**附录A**



**图A1 近代出口贸易与中国区域经济差距**

注：图A1 在地级市层面计算基尼系数和相关系数，出口数据来自杨端六和侯厚培（1931），1820—1910年、1953—2010年人口数据分别来自曹树基（2001）和新中国历次人口普查。

图A2（a）、（b）表明，近代出口较发达的口岸在1933年拥有较高的工业劳动力密度和工资水平，其当代劳动力流入比例与平均工资也相对较高。图A2（c）显示，在近代出口较发达的口岸，劳动密集型产业从近代存续至今的企业固定资产密度以及相应的技术工人密度均较高，说明近代出口贸易以企业为历史沉没投资的主要载体，为当代比较优势产业发展提供重要的人力资本要素支持。

****

**（a）1933年**

****

**（b）2005年**

****

**（c）2004年**

**图A2 近代出口贸易的长期影响机制**

注：图A2关注近代贸易数据可得的口岸地级市，横轴为1911—1924年平均出口总值。图（a）中劳动力密度和工资数据来自刘大钧（1937），其中工资记录在“省份-行业”层面，考虑到不同行业的工资可比性较差，本文仅取纺织业的最高、最低工资算术平均值，并将出口总值加总至省份层面。图（b）中，劳动力流入比例与工资数据来自2005年1%人口抽样调查。图（c）中，企业固定资产和技术工人数据来自2004年经济普查。

表A1 列举五个传统经济中心与近代主要口岸分离的省份，近代前（1776—1820年）这些省份的传统经济中心多数保持着相对较快的人口增长，而在1910—1953年间，口岸地区的人口增幅均超过前者。这一方面印证了国际贸易对国内经济空间变迁的推动作用，另一方面也说明口岸选址与地区初始经济发展水平的相关性较弱。

**表A1 传统区域经济中心与近代口岸的人口增幅对比**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **近代省份** | **城市类型** | **地级市** | **1776-1820年人口增幅** | **1910-1953年人口增幅** |
| 江苏省 | 传统经济中心 | 苏州市 | 15.6% | 23.8% |
| 近代主要口岸 | 上海市 | 15.1% | 64.6% |
| 直隶省 | 传统经济中心 | 北京市 | 94.2% | 27.8% |
| 近代主要口岸 | 天津市 | 19.2% | 42.0% |
| 山东省 | 传统经济中心 | 济南市 | 14.1% | 14.6% |
| 近代主要口岸 | 烟台市、青岛市 | 10.3% | 83.6% |
| 安徽省 | 传统经济中心 | 安庆市 | 22.5% | 9.5% |
| 近代主要口岸 | 芜湖市 | 25.2% | 34.8% |
| 四川省 | 传统经济中心 | 成都市 | 42.0% | 43.2% |
| 近代主要口岸 | 重庆市 | 32.9% | 53.1% |

注：数据来自曹树基（2001）和新中国人首次人口普查，本文将近代府人口匹配至当代地级市后计算其增幅。

**表A2 近代通商口岸信息**

|  |  |  |
| --- | --- | --- |
| **口岸所在府** | **所在省份** | **口岸名称（开埠年份）** |
| 太平府 | 安徽省 | 芜湖（1877） |
| 凤凰厅 | 奉天省 | 安东（1907） |
| 庄河厅 | 奉天省 | 大东沟（1907） |
| 奉天府 | 奉天省 | 大连（1899）、奉天（1908） |
| 营口厅 | 奉天省 | 牛庄（1861） |
| 福宁府 | 福建省 | 三都澳（1899） |
| 福州府 | 福建省 | 福州（1844） |
| 泉州府 | 福建省 | 厦门（1843）、鼓浪屿（1902） |
| 潮州府 | 广东省 | 汕头（1860） |
| 广州府 | 广东省 | 广州（1843）、九龙（1887）、拱北（1887）、三水（1897）、江门（1904） |
| 琼州府 | 广东省 | 琼州（1876） |
| 廉州府 | 广西省 | 北海（1877） |
| 南宁府 | 广西省 | 南宁（1907） |
| 太平府 | 广西省 | 龙州（1889） |
| 梧州府 | 广西省 | 梧州（1897） |
| 瑷珲厅 | 黑龙江省 | 爱珲（1907） |
| 胪滨府 | 黑龙江省 | 满洲里（1910） |
| 汉阳府 | 湖北省 | 汉口（1862） |
| 荆州府 | 湖北省 | 沙市（1896） |
| 宜昌府 | 湖北省 | 宜昌（1877） |
| 岳州府 | 湖南省 | 岳阳（1899） |
| 长沙府 | 湖南省 | 长沙（1904） |
| 宾州府 | 吉林省 | 哈尔滨（1907） |
| 宁安府 | 吉林省 | 绥芬河（1907） |
| 依兰府 | 吉林省 | 三姓（1909） |
| 新民府 | 吉林省 | 龙井村（1909） |
| 延吉府 | 吉林省 | 珲春（1910） |
| 江宁府 | 江苏省 | 江宁（1899） |
| 松江府 | 江苏省 | 上海（1843）、吴淞（1898） |
| 苏州府 | 江苏省 | 苏州（1896） |
| 镇江府 | 江苏省 | 镇江（1861） |
| 九江府 | 江西省 | 九江（1862） |
| 登州府 | 山东省 | 芝罘（1862）、威海卫（1898） |
| 济南府 | 山东省 | 济南（1906） |
| 胶州 | 山东省 | 青岛（1898） |
| 重庆府 | 四川省 | 重庆（1891） |
| 临安府 | 云南省 | 蒙自（1889） |
| 普洱府 | 云南省 | 思茅（1897） |
| 永昌府 | 云南省 | 腾越（1902） |
| 杭州府 | 浙江省 | 杭州（1902） |
| 宁波府 | 浙江省 | 宁波（1844） |
| 温州府 | 浙江省 | 温州（1877） |
| 天津府 | 直隶省 | 天津（1861） |
| 永平府 | 直隶省 | 秦皇岛（1901） |

注：口岸信息来自严中平（1955）。本文统计的均是在1911年前开埠、并且近代贸易数据可得的口岸所在府，其中，大连、威海卫、青岛等部分口岸的开放时间以其设立租借地的时间为准。此外，排除了新疆、西藏、内蒙古以及台湾省的口岸。1840-1911年间，中国共有67处开埠口岸（不含新疆、西藏、内蒙古、台湾口岸）。

**表A3 主要变量定义及描述性统计**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 变量定义 | 样本数 | 均值 | 标准差 | 最小值 | 最大值 |
| 人口密度（人/平方千米） | 1772 | 322.29 | 343.12 | 0.61 | 5537.28 |
| 出口可达性 | 246 | 739.60 | 576.11 | 207.92 | 4120.51 |
| 工具变量：条约口岸时空分布的综合影响 | 246 | 12.71 | 12.22 | 3.40 | 100.83 |
| 1933年工业企业密度（企业数/平方千米） | 205 | 0.26 | 1.12 | 0 | 11.21 |
| 近代存续至1985年的劳动密集型企业固定资产密度（万元/平方千米） | 246 | 0.47 | 2.59 | 0 | 36.91 |
| 近代存续至2004年的劳动密集型企业固定资产密度（万元/平方千米） | 246 | 0.09 | 0.38 | 0 | 4.02 |
| 经度 | 246 | 112.28 | 5.79 | 96.06 | 121.98 |
| 纬度 | 246 | 30.88 | 5.03 | 19.83 | 40.86 |
| 到省会的直线距离（千米） | 246 | 181.87 | 115.98 | 10 | 804.05 |
| 是否副省级城市或直辖市（“是”=1） | 246 | 0.06 | 0.24 | 0 | 1 |
| 到航运河流的直线距离（千米） | 246 | 227.51 | 211.36 | 0.03 | 1351.94 |
| 1949年是否已通铁路（“是”=1） | 246 | 0.56 | 0.50 | 0 | 1 |
| 经济开发区数量 | 984 | 2.16 | 3.90 | 0 | 41 |
| “入世”冲击 | 246 | 9.53 | 3.05 | 1.19 | 23.52 |

本文参考Jia（2014），运用多时点DID策略考察近代出口贸易对区域人口长期增长的动态影响，回归方程如式（A1）所示：

，

(A1)

其中，*It*为时期虚拟变量，*μi*和*μt*分别为地级市和年份固定效应，其余变量定义与式（4）相同。本文以1910年为基期，关注系数*βt*的变化趋势。

表A4列（1）回归结果显示，ln *Export\_accessi*×*It*的系数在1820—1880年均不显著，说明20世纪前发展尚不充分的出口贸易并未对区域人口增长造成显著影响；1953、1964、2000和2010四个年份的系数显著为正，说明近代出口贸易对区域人口长期增长具有显著的积极影响。1982、1990年的系数符号为正但不显著，可能是由于计划经济时期的经济内迁一定程度上弱化了近代出口的长期影响。总体来看，1953—2010年系数值变化趋势与表1保持一致。列（2）加入“省份×年份”固定效应以控制人口增长趋势的省际差异，列（3）则利用地理特征变量匹配所得的子样本缓解地理异质性的干扰，回归结果均保持稳健。

**表A4 近代出口贸易与区域人口长期增长：面板回归**

|  |  |
| --- | --- |
| 被解释变量 | ln *Densityi,t* |
|  | (1) | (2) | (3) |
|  | 基准DID回归 | 加入省份×年份固定效应 | 匹配样本 |
| ln *Export\_accessi*×*I*(1820) | 0.2018 | 0.2914 | 0.1980 |
|  | (0.1510) | (0.2365) | (0.1381) |
| ln *Export\_accessi*×*I*(1851) | -0.1527 | -0.1288 | -0.1452 |
|  | (0.1396) | (0.1422) | (0.1256) |
| ln *Export\_accessi*×*I*(1880) | 0.2061 | 0.2872 | 0.2186 |
|  | (0.1421) | (0.2285) | (0.1433) |
| ln *Export\_accessi*×*I*(1953) | 0.4968\*\* | 0.4997\*\* | 0.7127\*\*\* |
|  | (0.2060) | (0.1879) | (0.2033) |
| ln *Export\_accessi*×*I*(1964) | 0.4300\* | 0.4405\* | 0.6588\*\*\* |
|  | (0.2277) | (0.2155) | (0.2109) |
| ln *Export\_accessi*×*I*(1982) | 0.3811 | 0.4197\* | 0.6242\*\*\* |
|  | (0.2407) | (0.2132) | (0.2026) |
| ln *Export\_accessi*×*I*(1990) | 0.3707 | 0.4401\*\* | 0.5899\*\*\* |
|  | (0.2348) | (0.1936) | (0.2012) |
| ln *Export\_accessi*×*I*(2000) | 0.4596\*\* | 0.5665\*\*\* | 0.6046\*\*\* |
|  | (0.2196) | (0.1623) | (0.2095) |
| ln *Export\_accessi*×*I*(2010) | 0.5088\*\* | 0.6684\*\*\* | 0.6263\*\*\* |
|  | (0.2140) | (0.1607) | (0.2142) |
| 控制变量 | Y | Y | Y |
| 地级市固定效应 | Y | Y | Y |
| 年份固定效应 | Y | Y | Y |
| 省份×年份固定效应 | N | Y | N |
| 观测值 | 2460 | 2460 | 1530 |
| 调整*R*2 | 0.8551 | 0.8934 | 0.8573 |

注：（1）\*\*\* *p*<0.01, \*\* *p*<0.05, \* *p*<0.1；（2）限于篇幅，未报告控制变量结果。

表A5 以自然地理特征作为协变量，针对历史口岸与非口岸地区进行1:3近邻匹配。匹配后各协变量的均值标准化偏差均小于10，说明匹配样本具有良好的平衡性。

**表A5 匹配样本的平衡性检验**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| PSM：1:3近邻匹配 |  | 均值 | 均值标准化偏差 | *T*检验 |
| 协变量 |  | 口岸地区 | 非口岸地区 | *t*统计量 | *p*值 |
| 经度（取对数） | 匹配前匹配后 | 4.73924.7384 | 4.71154.7387 | 54.4-0.7 | 3.88-0.04 | 0.0000.967 |
| 纬度（取对数） | 匹配前匹配后 | 3.36793.3832 | 3.43673.3914 | -41.1-4.9 | -2.99-0.28 | 0.0030.781 |
| 地理面积（取对数） | 匹配前匹配后 | 9.03319.1362 | 9.45929.1156 | -55.22.7 | -4.100.16 | 0.0000.874 |
| 到海岸线直线距离（取对数） | 匹配前匹配后 | 4.19634.3899 | 5.42334.3278 | -76.53.9 | -5.650.21 | 0.0000.835 |
| 到航运河流直线距离（取对数） | 匹配前匹配后 | 3.84534.1563 | 4.93484.1224 | -56.51.8 | -4.410.10 | 0.0000.923 |
| 平均坡度 | 匹配前匹配后 | 1.79081.8893 | 2.74791.9603 | -48.4-3.6 | -3.31-0.24 | 0.0010.812 |
| 农业种植适宜度（取对数） | 匹配前匹配后 | 7.86417.9241 | 7.64847.9162 | 24.00.9 | 1.650.05 | 0.1000.956 |

表A6采取多种策略对工具变量进行外生性检验：Panel A在回归样本和工具变量*Trade\_endowi*的构建中均剔除选址内生性较强的条约口岸，回归结果与表2保持一致。Panel B进行“半简化式”回归，结果显示ln *Export*

*\_accessi*的系数仍保持显著为正，只是系数值相比于表1略有下降；而工具变量ln *Trade\_endowi*的系数不再显著，说明工具变量主要通过近代出口可达性影响区域人口长期增长。

Panel C则针对一系列反映地区初始经济社会特征的潜在遗漏变量进行证伪检验[[1]](#footnote-0)：列（1）以1820—1880年人口密度的对数差分作为被解释变量，该指标反映了近代出口贸易繁荣之前的区域经济增长趋势；列（2）以地级市到京杭大运河的直线距离作为被解释变量，京杭大运河作为贯穿中国南北的经济动脉，有力推动了国内市场整合，也增强了中央政府对地方经济发展的控制力（Jia，2014）；列（3）构建1820年土地税收规模指标，土地税收与地区农业经济发展水平及政府财政能力紧密相关；列（4）参考Banerjee *et al.*（2020）构建“冲繁疲难”的第一主成分指标[[2]](#footnote-1)，清代实施“冲繁疲难”制度作为地方治理的重要依据，该制度综合衡量了一地区的交通便利程度、政治地位、政府财政能力以及社会稳定性等政治经济特征，被赋予属性越多的地区越有可能受到中央政府的重点关注；列（5）参考Bai and Jia（2016），利用1820年佛寺数量作为地级市近代社会资本存量的代理指标。上述变量对ln *Trade\_endowi*回归的系数均不显著，说明这些初始经济特征造成遗漏变量偏误的可能性较小。

**表A6 工具变量的外生性检验**

|  |  |
| --- | --- |
| Panel A：工具变量和样本均剔除选址内生性较强的口岸 | 被解释变量：ln *Densityi,t* – ln *Densityi,*1910 |
| 1953年 | 1964年 | 1982年 | 1990年 | 2000年 | 2010年 |
|  | (1) | (2) | (3) | (4) | (5) | (6) |
| ln *Export\_accessi* | 1.1205\*\*\* | 1.0441\*\*\* | 0.7422\*\* | 0.8711\*\* | 1.1888\*\*\* | 1.2846\*\*\* |
|  | (0.3280) | (0.3822) | (0.3387) | (0.3590) | (0.4034) | (0.3783) |
| K-P Wald F 统计量 | 7.66 | 7.66 | 7.71 | 6.51 | 7.43 | 8.76 |
| 控制变量 | Y | Y | Y | Y | Y | Y |
| 省份固定效应 | Y | Y | Y | Y | Y | Y |
| 观测值 | 236 | 236 | 236 | 236 | 236 | 236 |
| 调整*R*2 | 0.5655 | 0.5636 | 0.5985 | 0.5741 | 0.4923 | 0.5193 |
| Panel B：半简化式回归 | 被解释变量：ln *Densityi,t* – ln *Densityi,*1910 |
|  | 1953年 | 1964年 | 1982年 | 1990年 | 2000年 | 2010年 |
|  | (1) | (2) | (3) | (4) | (5) | (6) |
| ln *Export\_accessi* | 0.4416\*\*\* | 0.3700\*\* | 0.4285\*\*\* | 0.3749\*\* | 0.4253\*\*\* | 0.5208\*\*\* |
|  | (0.1307) | (0.1515) | (0.1375) | (0.1463) | (0.1271) | (0.1323) |
| ln *Trade\_endowi* | 0.2851 | 0.2799 | 0.1299 | 0.2340 | 0.2752 | 0.1679 |
|  | (0.1729) | (0.1908) | (0.1631) | (0.1905) | (0.2159) | (0.2097) |
| 控制变量 | Y | Y | Y | Y | Y | Y |
| 省份固定效应 | Y | Y | Y | Y | Y | Y |
| 观测值 | 246 | 246 | 246 | 246 | 246 | 246 |
| 调整*R*2 | 0.5955 | 0.5967 | 0.6160 | 0.5995 | 0.5464 | 0.5710 |
| Panel C：证伪检验 | 1820-1880年人口密度增长 | 到京杭大运河距离 | 1820年土地税收规模 | “冲繁疲难”第一主成分 | 1820年佛寺数量 |
|  | (1) | (2) | (3) | (4) | (5) |
| ln *Trade\_endowi* | 0.0407 | -0.3985 | -0.1822 | -0.1719 | 1.1335 |
|  | (0.0422) | (0.2877) | (0.3494) | (0.4439) | (0.6894) |
| 控制变量 | Y | Y | Y | Y | Y |
| 省份固定效应 | Y | Y | Y | Y | Y |
| 观测值 | 246 | 246 | 246 | 246 | 246 |
| 调整*R*2 | 0.3286 | 0.7629 | 0.6353 | 0.1651 | 0.3673 |

注：（1）\*\*\* *p*<0.01, \*\* *p*<0.05, \* *p*<0.1；（2）Panel A为IV-2SLS回归，Panel B、Panel C为OLS回归，Panel C控制变量包含经纬度、到航运河流的直线距离、到省会直线距离；（3）限于篇幅，未报告第一阶段及控制变量结果。

国际贸易发展受到经济社会演化的深刻影响。本文所关注的1910—20年代正处于中国近代变革最为剧烈的时期之一，出口贸易经历了第一次世界大战等历史事件冲击，贸易产品结构、贸易方向均发生重大变化。因此，表A7从不同维度对近代出口贸易的长期影响展开异质性分析。列（1）和列（2）分别利用一战结束（1918年）前后的贸易数据构建出口可达性，回归结果显示两个时间段的出口均产生显著的积极影响，后者的影响强度相对较大，这可能是因为一战后劳动密集的轻工业资本积累加速，出口快速增长（Mitchener and Yan，2014）。列（3）—（5）利用三种代表性出口产品的贸易数量数据进行识别，其中棉纱（代表轻工业产品）出口的积极影响最为显著，豆类（代表农产品）和煤炭（代表工业原材料）的影响显著度相对较低，这与本文关于近代出口贸易通过工业化推动区域人口长期增长的理论分析相吻合。列（6）和列（7）则聚焦于出口目的地差异，发现中国对在华殖民国家[[3]](#footnote-2)的出口具有相对较强的积极影响。殖民国家与中国经济联系紧密，往来贸易规模较大，且工业化程度较高，其市场需求更有可能对中国工业投资增长起到引导作用。

**表A7 近代出口贸易的异质性分析**

|  |  |
| --- | --- |
| 被解释变量 | ln *Densityi,*2010 – ln *Densityi,*1910 |
| 出口类别 | 1911—1918年 | 1919—1924年 | 豆类 | 棉纱 | 煤炭 | 在华殖民国家 | 其他国家 |
|  | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| ln *Export\_accessi* | 1.0089\*\*\* | 1.9161\*\*\* | 4.4672\* | 1.4651\*\*\* | -90.0322 | 1.7163\*\*\* | 1.1479\*\* |
|  | (0.2650) | (0.5431) | (2.6181) | (0.5487) | (517.5043) | (0.6328) | (0.4471) |
| K-P Wald F 统计量 | 34.79 | 9.44 | 2.26 | 4.82 | 0.03 | 15.19 | 9.87 |
| 控制变量 | Y | Y | Y | Y | Y | Y | Y |
| 省份固定效应 | Y | Y | Y | Y | Y | Y | Y |
| 观测值 | 246 | 246 | 246 | 246 | 246 | 246 | 246 |
| 调整*R*2 | 0.5410 | 0.4912 | 0.0219 | 0.4879 | -2.9e+02 | 0.4773 | 0.5092 |

注：（1）\*\*\* *p*<0.01, \*\* *p*<0.05, \* *p*<0.1；（2）列（1）—（5）出口数据来自杨端六和侯厚培（1931）；列（6）、列（7）来自蔡谦和郑友揆（1936）；（3）限于篇幅，仅对2010年样本进行回归，未报告第一阶段及控制变量结果。

 表A8 Panel A将样本限定在近代口岸及其接壤的地级市中，直接考察口岸1911—1924年平均出口总值（*Exporti*）的长期影响[[4]](#footnote-3)。该指标能够排除距离因素影响，更准确地反映地区实际出口发展水平。回归结果与基准回归保持一致，说明近代出口贸易的长期影响受距离因素的干扰较小，出口不平衡发展在口岸之间也造成了长期经济差距。

Panel B将式（1）中的出口规模替换为1910年口岸人口规模，构建“国内市场可达性”指标（*DMA*1910*i*）进行安慰剂检验，该指标数值越大，说明地级市*i*越接近本土市场规模较大的口岸，国内贸易的区位优势越明显。回归结果显示，ln *DMA*1910*i*的系数在所有年份均不显著，说明出口可达性指标中的距离信息并未通过国内贸易影响区域经济长期增长。

Panel C控制了一系列反映制度、基础设施、文化观念等方面的变量（*CHi*），以排除口岸开埠的竞争性影响渠道干扰[[5]](#footnote-4)：列（1）和列（2）分别利用市场化指数和产权保护指数刻画当代地级市制度质量，以排除口岸制度优势的长期影响（Jia，2014）；列（3）和列（4）分别控制城市基础设施竞争力指数和半径450km内铁路里程总数，以排除口岸基础设施建设的长期影响；列（5）利用1920年教会学校数量控制地级市受西方文化的影响程度，列（6）则利用地级市居民对他人的平均信任程度控制近代口岸的观念塑造作用（张川川等，2021）。在控制上述变量后，ln *Export\_accessi*保持正向显著，说明近代出口贸易的长期影响没有受到这些竞争性渠道干扰。同时，回归结果也表明基础设施建设和居民间信任感的提升有利于区域经济增长。

**表A8 排除竞争性假说**

|  |  |
| --- | --- |
| Panel A：排除距离因素干扰 | 被解释变量：ln *Densityi,t* – ln *Densityi,*1910 |
|  | 1953年 | 1964年 | 1982年 | 1990年 | 2000年 | 2010年 |
|  | (1) | (2) | (3) | (4) | (5) | (6) |
| ln *Exporti* | 0.1475\*\*\* | 0.1585\*\*\* | 0.1449\*\* | 0.1488\*\*\* | 0.1858\*\*\* | 0.1971\*\* |
|  | (0.0493) | (0.0592) | (0.0573) | (0.0528) | (0.0694) | (0.0769) |
| K-P Wald F 统计量 | 5.95 | 5.95 | 5.10 | 5.72 | 6.24 | 6.44 |
| 控制变量 | Y | Y | Y | Y | Y | Y |
| 省份固定效应 | Y | Y | Y | Y | Y | Y |
| 观测值 | 112 | 112 | 112 | 112 | 112 | 112 |
| 调整*R*2 | 0.6173 | 0.6207 | 0.6778 | 0.6412 | 0.5482 | 0.5220 |
| Panel B：国内市场可达性 | 被解释变量：ln *Densityi,t* – ln *Densityi,*1910 |
|  | 1953年 | 1964年 | 1982年 | 1990年 | 2000年 | 2010年 |
|  | (1) | (2) | (3) | (4) | (5) | (6) |
| ln *DMA*1910*i* | 0.4034 | 0.4158 | 0.3663 | 0.2961 | 0.3005 | 0.3973 |
|  | (0.2432) | (0.2604) | (0.2360) | (0.2294) | (0.2361) | (0.2460) |
| 控制变量 | Y | Y | Y | Y | Y | Y |
| 省份固定效应 | Y | Y | Y | Y | Y | Y |
| 观测值 | 246 | 246 | 246 | 246 | 246 | 246 |
| 调整*R*2 | 0.5361 | 0.5580 | 0.5859 | 0.5627 | 0.4965 | 0.5223 |
| Panel C：其他竞争性渠道 | 被解释变量：ln *Densityi,*2010 – ln *Densityi,*1910 |
| *CHi*定义 | 市场化指数 | 产权保护指数 | 基础设施竞争力指数 | 铁路里程数 | 教会学校数量 | 平均信任度  |
|  | (1) | (2) | (3) | (4) | (5) | (6) |
| ln *Export\_accessi* | 1.2541\*\*\* | 1.0575\*\*\* | 1.0512\*\*\* | 1.4856\*\*\* | 1.2210\*\*\* | 1.7450\*\*\* |
|  | (0.2765) | (0.2209) | (0.2518) | (0.2452) | (0.2901) | (0.5015) |
| *CHi* | -0.1049 | -0.2235 | 0.2865\*\* | 0.4418\*\*\* | 0.0152 | 0.3945\* |
|  | (0.1548) | (0.1767) | (0.1213) | (0.0833) | (0.0215) | (0.2274) |
| K-P Wald F 统计量 | 12.71 | 18.40 | 8.18 | 12.23 | 12.01 | 12.41 |
| 控制变量 | Y | Y | Y | Y | Y | Y |
| 省份固定效应 | Y | Y | Y | Y | Y | Y |
| 观测值 | 245 | 104 | 246 | 246 | 246 | 114 |
| 调整*R2* | 0.5348 | 0.5988 | 0.5564 | 0.5534 | 0.5385 | 0.4347 |

注：（1）\*\*\* *p*<0.01, \*\* *p*<0.05, \* *p*<0.1；（2）表A8 Panel A的工具变量为ln *Port\_endowi*；（3）Panel B为OLS回归；（4）限于篇幅，Panel C仅对2010年样本进行回归，未报告第一阶段及控制变量结果。

表A9在纳入东北地级市样本后，利用两种替代性指标考察近代出口贸易对区域经济集聚水平的长期影响：其一是1953—2010年地级市人口密度水平值（*Densityi,t*），其二是2000—2012年地级市夜间灯光平均强度（*Lighti,t*）, 该指标能够比较准确地综合反映地区经济集聚水平与经济活跃程度。回归结果显示，近代出口贸易对不同时期的人口密度和夜间灯光强度均具有显著的积极影响，而且其对人口密度的影响强度时间趋势也与基准回归保持一致，该结果说明近代出口贸易在不同时期均显著促进了经济集聚。

**表A9 近代出口贸易对经济集聚水平的长期影响**

|  |  |
| --- | --- |
| Panel A | 被解释变量：ln *Densityi,t* |
|  | 1953年 | 1964年 | 1982年 | 1990年 | 2000年 | 2010年 |
|  | (1) | (2) | (3) | (4) | (5) | (6) |
| ln *Export\_accessi* | 0.8714\*\* | 0.8008\*\* | 0.7526\*\* | 0.6818\*\* | 1.0315\*\* | 1.1678\*\* |
|  | (0.3467) | (0.3460) | (0.2965) | (0.3178) | (0.5097) | (0.5167) |
| K-P Wald F 统计量 | 8.73 | 8.73 | 8.76 | 8.38 | 8.76 | 8.71 |
| 控制变量 | Y | Y | Y | Y | Y | Y |
| 省份固定效应 | Y | Y | Y | Y | Y | Y |
| 观测值 | 280 | 280 | 280 | 280 | 280 | 280 |
| 调整*R*2 | 0.6835 | 0.6544 | 0.6506 | 0.6639 | 0.6356 | 0.6511 |
| Panel B | 被解释变量：ln *Lighti,t* |
|  | 2000年 | 2004年 | 2008年 | 2012年 |
|  | (1) | (2) | (3) | (4) |
| ln *Export\_accessi* | 1.3165\*\* | 1.3362\*\* | 1.3562\*\*\* | 1.1681\*\*\* |
|  | (0.5474) | (0.5235) | (0.5107) | (0.4195) |
| K-P Wald F 统计量 | 8.76 | 8.65 | 8.60 | 8.71 |
| 控制变量 | Y | Y | Y | Y |
| 省份固定效应 | Y | Y | Y | Y |
| 观测值 | 280 | 280 | 280 | 280 |
| 调整*R*2 | 0.7240 | 0.7176 | 0.7218 | 0.7216 |

注：（1）\*\*\* *p*<0.01, \*\* *p*<0.05, \* *p*<0.1；（2）表A9控制变量不含1851—1910年人口密度的前定差分；（3）限于篇幅，未报告第一阶段及控制变量结果。

**表A10 近代出口贸易对地级市实际人均GDP的长期影响**

|  |  |
| --- | --- |
| 被解释变量 | ln *Per\_GDPi,t* |
|  | 1995年 | 2000年 | 2005年 | 2010年 |
|  | (1) | (2) | (3) | (4) |
| ln *Export\_accessi* | 1.6321\* | 1.7537\*\* | 1.9284\*\* | 2.1240\*\* |
|  | (0.8383) | (0.8621) | (0.9406) | (0.9898) |
| K-P Wald F 统计量 | 6.80 | 7.96 | 8.01 | 8.24 |
| 控制变量 | Y | Y | Y | Y |
| 省份固定效应 | Y | Y | Y | Y |
| 观测值 | 195 | 252 | 278 | 278 |
| 调整*R*2 | 0.1316 | 0.2514 | 0.2715 | 0.1717 |

注：（1）\*\*\* *p*<0.01, \*\* *p*<0.05, \* *p*<0.1；（2）由于《城市统计年鉴》数据缺失和行政区划调整，各年份样本量有所不同，样本含东北地区；（3）表A10控制变量不含1851—1910年人口密度的前定差分；（4）限于篇幅，未报告第一阶段及控制变量结果。

表A11汇报了排除后续事件冲击干扰的稳健性检验结果[[6]](#footnote-5)。我们重点关注的事件包括抗日战争、“三线建设”、当代出口贸易、当代政府财政支出、经济开发区政策以及“入世”冲击。其中，列（1）以各地级市在抗战期间卷入主要战役的时间长度（以月度为单位）作为战争破坏程度的代理指标（*WW*2）；列（2）以“是否三线建设重点项目所在地”（*TF*）刻画三线建设政策的影响；列（3）构建地级市当代出口贸易依存度指标（*Exp\_dep*，即出口总值占地级市GDP的比重）以辨析近代和当代出口贸易对区域经济增长的长短期影响；列（4）利用人均财政支出规模（*Finan\_exp*）刻画当代地方政府的财政能力；列（5）、（6）则分别在基准回归基础上考察经济开发区政策（*SEZ*）和“入世”（WTO）冲击对近代出口贸易影响的调节效应。Panel A为调节效应分析，结果显示ln *Export\_accessi*与抗日战争、“三线建设”、经济开发区政策以及人均财政支出的交互项均不显著，说明这些事件并不会边际改变近代出口贸易的长期影响；2010年出口贸易依存度则产生了正向调节效应，说明当代出口贸易发展能够进一步增强近代出口贸易对区域经济增长的积极影响；“入世”冲击对近代出口可达性产生了负向调节效应，其原因可能是“入世”所引致的进口自由化增加了本土厂商的竞争压力，降低了厂商的集聚激励，从而弱化了近代出口贸易的长期影响。Panel B则在基准回归中控制上述事件变量，其中*TFi*和*Exp\_depi*,2010的系数显著为正，说明“三线建设”政策和当代出口贸易发展也有利于区域经济增长。特别是在控制*Exp\_depi*,2010后，ln *Export\_accessi*系数值较之基准回归明显减小，说明近代和当代出口贸易的积极影响有部分重合，可能的解释是近代出口贸易通过历史沉没投资促进了当代地区的出口导向型经济发展。在表A11所有的回归中，ln *Export\_accessi*一次项系数均保持显著为正，这表明本文长期因果识别没有受到后续事件冲击的严重干扰。

**表A11 排除后续事件冲击干扰**

|  |  |
| --- | --- |
| 被解释变量 | ln *Densityi,*2010– ln *Densityi,*1910 |
| *Event*定义 | ln *WW*2*i* | *TFi* | *Exp\_depi,*2010 | ln *Finan\_expi,*2010 | ln *SEZi,t* | *WTOi* |
|  | (1) | (2) | (3) | (4) | (5) | (6) |
| Panel A：调节效应分析 |
| ln *Export\_accessi* | 1.1998\*\*\* | 1.3431\*\*\* | 1.1206\*\*\* | 1.2982\*\*\* | 2.0726\*\*\* | 2.0291\*\*\* |
|  | (0.3207) | (0.2923) | (0.4367) | (0.2442) | (0.7394) | (0.4394) |
| ln *Export\_accessi*×*Event* | 0.0590 | -0.1164 | -0.5008 | 0.0460 | -0.2963 | -0.0981\*\*\* |
|  | (0.0377) | (0.1304) | (0.5358) | (0.2900) | (0.2134) | (0.0247) |
| K-P Wald F 统计量 | 4.38 | 6.09 | 4.53 | 6.94 | 3.38 | 7.22 |
| 控制变量 | Y | Y | Y | Y | Y | Y |
| 省份固定效应 | Y | Y | Y | Y | Y | Y |
| 观测值 | 246 | 246 | 228 | 228 | 246 | 246 |
| 调整*R*2 | 0.5420 | 0.5417 | 0.5800 | 0.5559 | 0.4999 | 0.5587 |
| Panel B：控制后续历史事件 |
| ln *Export\_accessi* | 1.2641\*\*\* | 1.3056\*\*\* | 0.9484\*\*\* | 1.2821\*\*\* |  |  |
|  | (0.2827) | (0.2839) | (0.3091) | (0.2554) |  |  |
| *Event* | -0.0229 | 0.2400\*\*\* | 0.5468\*\*\* | -0.1185 |  |  |
|  | (0.0229) | (0.0843) | (0.1600) | (0.1091) |  |  |
| K-P Wald F 统计量 | 12.52 | 12.72 | 7.34 | 12.80 |  |  |
| 控制变量 | Y | Y | Y | Y |  |  |
| 省份固定效应 | Y | Y | Y | Y |  |  |
| 观测值 | 246 | 246 | 228 | 228 |  |  |
| 调整*R*2 | 0.5363 | 0.5445 | 0.5880 | 0.5580 |  |  |

注：（1）\*\*\* *p*<0.01, \*\* *p*<0.05, \* *p*<0.1；（2）限于篇幅，仅对2010年样本进行回归，未报告第一阶段及控制变量结果。

**表A12 其他稳健性检验**

|  |  |
| --- | --- |
| 被解释变量 | ln *Densityi,t* – ln *Densityi,*1910 |
|  | 1953年 | 1964年 | 1982年 | 1990年 | 2000年 | 2010年 |
|  | (1) | (2) | (3) | (4) | (5) | (6) |
| （1）仅保留解释变量，控制八大经济区固定效应 |
| ln *Export\_accessi* | 0.8968\*\*\* | 0.9487\*\*\* | 0.8271\*\*\* | 0.8568\*\*\* | 1.0196\*\*\* | 1.1106\*\*\* |
|  | (0.1501) | (0.1481) | (0.1424) | (0.1401) | (0.1813) | (0.1835) |
| K-P Wald F 统计量 | 25.78 | 25.78 | 25.78 | 25.78 | 25.78 | 25.78 |
| 控制变量 | N | N | N | N | N | N |
| 经济区固定效应 | Y | Y | Y | Y | Y | Y |
| 观测值 | 246 | 246 | 246 | 246 | 246 | 246 |
| 调整*R*2 | 0.5304 | 0.5461 | 0.5741 | 0.5534 | 0.4809 | 0.4649 |
| （2）对*Export\_accessi*进行前后5%缩尾处理 |
| ln *Export\_accessi* | 0.9988\*\*\* | 0.9013\*\*\* | 0.5077\* | 0.5905\* | 1.0091\*\*\* | 0.9311\*\* |
|  | (0.3007) | (0.3382) | (0.2894) | (0.3485) | (0.3698) | (0.4259) |
| K-P Wald F 统计量 | 18.93 | 18.93 | 20.43 | 15.18 | 17.96 | 21.13 |
| 控制变量 | Y | Y | Y | Y | Y | Y |
| 省份固定效应 | Y | Y | Y | Y | Y | Y |
| 观测值 | 222 | 222 | 222 | 222 | 222 | 222 |
| 调整*R*2 | 0.5559 | 0.5658 | 0.6062 | 0.5804 | 0.5009 | 0.5335 |
| （3）解释变量替换为距最近前五口岸的出口可达性 |
| ln *Export\_accessF*5*i* | 0.4177\*\*\* | 0.3969\*\*\* | 0.2641\*\* | 0.3329\*\* | 0.4445\*\*\* | 0.4901\*\*\* |
|  | (0.1313) | (0.1499) | (0.1272) | (0.1532) | (0.1267) | (0.1238) |
| K-P Wald F 统计量 | 19.45 | 19.45 | 25.25 | 21.37 | 19.68 | 19.62 |
| 控制变量 | Y | Y | Y | Y | Y | Y |
| 省份固定效应 | Y | Y | Y | Y | Y | Y |
| 观测值 | 246 | 246 | 246 | 246 | 246 | 246 |
| 调整*R*2 | 0.4589 | 0.4747 | 0.5482 | 0.4866 | 0.3537 | 0.3434 |
| （4）出口可达性和样本均剔除选址内生性较强的口岸 |
| ln *Export\_accessi* | 1.7080\*\* | 1.6321\*\* | 1.0241 | 1.3380\* | 1.8884\*\* | 2.1620\*\* |
|  | (0.7712) | (0.7558) | (0.6556) | (0.8041) | (0.8989) | (1.0630) |
| K-P Wald F 统计量 | 7.50 | 7.50 | 7.88 | 6.03 | 7.41 | 7.47 |
| 控制变量 | Y | Y | Y | Y | Y | Y |
| 省份固定效应 | Y | Y | Y | Y | Y | Y |
| 观测值 | 236 | 236 | 236 | 236 | 236 | 236 |
| 调整*R*2 | 0.4995 | 0.5083 | 0.5768 | 0.5202 | 0.3355 | 0.2695 |
| （5）解释变量替换为1914-1920年进、出口可达性的年平均增长率 |
| *Exp\_growthi* (1914-1920) | 0.2432\*\* | 0.2532\*\* | 0.2382\*\* | 0.2259\* | 0.2393\*\* | 0.2689\*\* |
| (0.1169) | (0.1200) | (0.1097) | (0.1108) | (0.1086) | (0.1095) |
| 控制变量 | Y | Y | Y | Y | Y | Y |
| 省份固定效应 | Y | Y | Y | Y | Y | Y |
| 观测值 | 246 | 246 | 246 | 246 | 246 | 246 |
| 调整*R*2 | 0.5578 | 0.5691 | 0.5925 | 0.5714 | 0.4981 | 0.5163 |
| *Imp\_growthi* (1914-1920) | 0.0002 | -0.0164 | -0.0049 | -0.0115 | 0.0016 | 0.0234 |
| (0.2400) | (0.2313) | (0.2298) | (0.2165) | (0.2285) | (0.2450) |
| 控制变量 | Y | Y | Y | Y | Y | Y |
| 省份固定效应 | Y | Y | Y | Y | Y | Y |
| 观测值 | 246 | 246 | 246 | 246 | 246 | 246 |
| 调整*R*2 | 0.5452 | 0.5556 | 0.5793 | 0.5586 | 0.4829 | 0.4982 |
| （6）空间误差模型回归 |
| ln *Export\_accessi* | 0.6921\*\*\* | 0.6150\*\*\* | 0.5588\*\*\* | 0.5638\*\*\* | 0.6296\*\*\* | 0.6718\*\*\* |
| (0.1420) | (0.1427) | (0.1383) | (0.1161) | (0.1208) | (0.1231) |
| 控制变量 | Y | Y | Y | Y | Y | Y |
| 省份固定效应 | N | N | N | N | N | N |
| 观测值 | 246 | 246 | 246 | 246 | 246 | 246 |
| Log-likelihood | -153.52 | -156.34 | -137.96 | -132.74 | -143.70 | -149.38 |

注：（1）\*\*\* *p*<0.01, \*\* *p*<0.05, \* *p*<0.1；（2）限于篇幅，未报告第一阶段及控制变量结果。

表A13实证检验近代出口贸易能否在短期内通过劳动密集型产业投资激励人力资本投资。我们利用陈元晖（1991）的数据计算地级市*i*在1911—1926年间累计建立的新式学堂数量（*Nschooli,*1911-1926）度量近代人力资本投资，新式学堂是近代中国为适应工业化需求而建立的教育机构，该为劳动密集型产业发展提供了重要的人力资本支持（欧阳峣和易思维，2021）。由于该数据截至1926年，表A13改用杜恂诚（1991）的数据计算地级市*i*在1911—1924年间劳动密集型企业新增注册资本的累计金额（ln *Linvesti,*1911-1924）度量该产业投资水平。回归结果显示，近代出口贸易对劳动密集型产业投资（列（1））和人力资本投资（列（2））均有显著的积极影响，列（3）中交互项ln *Export\_accessi*×ln *Linvesti,*1911-1924系数显著为正，说明近代出口贸易可通过劳动密集型产业投资激励相应的人力资本投资。表A13的回归结果在短期视角下为历史沉没投资如何塑造地区要素禀赋补充了经验证据。

**表A13 近代出口贸易、劳动密集型产业投资与近代人力资本投资**

|  |  |  |  |
| --- | --- | --- | --- |
| 被解释变量 | ln *Linvesti,*1911-1924 | ln *Nschooli,*1911-1926  | ln *Nschooli,*1911-1926  |
|  | (1) | (2) | (3) |
| ln *Export\_accessi* | 6.4331\*\*\* | 2.6340\*\* | 0.4632 |
|  | (2.0683) | (1.0562) | (1.0176) |
| ln *Export\_accessi*×ln *Linvesti,*1911-1924 |  |  | 0.0144\*\*\* |
|  |  |  | (0.0051) |
| ln *Linvesti,*1911-1924 |  |  | 0.2016 |
|  |  |  | (0.4991) |
| K-P Wald F 统计量 | 12.77 | 12.77 | 6.16 |
| 控制变量 | Y | Y | Y |
| 省份固定效应 | Y | Y | Y |
| 观测值 | 246 | 246 | 246 |
| 调整*R*2 | 0.4537 | 0.2064 | 0.2922 |

注：（1）\*\*\* *p*<0.01, \*\* *p*<0.05, \* *p*<0.1；（2）表A13控制变量不含1851—1910年人口密度的前定差分；（3）限于篇幅，未报告第一阶段及控制变量结果。

**附录B 指标测算方法**

**1. 产权保护指数：**世界银行2005年的中国企业抽样调查数据询问了企业“在法律纠纷中，合同权利和财产受到保护的案件占比”。本文在地级市层面上计算该指标的算术平均值度量该地级市的产权保护水平。

**2. 基础设施竞争力指数：**倪鹏飞主编的《中国城市竞争力报告》（2010年版）提供了2009年中国56个主要城市的基础设施竞争力指标。本文以地级市*i*到这些主要城市的直线距离为权重，对该指标进行加权平均，得到每个地级市的相应变量。

**3. 铁路里程数：**本文参考Baum-Snowet al.（2020）,以地级市*i*的几何中心为圆心，利用Baum-Snow *et al.*（2017）所提供的中国铁路网络矢量地图，计算其半径450km范围内的铁路里程数总和。

**4. 居民对他人的平均信任度：**中国综合社会调查（CGSS）询问受访者“总的来说，您是否同意在这个社会上，绝大多数人都是可以信任的？”，回答选项取值在1-5之间，数值越大表示信任程度越高。本文在地级市层面上计算该指标的算术平均值度量当地居民对他人的平均信任度。

**5. 地区收入溢价：**鉴于现有的劳动力收入调查数据既呈现了各地区劳动力收入水平的整体差距，又与劳动力年龄、学历等个体特征以及行业结构的空间差异相关，本文参考Dix-Carneiro and Kovak（2017），基于2005年全国1%人口抽样调查数据构建“地区收入溢价”（Regional Earnings Premia）指标[[7]](#footnote-6)，具体构建方法如下：

 ， (B1)

其中，*Earningijm*表示定居于地级市*i*、从事行业*m*的个体劳动者*j*的月收入水平，*Xj*为劳动者的个体特征向量，包含年龄、性别和受教育年数等变量；*θm*为劳动者所属行业的固定效应，*μi*为地级市固定效应。本文通过回归式（B1）得到地区固定效应的拟合值，该指标即为排除劳动者个体特征差异与行业结构差异后的地区劳动力收入溢价（*Premiai*）。为确保被解释变量反映劳动者收入，本文将式（B1）的回归样本限定在18-65岁之间的劳动适龄人口。

**6. 企业管理效率：**根据孙浦阳等（2018），企业管理效率测算公式如下：

， (B2)

其中，*MCj*为企业*j*的管理费用，*lj*、*Expj*和*Markupj*分别为企业雇佣规模，出口额以及价格加成率（根据“企业总收入/（企业总收入-利润总额）”得到），*θj,K*为企业所属行业的固定效应。回归（B2）式得到残差*εi*即为企业管理效率的对数值。由于该指标取值越高表示管理效率越低，本文在实证中取-*εi*作为被解释变量。

**7. 区位熵：**计算公式为：

* ，*  (B3)

其中*Laborim*表示地区*i*从事产业*m*的劳动力数量。该指标计算产业*m*从业人数占地级市*i*总工人数之比相对于全国的比值，从而反映该产业的集聚程度。

**参考文献：**

[1] Bai, Y., and R. Jia, “Elite Recruitment and Political Stability: The Impact of the Abolition of China’s Civil Service Exam”, *Econometrica*, 2016, 84 (2), 677–733.

[2] Banerjee, A., E. Dufloa, and N. Qian, “On the Road: Access to Transportation Infrastructure and Economic Growth in China”, *Journal of Development Economics*, 2020, 145, No.102442.

[3] Baum-Snow, N., L. Brandt, J. Henderson, M. Turner, and Q. Zhang, “Roads, Railroads, and Decntralization of Chinese Cities”, *Review of Economics & Statistics*, 2017, 99(3), 435-448.

[4] Baum-Snow, N., L. Brandt, J. Henderson, M. Turner, and Q. Zhang, “Does Investment in National Highways Help or Hurt Hinterland City Growth?”, *Journal of Urban Economics*, 2020, 115, No. 103124.

[5] 蔡谦、郑友揆，“中国各通商口岸对各国进出口贸易统计”，《国立中央研究院社会科学研究所丛刊》，1936年第5期，第24—27页。

[6] 陈元晖，《中国近代教育史资料汇编》。上海：上海教育出版社，1991年。

[7] 杜恂诚，《民族资本主义与旧中国政府（1840—1937）》。上海：上海社会科学院出版社，1991年。

[8] 樊纲、王小鲁、朱恒鹏，《中国市场化指数：各地区市场化相对进程2009年报告》。北京：经济科学出版社，2010年。

[9] 梁方仲，《中国历代户口、田地、田赋统计》。北京：中华书局，2008年。

[10] Mitchener, K., and S. Yan, “Globalization, Trade, and Wages: What Does History Tell Us about China? ”, *International Economic Review*, 2014, 55(1), 131-167.

[11] 倪鹏飞，《中国城市竞争力报告》。北京：社会科学文献出版社，2010年。

[12] 《三线建设》编写组，《三线建设》。北京：国务院三线建设调整改造规划办公室，1991年。

[13] 武月星，《中国抗日战争史地图集》。北京：中国地图出版社，1995年。

[14] 中华续行委办会调查特委会（编），《1901—1920 年中国基督教调查资料》（蔡詠春译），北京：中国社会科学出版社，2007 年。

**注：该附录是期刊所发表论文的组成部分，同样视为作者公开发表的内容。如研究中使用该附录中的内容，请务必在研究成果上注明附录下载出处**。

1. 表A6 PanelC的被解释变量的数据来源如下：1820-1880年人口密度来自曹树基（2001）；到京杭大运河距离根据CHGIS矢量地图测算；1820年土地税收规模来自梁方仲（2008）；清末各府“冲繁疲难”名录来自《清史稿》，1820年佛寺数量来自哈佛大学Worldmap数据库。 [↑](#footnote-ref-0)
2. “冲繁疲难”最早提出于1731年（雍正六年），“冲”指处于交通要冲位置，“繁”指行政事务繁杂，“疲”指财政困难，“难”指社会动荡。清政府根据实际情况将这四种属性赋予每个州府，并据此任命地方官员。由于该制度中每一项属性对经济增长的影响方向难以确定，本文构建第一主成分指标。第一主成分是通过线性组合，将多维属性映射到一维，最大化表征所有个体特征差异的一种主成分分析指标。 [↑](#footnote-ref-1)
3. 指英、法、德、意、日、美、俄、比利时、奥匈帝国这些对中国大陆实施过殖民侵略或设有租界的国家。 [↑](#footnote-ref-2)
4. 接壤城市作为口岸直属腹地，与口岸市场联系十分紧密，故本文将每个口岸及其接壤城市视作一个整体，统一以该口岸的出口规模赋值。当一个非口岸地级市同时与多个口岸接壤时，我们按照各口岸的1910年人口规模为权重，计算加权平均的出口规模赋予该地级市。 [↑](#footnote-ref-3)
5. *CHi*变量的数据来源如下：市场化指数来自樊纲等（2010）；产权保护指数根据2005年世界银行的中国企业调查数据测算；基础设施竞争力指数根据倪鹏飞（2010）测算；铁路里程数根据Baum-Snow et al.（2017）提供的2010年铁路网络矢量图测算；教会学校数量来自《1901—1920年中国基督教调查资料》；平均信任度根据中国综合社会调查（CGSS）2010年数据测算。部分指标测算方法见附录。 [↑](#footnote-ref-4)
6. 后续历史事件变量的数据来源如下：抗日战争主要战役的时间和地点信息来自武月星（1995）；三线建设重点项目目录来自（《三线建设》编写组，1991）；地级市当代出口贸易数据来自区域统计年鉴，财政支出数据来自城市统计年鉴。 [↑](#footnote-ref-5)
7. 从1990年始，我国每5年展开一次微观人口抽样调查，但仅有2005年的调查数据中公开了个体收入信息。 [↑](#footnote-ref-6)