说明文档

**（一）文件说明**

本论文使用的数据代码文件包命名为“2024-01034+数据”，里面有**“2024-01034+程序代码”“2024-01034+日志文件”“2024-01034+说明文档”3**个子文件。

**（1）“2024-01034+程序代码”**里面包括2个文件，分别是“statistic\_vertical\_work”和“statistic\_vertical\_figure”文件：

“statistic\_vertical\_work”是表格运行程序，主要生成正文的表1-表6，在线附录II的表II1、在线附录III的表III1和表III2、在线附录IV的表IV1和表IV2的运行结果。

“statistic\_vertical\_figure”是图运行程序，主要生成正文的图1，以及在线附录III的图III1。

注：附录I的图I1、图I2和图I3，使用Excel界面进行手动绘制，因此无绘制程序。

**（2）“2024-01034+日志文件”**中的“Log\_正文\_表1”是正文中的表1“基准回归结果”的输出结果，“Log\_正文\_表2”是正文中的表2“准自然实验”的输出结果，“Log\_正文\_表3”是正文中的表3“工具变量估计”的输出结果，“Log\_正文\_表4”是正文中的表4“机制检验：中外企业协作与市场竞争”的输出结果，“Log\_正文\_表5”是正文中的表5“机制检验：信息传递”的输出结果，“Log\_正文\_表6”是正文中的表6“影响路径”的输出结果。“Log\_正文\_图1”是正文中图1“事前同趋势检验”的输出结果。

“Log\_附录\_表II1”是附录II中表II1“主要变量的描述性统计”的输出结果。“Log\_附录\_表III1”是附录III中表III1“稳健性检验（I）”的输出结果。“Log\_附录\_表III2”是附录III中表III2“稳健性检验（II）”的输出结果。“Log\_附录\_图III1”是附录III中图III1“行业国有企业数量占比与标准开放度的相关性”的输出结果。“Log\_附录\_表IV1”是附录IV中表IV1“进一步分析”的输出结果。“Log\_附录\_表IV2”是附录IV中表IV2“拓展研究”的输出结果。

**（3）“2024-01034+说明文档”**即本说明文档。

**（二）数据来源**

**本文使用的数据包括五大部分，分别是①国家标准信息数据库、②外商投资审批（备案）数据库、③中国工业企业数据库、④中国海关数据库、⑤专利数据库。**

①国家标准信息数据库来自全国标准信息公共服务平台，本文访问该平台并获取了1958年1月至2023年12月标准化组织颁布的各项国家标准的标准号、标准名称、发布日期、实施日期、废止日期、起草单位和采标情况等。其中，起草单位包括政府部门、各类企业、科研院所和标准化协会等各类主体，据此可以识别参与制定我国国家标准的主体结构。

②外商投资审批（备案）数据库来自商务部外商投资综合管理平台，本文获取了1978—2019年政府审批（备案）通过的在华外商投资项目数据，包括在华外资持股企业的名称、地址、行业、营业期限、投资者名称及其国籍、投资时间和投资金额等。不同于大部分以企业自愿上报形式进行数据采集的微观数据库，外商投资审批（备案）数据库是我国外商投资逐案审批（备案）制度下的时代产物 ，因此该数据库的数据质量和信息精度较高，具有权威性。

③中国工业企业数据库来自国家统计局，样本年份跨度为2000—2013年，该数据库提供了国有企业及规模以上工业企业的基础信息和财务信息，例如，企业名称、成立时间、资产、负债和利润状况等。

④中国海关数据库来自海关总署，样本年份跨度为2000—2013年，该数据库记载了HS6位编码商品的进出口状态、企业名称、贸易额、贸易伙伴、贸易方式和贸易时间等。

⑤专利数据库来自国家知识产权局，样本年份跨度为2000—2013年，该数据库记录了专利案件的申请及授权信息，包括主体名称、申请（授权）时间、所属类型（发明、实用新型或外观设计专利）等。

**本文基于如下步骤处理以上数据库，进而得到本文所用数据：**

**首先，处理数据并得到本文的基准回归数据（即“2024-01034+程序代码”中的“workdata.dta”数据）。**第一步，识别标准起草单位是否为外资持股企业。本文将起草单位名称与外资持股企业名录进行匹配，得到外资持股企业参与制定国家标准的频次，并据此测算行业层面的标准开放度指标。第二步，本文借鉴Liu and Qiu（2016）匹配工业企业和专利数据库，借鉴何小钢等（2023）匹配工业企业和海关数据库。第三步，数据库合并与清洗。一是合并前述面板数据，基于实收资本和登记注册类型识别并保留**本土企业**；二是统一各年四分位行业编码；三是删除存在质量问题的2010年企业样本；四是借鉴余淼杰和梁中华（2014）、聂辉华等（2009）删除财务数据存在明显问题、不符合规模以上标准的企业样本。据此，得到“workdata.dta”数据。

**其次，处理数据并得到本文拓展研究部分“附录IV\_表IV2\_拓展研究”第（4）列和第（5）列回归所用数据（即“2024-01034+程序代码”中的“Foreignworkdata.dta”数据）。**第一步，识别标准起草单位是否为外资持股企业。本文将起草单位名称与外资持股企业名录进行匹配，得到外资持股企业参与制定国家标准的频次，并据此测算行业层面的标准开放度指标。第二步，数据库合并与清洗。一是合并工业企业与行业层面的标准开放度数据，借鉴杨志浩和洪俊杰（2023）的做法识别并保留**外资持股企业**；二是统一各年四分位行业编码；三是删除存在质量问题的2010年企业样本；四是借鉴余淼杰和梁中华（2014）、聂辉华等（2009）删除财务数据存在明显问题、不符合规模以上标准的企业样本。据此，得到“Foreignworkdata.dta”数据。

**最后，处理数据并得到本文绘制“附录III\_图III1\_行业国有企业数量占比与标准开放度的相关性”所用数据（即“2024-01034+程序代码”中的“Figuredata.dta”数据）。**第一步，识别标准起草单位是否为外资持股企业。本文将起草单位名称与外资持股企业名录进行匹配，得到外资持股企业参与制定国家标准的频次，并据此测算行业层面的标准开放度指标。第二步，计算行业内的国有企业数量占所有企业数量的比重。第三步，合并前述数据，得到“Figuredata.dta”数据。

**（三）Dofile里面使用的指标说明**

LnInnovation：本土企业创新，使用企业发明专利申请量的对数衡量（取对数前加1处理）。

Innovation：企业发明专利申请量。

AInnovation：对企业发明专利申请量进行反双曲正弦变换。

S：标准开放度，使用外资持股企业参与制定的处于有效期（标准发布日至废止日）的我国国家标准总频次与行业内企业数量的比值衡量。

Lnage：企业年龄，使用样本年份与企业成立年份差值的对数衡量。

Lnemploy：企业规模，使用企业职工人数的对数衡量。

TFP：劳动生产率，使用企业工业总产值与职工人数的比值衡量。

LA：资产负债率，使用企业总负债与总资产的比值衡量。

Lnpergdp：经济发展水平，使用各省人均GDP的对数衡量。

Secondgdp\_weight：第二产业占比，使用各省第二产业GDP占比衡量。

Thirdgdp\_weight：第三产业占比，使用各省第三产业GDP占比衡量。

Foreignweight：外资进入强度，使用各省外资企业数量与GDP比值衡量。

LSN：本土企业参与标准制定，使用本土企业参与制定且处于有效期的国家标准数量衡量。

year：样本年份，本文使用的样本年份跨度为2000——2013年。

SD：双重差分项，当企业的样本年份处于标准专项试点年份及之后年份时，SD取值为1，否则为0。

FolseSD：前置一期的双重差分项，当企业的样本年份处于虚设的标准专项试点年份及之后年份时，FolseSD取值为1，否则为0。其中，虚设的标准专项试点年份是指标准专项试点年份的前一年。

USAposition：美国国际投资头寸，使用各行业的美国国际投资头寸规模衡量。

Exp\_Imp\_firm：是否进出口企业，是则取值为1，否则为0。

Lnxiezuo：共同参标规模，使用本土企业与外资持股企业共同参与制定国家标准数量的对数衡量。

Coop：共同参与标准制定（虚拟变量），当样本所处年份大于本土企业与外资持股企业共同参与国家标准制定的年份时，该变量取值为1，否则为0。

BZLJ：本土企业是否参与过我国国家标准制定，是则取值为1，否则为0。

Marketshare：市场占有率，使用企业国内销售产值占所在行业国内销售总产值的比重衡量。

Imppagerank：中间品进口贸易网络中心度，基于pagerank算法和中间品贸易数据测算。

Exppagerank：最终品出口贸易网络中心度，基于pagerank算法和最终品贸易数据测算。

exp\_sdprice：出口环节信息摩擦度，使用出口产品价格的方差衡量。

imp\_sdprice：进口环节信息摩擦度，使用进口产品价格的方差衡量。

S\_shifei：调整有效期后的标准开放度，使用外资持股企业参与制定的处于有效期（标准实施日至废止日）的我国国家标准总频次与行业内企业数量的比值衡量。

S\_gb2：二分位行业的标准开放度，使用外资持股企业参与制定的处于有效期（标准发布日至废止日）的我国国家标准总频次与行业内企业数量的比值衡量。

S\_chanpin：使用产品标准测算的标准开放度，使用外资持股企业参与制定的处于有效期（标准发布日至废止日）的我国产品国家标准总频次与行业内企业数量的比值衡量。

S\_gbmean：考虑行业外资市场准入差异的标准开放度，使用外资持股企业参与制定的处于有效期（标准发布日至废止日）的我国国家标准总频次与行业内外资持股企业数量的比值衡量。

S\_value：考虑企业规模的标准开放度，使用参与制定我国国家标准的外资持股企业的总规模与该行业所有企业总规模的比值衡量，

S\_qiangzhi：强制性标准领域的标准开放度，使用外资持股企业参与制定的处于有效期（标准发布日至废止日）的我国强制性国家标准总频次与行业内企业数量的比值衡量。

S\_tuijian：推荐性标准领域的标准开放度，使用外资持股企业参与制定的处于有效期（标准发布日至废止日）的我国推荐性国家标准总频次与行业内企业数量的比值衡量。

S\_standard：归一化处理后的标准开放度，对外资持股企业参与制定的处于有效期（标准发布日至废止日）的我国国家标准总频次进行归一化处理，得到归一化处理后的标准开放度。

l3.S：标准开放度的滞后3期。

l2.S：标准开放度的滞后2期。

l.S：标准开放度的滞后1期。

Lnfmshouquan：本土企业发明专利授权量，使用企业发明专利授权量的对数衡量（取对数前加1处理）。

Lnbzfmzls：本土企业标准相关的发明专利申请量，使用企业与标准相关的发明专利申请量的对数衡量（取对数前加1处理）。

Lnvalue：企业的工业总产出，使用企业工业总产值的对数衡量。

Lnnew：企业的新产品产出，使用企业新产品产值的对数衡量。

Newweight：新产品占比，使用新产品产值与工业总产值的比重衡量。

Lnhomesale：国内销售额，使用企业工业销售产值与出口交货值之差的对数衡量。

Homeweight：内销份额，使用国内销售额与工业销售产值之比衡量。

cls2002：四分位行业代码。

cls2002\_gb2：二分位行业代码。

area\_code4：城市代码。

Wage:职工平均工资，使用各省份职工平均工资衡量。

SY：上游标准开放度，利用某行业使用的上游行业中间投入品占其使用的中间投入品总额的比重，对行业的标准开放度进行加权，得到上游行业的标准开放度。

SX：下游标准开放度，利用某行业出售给下游行业中间投入品占其出售的中间投入品总额的比重，对行业的标准开放度进行加权，得到下游行业的标准开放度。

zwfirm：在位企业，参考高凌云和易先忠（2019）识别在位企业，属于在位企业取值为1，否则为0。

newfirm：初创企业，参考高凌云和易先忠（2019）识别初创企业，属于初创企业取值为1，否则为0。

tcfirm：即将退市企业，参考高凌云和易先忠（2019）识别即将退市企业，属于即将退市企业取值为1，否则为0。

SU：国内外标准统一度，使用企业贸易结构加权对标国际标准的国家标准（商品层面），得到企业面临的国内外标准统一度。

Lnre：外资参与标准制定，使用外资持股企业参与制定且处于有效期的国家标准数量的对数衡量。

Addrate：外商追加投资率，使用当期外商追加投资额与上一期外商投资额之比衡量。

pre3:当样本年份等于(试点年份-3)时，取值为1，否则为0。

pre2:当样本年份等于(试点年份-2)时，取值为1，否则为0。

pre1:当样本年份等于(试点年份-1)时，取值为1，否则为0。

current:当样本年份等于试点年份时，取值为1，否则为0。

post1:当样本年份等于(试点年份+1)时，取值为1，否则为0。

post2:当样本年份等于(试点年份+2)时，取值为1，否则为0。

post3:当样本年份等于(试点年份+3)时，取值为1，否则为0。

post4:当样本年份等于(试点年份+4)时，取值为1，否则为0。

post5:当样本年份等于(试点年份+5)时，取值为1，否则为0。

post6:当样本年份等于(试点年份+6)时，取值为1，否则为0。

post7:当样本年份等于(试点年份+7)时，取值为1，否则为0。

post8:当样本年份等于(试点年份+8)时，取值为1，否则为0。

post9:当样本年份等于(试点年份+9)时，取值为1，否则为0。

post10:当样本年份等于(试点年份+10)时，取值为1，否则为0。

post11:当样本年份等于(试点年份+11)时，取值为1，否则为0。

state\_weight：行业内国有企业数量占比，使用行业内国有企业数量占行业内企业总数量的比重衡量。