹配，匹配依据的协变量与估计平均效应保持一致，进而估计出动态效应。估计模型如下所示：

$$\begin{aligned} \Delta \ln y_{ct} = & \sum_{k=-5, k \neq -1}^{10} \alpha_k \Delta \mathbf{1}\{t - t_R^0 = k\} \\ & + b \Delta \ln k_{ct} + c \Delta \ln z_{ct} + \beta \Delta (R_{ct} \times \ln k_{ct}) + \gamma \Delta (R_{ct} \times \ln z_{ct}) + \Delta v_t + \Delta \varepsilon_{ct} \end{aligned} \quad (9)$$

$$\begin{aligned} \Delta \ln y_{ct} = & \alpha \Delta R_{ct} + b \Delta \ln k_{ct} + c \Delta \ln z_{ct} \\ & + \sum_{k=-5, k \neq -1}^{10} \beta_k \Delta \mathbf{1}\{t - t_R^0 = k\} \ln k_{st} + \sum_{k=-5, k \neq -1}^{10} \gamma_k \Delta \mathbf{1}\{t - t_R^0 = k\} \ln z_{st} + \Delta v_t + \Delta \varepsilon_{ct} \end{aligned} \quad (10)$$

其中， t_R^0 是县域 c 设立金融机构的第一年，虚拟变量 $\{t - t_R^0 = k\}$ 表示的是 t 相对于 t_R^0 的时间间隔为 k 年，估计出来的 α_k 、 β_k 捕捉了金融机构设立后随着时间变化的农业劳动生产率和资本密集度的边际回报率。同样地，在基准的事件研究法的基础上，本文用 Brantly Callaway & Pedro HC Sant'Anna (2021), Clément De Chaisemartin & Xavier d'Haultfoeuille (2020), John Gardner (2022), Liyang Sun & Sarah Abraham (2021)的稳健估计量来检验金融机构设立对于农业劳动生产率的有效性。

本文的第三步是探究该影响的异质性。本文先通过分样本回归，探究贫困县和非贫困县,粮食与非粮食主产区在金融机构进入时是否存在明显差异。在本样本中贫困县有622个,粮食主产区有520个。通过核心解释变量与分组变量交乘来进行异质行分析。但在分析但在分析影响渠道方面显得比较单一，无法同时全方位捕捉各种因素的异质性影响。实践中常用的事后通过寻找分组或者交乘项进行异质性分析的方法可能导致过度拟合问题 (Jonathan MV Davis & Sara B Heller, 2020; Victor Chernozhukov et al., 2018),而且这种做法只能检验单一维度线性关系中的异质性。有鉴于此，本研究也采用最近发展的广义随机森林方法(Generalized Random Forest, GRF)，分析不同特征的县域在金融机构设立对劳动生产率的影响中是否存在差异。

(二) 数据描述

本文的数据来源涵盖中国1474个农业县1993-2016年的农业生产、金融机构的设立和

数量、县域主要经济和财政等变量。以 1993 年为起点的原因是，20 世纪 80-90 年代，随着经济体制市场化改革的推进，一批股份制商业银行诞生，到 1992 年末，初步形成了多样化、多层次的金融体系。县域的数据来源有县域和省级统计年鉴、《全国农产品收益资料汇编》等。核心变量的具体计算方法如下：

1. 农业生产数据

(1) 农业产出数据。本文所使用的农业产出，主要包含了六大类农产品，即粮棉油糖果菜主要的农作物，以及四大类畜产品，即猪牛羊禽。

(2) 资本投入产量。本文用来衡量资本服务的变量有两个思路，一是用县级的农业机械总动力作为农业生产中资本服务的近似，二是用永久库存法 (Perpetual inventory method, PIM) 来构建县级的存量。本文假设各个县的资本品类型同质，且具有相同的效率。通过固定资产折旧价值除以不变的折旧率来估计当年的固定资产投资总价值。永久库存法以不变价格衡量固定资产投资累计换算为资本存量，计算的方式如下：

$$K_t = \sum_{\tau=0}^{\infty} d_{\tau} I_{t-\tau} \quad (11)$$

其中， K_t 表示的是在 t 时期末的固定资产存量， d_{τ} 是使用年份为第 τ 年的固定资产的相对效率。 $I_{t-\tau}$ 表示 $t-\tau$ 时的固定资产投资。目前文献中常用三种相对资本效率模式来对固定资产存量进行估计，即递减平衡模式、一字型模式和直线型模式。本文选取集几何递减的资本效率模式对固定资产存量进行估计。因此，资本存量的公式可以改写为如下。

$$K_t = K_{t-1} + I_t - \delta K_{t-1} \quad (12)$$

δ 为折旧率，由于缺乏固定资产存量的基准数据，本文选取 1977 年作为基准年份，根据 Paul Schreyer (2009) 的衡量方法，以 2% 的投资回报收益以及 25% 的折旧率对固定资产投资和资本存量基准进行估计，运用几何递减的资本效率来估计每年的资本服务价值，最后选去省级固定资产投资价格指数对资本服务的价值进行平减。

(3) 劳动投入数据。本文基于农本数据，统计各个县的六大类农产品粮棉油糖果菜，以及四大类畜产品猪牛羊禽的平均劳动投入，基于各县主要农作物的产量，可以计算出对应县在当年整体的劳动投入。

(4) 土地投入数据。本文收集了用于生产六大类农作物产品和四大类畜产品的土地使用机会成本 (土地租赁价格)，用于农业生产所需土地服务的总价值和土地服务的总数量。首先，将每种农作物的总产出数量作为权重，乘上每单位产出的土地使用机会成本，即土地租金，从而得到每种农产品的土地服务总价值。而后，根据土地使用的比边际成本，计算价格指数，最后，通过土地服务总价值除上总的土地价格指数可以得到土地服务的数量。

2. 金融机构数据

本文使用的金融机构数据主要来自于中国银保监会管理委员会，涵盖了全国 342 个城市区域和 2814 个县域 1948-2021 年金融机构的相关统计情况，包含金融机构的分支机构名称、设立时间以及地理位置。本文根据民政部对行政区域的编码规则，筛选出除市辖区外的县级行政区，并且根据各地的农业产出占比，将样本集中在农业县。

3. 县域经济数据

本文使用的县域经济数据主要来自于中国县域统计年鉴。包含县域基本的经济、财政发展等变量，来作为实证部分的匹配依据和控制变量。

在本文中，县域的平均农业劳动生产率和增长率的变化如图 所示。由图可知，在过去 1993-2016 的 24 年时间里，农业县的农业劳动生产率增长了近 7 倍，年均增长率为 8.4%。

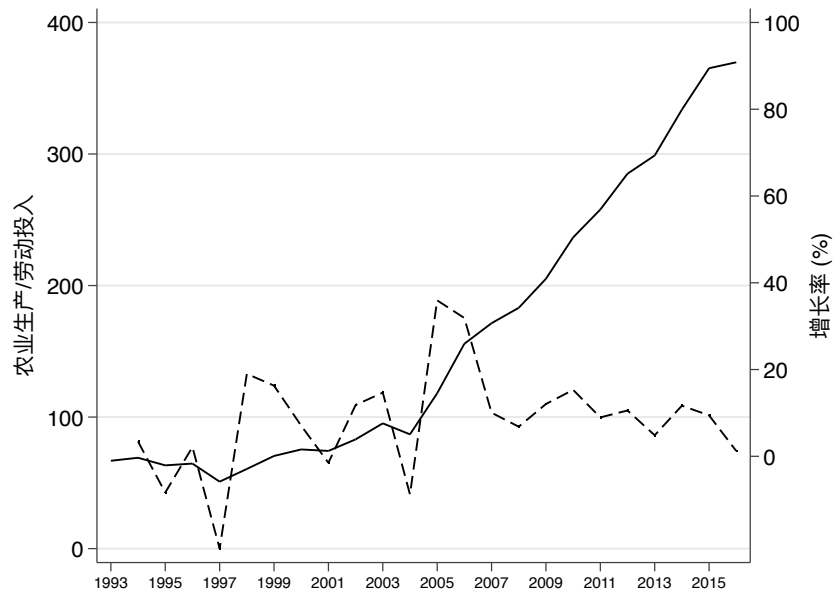
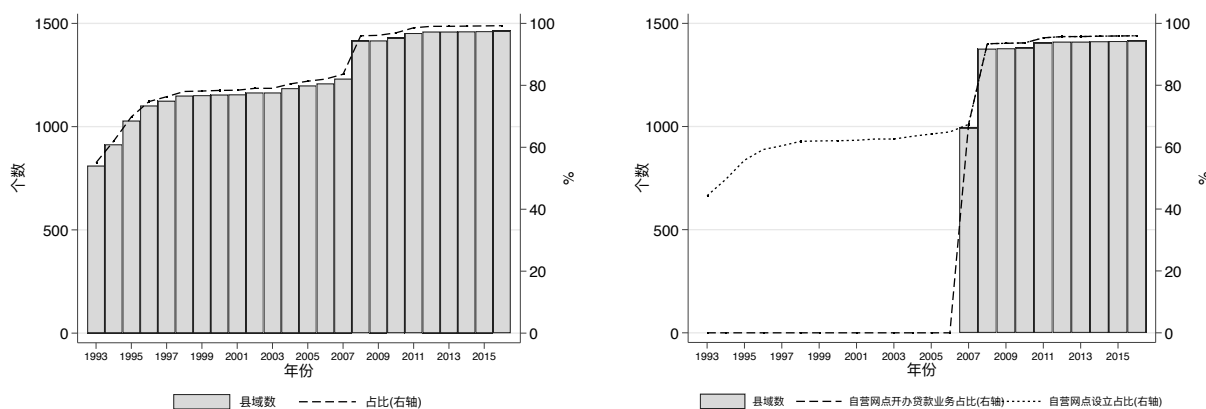


图 1 1993-2016 年农业劳动生产率

正如前文所说，邮储银行最早在 1986 年就开始在县域开展回笼资金的业务，在本样本开始前，即 1993 年前已经有 780 个县设立了邮储银行的分支机构。剩下的 683 个县在 1993 年至 2016 年期间陆续在县域设立邮储银行网点，样本中剩下 11 个县在 2016 年后才拥有邮储银行的网点，这 11 个县主要集中在西藏自治区和青海省。参考 Thorsten Beck et al. (2010)对于美国银行业改革的研究，本文根据邮储银行在每个县域第一次设立分支机构的时间是指虚拟变量，该虚拟变量在每个县域第一次设立的时间及其之后的年份设置为 1，其余年份为 0。在 1993-2016 年共有 683 个县设立网点。值得注意的是，2007 年以前，邮

储银行只开展和储蓄有关的业务，2007 年国务院和监管部门推进邮政储蓄改革的时候推进了邮储银行“自营+代理”的模式，自营网点的业务包括向客户提供包括各类贷款、存款及中间业务产品与服务在内的全面金融服务。与之相对的代理网点数量，约占总网点数量的 80%左右，其业务主要包括从事个人储蓄存款、结算类、代理类金融服务，与自营网点适用统一的业务管理制度，原则上不开办对公存款业务和资产业务。直到 2019 年 5 月，在银保监会的支持下，邮储银行代理网点才开始开展小额贷款试点业务，首批试点集中在浙江、安徽、福建、湖南、广东、云南等 6 个省份。图展示了具体在本样本时间内的邮储银行进入县域的时间点分布。可以看出在 20 世纪 90 年代前期以及 2007-2008 年是邮储银行在县域设立网点比较集中的年份。邮储银行在县域及以下网点数量占比 70%左右，营业网点已覆盖中国 99%的县。



(1) 县域设立邮储银行个数和占比

(2) 邮储银行在县域开展贷款业务分布

图 2 1993-2016 邮储银行进入县域时间分布

正如前文所说，尽管邮储在一些地区设立了网点，但是直到 2007 年才开展贷款业务。上图显示的是县域设立网点的的时间和分布，并未分离自营网点和代理网点。因此，本文进一步将目标定位于自营网点。由于无法获得 2007 年前成立的网点何时开展了贷款的业务，本文以下两种方式来处理。一是将 2007 年前后作为邮储银行整体的变化，粗略估计 2007 年邮储银行正式成立，开展贷款业务对于农业劳动生产率以及资本投入的影响。二是将目光转向于自营网点，将县域内自营网点开展贷款业务的时间设置成虚拟变量，更加准确估计贷款业务的开设对于农业劳动生产率的影响和机制分析。图显示了自营网点在县域开展贷款业务的时间。正如图中所示，点线所代表的是真实的自营网点在当地设立的时间，但 2007 年前其未开展贷款业务，虚线是其开展贷款业务的时间分布。

本文最终选取的样本涵盖了 1474 个县 1993-2016 年。其主要变量的描述数据如表 1 所示。

表 1 描述性统计：金融机构开设前 vs. 金融机构开始后

变量	含义	均值/标准差	全部样本	进入前	进入后
Agr. Labor Productivity	农业劳动生产率，用农业产出/劳动投入（天）表示	Mean	160.02	115.93	168.32
		S.D.	225.82	151.69	236.28
Capital intensity	资本密集度，用资本投入/劳动投入（天）来表示	Mean	0.48	0.21	0.53
		S.D.	0.91	0.50	0.96
Land intensity	土地劳动比，用土地投入/劳动投入（天）来表示	Mean	5.07	2.67	5.53
		S.D.	6.24	5.53	6.26
GDP	总产出（万元）	Mean	756824.77	258209.09	850795.03
		S.D.	4844589.34	537898.44	5270977.67
First GDP	第一产业产出（万元）	Mean	38.69	42.06	38.06
		S.D.	26.48	32.48	25.14
Fiscal expenditure	财政支出（万元）	Mean	112939.62	53029.36	124230.45
		S.D.	128750.90	64141.17	134617.29
Fiscal income	财政收入（万元）	Mean	86292.79	18935.04	98987.19
		S.D.	132767.49	31642.36	140515.01
Fiscal self sufficiency	财政自给率，用财政收入/财政支出比值来表示	Mean	34697.24	9280.81	39487.28
		S.D.	97363.41	23564.46	104961.51
PSBC number per 10,000 capita	每万人所拥有的邮储银行机构数量	Mean	0.19	0.00	0.23
		S.D.	0.22	0.00	0.22
All other banks number per 10,000 capita	每万人所拥有的除邮储银行之外的其他国有银行机构数	Mean	0.28	0.34	0.27
		S.D.	0.51	0.66	0.48
Observations(N)			35352.00	5606.00	29746.00

五、金融机构发展和农业劳动生产率

1. 金融机构设立（开展贷款业务）与农业劳动生产率

在具体回归时，本文选择控制时间和县域层面的双向固定效应来解决那些不随时间变动但随个体而异的遗漏变量问题，以及随时间变动但不随个体变动的遗漏变量问题，并且使用了县级层面的聚类稳健的标准误控制异方差的问题。表 2 报告了基准的回归结果。前文在描述邮储银行发展背景中强调，2007 年前，邮储的业务“只存不贷”，到了 2007 年邮储银行正式从邮政集团中剥离出来后，业务范围逐渐从个人存款拓展至个人贷款和对公业务。理论上来看，存款业务会吸收当地储蓄，而贷款业务会提升当地可利用的资金，进而对生产的投入带来影响。因此本文也观察了 2007 年后邮储银行开展贷款业务之后对于县

域农业劳动生产率的影响。表 2 展示的是 1993-2016 年间邮储银行网点设置对于农业劳动生产率的影响，并未区分最早设立的网点是“自营网点”还是“代理网点”，表 3 对比了两种方式对于开展贷款业务处理的影响。第一种方式是直接加入了 2007 年前后的虚拟变量 *post*，在模型回归中形成“三重查分”，关注各核心变量与 *post* 相乘后的影响，即 $\alpha^{PSBC} \times post$ 、 $\beta^{PSBC} \times post$ 的系数。其反映的是，2007 年开展贷款业务之后，比没有开展贷款业务之前，网点设立对于农业劳动生产率、以及资本密集度的影响。第二种方式是关注于自营网点的贷款业务，此时的 α^{PSBC} 、 β^{PSBC} 衡量的是开设贷款业务对于农业劳动生产率的直接影响和间接影响。

表 2 金融机构设立对农业劳动生产率的影响

	全样本					
	TWFE (1)	TWFE (2)	TWFE (3)	TWFE+PSM (4)	TWFE+PSM (5)	TWFE+PSM (6)
Dependent variable: logged agricultural labor productivity(ln_yl)						
b₀	0.056*** (0.011)	0.059*** (0.011)	0.045*** (0.017)	0.059*** (0.012)	0.060*** (0.013)	0.069** (0.029)
c₀	0.292*** (0.027)	0.289*** (0.028)	0.335*** (0.034)	0.313*** (0.038)	0.312*** (0.038)	0.344*** (0.062)
α^{PSBC}		0.076*** (0.020)	0.153*** (0.056)		0.016 (0.032)	0.009 (0.103)
β^{PSBC}			0.019 (0.016)			-0.009 (0.028)
γ^{PSBC}			-0.054** (0.026)			-0.034 (0.052)
Fiscal self sufficiency		-0.019 (0.017)	-0.018 (0.016)		-0.077*** (0.020)	-0.076*** (0.019)
GDP		0.057 (0.069)	0.056 (0.069)		0.053 (0.091)	0.052 (0.091)
Constant	4.330*** (0.036)	3.630*** (0.832)	3.583*** (0.834)	4.277*** (0.050)	3.866*** (1.108)	3.881*** (1.114)
R-squared	0.77	0.77	0.77	0.77	0.77	0.77
No. of observations	35208	35174	35174	23856	23822	23822
No. of counties	1467	1467	1467	994	994	994

注：“TWFE”表示双向固定效应模型，“TWFE+PSM”表示的是匹配后的样本的双向固定效应模型。括号中为聚类稳健标准误；***、**、*分别表示 1%、5%、10%的水平上显著。

表 3

基准回归结果：全样本 vs. 自营网点

	2007 年前后对比		自营网点开办贷款业务	
	TWFE	TWFE+PSM	TWFE	TWFE+PSM
	(1)	(2)	(3)	(4)
Dependent variable: logged agricultural labor productivity(ln_yl)				
b_0	0.059*** (0.017)	0.083*** (0.029)	0.044*** (0.012)	0.048*** (0.012)
c_0	0.309*** (0.033)	0.318*** (0.061)	0.328*** (0.028)	0.335*** (0.029)
α^{PSBC}	-0.016 (0.060)	-0.129 (0.105)	0.168*** (0.050)	0.195*** (0.047)
β^{PSBC}	-0.022 (0.017)	-0.044 (0.029)	0.095*** (0.022)	0.098*** (0.023)
γ^{PSBC}	0.026 (0.026)	0.022 (0.052)	-0.103*** (0.019)	-0.106*** (0.019)
$\alpha^{PSBC} \times post$	0.212*** (0.058)	0.192** (0.081)		
$\beta^{PSBC} \times post$	0.102*** (0.023)	0.091*** (0.028)		
$\gamma^{PSBC} \times post$	-0.108*** (0.019)	-0.082*** (0.023)		
Fiscal self sufficiency	-0.018 (0.016)	-0.075*** (0.019)	-0.015 (0.016)	-0.022 (0.018)
GDP	0.058 (0.068)	0.055 (0.090)	0.060 (0.068)	0.061 (0.076)
Constant	3.621*** (0.824)	3.907*** (1.101)	3.603*** (0.819)	3.582*** (0.907)
R-squared	0.78	0.77	0.78	0.77
No. of observations	35174	23822	35174	34080
No. of counties	1467	994	1467	1421

注：“TWFE”表示双向固定效应模型，“TWFE+PSM”表示的是匹配后的样本的双向固定效应模型。括号中为聚类稳健标准误；***、**、*分别表示 1%、5%、10%的水平上显著。

表 2 的第 (1) 列显示，未加入控制变量时，资本密集度和土地劳动投入比的系数分别为 0.056 和 0.292，并且在 1%的水平上显著，与模型的结果保持一致，即资本劳动比、土地劳动比是农业劳动生产率的重要驱动力，且其比值接近 2:8。第 (2) 列是加入控制变

量以及金融机构设立时间的虚拟变量，结果显示，金融机构设立对于农业劳动生产率有正向的影响。第（3）列加入了金融机构设立的虚拟变量，以及资本密集度与金融机构设立之交乘项，也是本文重点关注的模型。结果显示，金融机构的设立对资本投入并不会产生影响，即间接影响并不显著，但会直接提升当地的农业劳动生产率。（4）-（6）列是对匹配后的 994 个样本进行了分析，金融机构设立带来的直接影响因为在匹配后消失，也说明了在之前存在样本选择偏误的问题存在。为了进一步分析贷款业务开展的影响，本文对比了两种处理方式下，邮储银行开展贷款业务对于农业劳动生产率的直接和间接影响。结果在表 3 中显示。

表 3 的第（1）-（2）列展示了直接以 2007 年前后为界限，邮储银行贷款业务开展对于农业劳动生产率的直接和间接影响，结果 2007 年开展贷款业务后，会对农业劳动生产率带来直接的影响（0.212 vs.0.192），并且是通过提升资本密集度来提升劳动生产率（ $\beta^{PSBC} \times post$ 的系数为正）。第（2）列是匹配后的结果，与基础结果保持一致。第二种方式，以自营网点开办贷款业务作为外生冲击，但就贷款业务开办对于农业劳动生产率的直接和间接影响，结果如（3）和（4）列所示，此时的 α^{PSBC} 表示的是邮储银行网点开设贷款业务对于农业劳动生产率的影响，均为正向显著（0.168 vs. 0.195）， β^{PSBC} 是邮储银行网点开设贷款业务对于资本投入的影响（0.095 vs. 0.098）。两种方式的结果互相印证，肯定了贷款业务的开展对于资本投入有提升作用，进而提升县域农业劳动生产率。

基准模型的结果与已有的文献微观结果保持一致，特别是在改革开放后农村经济快速发展的时期，农村资金源源不断向外流。农民以净存款人的身份为其他经济部门贡献金融剩余（孙同全 and 潘忠, 2019）。20 世纪 90 年代的农户调研数据显示，1996-1998 年上半年农户从金融机构得到的贷款金额占存款的 13.99%（何广文, 2001）。在金融机构大量撤离农村的高峰期，1997 年底在湖北平均每户得到的贷款仅为前两年存款的 1/40（张元红, 1999）。从个体需求来看，大部分农户是农村金融机构的纯存款人而非贷款人，农户借款的来源大部分以非正规信贷为主。从金融机构供给角度来看，金融抑制问题在这一期间十分突出，信贷利率长期受到关注，被人为压低，导致金融机构不能实行差别化利率，金融机构在农村地区只能完成政府制定的贷款业务，缺少商业利润，农村信贷资金大量流出农村（徐忠 and 程恩江, 2004）。而在金融结构开展贷款业务后，文献认为金融机构贷款供给的增加会显著促进经济增长林毅夫 and 孙希芳, 2008。农村金融机构准入条件的放松、农业新型金融机构的改革也有利于提高居民的收入水平（张晓云 et al., 2016；张珩 et al., 2022）。

2. 金融机构发展程度与农业劳动生产率

更进一步，本文构建了连续变量的冲击，即每万人所拥有的可以开展贷款业务的银行

网点数量，来探究网点的数量的增加是否会影响农业劳动生产率。结果如表 4 所示。相比县域是否设立金融机构网点，邮储自营网点数量的增加并不会直接影响农业劳动生产率与资本投入。

表 4 网点数量对农业劳动生产率的影响

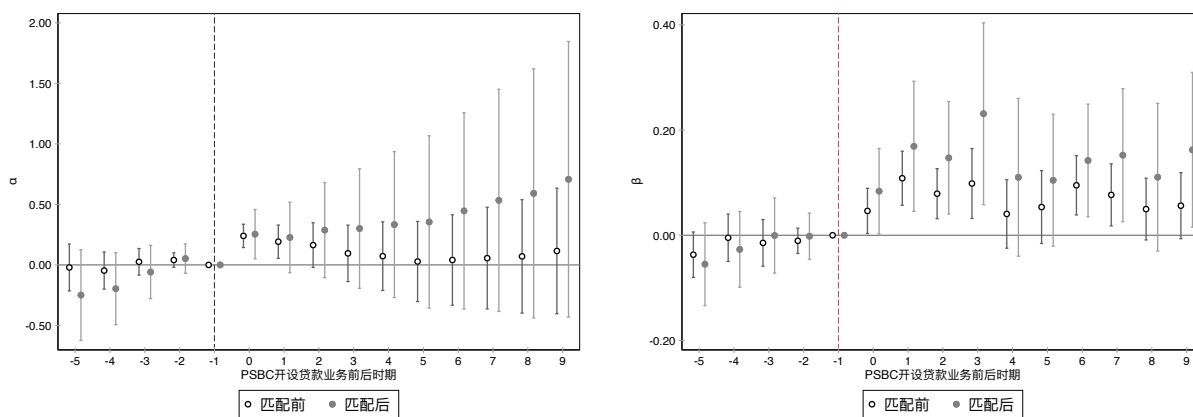
	TWFE	TWFE	TWFE+PSM	TWFE+PSM
	(1)	(2)	(3)	(4)
Dependent variable: logged agricultural labor productivity(ln_yl)				
b_0	0.040*** (0.012)	0.040*** (0.012)	0.046*** (0.013)	0.046*** (0.013)
c_0	0.320*** (0.028)	0.320*** (0.028)	0.337*** (0.029)	0.338*** (0.029)
α^{PSBC}	0.180*** (0.049)	0.175*** (0.055)	0.214*** (0.048)	0.209*** (0.053)
β^{PSBC}	0.097*** (0.022)	0.100*** (0.024)	0.093*** (0.023)	0.097*** (0.025)
γ^{PSBC}	-0.104*** (0.019)	-0.099*** (0.023)	-0.110*** (0.019)	-0.105*** (0.023)
$\alpha^{PSBC} \times$ PSBC number per 10,000 capita		-0.004 (0.024)		-0.004 (0.024)
$\beta^{PSBC} \times$ PSBC number per 10,000 capita		0.002 (0.006)		0.003 (0.006)
$\gamma^{PSBC} \times$ PSBC number per 10,000 capita		0.004 (0.011)		0.004 (0.011)
Fiscal self sufficiency	-0.019 (0.016)	-0.019 (0.016)	-0.029 (0.018)	-0.029 (0.018)
GDP	0.064 (0.069)	0.064 (0.069)	0.065 (0.077)	0.065 (0.077)
Existing number of all other banks per 10,000 capita	0.027*** (0.009)	0.027*** (0.009)	0.034*** (0.009)	0.035*** (0.009)
Constant	3.597*** (0.822)	3.600*** (0.823)	3.585*** (0.919)	3.588*** (0.919)
R-squared	0.78	0.78	0.77	0.77
No. of observations	35321	35315	34205	34199
No. of counties	1474	1474	1426	1426

注：“TWFE”表示双向固定效应模型，“TWFE+PSM”表示的是匹配后的样本的双向固定效应模型。括号中为聚类稳健标准误；***、**、*分别表示 1%、5%、10%的水平上显著。

六、稳健型检验

1. 平行趋势与动态效应检验

渐进双重差分法的使用必须满足，在金融机构开设贷款业务发生之前，处理组和控制组的变动趋势是平行的，不存在趋势性的差异。为了验证平行趋势是否满足，并进一步考察金融机构开展贷款业务对于农业劳动生产率的影响，本文采用了事件研究方法。首先用双向固定效应的方法来观测核心变量 R_{ct} 的动态变化。结果显示，在尚未匹配之前，在 95% 的置信空间下，尚未匹配的样本中确实存在平行趋势不满足的情况。然而当按照前文所示，将样本在金融机构开设贷款业务发生之前进行匹配，即可发现，金融开设贷款业务发生前 5 年，政策变量 R_{ct} 的稳健估计量大部分在 0 附近且在 5% 的显著性水平下不显著，表明平行趋势假定得到满足。在金融机构开设贷款业务之后，变量 R_{ct} 与 $R_{ct} \times lnk_{ct}$ 的稳健估计量在 5% 的显著水平下为正，但随着时间的推移效果消失。这与平均意义下的雾过保持一致，即金融机构开展贷款业务对于县域的农业劳动生产率有显著的促进作用，该结果会随着时间的推移逐渐减弱。与此同时，金融机构开展贷款业务是通过增加当地的资本投入来影响农业劳动生产率的，但该结果存在较强的持续性。



(1) 直接影响的动态效果

(2) 间接影响的动态效果

图 3 传统金融机构发展对于农村经济转型的动态影响

注：本文考虑了一个 14 年的时间窗口，5 年前和 9 年后，原因是因为邮储银行在 2007 年才开始贷款业务，样本结束的年份为 2016 年，这其中是 9 年的时间。将金融机构开设贷款业务的前一年作为基准年份，蓝线和红线分别代表了没有匹配和匹配后的 95% 的置信区间。

2. 同质性处理效应假设检验

本文首先采用 Andrew Goodman-Bacon (2021) 提出的分解方法对于本文中邮储银行开展贷款业务的时间进行偏误诊断，Goodman-Bacon 认为，所有的 2×2 DID 系数的权重为正，但部分系数符号为负，导致多时点 DID 估计系数存在偏差。他们将双向固定效应估

计量分解为“先处理组 (Early-treated) vs 后处理组 (Late-treated)”、“后处理组 (Late-treated) vs. 先处理组 (Early treated)”,其中,“后处理组 (Late-treated) vs. 先处理组 (Early-treated)”的对照组中包含了处理效应,由表的结果可知,“后处理组 (Late-treated) vs. 先处理组 (Early-treated)”的平均处理效应为-0.092,负权重占 TWFE 的总量的 26.8%,相对来说是较高的比例,说明由异质性处理效应带来的偏误较大,估计结果不稳健。由于上述 Bacon 分解的方法对于数据的平衡性要求较高,仅在具有强平衡面板数据的情况下有效,对于具有不平衡面板数据的情况,为了进行 Bacon 分解并找到附加每个系数的权重。本文进一步地使用 Clément De Chaisemartin & Xavier d'Haultfoeuille (2020)提出的方法考察双向固定效应的权重是否稳健,De Chaisemartin 和 D'Haultfoeuille 认为多时点 DID 系数是组别-时间处理效应的加权平均,组别-时间处理效应为正,但部分权重为负,导致了最终估计结果有偏。结果显示在 13593 个权重中,有 8463 个权重为正,而多达 5130 个权重为负。度量估计量在异质性处理效应下稳健性的两个指标,分别为 0.0462 和 0.0783。这两个系数估计在异质处理下接近于 0,说明双向固定效应估计结果并不稳健,存在异质性的处理效应。

表 5 Bacon 分解结果

	alpha	权重
后处理组 vs 先处理组 (Earlier T vs. Later C)	-0.092	0.268
处理组内 (Within)	0.451	0.458
处理组 vs 从未处理组 (T vs. Never treated)	-0.028	0.274

为了处理以上双向固定效应模型中可能存在的“负权重”问题,本文进一步采用 Brantly Callaway & Pedro HC Sant’Anna, 2021, Liyang Sun & Sarah Abraham (2021), John Gardner, 2022 以及 Clément De Chaisemartin & Xavier d’Haultfoeuille (2020)提出的四类稳健估计量检验金融机构开设贷款业务对于农业劳动生产率的直接和间接影响。

由 Callaway & Sant’Anna (以下简称“CS2021”)、Sun & Abraham (以下简称“SA2021”)开发这两个估计方法都是依赖于首先估计单个队列-时间-特异性处理效应,具体估计方法如下式所示:

$$ATT(g, t) = \mathbb{E}[Y_{i,\tau}(1) - Y_{i,\tau}(0) | E_i = g] \text{ (CS2021)} \quad (13)$$

$$CATT(g, l) = \mathbb{E}[Y_{i,g+l} - Y_{i,g+l}^\infty | E_i = g] \text{ (SA2021)} \quad (14)$$

考虑处理效应的异质性，然后将其汇总以产生总体处理效应的度量。二者在灵活性、协变量的调节、控制组的选择和推断方面存在方法上的差异。CS2021 估计值的变体可以归结为通过简单的 $2 \times 2s$ 和干净的控制来估计队列-时间-特定的处理效应，例如，一个在时间 g 接受处理的处理效应可以通过 $y_{it} = \alpha_1^{g,t} + \alpha_2^{g,t} \cdot \mathbb{I}\{E_i = g\} + \alpha_3^{g,t} \cdot \mathbb{I}\{t = \tau\} + \beta^{g,t} \cdot (\mathbb{I}\{E_i = g\} \times \mathbb{I}\{t = \tau\})$ 来估计，这些观测值来自于的被处理集合，或者来自于一组干净的控制集合。CS 允许将尚未处理的、最后处理的或从未处理的作为干净的对照，并在无预期条件与无平行趋势下估计有效。而 SA (2021) 提出的是一个完全参数化的基于回归的估计，该估计使用在相对时间指标和队列指标中饱和的交互式模型设定，即联合估计全部的群组特定的相对时间处理效应，即 $y_{it} = \alpha_i + \lambda_t + \sum_{g \in C} \sum_{l \neq -1} \mu_{g,l} (\mathbb{I}\{E_i = g\} \cdot D_{it}^l) + \epsilon_{it}$ 。在无条件的平行趋势和无预期的情况下，估计是一致的。与 CA2021 稍有不同的是，SA2021 在估计的过程中，放弃了那些总是被处理 (Always-treated) 的个体，只留下了那些从未被处理或者最后被处理的个体。

Gardner(以下简称 G2021)开发的两阶段回归方法 (Two-Stage Regression Approach)，采取了“插补”的形式，通过两步获得无偏的估计系数。第一步，保留 $D_{gp} = 0$ 的样本，用 Y_{gpit} 对 λ_g 和 γ_p 进行回归；第二个阶段，在全样本里将 $Y_{gpit} - \hat{\lambda}_g - \hat{\gamma}_t$ 对 D_{gp} 进行回归，最终估计系数识别的是受处理个体的平均处理效应 (王鹏超和韩立彬，2023)。De Chaisemartin 和 d'Haultfoeuille (2020) 提出了新的估计量，以下简称为 CH2020，即将政策处理从无到有、从有到无两个正负方向的效果进行加权平均，而将两期的处理不变 (如：两期都是 1 或两期都是 0) 的样本作为控制组。对比多期 DID 模型，仅考虑了政策从无到有的情况。 $W_{TC} = \sum_{t=1}^T \left(\frac{N_{1,0,t}}{N_S} DID_{+,t} + \frac{N_{0,1,t}}{N_S} DID_{-,t} \right)$ ， $DID_{+,t}$ 指的是政策从 0 到 1 的平均值减去两期都是 0 的平均结果， $DID_{-,t}$ 指的是两组都是 1 的平均结果减去政策从 0 到 1 的平均结果。CH2020 的共同趋势假设相对前面三种方法较为严格，其假设两期都是 0 或者两期都是 1 的组在 t 和 $t-1$ 都无政策的组有共同趋势，在 t 和 $t-1$ 都有政策的组有共同趋势。

值得注意的是，本文关注金融机构开设贷款业务对于农业劳动生产率的直接和间接影响，直接影响可以根据上述方法直接估计出，但间接影响暂时无法用上述方法解决，只有 G2021 的方法提供了观测间接影响的途径。本文试图基于 CS2021 的方法，两步来估计间接的影响。本文在第一步将分散的处理过程分割成一组子集，即每个组在每个队列-时间的平均处理效应，即 $ATT(g, t)$ ， $ATT(g, t) = [EY(g)_t - EY(NT)_t] - [EY(g)_{t-1} - EY(NT)_{t-1}]$ 。其中，NT 代表的是“尚未处理组 (Not-yet treated)”，可以作为一个良好的对照组， $EY(g)_t$ 是在时间 t 组别 g 的预期结果值，第一个括号里计算了时间 t 时的结果的差异，第二个括号计算了时间 $t-1$ 时的结果差异，即冲击发生之前的时期。通过每个组别

的加总可以得到直接影响的估计，以及 $\widehat{ATT}(g, t)$ 。第二步在第一步的残差的基础上，再估计出此时的间接影响。

表 6 展示了 Callaway & Sant’Anna (CS2021)、Sun & Abraham (SA2021)、Gardner, (G2021) 以及 De Chaisemartin & d’Haultfoeuille (CH2020) 提出的四类稳健估计量，并汇报了各个估计量的标准误、t 值/z 值，95%的置信空间与观测值。由于各个方法在具体估计时使用的方法不同，会自动减掉某些观测值，使得表中的观测值数量也略有不同。从表中的结果可知，金融机构开设贷款业务对于农业劳动生产率的直接影响均为正，SA2021 与 CS2021 的估计量与 TWFE 的结果接近，且显著。而 CH2020 和 G2021 的方法估计出来结果偏小，且不显著。同样地，对于间接影响的估计，结果显示，CS2021 和 G2021 的方法都表示资本密集度的间接效果较为显著，然而土地的作用存在相反的结果，并不稳定。总的来说，本文用其他稳健估计量对近期对于传统的双向固定效应模型方法的质疑进行了回应，从平均意义上来说，金融机构开展贷款业务会显著提升农业劳动生产率的水平，并且主要是通过提升县域的资本投资这一渠道来实现的。

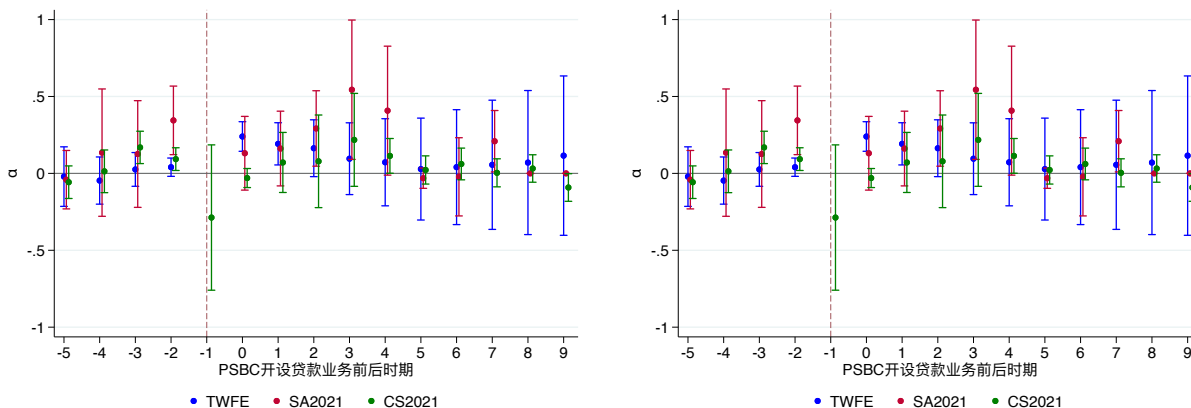
表 6 稳健估计量

Panel A.	α	Std. err	t/z	95% confidence interval		Num. of Obs.
TWFE	0.168	0.050	3.37	0.070	0.265	35174
CS2021	0.116	0.048	2.41	0.022	0.211	34003
SA2021	0.150	0.047	3.18	0.057	0.242	32568
G2021	0.059	0.070	0.84	-0.078	0.197	35174
CH2020	0.014	0.019		-0.023	0.052	1408 swithers
Panel B.	β	Std. err	t/z	95% confidence interval		Num. of Obs.
TWFE	0.095	0.022	4.34	0.052	0.137	35174
CS2021	0.106	0.051	2.09	0.006	0.207	34058
G2021	0.051	0.031	1.65	-0.009	0.112	35174
Panel C.	γ	Std. err	t/z	95% confidence interval		Num. of Obs.
TWFE	-0.103	0.019	-5.53	-0.140	-0.067	35174
CS2021	0.107	0.049	2.18	0.011	0.202	34068
G2021	-0.067	0.021	-3.21	-0.108	-0.026	35174

注：“TWFE”代表的是双向固定效应模型，“CS2021”是 Callaway and Sant’Anna (2021) 提出的估计量，

“SA2021”是 Sun and Abraham (2021) 提出的估计量, “G2021”是 Gradner (2021) 的估计量, “CH2020”是 De Chaisemartin 和 d’Haultfoeuille (2020)提出的估计量, 不报告 t 值或 z 值的估计量。

与前文在估计静态的平均效果保持一致, 本文在事件研究的过程中, 除了包括原来的 TWFE OLS 的估计外, 也包括了最新的渐进 DID 或者多期 DID 的稳健估计方法。图 4 分别展示了金融机构开设贷款业务的直接动态影响与间接动态影响的结果。与基准回归结果一致, 即金融机构开设贷款业务会直接影响农业劳动生产率。



(1) 直接影响的动态效果

(2) 间接影响的动态效果

图 4 传统金融机构发展对于农村经济转型的动态影响: 稳健估计量

注: 本文考虑的是 5 年前和 9 年后的时间窗口。“TWFE”代表的是双向固定效应模型, “CS2021”是 Callaway and Sant’Anna (2021) 提出的估计量, “SA2021”是 Sun and Abraham (2021) 提出的估计量, “G2021”是 Gradner (2021) 的估计量, 线表示 95%的置信空间。

七、异质性分析

本文想要理解异质性的来源, 采用了 Susan Athey and Stefan Wager (2021)提出的广义随机森林方法 (Generalized Random Forest, GRF) 来确定具有何种特征的县域会和金融机构开设贷款业务有更大的相关性。随机森林(random forest)是一种用以估计条件期望 $\mu(x) = \mathbb{E}[Y_i|X_i = x]$ 的统计学习方法(Leo Breiman, 2001)。随机森林顾名思义是用随机的方式建立一个森林, 森林里面有很多的决策树组成。每有一个新的样本进入时, 森林中的每一颗决策树将判断样本的分类。Susan Athey & Stefan Wager (2021)在随机森林的基础上, 提出了广义随机森林 (Generalized Random Forest, GRF)。这种方法的目的是, 评估异质性的效果。从计量经济学角度讲, 这是一种非参的估计方法。GRF 与局部最大似然估计 (local maximum likelihood estimation) 想法有相似性, 然而与传统的核权重 (kernel

weighting) 不同, GRF 方法中, 点与点之间的相似度权重不取决于所有维度协变量之间平均距离的最小化, 而是取决于他们在随机森林递归分组中出现在同一片叶结点(leaf)上的频率。因此 GRF 能够通过数据驱动发现异质性, 并克服维度诅咒, 实现高维度异质性分析。从机器学习算法角度讲, GRF 基于随机森林算法, 但与机器学习中常用的回归森林(regression forest) 算法目标不同。回归森林的分割以最优化预测准确度为目标, 而 GRF 的分割以最大化目标参数(例如因果系数)的异质性为目标。

为了避免过度拟合, GRF 提出了一种可信的(honest)纠偏方法, 即将训练集(training set) 随机分为两组——一组用于根据最大化异质性原则来将观测值分组, 另一组根据获得的分组原则来估计处理效应。在大样本中多次重复以上步骤, 即可获得稳定的分组结构, 并通过多次的平均值获得无偏的异质性处理效应。最后使用测试集(test set)来检验算法是否能够在从未用来训练树分割结构和估计处理效应的数据集上成功发现处理效应的异质性。

GRF 在偏差和方差等方面, 都要比近邻匹配的方式表现更优。并且为三种统计方法, 即非参分位数回归、条件平均部分效应估计、工具变量估计异质性处理效应, 提供更有效的分析。下面我们介绍具体的实现步骤如下:

假设我们有一组数据 $(X_i, O_i) \in \mathcal{X} \times \mathcal{O}, i = 1, \dots, n$. 其中 X_i 代表的是一组协变量我们希望用于估计真实分布的参数, O_i 表示的是一组可以观察到的用于预测分布参数的变量。我们想要预测的是 $\theta(\cdot)$, 关于 x 的一个函数。参数被定义成“局部估计方程”的形式:

$$\mathbb{E}[\psi_{\theta(x), v(x)}(O_i) | X_i = x] = 0 \text{ for all } x \quad (15)$$

其中 $\psi(\cdot)$ 是打分函数, $v(x)$ 是冗杂参数。如果我们假设条件分布 $O_i | X_i = x$ 的密度为 $f_{\theta(x), v(x)}(\cdot)$, 那么 $\psi_{\theta(x), v(x)}(O) = \nabla \log(f_{\theta(x), v(x)}(O_i))$, 这就定义了局部最大似然参数 $((\theta(x), v(x)))$ 。

经典的随机森林是由一颗颗随机树组成的, 每棵小树可以由 T_1, \dots, T_B 来表示。为了获得对新的点 x 的预测, 我们需要将 x 逐步推到每棵小树上, 直到它到达某个与预测相关联的终端节点。假设在树 b 得到的预测值为 $\widehat{\mu}_b$, 那么随机森林对于 x 的预测值就是均值 $\frac{1}{B} \sum_{b=1}^B \widehat{\mu}_b$ 。而在广义随机森林模型中, 在我们将 x 逐步推到每棵小树上的时候, 我们不是寻找与终端节点相关的预测, 而是看哪些训练集与 x 落入到了相同的终端节点中。对于每棵树 T_b , 用 $L_b(x)$ 来表示一系列同 x 落入到相同节点的训练集。对于每个 $i = 1, \dots, n$, 可以定义:

$$\alpha_{b_i}(x) = \frac{1\{X_i \in L_b(x)\}}{|L_b(x)|}$$

$$\alpha_i(x) = \frac{1}{B} \sum_{b=1}^B \alpha_{b_i}(x)$$

(16)

可以将 $\alpha_i(x) \in [0,1]$ 视为训练样本集样本 X_i 与 x 的相似度。 X_i 和 x 越经常落入相同的节点中，那 $\alpha_{b_i}(x)$ 越接近 1。值得注意的是， $\sum_{i=1}^n \alpha_i(x) = 1$ ，可以将 α_i 视为权重函数，该权重表示我们在预测 x 的时候对于每个训练样本的重视程度。一旦当 α_i 确定之后，通过求解最小化问题，可以得到 $\hat{\theta}(x)$:

$$(\hat{\theta}(x), \hat{\nu}(x)) \in \operatorname{argmin}_{\theta, \nu} \|\sum_{i=1}^n \alpha_i(x) \psi_{\theta, \nu}(O_i)\|_2 \quad (17)$$

在 GRF 中，一个关键问题是如何决定每棵树的增长方向。和经典的随机森林树方法稍有不同的地方在于判断标准。对于数据集 J 中的样本数据，以及协变量空间 $N \subseteq \mathcal{X}$ ，定义估计量为：

$$(\hat{\theta}_N, \hat{\theta}_N)(J) \in \operatorname{argmin}_{\theta, \nu} \|\sum_{\{i \in J; X_i \in N\}} \psi_{\theta, \nu}(O_i)\|_2 \quad (18)$$

假设有同样的父辈节点 P ，想把该节点分成两个子节点 C_1 和 C_2 。一个合理的判断标准就是最小化 θ 的预期平方损失（error loss），但由于 $\theta(x)$ 是通过矩估计的方式来识别的，而不是最小化损失，所以很难对其进行无偏估计。Athey et al. (2019)通过最大化

$$\Delta(C_1, C_2) = \frac{n_{C_1} n_{C_2}}{n_p^2} [\hat{\theta}_{C_1}(J) - \hat{\theta}_{C_2}(J)]^2 \quad (19)$$

其中 n_{C_1} ， n_{C_2} ， n_p 是进入到节点 C_1 ， C_2 以及 P 的点数量。

对于所有的拆分，都需要计算 $\Delta(C_1, C_2)$ ，为避免这种情况的产生，需要对 $\hat{\theta}_{C_1}$ 和 $\hat{\theta}_{C_2}$ 进行梯度的估计。

为防止过度拟合，本文采取 Susan Athey and Stefan Wager (2021)的可信方法(honest approach)。具体而言包括以下几个步骤：（1）是在总样本中不放回地随机取出 20%的字样本；（2）用回归树的方法根据 X 的取值来将观测值分组；（3）当分组完成后，把剩下的一半子样本根据当前组别分组；（4）在每组里面用新的样本来计算处理效应（ $\hat{\tau}_b = \bar{y}_{T,b} - \bar{y}_{C,b}$ ）；（5）再用剩下的 80%的样本，根据样本的 X 取值将样本逐一分配到对应的组别中，用该组别对应的 $\hat{\tau}_b$ 来估计预测的处理效应。（6）用大量样本重复以上步骤，以平均值的方式获取到对应的处理效应。

由此，本文可以识别出处理效应在不同类型群体之间的分布。更进一步，按照 GRF 分组中协变量的重要性排序，即可以得到对处理效应异质性影响的重要性排序。保持其他变量不变，预测条件平均处理效应(CATE)如何随着某一变量的变化而改变，可以帮助量化不同维度异质性的影响模式。本文借助 GRF 的方法揭示金融机构开设贷款业务影响农业劳动生产率的效果在不同特征的县域之间的差异性。

在选择县域特征的时候，本文将基础模型回归中的协变量涵盖进来，具体包括了县域的生产总值（total gdp）、财政赤字强度(fiscal self sufficiency)、县域除邮储银行之外的其

他国有银行每万人所拥有的网点数量(other bank number)。本文根据 Athey (2019) 的方法，下图展示了构建的 GRF 中其中一棵树的示例，方框中显示的是落入该叶子中的观测量，箭头上的数值代表的是协变量满足的条件。

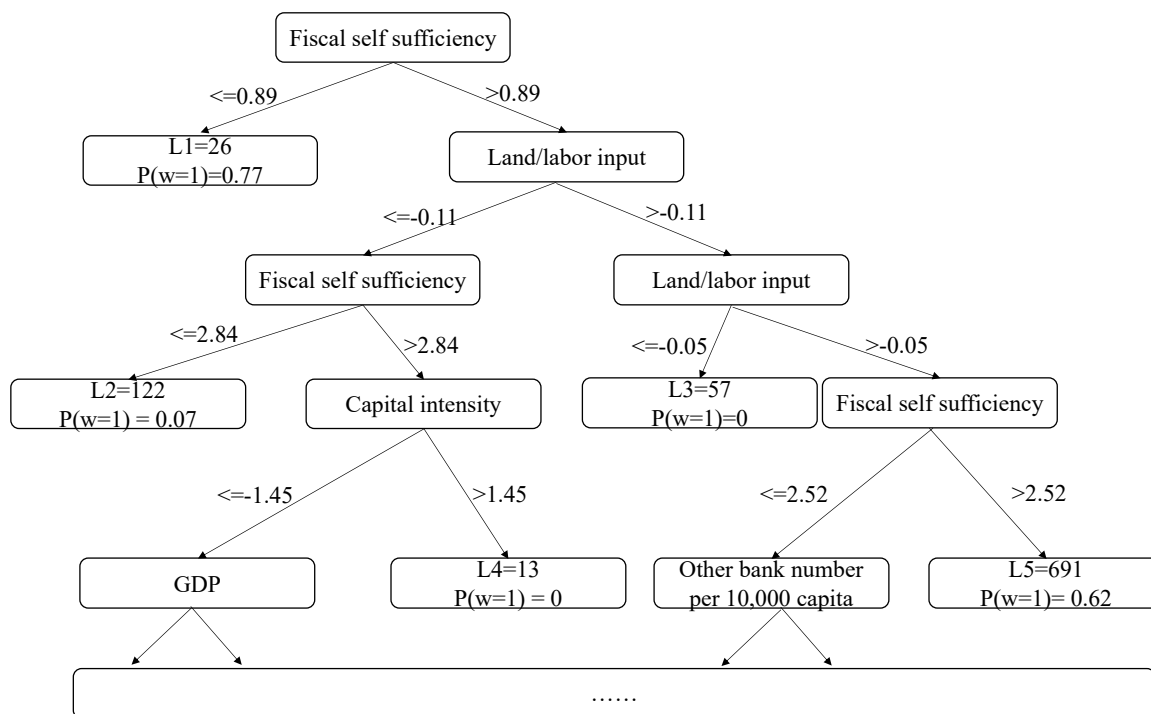


图 5 决策树示例

值得注意的是，图中树的叶子数和协变量数不相同，GRF 对每个协变量在森林中的“重要性”进行了排序，即变量在森林中作为分割变量使用的频率。其相对重要程度在表中显示。可以看出，当地的财政状况、土地劳动投入比、每万人拥有的其他国有银行数量是相对较为重要的变量。

表 7 县域各经济指标的重要性

变量	含义	重要程度
log_fiscself	财政收入/财政支出（对数形式）	0.3901
log_ll	土地投入/劳动力投入	0.1952
log_number_all_otherbank	每万人所拥有的除邮储银行之外的其他国有银行数量	0.1015
log_kl	资本投入/劳动投入	0.0500
log_gdp	县域生产总值	0.0441
FR_kl_new	虚拟变量 FR_PSBC 与资本劳动比的乘积	0.1424
FR_ll_new	虚拟变量 FR_PSBC 与土地劳动比的乘积	0.0767

在生成了一组树之后，本文使用森林根据协变量的特定值来创建治疗效应的预测。例

如，如果条件于 x_p ，那么在第 b 颗树，就可以计算出训练样本中第 i 个观测值的树权重 $\alpha_{bi}(x_p)$:

$$\alpha_{bi}(x_p) = \frac{1(\{x_i \in \text{leaf}_b(x_p)\})}{(\#x_i \in \text{leaf}_b(x_p))} \quad (20)$$

在此，本文使用图 中的示例树来说明如何计算该树的权重。比如用一个向量来描述一个县对应的各个特征： $x_m = (3, -0.2, 2.6)$ ，分别对应的是该县的财政状况、土地/劳动投入比，资本/劳动投入比，由图中的树可知，该样本落入了图中 L4 的位置，共有 13 个样本。这些观测值的权重为 $\alpha_{ET,i}(x_m) = \frac{1}{12}$ ，其他没有落入该叶子中的观测值的权重赋予 0。接着，对森林中的每颗树 B 进行树权重的计算，因此，第 i 个观测值的最终森林权重 $\alpha_i(x_p)$ 为:

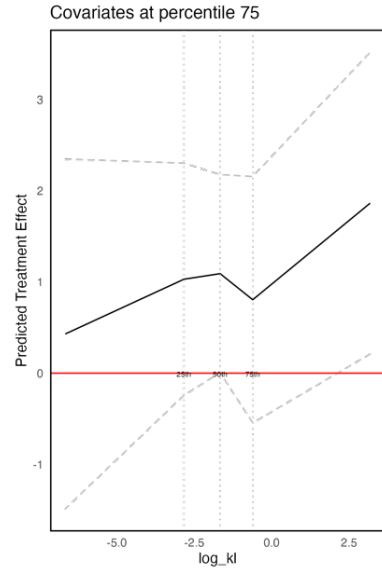
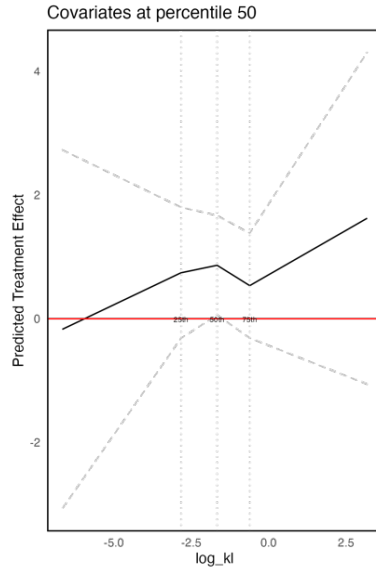
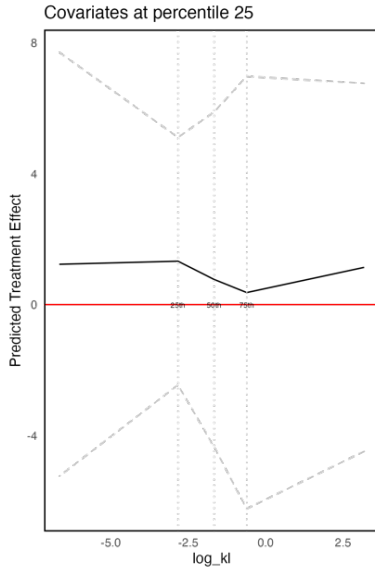
$$\alpha_i(x_p) = \frac{\sum_{b=1}^B \alpha_{bi}(x_p)}{B} \quad (21)$$

权重较大的观测值与 x_p 更相似，因为它们通常与 x_p 落在同一个叶子结点上，可以利用这些权重来估计 $\tau(x_p)$ ，在给定 x_p 条件下的处理效应，通过求解以下的加权矩方程来获得。

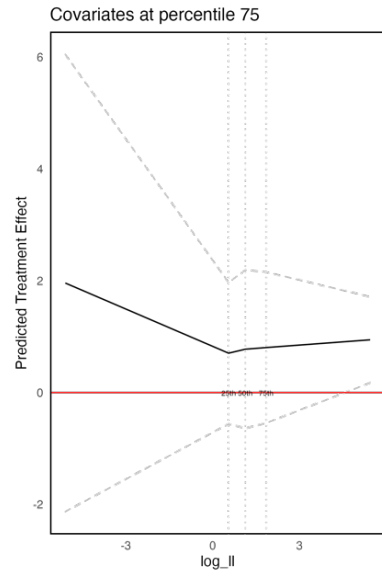
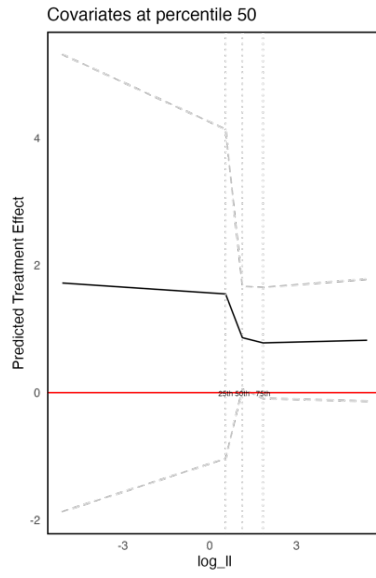
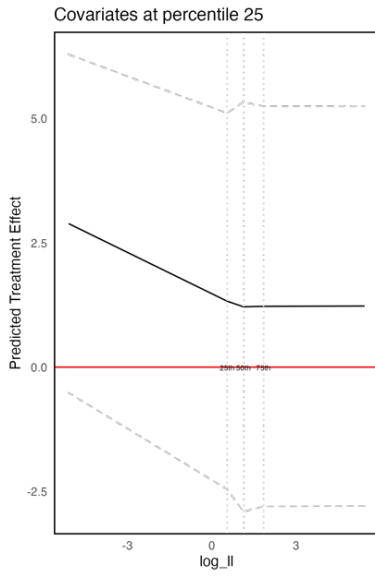
$$\sum_i^n \alpha_i(x_p) \psi_{\tau(x_p)}(Y, W) = \sum_i^n \alpha_i(x_p) [w_i - \bar{w}_\alpha] [(y_i - \bar{y}_\alpha) - \tau(x_p)(w_i - \bar{w}_\alpha)] = 0 \quad (20)$$

其中， $\bar{w}_\alpha = \sum_{i=1}^n \alpha_i(x_p) w_i$ ， $\bar{y}_\alpha = \sum_{i=1}^n \alpha_i(x_p) y_i$ 。从上述表达式中也可以看出 $\hat{\tau}(x_p)$ 类似于加权最小二乘估计量。其方差表达式为 $Var(\hat{\tau}(x_p)) = \frac{var[\sum_{i=1}^n \alpha_i(x_p) \psi_{\tau(x_p)}(Y, W)]}{E(W^2|X = x_p) - E(W|X = x_p)^2}$ 。

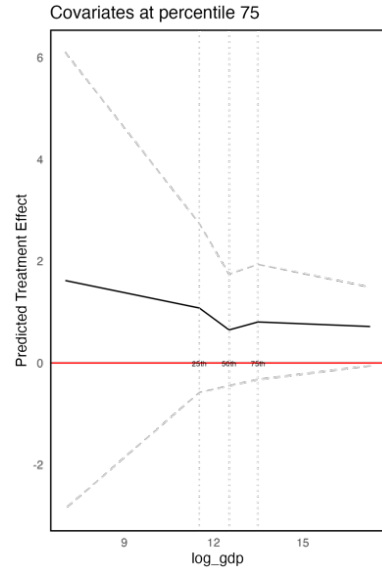
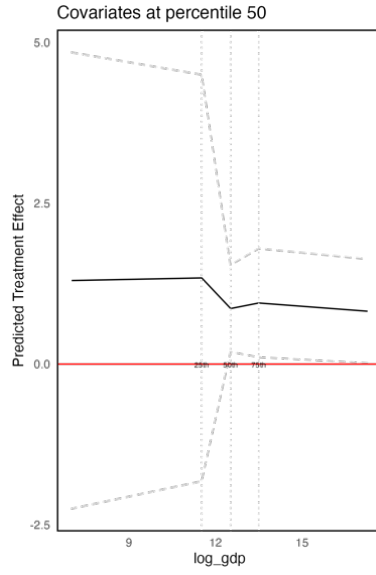
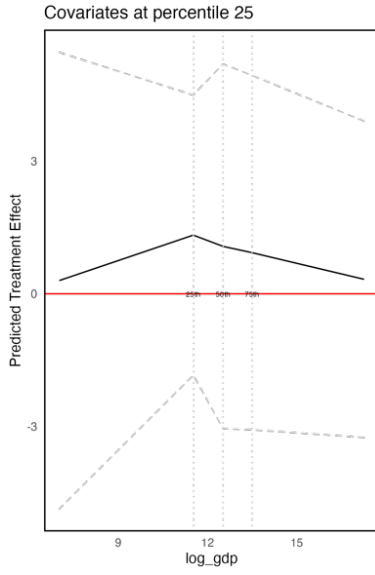
图 6 是对 GRF 结果的总结，从图中可以看出，每一行展示的是每个协变量得出的具体结果。每一列是将其他协变量控制在不同百分位上的时候预测出来的处理效应。黑色实线代表的是估计出来的系数，虚线代表的是 95% 的置信空间。第一列是将除了横坐标上的其他协变量控制在 25% 的百分位上，第二、三列分别是控制在 50%、75% 的百分位上。从图中的对比结果可以看出，在每个变量的中间的图会看到显著的影响，即当其他变量控制在 50% 的水平上的时候，金融机构开设贷款业务会有较为明显的预测效果。以县域 GDP 生产总值为例，当其他的变量在 50% 的水平上时，当县域 GDP 水平处于 50% 以上水平时，金融机构开设贷款业务会显著提升该县域的农业劳动生产率。其他变量可做类似的解释分析。总的来看，金融机构开设贷款业务对那些经济发展水平在中等水平、财政自足状况良好、金融业务开展有一定水平的县域的农业劳动生产率的提升有较大的促进作用。该结论与朱一鸣和王伟 (2017)、王雪和何广文 (2019) 保持一致，即在经济、金融基础设施有一定基础的县域内，银行金融服务的作用才能被显现出来。



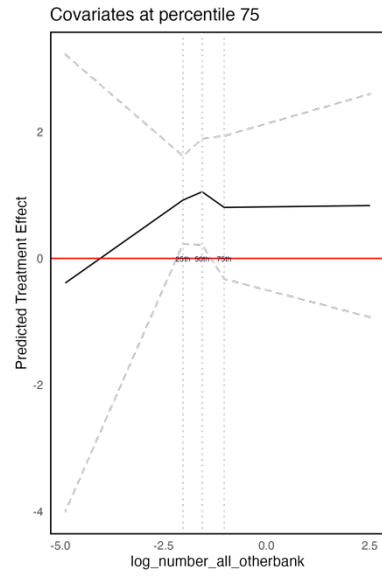
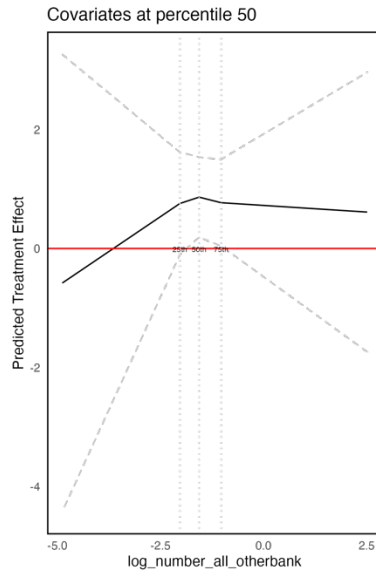
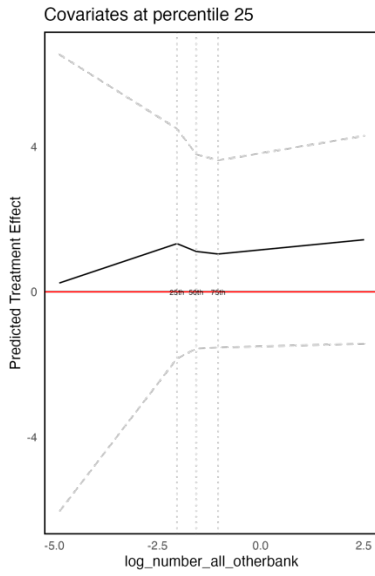
(1) 资本劳动比



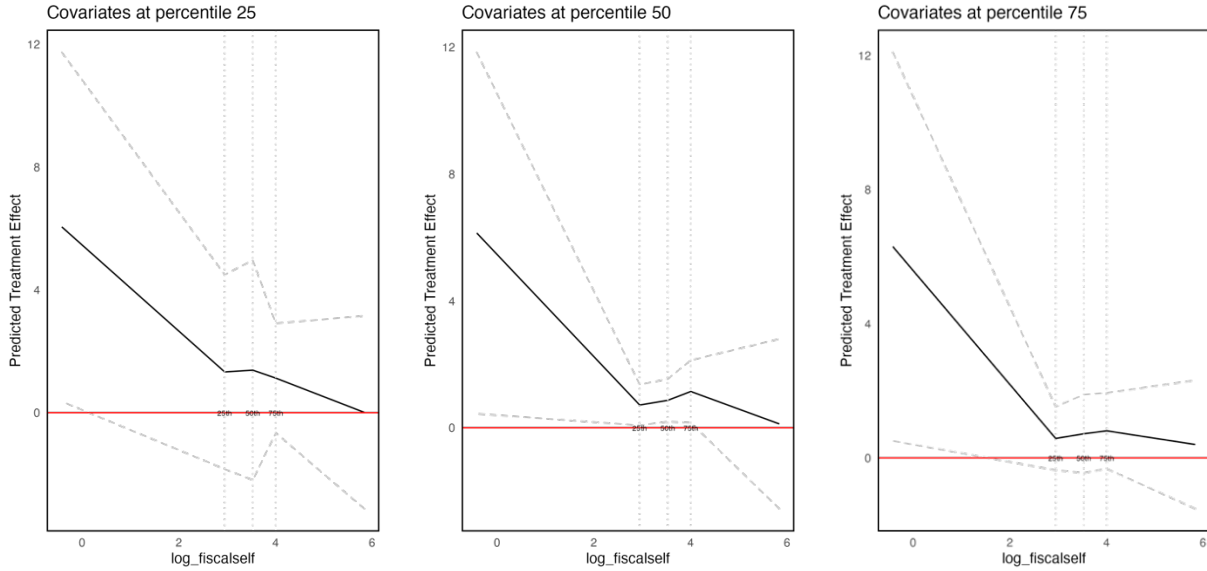
(2) 土地劳动比



(3) GDP 生产总值



(4) 县域其他国有银行数量



(5) 财政自给程度

图 6 广义随机森林分析异质性结果

八、结论

本文以邮储银行在县域设立和开展贷款业务的时间为外生冲击，探究了金融机构的发展对于农业劳动生产率的影响和作用机制，并讨论了该影响在不同特征区域间的差异。得出以下结论：

金融机构的设立对于农业劳动生产率表现出显著的正向作用，特别是在贷款业务开展之后，这一影响跟更加显著，且传统金融机构对农业劳动生产率的影响主要是通过影响当地的资本密集度，即提升县域投资来提升农业劳动生产率，然而相比县域是否设立金融机构网点，邮储自营网点数量的增加并不会直接影响农业劳动生产率与资本投入。这与以往文献的结论相似，即金融机构的设置，通过提高当地的贷款供给，促进县域的投资和经济发展。这一结果在解决了“样本选择偏误”和“负权重”后依旧稳健。

进一步，本文使用广义随机森林（GRF）的方法，分析了金融机构设立和发展贷款业务对于不同经济特征的县域作用，结果显示，金融机构开设贷款业务对那些经济发展水平在中等水平、财政自足状况良好、金融业务开展有一定水平的县域的农业劳动生产率的提升有较大的促进作用。即金融机构的作用要在一定经济水平的县域中才能发挥出来，进一步提升当地的收入水平，形成良好的互相促进局面。

参考文献

- 何广文. 2001: 《中国农村金融供求特征及均衡供求的路径选择》, 《中国农村经济》第10期.
- 何广文、刘甜. 2018: 《基于乡村振兴视角的农村金融困境与创新选择》, 《学术界》第10期.
- 黄季焜. 2018: 《四十年中国农业发展改革和未来政策选择》, 《农业技术经济》第3期.
- 林毅夫、孙希芳. 2008: 《银行业结构与经济增长》, 《经济研究》第9期.
- 吕勇斌、赵培培. 2014: 《我国农村金融发展与反贫困绩效: 基于2003—2010年的经验证据》, 《农业经济问题》第35期.
- 孙同全、潘忠. 2019: 《新中国农村金融研究70年》, 《中国农村观察》第6期.
- 谭词、杨军、孙嘉泽. 2022: 《数字乡村背景下农村电子商务如何提高农业劳动生产率——基于电子商务进农村综合示范县的准自然实验》, 《世界农业》第8期.
- 汪昌云、钟腾、郑华懋. 2014: 《金融市场化提高了农户信贷获得吗?——基于农户调查的实证研究》, 《经济研究》第10期.
- 王定祥、李伶俐、冉光和. 2009: 《金融资本形成与经济增长》, 《经济研究》第9期.
- 王雪、何广文. 2019: 《县域银行业竞争与普惠金融服务深化——贫困县与非贫困县的分层解析》, 《中国农村经济》第4期.
- 温铁军. 2013: 《农民专业合作社发展的困境与出路》, 《湖南农业大学学报(社会科学版)》第4期.
- 薛超、史雪阳、周宏. 2020: 《农业机械化对种植业全要素生产率提升的影响路径研究》, 《农业技术经济》第10期.
- 薛国琴. 2018: 《我国农业劳动生产率提高: 效应与调适——基于uv曲线移动的视角》, 《浙江社会科学》第3期.
- 徐忠、程恩江. 2004: 《利率政策、农村金融机构行为与农村信贷短缺》, 《金融研究》第12期.
- 姚耀军、李明珠. 2014: 《中国金融发展的反贫困效应: 非经济增长视角下的实证检验》, 《上海财经大学学报》第1期.
- 张爱英、孟维福. 2021: 《普惠金融、农业全要素生产率和城乡收入差距》, 《东岳论丛》第9期.
- 张珩、程名望、罗剑朝、李礼连. 2022: 《破解地方金融机构支持县域经济发展之谜》, 《财贸经济》第2期.
- 张晓云、范香梅、辛兵海. 2016: 《机构准入、金融包容与收入分配》, 《中国农村观察》第6期.
- 张元红. 1999: 《农民的金融需求与农村的金融深化——以湖北汉川福星村为例》, 《中国农村观察》第1期.
- 周振; 伍振军、孔祥智. 2015: 《中国农村资金净流出的机理、规模与趋势: 1978~2012年》, 《管理世界》第1期.
- 周颖辉、彭建刚. 2010: 《论中国邮政储蓄银行的核心竞争力》, 《软科学》第9期.
- 朱一鸣、王伟. 2017: 《普惠金融如何实现精准扶贫?》, 《财经研究》第10期.
- Abiad, Abdul and Ashoka Mody. 2005. "Financial Reform: What Shakes It? What Shapes It?" *American Economic Review*, 95(1), 66-88.
- Aghion, Philippe and Patrick Bolton. 1997. "A Theory of Trickle-Down Growth and Development." *The review of economic studies*, 64(2), 151-72.
- Andrews, Dan and Federico Cingano. 2014. "Public Policy and Resource Allocation: Evidence from Firms in OECD Countries." *Economic Policy*, 29(78), 253-96.
- Athey, Susan and Stefan Wager. 2021. "Policy Learning with Observational Data." *Econometrica*, 89(1), 133-61.
- Banerjee, Abhijit; Esther Duflo and Garima Sharma. 2021. "Long-Term Effects of the Targeting the Ultra Poor Program." *American Economic Review: Insights*, 3(4), 471-86.
- Banerjee, Abhijit; Paul Niehaus and Tavneet Suri. 2019. "Universal Basic Income in the Developing World." *Annual Review of Economics*, 11, 959-83.
- Banerjee, Abhijit V and Andrew F Newman. 1993. "Occupational Choice and the Process of Development." *Journal of political Economy*, 101(2), 274-98.
- Beck, Thorsten; Ross Levine and Alexey Levkov. 2010. "Big Bad Banks? The Winners and Losers from Bank Deregulation in the United States." *The Journal of Finance*, 65(5), 1637-67.
- Benmelech, Efraim; Carola Frydman and Dimitris Papanikolaou. 2019. "Financial Frictions and Employment During the Great Depression." *Journal of Financial Economics*, 133(3), 541-63.
- Borio, Claudio EV; Robert N McCauley; Patrick McGuire and Vladyslav Sushko. 2016. "Covered Interest Parity Lost: Understanding the Cross-Currency Basis." *BIS Quarterly Review September*.
- Borusyak, Kirill; Xavier Jaravel and Jann Spiess. 2021. "Revisiting Event Study Designs: Robust and Efficient Estimation." *arXiv preprint arXiv:2108.12419*.
- Bravo-Biosca, Albert; Chiara Criscuolo and Carlo Menon. 2016. "What Drives the Dynamics of Business Growth?" *Economic Policy*, 31(88), 703-42.
- Breiman, Leo. 2001. "Random Forests." *Machine learning*, 45, 5-32.
- Broderick, Tamara; Ryan Giordano and Rachael Meager. 2020. "An Automatic Finite-Sample Robustness Metric: Can Dropping a Little Data Change Conclusions." *arXiv preprint arXiv:2011.14999*, 16.
- Brown, James R; Gustav Martinsson and Bruce C Petersen. 2012. "Do Financing Constraints Matter for R&D?" *European Economic Review*, 56(8), 1512-29.
- Callaway, Brantly and Pedro HC Sant'Anna. 2021. "Difference-in-Differences with Multiple Time Periods." *Journal of Econometrics*, 225(2), 200-30.
- Carter, Colin A; CHEN Jing and CHU Baojin. 2003. "Agricultural Productivity Growth in China: Farm Level Versus Aggregate Measurement." *China Economic Review*, 14(1), 53-71.
- Carter, Michael R and Pedro Olinto. 2003. "Getting Institutions "Right" for Whom? Credit Constraints and the Impact of Property Rights on the Quantity and Composition of Investment." *American Journal of Agricultural Economics*, 85(1), 173-86.
- Chernozhukov, Victor; Mert Demirer; Esther Duflo and Ivan Fernandez-Val. 2018. "Generic Machine Learning Inference on Heterogeneous Treatment Effects in Randomized Experiments, with an Application to Immunization in India," National Bureau of Economic Research.
- Chodorow-Reich, Gabriel. 2014. "The Employment Effects of Credit Market Disruptions: Firm-Level Evidence from the 2008-9 Financial Crisis." *The Quarterly Journal of Economics*, 129(1), 1-59.

- Davis, Jonathan MV and Sara B Heller.** 2020. "Rethinking the Benefits of Youth Employment Programs: The Heterogeneous Effects of Summer Jobs." *Review of economics and statistics*, 102(4), 664-77.
- De Chaisemartin, Clément and Xavier d'Haultfoeuille.** 2020. "Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects." *American Economic Review*, 110(9), 2964-96.
- Dollar, David and Aart Kraay.** 2002. "Growth Is Good for the Poor." *Journal of Economic Growth*, 7, 195-225.
- Duygan-Bump, Burcu; Alexey Levkov and Judit Montoriol-Garriga.** 2015. "Financing Constraints and Unemployment: Evidence from the Great Recession." *Journal of Monetary Economics*, 75, 89-105.
- Evans, David S and Boyan Jovanovic.** 1989. "An Estimated Model of Entrepreneurial Choice under Liquidity Constraints." *Journal of political Economy*, 97(4), 808-27.
- Foltz, Jeremy D.** 2004. "Credit Market Access and Profitability in Tunisian Agriculture." *Agricultural Economics*, 30(3), 229-40.
- Fonseca, Julia and Bernardus Van Doornik.** 2022. "Financial Development and Labor Market Outcomes: Evidence from Brazil." *Journal of Financial economics*, 143(1), 550-68.
- Gardner, John.** 2022. "Two-Stage Differences in Differences." *arXiv preprint arXiv:2207.05943*.
- Gatti, Roberta and Inessa Love.** 2008. "Does Access to Credit Improve Productivity? Evidence from Bulgaria 1." *Economics of Transition*, 16(3), 445-65.
- Ghatak, Anita and Ferda Halicioglu.** 2007. "Foreign Direct Investment and Economic Growth: Some Evidence from across the World." *Global Business and Economics Review*, 9(4), 381-94.
- Goodman-Bacon, Andrew.** 2021. "Difference-in-Differences with Variation in Treatment Timing." *Journal of Econometrics*, 225(2), 254-77.
- Guirkinger, Catherine and Stephen R Boucher.** 2008. "Credit Constraints and Productivity in Peruvian Agriculture." *Agricultural Economics*, 39(3), 295-308.
- Ito, Junichi.** 2010. "Inter-Regional Difference of Agricultural Productivity in China: Distinction between Biochemical and Machinery Technology." *China Economic Review*, 21(3), 394-410.
- Kaboski, Joseph P and Robert M Townsend.** 2012. "The Impact of Credit on Village Economies." *American Economic Journal: Applied Economics*, 4(2), 98-133.
- Karlan, Dean and Jonathan Zinman.** 2010. "Expanding Credit Access: Using Randomized Supply Decisions to Estimate the Impacts." *The Review of Financial Studies*, 23(1), 433-64.
- Levine, Ross and Sara Zervos.** 1998. "Stock Markets, Banks, and Economic Growth." *American Economic Review*, 537-58.
- Lloyd-Ellis, Huw and Dan Bernhardt.** 2000. "Enterprise, Inequality and Economic Development." *The review of economic studies*, 67(1), 147-68.
- Marconi, Daniela and Christian Upper.** 2017. "Capital Misallocation and Financial Development: A Sector-Level Analysis." *Bank of Italy Temi di Discussione (Working Paper) No*, 1143.
- Mundlak, Yair.** 2005. "Economic Growth: Lessons from Two Centuries of American Agriculture." *Journal of economic literature*, 43(4), 989-1024.
- Pardey, Philip G and Julian M Alston.** 2019. "Transforming Traditional Agriculture Redux."
- Peek, Joe and Eric S Rosengren.** 2000. "Collateral Damage: Effects of the Japanese Bank Crisis on Real Activity in the United States." *American Economic Review*, 91(1), 30-45.
- Presbitero, Andrea F and Alberto Zazzaro.** 2011. "Competition and Relationship Lending: Friends or Foes?" *Journal of financial Intermediation*, 20(3), 387-413.
- Roth, Jonathan; Pedro HC Sant'Anna; Alyssa Bilinski and John Poe.** 2023. "What's Trending in Difference-in-Differences? A Synthesis of the Recent Econometrics Literature." *Journal of Econometrics*.
- Schreyer, Paul.** 2009. *Measuring Capital: Oecd Manual 2009*. OECD.
- Sun, Liyang and Sarah Abraham.** 2021. "Estimating Dynamic Treatment Effects in Event Studies with Heterogeneous Treatment Effects." *Journal of Econometrics*, 225(2), 175-99.
- Valickova, Petra; Tomas Havranek and Roman Horvath.** 2015. "Financial Development and Economic Growth: A Meta-Analysis." *Journal of economic surveys*, 29(3), 506-26.