

要素禀赋结构与早期工业建设的长期影响

张皓辰 王 歆 林毅夫*

摘要: 本文基于 1985 年全国工业普查微观数据, 利用我国改革开放之前的工业建设在产业结构和地理分布上的特征, 讨论这些工业投资如何对当地的产业发展产生长期影响, 并从要素禀赋结构的角度分析该影响的异质性。本文的研究结果显示, 一个地区某个行业在改革开放前的工业投资规模, 对本地区同行业的长期规模具有显著的正向影响; 更重要的是, 工业投资所在行业的资本密集度越是与当地的资本丰裕程度相契合, 该正向影响越强烈。

关键词: 早期工业建设; 要素禀赋结构; 长期影响

DOI: 10.13821/j.cnki.ceq.2023.04.05

一、引 言

新中国成立七十多年来, 我国在经济建设上取得了举世瞩目的成就, 其中工业建设在整个经济建设中发挥着主导和支柱作用。回顾我国社会主义经济建设的历史征程, 如何对新中国成立以来工业建设进行科学的评价, 探讨其中正反两方面的经验, 从而更好地利用这笔宝贵历史财富, 是一个具有重大现实意义的历史命题。本文通过对我国早期工业建设的分析, 探讨了“政府指导产业布局、实施产业政策应该遵循什么原则”这一重要问题。具体来说, 本文以产业结构与其所在地区要素禀赋结构的匹配度^①作为分析的切入点, 首次基于我国 1985 年的工业普查微观数据构建出对改革开放前各地区各行业在不同历史时期的工业建设情况的详细度量, 探讨了改革开放前的早期投资对于地方产业发展的长期影响, 以及匹配度差异所造成的该影响的异质性。

从理论上来说, 企业的盈利能力与市场竞争力取决于其生产经营的总成本, 而该成本受到投入要素价格、生产率以及交易费用等因素的影响。新经济地理学强调规模经济和集聚效应的存在为短期政策干预产生长期影响提供了可能性 (Krugman, 1991); 新结构经济学则指出, 如果企业或行业对要素投入的选择偏离企业所在地的要素禀赋结构所决定的比较优势, 即使在受到政策保护的有限时间内该行业发展了一定规模, 它仍然可能由于要素成本过高而在整体上缺乏市场竞争力, 无法实现长期发展 (Lin, 2011)。本文的实证结果为该假说提供了支持: 总体上改革开放之前在各地的工业投资对当地同

* 张皓辰、王歆、林毅夫, 北京大学新结构经济学研究院。通信作者及地址: 王歆, 北京市海淀区颐和园路 5 号北京大学新结构经济学研究院, 100871; 电话: 15810480552; E-mail: xin.wang@nsd.pku.edu.cn。本研究得到国家自然科学基金专项项目 (72141301) 和国家社会科学基金青年项目 (21CJL011) 的资助。作者感谢主编和三位匿名审稿人的宝贵意见。文责自负。

① 关于匹配度的详细讨论见 Ju et al. (2015) 以及后文的具体讨论。

行业经济的长期发展产生了显著的正向影响,这种影响体现为当地受影响的行业在全国同行业中具有更高的占比,说明早期工业建设在长期影响了各行业的空间布局。但更重要的是,该影响存在显著的地区和行业差异:给定其他条件不变,越是符合所在地区由要素禀赋结构所决定的比较优势的早期投资,越能够对当地产生显著的长期影响,即在资本(劳动)更加丰裕的地区对更加资本(劳动)密集的产业进行的早期投资,对该行业的长期发展有更大的影响。

当然,需要承认的是,在本文的研究场景中,难以找到一个足够外生的准自然实验来提供清晰的因果识别策略。为减少遗漏变量偏误对本文的主要结果造成的干扰,本文从对“三线建设”子样本的分析以及广义精确匹配方法的使用入手,部分地解决潜在的内生性问题,并从多个方面验证了主要结果的稳健性。此外,本文还从溢出效应的角度对早期工业建设产生长期影响的机制进行了进一步分析,加深了对长期影响作用机制的理解。

从总体上看,与现有研究相比,本文的贡献主要体现在两个方面。第一,关于我国早期工业建设历程,目前已经有不少文献从产业结构的角度进行回顾(管汉晖等,2020;林毅夫等,1994;张军,2002;郭凯明等,2017;程名望等,2019)。这些文章在实证上大多在较为宏观的层面,依据加总的地区或行业数据进行描述分析,或结合宏观模型进行结构性估计与定量核算。与此不同的是,本文利用1985年的全国工业普查资料,充分挖掘其中《全国大中型工业企业名录》所包含的地理位置和产品信息,得以考察各地级市层面细分工业行业的建设历史和结构变迁,并挖掘其中丰富的地区和行业异质性,在数据和关键指标的测度上迈出重要一步,以求对我国的早期工业建设进行更为细致和全面的评价。

第二,本文在讨论中强调不同行业和地区的差异,以及产业结构与地区要素禀赋结构的匹配度的重要性。在地区层面,关于政府主导的投资行为及其长期影响,文献中已有较多讨论(Rosenstein-Rodan,1943;Murphy et al.,1989;Bartelme et al.,2019;Kline and Moretti,2014)。Fan and Zou(2021)基于对我国20世纪六七十年代“三线建设”的研究,发现一个地方早期得到更多投资的产业在“三线建设”结束后的二十年之后仍有更大的规模和更快的发展;李艳和杨汝岱(2018)发现政府主导的建设留下的国有企业可能导致地方经济对国企形成依赖,对非国有企业扩大规模以及新企业的进入产生不利影响,不利于地方资源配置效率的提高。此外,不少文献研究了我国的各种区域性政策对地区产业发展的影响(陈钊和熊瑞祥,2015;董香书和肖翔,2017;Jia et al.,2020)以及政府对特定行业发展的政策对地区产业发展或企业发展的效果(陈钊和熊瑞祥,2015;李力行和申广军,2015;赵婷和陈钊,2020)。本文则从要素禀赋结构的视角考察我国早期工业建设作为一种地区发展政策和产业政策的效果及其异质性,为这一支文献提供了新的研究情境和经验证据。

在现有文献中,分析一个地区早期工业建设长期影响的实证文章较少,而对该影响异质性讨论的研究更是缺乏。一方面,这是由于大部分投资都是市场自发形成的,市场机制会驱动企业自发地选择回报率较高的产业和地区进行投资,这导致前后时间上的相关性大部分被地区本身的经济特征所驱动,从而难以进行因果推断;另一方面,要在实证上检验关于产业结构与要素禀赋结构的匹配度对早期投资长期影响的调节机制的假说,还需要在早期投资与要素禀赋结构的匹配度上有足够的跨地区和跨行业的差异,而

市场驱动形成的投资则缺乏这种差异。

我国在改革开放前所实施的重工业优先发展的战略为我们提供了实证检验的契机，让本文能够在实证上较好地克服上面提到的两个方面的挑战。具体而言，第一，这一时期的工业布局总体来说并不是市场自发投资行为，而是根据国家战略目标在指令性计划的体制下完成，这些政策的制定是以国家需要而非以各地经济利益最大化为目标。例如在1964—1978年间，我国出于备战的考虑，进行了大规模的“三线建设”，在西部内陆地区建设了一批主要服务于国防的工业项目。第二，我国各省的要素禀赋在这一时期之初有较大差异，中西部地区相对于东部沿海和东北地区而言，资本更加稀缺。因此无论是“大跃进”时期的全国兴建的钢铁、机械等重工业，还是“三线建设”时期在西部建立的国防工业项目，都使得各地的早期投资与其要素禀赋结构的匹配度上存在较大差异。

本文的其余部分安排如下：第二部分介绍本文的数据来源和变量的构建；第三部分介绍实证方法和主要回归结果，包括对早期工业建设长期影响的总体估计，关于要素禀赋结构匹配度的异质性分析，以及对潜在内生性问题的尝试解决；第四部分为稳健性检验和其他拓展，包括对早期工业建设溢出效应的分析；第五部分为结论。

二、数据介绍

（一）数据来源和主要变量

1. 1985年工业普查资料

本文所依据的第一部分数据是《中华人民共和国1985年工业普查资料》中的《全国大中型工业企业名录》（以下简称“名录”），这份名录包括了7592个大中型工业企业的名称、地址、开工年份、产品名称信息，以及1985年的工业总产值、年末职工人数、固定资产等信息。为了进行时间维度的刻画，考察1985年之前不同年份我国工业建设的情况，我们基于每个企业开工年份的信息，通过向前追溯队列的方法，来推算每个年份上的各地区和各行业的工业建设规模。^①同时，为了详细刻画产业结构与地区要素禀赋结构的关系，我们需要确定各企业所在的细分行业，而原始数据仅提供了企业的2位码行业，不足以考察不同行业生产技术在要素密集度上的差异。因此我们根据企业名称和产品名称，并结合原始数据中的行业分类，进一步确定了各企业所在的3位码行业^②。此外，数据中还包括了企业的详细地址信息（通常具体到街道或乡镇），作者手动将其匹配到所在的地级市代码。

2. 规模以上工业企业数据库

为考察改革开放前的工业投资对地区经济发展的长期影响，我们利用1998年及之

^① 我们基于“名录”的微观数据进行的对历史工业投资规模的推断，与历史上真实的各时间段的工业投资规模的加总数据中反映的省份和2位码行业间的分布模式基本一致。详见本文附录A的详细讨论，篇幅所限，附录未在正文列示，感兴趣的读者可在《经济学》（季刊）官网（<https://ceq.ccer.pku.edu.cn>）下载。

^② 这里的3位码行业，指的是国民经济行业分类（1994年）的行业子类。1985年“名录”的原始数据中共有39个大类行业（按照1985年时的行业划分标准的2位码行业），经我们进一步细分后，有157个3位码行业。

后年份的规模以上工业企业数据库^①。工业企业数据库由国家统计局构建,包含所有的国有企业和年主营业务收入在500万元以上的非国有企业的数据库,涵盖了企业的基本信息和较详细的财务数据,包括企业的固定资产、从业人数、产值、出口额、利润、增加值等。在回归分析中,我们将1998年工业企业数据库当中的企业数据加总到地区和行业层面,并剔除掉1985年在当地该行业还没有大中型工业企业的那些地区和行业的样本。^②在加总之前,我们剔除了从业人数在8人以下或关键变量缺失的企业(Brandt et al., 2012),并对总资产和总产值等结果变量删掉最大和最小1%的极端值,以避免异常值对回归结果的干扰。

3. 地区要素禀赋结构的度量

要素禀赋结构指的是资本、劳动力和自然资源等生产要素的相对丰裕程度。本文主要的研究对象是1953—1978年的工业投资,因此我们使用1952年的资本和劳动来度量地区要素禀赋。具体而言,我们根据张军等(2004)的方法,得到对1952年及其之后各年份省级物质资本存量的估算,并从《新中国六十年统计资料汇编》中获得各省在新中国成立以来各年份的就业规模信息,各地区的要素禀赋结构以资本和劳动的比值来度量。

当然,使用1952年的省级资本劳动比度量要素禀赋结构,可能引起几个潜在问题:第一,生产要素在一国内部的地区之间具有流动性,这可能弱化地区间的禀赋结构差异对工业建设长期表现的影响。对此我们指出,本文所关注的历史阶段,即我国的计划经济时期,地区间的要素流动受到严格限制(Boyreau-Debray and Wei, 2005; Gordon and Li, 2003; Li, 2010; Chan et al., 2011; Lai et al., 2013; Chan et al., 2013)。第二,本文选取1952年的一个截面来刻画各地区的要素禀赋结构,考察其对1953—1978年整个时间段内工业投资长期表现的影响,这一实证设定依赖于一个假设,即在这段时间内禀赋结构的地理分布是基本稳定的。对此我们指出,改革开放前的诸多工业化运动,比如将沿海地区的工厂迁到中西部,确实提高了流入地的物质资本和人力资本的禀赋,但这只是“量”的层面的变化,并不能从根本上,即“质”的层面上改变流入地资本相对短缺的情况。为了说明这一点,我们将若干年份(1952年、1958年、1965年、1978年和1985年)的各省要素禀赋结构(劳均资本)的排序进行比较,发现不同年份之间各省劳均资本的排名的相关系数都在0.7以上,多数年份之间在0.8以上。

4. 其他数据

除了前述三个方面的主要数据之外,本文用到的其他数据还包括:(1)《中国工业经济统计资料(1949—1984)》;(2)各城市层面的地理信息,包括各城市的地理位置(经纬度)和坡度;(3)我国1982年人口普查微观数据,来自IPUMS International数据库。

(二) 主要变量的描述统计

在回归分析中,我们主要关注改革开放之前的工业投资对本地区本行业发展的长期影响。我们在地区和行业层面定义早期工业投资规模,对于一个地区一个行业“1953—

^① 文中的主要结果是基于1998年的数据,但我们也利用2001年、2004年和2007年的数据对基准回归结果的稳健性进行了检验。

^② 对于改革开放之后在当地新出现的行业,其发展可能受到技术变革、外资引进等诸多因素的影响,故并未包含在我们的回归中。

1978年的投资”定义为1985年“名录”中在1953—1978年成立的企业的固定资产之和。类似地，可以定义其他时段的投资规模。在表1中，我们展示了回归中用到的主要变量的描述性统计信息，其中每个观测值是一个地区的一个3位数行业。1998年信息由工业企业数据库加总得到。根据前面所述的方式清理样本之后，我们在地区-行业层面样本共有13 671个，其中平均国有企业占比28%，企业平均年龄为12.24年，在1998年平均的从业人数为1 351.86人，平均的年产值为11.98亿元。

表1 主要变量的描述性统计

变量	观测数	均值	标准差	最小值	最大值
1952年之前的投资(万元)	13 671	2041.58	10 973.10	0	306 637
1953—1978年的投资(万元)	13 671	2937.50	18 744.05	0	599 660
1998年资产总计(万元)	13 671	177 425.60	496 571.30	0	1 564 255
1998年从业人数	13 671	1 351.86	3 641.80	0	147 819
1998年总产值(万元)	13 671	119 792	400 201.50	0	17 177 782
1998年国有企业占比	13 671	0.28	0.38	0	1
1998年企业平均年龄	13 671	12.24	13.178	0	118

数据来源：工业企业数据库(1998年)；1985年工业普查资料中的《全国大中型工业企业名录》。

注：每个观测值代表一个地区的一个行业。这里的地区是地级市，行业是3位数制造业行业。1998年信息由工业企业数据库加总得到。这里在1998年工业企业数据库的原数据基础上，进行正文所述的样本筛选，并在企业层面对总资产、产值、销售额、利润等与企业规模相关的变量删掉最大和最小1%的极端值之后再加总。

图1分地区展示了不同资本密集度分组^①在各地区^②内部占比的演变过程。我们发现不同地区发展的产业结构有很大的差异，且该结构在本文关注的时段内发生了剧烈变化。

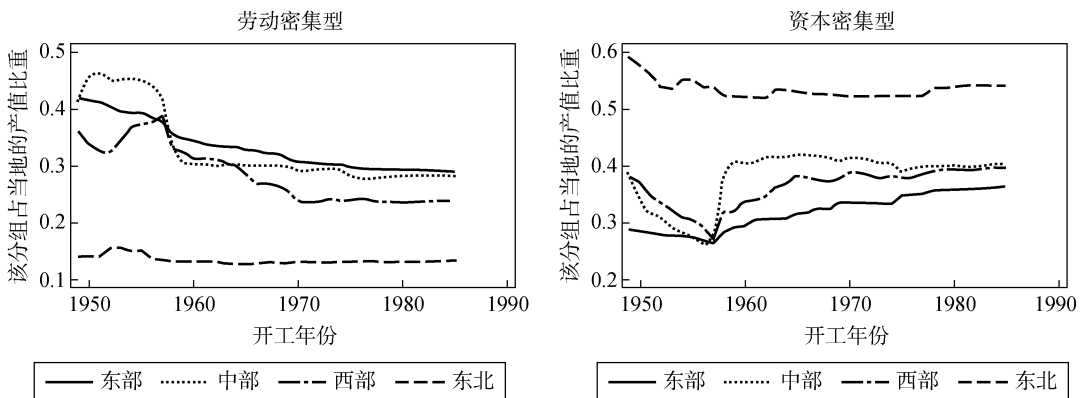


图1 各地理区域内产业结构变化

① 这里的“资本密集型”和“劳动密集型”的划分标准是：将全国所有3位数制造业行业的资本密集度排序，资本密集度最高的25%的行业作为“资本密集型”一组，将资本密集度最低的25%的行业作为“劳动密集型”一组。

② 将本文样本分成四大地理区域的方式具体如下。东部地区包括：北京、天津、河北、山东、江苏、上海、浙江、福建、广东、广西、海南；中部地区包括：内蒙古、山西、河南、湖北、湖南、安徽、江西；西部地区包括：陕西、甘肃、宁夏、云南、贵州、四川、重庆、新疆、青海；东北地区包括：黑龙江、吉林、辽宁。

三、实证方法和主要结果

(一) 早期工业建设长期影响的总体估计

首先我们考察平均而言早期工业投资对当地行业发展是否存在长期影响,为下一步从要素禀赋结构的视角考察长期影响的异质性提供基准。具体地,在“地区×行业”层面,我们采用的回归方程是:

$$OutcomeShare_{sc} = \beta_1 InvestShare_{sc} + \gamma X_{sc} + \delta_c + \varepsilon_{sc}, \quad (1)$$

其中, s 代表(3位数)行业, c 代表地区(地级市)。因变量 $OutcomeShare_{sc}$ 是1998年 c 地区 s 行业的规模占全国范围内 s 行业规模的比例,其中规模的度量指标包括资产、从业人数和产值,由本地区本行业中所有企业的数据加总而得。主要的自变量 $InvestShare_{sc}$ 表示 c 地区 s 行业在1953—1978年间建设的大中型企业在1985年的总资本占全国范围内 s 行业在这一时间段内开工的企业的总资本的比例,作为对1953—1978年间投资的相对规模的度量。之所以将1953—1978年作为本文主要关注的历史时段,是由于政府主导的工业建设开始于1953年的“一五”计划,而1978年之后计划经济体制逐渐向商品经济转轨,政府对经济的指令性干预逐渐减少。控制变量 X_{sc} 包括本地区本行业在1952年之前开工的大中型企业在1985年的总资本在全国同行业中的占比、本地区1998年国有企业所占的比重和企业的平均年龄, δ_c 为地区的固定效应, ε_{sc} 是残差项。需要指出的是,由于回归中控制了地区固定效应,估计式(1)相当于是在同一个地区内部的不同行业之间进行比较,即这里的 β_1 表示,相对于本地区的其他行业,如果本地区本行业在1953—1978年的投资在全国的占比比本地区其他行业高出1个百分点,则平均而言本地区本行业的1998年在全国的占比比本地区其他行业平均高出 β_1 个百分点。

基准回归的结果展示在表2中,第(1)—(3)列分别用总资产、从业人数和产值来度量行业规模。我们发现改革开放前的工业投资对当地本行业后来发展的规模有显著的正向长期影响。给定其他条件不变,平均而言,相比于本地区的其他行业,如果该地区某行业在1953—1978年间的投资在全国同行业的占比上升1个百分点,则本地区的该行业在1998年的同行业中的总资产占比会提高0.217个百分点,从业人数上升0.222个百分点,产值上升0.216个百分点。这样的结果具有经济和统计意义上的显著性。

表2 基准回归结果

因变量: 本地区本行业在全国同行业中的占比 (%)	总资产 (1)	从业人数 (2)	产值 (3)
本地区本行业1953—1978年的投资在全国占比 (%)	0.217*** (0.024)	0.222*** (0.018)	0.216*** (0.021)
本地区本行业1952年之前的投资在全国占比 (%)	0.071*** (0.023)	0.046** (0.020)	0.038** (0.018)
国有企业占比 (%)	0.001* (0.049)	-0.000 (0.057)	-0.002** (0.061)

(续表)

因变量：本地区本行业在全国同行业中的占比 (%)	总资产 (1)	从业人数 (2)	产值 (3)
企业平均年龄	0.013*** (0.001)	0.020*** (0.002)	0.013*** (0.002)
地区固定效应	控制	控制	控制
观测数	13 670	13 670	13 670
R ²	0.397	0.332	0.365

注：表格中报告的是回归系数，括号里是在地级市层面聚类的稳健标准误。*、**和*** 分别代表 10%、5%和 1% 的显著性水平。

(二) 要素禀赋结构与长期影响的异质性

Ju et al. (2015) 指出了产业变迁过程中遵循的匹配度原则，即一个地区发展的产业的资本密集度如果与当地的资本丰裕程度相匹配，则能够取得更好的发展。为了考察要素禀赋结构决定的比较优势方面的差异如何导致早期投资长期影响的异质性，我们首先定义一个关于行业的要素投入结构（资本密集度）与地区的要素禀赋结构（资本丰裕度）之间的匹配度的指标（congruence index）^①：

$$congruence_{sc} = - \left| \log\left(\frac{K_s}{L_s}\right) - \log\left(\frac{\overline{K}_c}{\overline{L}_c}\right) \right|, \quad (2)$$

其中， s 代表行业， c 代表地级市， $\frac{K_s}{L_s}$ 是 s 行业的资本密集度，是用“名录”数据中 1985 年这个行业全国的总资本除以总劳动得到的，是一个全国平均的行业资本密集度。 $\frac{\overline{K}_c}{\overline{L}_c}$ 是 c 城市所在省份在 1952 年的资本丰裕程度，代表当地要素禀赋结构。^②

为了考察要素禀赋结构匹配度如何影响早期工业投资在行业相对规模的长期作用，我们采用如下交叉项形式的回归模型：

$$OutcomeShare_{sc} = \beta_1 \times InvestShare_{sc} \times congruence_{sc} + \beta_2 InvestShare_{sc} + \beta_3 \times congruence_{sc} + \gamma \mathbf{X}_{sc} + \delta_c + \varepsilon_{sc}, \quad (3)$$

其中 $InvestShare_{sc}$ 代表 c 城市的 s 行业在我们关注的历史阶段的投资在全国范围内同行业中的占比，定义方式与基准回归方程 (2) 相同，其他控制变量也与基准回归相同。我们将 $congruence_{sc}$ 和 $InvestShare_{sc}$ 均做了标准化处理，使其在回归样本中均值为 0，标准差为 1，以便解读回归系数。

表 3 展示了这部分的回归结果。我们发现方程 (3) 中的系数 β_1 显著为正，说明早期

① 作为稳健性检验，作者还采用了 Ju et al. (2015) 构建的匹配度指标， $congruence_{sc} = - \left| \frac{\frac{K_s}{L_s} - \frac{\overline{K}_c}{\overline{L}_c}}{\frac{\overline{K}_c}{\overline{L}_c}} \right|$ ，基于

这个指标所得到的实证结果与正文中指标十分相似。

② 资本密集度和地区资本丰裕程度的测度都统一到 1985 年价格。

投资的行业的资本密集度与当地的要素禀赋程度的契合度越高,则早期工业投资能在当地同行业的相对规模上产生更强的长期影响。具体地,以第(1)列为例,如果匹配度提高1个标准差,则早期工业投资在全国同行业中的占比提高1个百分点带来的影响会上升0.068个百分点,约占表2中所示的基准结果中平均影响(0.22个百分点)的31%,这说明匹配度指标的调节作用具有经济上的显著性。此外,在方程(3)中, β_2 代表样本中 $congruence_{sc}$ 的均值水平对应的地区和行业中 $InvestShare_{sc}$ 的边际影响,这里显著为正,符合回归设定的预期。同时我们发现在所有回归中,系数 β_3 都为正,说明在早期投资相对规模取样本均值对应的地区和行业中,早期投资的匹配度本身对本行业的长期发展具有一定的正向作用。

表3 关于要素禀赋结构匹配度的异质性分析

因变量:本地区本行业在全国同行业中的占比(%)	总资产 (1)	从业人数 (2)	产值 (3)
$InvestShare_{sc} \times congruence_{sc}$	0.068*** (0.017)	0.048*** (0.012)	0.049*** (0.016)
$InvestShare_{sc}$	0.068*** (0.017)	0.048*** (0.012)	0.049*** (0.016)
$congruence_{sc}$	0.084* (0.047)	0.093** (0.043)	0.056 (0.047)
控制变量	控制	控制	控制
地区固定效应	控制	控制	控制
观测数	12 844	12 844	12 844
R^2	0.380	0.310	0.368

注:表格中报告的是回归系数,括号里是在地级市层面聚类的稳健标准误。*、**和***分别代表10%、5%和1%的显著性水平。

为了更好地说明这种要素禀赋结构匹配度的调节机制在经济意义上的显著性,我们以方程(3)的回归模型为基础,计算出每个地区每个行业早期投资的作用强度大小。具体地,对方程(3)的估计得到回归系数 $\hat{\beta}_1$ 、 $\hat{\beta}_2$ 、 $\hat{\beta}_3$,则历史投资规模的对数的变动对因变量 Y_{sc} 的边际影响的估计值为:

$$\widehat{Effect}_{sc} = \hat{\beta}_1 \times congruence_{sc} + \hat{\beta}_2.$$

以表3第(1)列的回归为例,我们可以计算得到早期工业建设的投资规模对当地同行业1998年总资产规模的影响在不同地区和不同行业的大小的估计 \widehat{Effect}_{sc} ,这一估计值的均值为0.146,10%分位数为0.067,90%分位数为0.339,说明在不同地区和行业间影响的大小确实存在较大的差异性。更加具体地,以1952年省级资本丰裕程度在全国排名较低的安徽省为例,典型的资本密集型的水泥制造业的作用大小的估计值为0.050,炼钢业为0.061,而典型的劳动密集型行业如服装制造业为0.137,棉纺织业为0.146。再从一个产业在不同地区间作用大小的横向对比来看,以资本密集型的炼钢业为例,其在资本丰裕程度较高的辽宁省的作用大小为0.217,而在资本丰裕程度较低的云南省则只有0.031,相差2.5个标准差以上,这样的差异具有高度的经济显著性。由

此看来，从地区要素禀赋结构和行业要素投入结构的匹配度入手，确实有助于挖掘早期工业建设的长期表现在地区和行业间的异质性。

(三) 对潜在内生性问题的尝试解决

由于早期建设的选址决策本身可能内生于各地区的条件和不同行业在各地发展的适宜程度，早期投资的相对规模和回归中的残差项可能相关，如果这种相关性是正确的，即有一些地区和行业层面的不可观测的因素使得某地区的某行业得到相对更多的早期工业投资，且这些因素也导致本地区本行业相对于其他行业在改革开放之后有更大的相对规模，则上述分析高估了早期工业建设产生的长期影响。本小节将使用多种方法来解决这一潜在问题。

1. 基于“三线建设”小样本的分析

在本节中，作者使用“三线建设”的子样本进行分析。之所以这样做，是因为“三线建设”的项目本身并不是出于发展地方经济的目的而建设的，而是出于国防和备战的考虑。因此“三线建设”时期各地区各行业的工业投资规模受到当地禀赋条件的影响较小，而且更多的是由国家在全国集中调配资源而非利用本地资源进行建设，对于一个地区来说较为外生，与其他影响早期工业投资长期影响的不可观测因素的相关性可能较低。从“名录”中企业的资本密集度来看，在“三线建设”之外的时期和三线之外的地区，企业的资本劳动比和所在地区的资本丰裕程度正相关，而在“三线建设”时期，二者的相关性则为负，这种负向关系主要表现在三线地区内部，这与当时的历史背景相符合。

在下面的分析中，作者只关注“三线建设”区域内的样本，且只关注在“三线建设”期间（1964—1978年）开工的企业规模对当地经济发展的长期影响的异质性，所使用的回归模型与表3的模型相同，这部分结果展示在表4当中。在表4的各列中，交叉项的系数都显著为正。这样的结果印证了本文主要结论的稳健性。

表4 基于“三线建设”子样本的回归结果

因变量：本地区本行业在全国同行业中的占比（%）	总资产	从业人数	产值
	(1)	(2)	(3)
$InvestShare_{sc} \times congruence_{sc}$	0.094** (0.044)	0.129** (0.062)	0.047* (0.025)
$InvestShare_{sc}$	0.161*** (0.047)	0.196** (0.078)	0.077* (0.044)
$congruence_{sc}$	0.010* (0.005)	0.010** (0.004)	0.010** (0.004)
控制变量	控制	控制	控制
地区固定效应	控制	控制	控制
观测数	1 319	1 319	1 319
R^2	0.142	0.123	0.145

注：表格中报告的是回归系数，括号里是在地级市层面聚类的稳健标准误。*、**和***分别代表10%、5%和1%的显著性水平。

2. 基于广义精确匹配方法的分析

为了进一步解决工业建设中企业选址决策的内生性问题对本文估计结果的干扰,在这里,作者使用广义精确匹配(coarsened exact matching,简称CEM)的方法进行稳健性检验。这种方法的主要思想是,对于每一个城市和产业,匹配与其具有相似特征的城市中的相似产业,经过匹配之后,在具有相似城市和产业特征的每一个组的内部考察这些城市和产业之间的早期工业建设情况的差异,估计早期工业建设的长期影响及其在要素结构匹配度上的异质性。

具体地,我们考察如下几个维度的城市特征:(1)是否沿海岸线;(2)是否沿国境线;(3)是否为省会城市;(4)到省会城市的距离,这是一个连续变量,我们按照所有城市的四分位数值将其分成四组,小于或等于25%分位数的为一组、在25%和50%分位数(含50%分位数)之间的为一组,50%到75%分位数之间(含75%分位数)的为一组,大于75%分位数的为一组;(5)城市几何中心到海岸线的距离,同样按(4)中所述的方式分为四组;(6)城市的平均崎岖程度,根据城市中地理范围内的海拔的变异程度计算得到,也是一个连续变量,同样按照(4)中所述的方式分为四组;(7)城市在1952年之前是否有大中型工业企业。根据该分组方式,可以将数据划分成众多的栅格^①,匹配后每一个栅格内部的不同观测所在的城市在用于匹配的特征上高度相似,而文中关注的主要解释变量,如历史投资规模、地区资本丰裕度和行业资本密集度等,变异性下降不多,仍可为我们的统计推断提供支持。

在回归中,我们控制CEM栅格的固定效应,回归的结果如表5所示。可以发现,回归系数的符号、大小和显著性水平都与基准分析中表3的结果非常接近,从而进一步印证了本文主要结果的稳健性。

表5 基于广义精确匹配方法的回归结果

因变量:本地区本行业在全国同行业中的占比(%)	总资产	从业人数	产值
	(1)	(2)	(3)
$InvestShare_{sc} \times congruence_{sc}$	0.092*** (0.016)	0.068*** (0.014)	0.073*** (0.022)
$InvestShare_{sc}$	0.098*** (0.013)	0.130*** (0.024)	0.136*** (0.023)
$congruence_{sc}$	0.120** (0.056)	0.155** (0.061)	0.099 (0.061)
控制变量	控制	控制	控制
地区固定效应	控制	控制	控制
CEM栅格固定效应	控制	控制	控制

^① 在本文回归用到的数据中,共有3384个栅格中有数据观测。我们去掉那些在一个栅格中只有一个“城市×行业”对的观测,这些观测没有在CEM过程中匹配到合适的对照组,经过这一样本限制的过程,用于CEM的回归的栅格有3168个(93.6%),用到的观测值“地区×行业”层面是基准的异质性回归(如表5所示)的94.8%,具有较高的匹配率。

(续表)

因变量：本地区本行业在全国同行业中的占比 (%)	总资产 (1)	从业人数 (2)	产值 (3)
观测数	12 170	12 170	12 170
R ²	0.560	0.511	0.538

数据来源：工业企业数据库（1998年）；1985年工业普查资料中的《全国大中型工业企业名录》。

注：表格中报告的是回归系数，括号里是在地级市层面聚类的稳健标准误。*、**和***分别代表10%、5%和1%的显著性水平。

四、稳健性和其他拓展

（一）关于要素禀赋结构调节作用的稳健性检验

1. 加入其他维度的要素禀赋结构

我们在回归方程（3）中，加入人力资本禀赋和自然资源禀赋的匹配度对早期投资长期影响的调节作用。我们基于1982年人口普查数据（全国1%样本），计算每个地级市在1952年15岁以上的受过高中教育的人口比例以衡量人力资本丰裕度，并用每个3位数行业的就业人员中受过高中教育的比例度量每个行业人力资本密集度。对于矿产资源，参考Fan and Zou（2021），我们用1982年人口普查中各地级市的采掘业就业占比来度量一个地区的矿产资源丰裕程度 $Mine_c$ ，并利用国家统计局发布的投入产出表（2007年），计算每个行业的投入品中来自采掘业的比例，作为对一个行业使用矿产资源密集度 $ResInt_s$ 的度量。

表6中结果显示，一方面，我们确实发现人力资本和自然资源两个维度的要素禀赋的匹配度也对于早期工业投资的长期影响存在显著正向的调节作用；另一方面，加入这两个维度的交互项之后，本文重点关注的物质资本的要素禀赋结构的匹配度的调节作用在系数符号和显著性上依然稳健，这说明另外两个维度的要素禀赋结构的调节作用不会与本文重点关注的物质资本的作用相混杂，支持了本文主要结论的稳健性。

表6 稳健性检验：加入人力资本和矿产资源两个维度的要素禀赋结构

因变量：本地区本行业在全国同行业中的占比 (%)	总资产 (1)	从业人数 (2)	产值 (3)
A. 物质资本			
$InvestShare_{sc} \times congruence_{sc}$	0.088*** (0.019)	0.043* (0.022)	0.052** (0.022)
B. 人力资本			
$InvestShare_{sc} \times congruence_{sc}^{HC}$	0.101 (0.061)	0.122* (0.066)	0.122** (0.062)
C. 矿产资源			
$InvestShare_{sc} \times ResInt_s \times Mine_c$	0.104*** (0.030)	0.097*** (0.026)	0.108*** (0.030)

(续表)

因变量：本地区本行业在全国同行业中的占比(%)	总资产	从业人数	产值
	(1)	(2)	(3)
交叉项的低阶项	控制	控制	控制
控制变量	控制	控制	控制
地区固定效应	控制	控制	控制
观测数	12 844	12 844	12 844
R^2	0.393	0.324	0.386

注：表格中报告的是回归系数，括号里是在地级市层面聚类的稳健标准误。*、**和***分别代表10%、5%和1%的显著性水平。回归中控制了所有交互项对应的低阶项，此处只汇报了关键的交互项的系数。 $congruence_{it}^{HC}$ ， $ResInt_{it}$ 和 $Mine_{it}$ 都在样本范围内进行了标准化处理，使其均值为0，标准差为1。

2. 对本文主要结论的其他稳健性检验

除了上述分析之外，作者还对本文的主要结论进行了其他的稳健性检验，主要结果均稳健^①，具体包括以下几个方面：

(1) 考虑到东北地区可能存在一定的特殊性，比如在新中国成立之前，东北就具有了较高的工业化水平，其资本丰裕程度高的原因与当地自然资源的开发以及日伪时期的工业建设有关，为了排除我们观察到的要素禀赋结构匹配程度的调节机制主要由东北地区的特殊性所驱动，我们去掉东北地区进行表3所示的回归分析。

(2) 可能与要素禀赋结构的调节效应相混杂的一个机制是集聚效应，即资本丰裕程度较高的地区往往已经具有较多的重工业，这时再投资重工业，则可能由于同行业的集聚经济而使得这些重工业投资有更强的长期影响。为了排除本文的主要结果是由集聚效应所驱动的这种情况，我们在方程(3)的回归的基础上加入早期工业投资相对规模与表示本地区本行业在1952年之前是否有大中型工业企业的虚拟变量之间的交互项。

(3) 在目前回归结果的基础上，我们还参考现有文献(李艳和杨汝岱, 2018; 刘冲等, 2020)，加入更多地区-行业层面的控制变量，这些变量根据工业企业数据库计算得到，包括地区-行业层面的企业数量、Herfindahl-Hirschman指数(HHI)、平均资本劳动比、平均资产负债率和出口企业比例。

(4) 在本文的基准分析中，对当代的行业空间分布的度量基于1998年的工业企业数据库，作为稳健性检验，本文还使用了后续若干年份(2001年、2004年和2007年)的工业企业数据库进行度量。

(二) 关于早期工业建设溢出效应的拓展分析

在目前的回归中，我们估计得到的早期工业建设的长期影响，事实上可能包含了多个方面的效应，为了更全面地刻画早期工业建设产生长期影响的机制，本小节提供一些额外的经验证据。

1. 对改革开放之后进入的非国有部门发展的外溢效应

在本节中，我们通过将1978年后成立的非国有企业的规模加总到地区和行业层面

^① 具体结果详见本文附录B部分，篇幅所限，附录未在正文列示，感兴趣的读者可在《经济学》(季刊)官网(<https://ceq.ccer.pku.edu.cn>)下载。

进行考察，以区分1978年前建成的大中型工业企业由于自身的存续和规模变化而对整体行业规模产生的影响和早期工业建设对非国有部门的外溢效应。在表7中我们报告了这部分结果，其中使用的回归模型与前面给出的方程(1)和(3)相同。地区和行业层面的被解释变量是由本地区本行业1998年截面上的所有1978年之后成立的非国有企业的规模(总资产、从业人数和产值)加总得到的，仍表示本地区本行业的规模在全国同行业中的占比。结果显示，早期工业建设的确对行业内的非国有企业总体规模有显著正影响，且要素禀赋匹配度的调节效应仍存在。

表7 早期工业投资对改革开放后建立的非国有企业发展的外溢效应

因变量：本地区本行业在全国同行 业中的占比(%)	总资产		从业人数		产值	
	(1)	(2)	(3)	(4)	(5)	(6)
$InvestShare_{sc}$	0.194*** (0.018)	0.154*** (0.030)	0.206*** (0.014)	0.176*** (0.032)	0.202*** (0.018)	0.184*** (0.029)
$InvestShare_{sc} \times congruence_{sc}$		0.030* (0.016)		0.034** (0.016)		0.035** (0.014)
$congruence_{sc}$		0.051 (0.045)		-0.001 (0.042)		0.017 (0.044)
控制变量	控制	控制	控制	控制	控制	控制
地区固定效应	控制	控制	控制	控制	控制	控制
观测数	12 433	11 672	12 433	11 672	12 433	11 672
R^2	0.423	0.425	0.365	0.368	0.403	0.415

注：回归中地区和行业层面各因变量的构建是基于1978年之后成立的非国有企业的样本。表格中报告的是回归系数，括号里是在地级市层面聚类的稳健标准误。*、**和***分别代表10%、5%和1%的显著性水平。

2. 对上下游行业的影响

本文前面的内容考察的都是早期工业建设对本地区同行业的长期影响，在本节中，我们将分析拓展到与早期工业投资所在行业有投入产出关联的其他行业，分析早期工业建设对上下游行业的影响。我们参考文献中现有的做法(Javorcik, 2004)，使用国家统计局2007年的投入产出表中两两行业间的直接消耗系数构造权重，得到每个地区每个行业所对应的上游行业的平均早期投资相对规模和下游行业的平均早期投资相对规模， $UpShare_{sc}$ 和 $DownShare_{sc}$ 。^①在基准回归方程(1)的基础上，将这两个变量加入回归中，其他设定与方程(1)相同，得到的回归结果如表8所示。表8的结果表明，第一，加入其他行业的溢出效应之后，本行业长期影响的系数保持稳健，与表2的基准结果相比系数大小十分接近；第二，早期工业建设对其他行业的长期影响主要表现为上游对下游的前向溢出效应，而非下游对上游的后向溢出效应，表8中 $UpShare_{sc}$ 的系数显著为正，且系数大小高于本行业的作用大小，而 $DownShare_{sc}$ 的系数不显著且系数大小很接近0。这样的结果总体上符合预期，因为我国在改革开放前工业建设的重点多为上游行业，而在改革开放之后面向消费者市场的下游行业得到了快速发展。

^① 具体定义为： $UpShare_{sc} = \sum_r \omega_{s \leftarrow r} InvestShare_{rc}$ ， $DownShare_{sc} = \sum_r \omega_{s \rightarrow r} InvestShare_{rc}$ ，其中 s 和 r 代表两个行业， $\omega_{s \leftarrow r}$ 表示 s 行业的中间投入品中 r 行业的占比， $\omega_{s \rightarrow r}$ 表示 s 行业用作其他行业中间投入品的产品中 r 行业的占比。

表8 早期工业投资规模对本地上下游行业的长期影响

因变量:本地区本行业在全国同行业中的占比(%)	总资产	从业人数	产值
	(1)	(2)	(3)
$InvestShare_{sc}$	0.205*** (0.020)	0.208*** (0.016)	0.204*** (0.021)
$UpShare_{sc}$	0.380*** (0.134)	0.353*** (0.107)	0.323*** (0.095)
$DownShare_{sc}$	-0.039 (0.080)	-0.005 (0.056)	-0.025 (0.072)
控制变量	控制	控制	控制
地区固定效应	控制	控制	控制
观测数	13 670	13 670	13 670
R^2	0.408	0.351	0.377

注:表格中报告的是回归系数,括号里是在地级市层面聚类的稳健标准误。*、**和***分别代表10%、5%和1%的显著性水平。

3. 对地区总体工业规模的长期影响

在地区层面早期工业建设的总体影响既包括了前面已经讨论的正向效应——这包括行业内的效应和行业间的效应,同时也包括了可能存在的负向的挤出效应^①,表9对这种总体影响进行了探讨。前三列的结果显示,一个地区早期的工业投资的相对规模对当地1998年的工业规模在全国的占比具有显著的正向影响;后三列的结果进一步表明,这种影响的大小受到地区平均的匹配度指标的调节,早期投资的要素禀赋匹配度越高的地区,其早期投资所产生的长期影响越大。这样的结果表明,从溢出效应和挤出效应的加总来看,早期工业投资对本地区会产生正向的长期影响,且这种影响在总体上会受到要素禀赋结构匹配度的调节,这与本文前面在地区和行业层面的分析结果相印证。

表9 加总到地区层面的估计结果

因变量:本地区在全国的占比(%)	总资产	从业人数	产值	总资产	从业人数	产值
	(1)	(2)	(3)	(4)	(5)	(6)
$InvestShare_c$	0.255* (0.132)	0.303*** (0.091)	0.322*** (0.133)	0.480*** (0.132)	0.372*** (0.084)	0.451*** (0.141)
$InvestShare_c \times congruence_c$				0.591*** (0.136)	0.439*** (0.087)	0.529*** (0.146)
$congruence_c$				0.126*** (0.043)	0.101*** (0.027)	0.129*** (0.046)
控制变量	控制	控制	控制	控制	控制	控制
观测数	230	230	230	230	230	230
R^2	0.267	0.408	0.311	0.332	0.479	0.358

注: $Congruence_c$ 代表地区内各产业匹配度的平均值。控制变量包括地区内的国有企业占比、平均的企业年龄和1952年之前投资在全国的占比。表格中报告的是回归系数,括号里是异方差稳健标准误。*、**和***分别代表10%、5%和1%的显著性水平。

① 早期的工业企业可能挤占了当地的资源,而挤出了本行业的其他企业或本地的其他行业的发展。

五、结 论

本文从1985年全国工业普查资料中的《全国大中型工业企业名录》这一微观数据出发，基于新中国改革开放之前的工业建设在产业结构和地理分布上的特征事实，讨论这些早期的工业投资如何对当地的产业发展产生长期影响，并从要素禀赋结构的角度分析这种影响的异质性。本文的研究显示，一个地区某行业的早期工业建设规模在全国范围内同行业的占比越高，则在长期内当地的该行业在全国范围内同行业中所占的比重越大，说明早期工业建设对行业的空间布局产生了持久的影响。同时，早期工业投资的资本密集度越是与当地的资本丰裕程度相契合，则在长期内对本行业的相对规模有更强的正向影响，这为地区产业长期发展所要遵循的匹配度原则提供了微观的实证支持。

这项研究为评价我国改革开放前三十年的工业建设提供了新的视角和经验证据，也为我国在新发展阶段的产业发展提供了政策启示：政府在制定产业政策时，既要考虑对当下经济的拉动作用，也要考虑这个产业是否与当前的发展阶段以及地方的要素禀赋条件相适应，避免一些违背当地比较优势的短视的投资行为，使得政策扶持的产业在长期中取得持续的发展；在产业的空间布局上，要注重甄别各地区的禀赋条件所决定的潜在比较优势，从而实现地区间优势互补，实现区域间协调发展和共同富裕。

参 考 文 献

- [1] Bartelme, D., A. Costinot, D. Donaldson and A. Rodriguez-Clare, "The Textbook Case for Industrial Policy: Theory Meets Data", NBER Working Paper, 2019, No. 26193.
- [2] Boyreau-Debray, G., and S-J. Wei, "Pitfalls of a State-Dominated Financial System: The Case of China", Working Paper No. W11214, National Bureau of Economic Research, New York, 2005.
- [3] Brandt, L., J. Van Biesebroeck, and Y. Zhang, "Creative Accounting or Creative Destruction: Firm Level Productivity Growth in Chinese Manufacturing", *Journal of Development Economics*, 2012, 97 (2), 339-351.
- [4] Chan, K. S., V. Q. Dang, J. T. Lai, and I. K. Yan, "Regional Capital Mobility in China: 1978-2006", *Journal of International Money and Finance*, 2011, 30 (7), 1506-1515.
- [5] 陈钊、熊瑞祥：“比较优势与产业政策效果——来自出口加工区准实验的证据”，《管理世界》，2015年第8期，第67—80页。
- [6] 程名望、贾晓佳、仇焕广，“中国经济增长（1978—2015）：灵感还是汗水？”，《经济研究》，2019年第7期，第30—46页。
- [7] 董香书、肖翔，“‘振兴东北老工业基地’有利于产值还是利润？——来自中国工业企业数据的证据”，《管理世界》，2017年第7期，第24—34页。
- [8] Fan, J., and B. Zou, "Industrialization from Scratch: The 'Third Front' and Local Economic Development in China's Hinterland", *Journal of Development Economics*, 2021, 152, 102698.
- [9] Gordon, R. H., and W. Li, "Government as a Discriminating Monopolist in the Financial Market: The Case of China", *Journal of Public Economics*, 2003, 87 (2), 283-312.
- [10] 郭凯明、杭静、颜色：“中国改革开放以来产业结构转型的影响因素”，《经济研究》，2017年第3期，第32—46页。
- [11] 管汉晖、刘冲、辛星，“中国的工业化：过去与现在（1887—2017）”，《经济学报》，2020年第7卷第3期，第202—238页。
- [12] Javorcik, B. S., "Does Foreign Direct Investment Increase the Productivity of Domestic Firms? In Search of Spill-

- overs through Backward Linkages”, *American Economic Review*, 2004, 94 (3), 605-627.
- [13] Jia, J., G. Ma, C. Qin, and L. Wang, “Place-Based Policies, State-Led Industrialisation, and Regional Development: Evidence from China’s Great Western Development Programme”, *European Economic Review*, 2020, 123, 103398.
- [14] Ju, J., J. Y. Lin, and Y. Wang, “Endowment Structures, Industrial Dynamics, and Economic Growth”, *Journal of Monetary Economics*, 2015, 76, 244-263.
- [15] Kline, P., and E. Moretti, “Local Economic Development, Agglomeration Economies, and the Big Push: 100 Years of Evidence from the Tennessee Valley Authority”, *Quarterly Journal of Economics*, 2014, 129 (1), 275-331.
- [16] Krugman, P., “Increasing Returns and Economic Geography”, *Journal of Political Economy*, 1991, 99 (3), 483-499.
- [17] Lai, J. T., P. D. McNelis, and I. K. Yan, “Regional Capital Mobility in China: Economic Reform with Limited Financial Integration”, *Journal of International Money and Finance*, 2013, 37, 493-503.
- [18] Li, C., “Savings, Investment, and Capital Mobility within China”, *China Economic Review*, 2010, 21 (1), 14-23.
- [19] Lin, J. Y., “New Structural Economics: A Framework for Rethinking Development”, *World Bank Research Observer*, 2011, 26 (2), 193-221.
- [20] 李力行、申广军, “经济开发区、地区比较优势与产业结构调整”, 《经济学》(季刊), 2015年第14卷第3期, 第885—910页。
- [21] 李艳、杨汝岱, “地方国企依赖、资源配置效率改善与供给侧改革”, 《经济研究》, 2018年第2期, 第80—94页。
- [22] 林毅夫、蔡昉、李周, 《中国的奇迹: 发展战略与经济改革》。上海: 上海人民出版社, 1994年。
- [23] 刘冲、吴群锋、刘青, “交通基础设施、市场可达性与企业生产率——基于竞争和资源配置的视角”, 《经济研究》, 2020年第7期, 第140—158页。
- [24] Murphy, K. M., A. Shleifer, and R. W. Vishny, “Industrialization and the Big Push”, *Journal of Political Economy*, 1989, 97 (5), 1003-1026.
- [25] Rosenstein-Rodan, P. N., “Problems of Industrialization of Eastern and Southeastern Europe”, *Economic Journal*, 1943, 53, 202-211.
- [26] 张军, “资本形成、工业化与经济增长: 中国的转轨特征”, 《经济研究》2002年第6期, 第3—13页。
- [27] 张军、吴桂英、张吉鹏, “中国省际物质资本存量估算: 1952—2000”, 《经济研究》2004年第10期, 第35—44页。
- [28] 赵婷、陈钊, “比较优势与产业政策效果: 区域差异及制度成因”, 《经济学》(季刊), 2020年第3期, 第777—796页。

Factor Endowment Structure and the Long-Run Impact of the Early-Stage Industrial Constructions in China

ZHANG Haochen WANG Xin* LIN Justin Yifu
(Peking University)

Abstract: Based on the micro-data of the National Industrial Census in 1985, we exploit variations in the industrial structure and geographical distribution of the industrial construction before 1978 to discuss how these industrial investments affect the long-term development of local industrial structure and how factor endowment structure shapes the magnitude of this impact. Our results show that the industrial construction in a certain industry before 1978 has significantly positive impact on the long-term scale of that industry. More importantly, deviation from the comparative advantage determined by factor endowment structure significantly lowers this long-run impact.

Keywords: early-stage industrial constructions; factor endowment structure; long-run impact

JEL Classification: O14, R11, R12

* Corresponding Author: Wang Xin, Institute of New Structural Economics, Peking University, No.5 Yiheyuan Road, Beijing 100871, China; Tel: 86-15810480552; E-mail: xin.wang@nsd.pku.edu.cn.