**说明文档**

**（一）文件说明**

本论文使用的数据代码压缩包命名为“2023-01204\_数据”，解压之后里面有“2023-01204+程序代码”、“2023-01204+日志文件”、“2023-01204+maindata”、“2023-01204+说明文档”和“生成文章图表的截图”这5个部分。

（1）“2023-01204+程序代码”里面包括1个文件，是名为“all\_work”的do文件，这一文件已经包含了本文所有图表（含附录）的运行文件，可生成本文的所有结果（含附录）。

（2）“2023-01204+日志文件”的说明如下：

**正文部分对应日志：**“描述性统计”是表1中变量的描述性统计，“table2”到“table6”分别是正文中表2到表6的运行结果输出，“parallel\_entry”和“parallel\_exit”是平行趋势检验（正文中图1）的输出结果。

**附录部分对应日志：**“tableⅣ1”到“tableⅣ4”是附录中表Ⅳ1到表Ⅳ4的运行结果，“stackdid”是通过堆叠DID构建的dta格式数据样本集（构建过程见all\_work文件），“simulations1”到“simulations4”是安慰剂检验构建的4个绘图样本集， “beta\_density\_plot1”到“beta\_density\_plot4”是安慰剂检验的4幅图（即附录图Ⅳ1）的输出结果，“stackedDID\_entry”和“stackedDID\_exit”是附录图Ⅳ2中的2幅图的输出结果。

需要说明的是，安慰剂检验汇报的是随机抽样运行500次的结果。

（3）“2023-01204+maindata”是本文实证分析使用的数据集，已经包含了文中实证分析所需的全部基础数据。

（4）“2023-01204+说明文档”即是本说明文档。

（5）“生成文章图表的截图”是“2023-01204+程序代码”的运行结果截图。

**（二）数据来源**

本文使用的数据来源包括两部分：

（1）中国工商注册企业数据，来源于RESSET数据。这一数据库包含中国（不含港澳台）所有工商主体的注册登记信息，包括企业名称、注册时间、经营状态、注销时间（如有）、注册地址、注册登记行业、企业类型、注册资本额等信息。本文在省份-行业-年份面板维度统计出注册和注销企业数量进行实证分析。同时，对不同所有制企业的分类研究以及企业规模分布的分析使用了企业类型、注册资本额这两个信息。

（2）中国企业投融资数据。来源于RESSET数据，包含中国（不含港澳台）各类融资事件信息，融资类型覆盖战略融资、收并购、风险投资等各种类型的融资事件，包含融资企业名称、投资机构、融资金额、融资时间等信息。本文按融资企业名称与工商注册企业信息匹配后，依据融资时间统计出省份-行业当年私营或国有企业融资项目的数量进行分析。

**（三）数据集中的变量与指标说明**

pro：省份代码（31个省/直辖市/自治区），与中文变量“省份”相对应

indus：行业代码，分类标准为《国民经济行业分类（2011）》中的96个行业大类

year：年份代码，0到19分别代表2000到2019年

pro\_indus：省份-行业序号，每一个省份-行业有独特的编码

reform\_num：该省份该行业当年实施混合所有制改革的国企数量，来源于手工整理

skew：该省份该行业当年存续企业注册资本规模分布的偏度系数，基于RESSET工商注册企业数据库计算得到

kur：该省份该行业当年存续企业注册资本规模分布的峰度系数，基于RESSET工商注册企业数据库计算得到

entry：该省份该行业当年新注册企业数，基于RESSET工商注册企业数据库

exit：该省份该行业当年注销企业数，基于RESSET工商注册企业数据库

exist：该省份该行业当年年末的存续企业数，基于RESSET工商注册企业数据库

entrysoe：该省份该行业当年新注册国有企业数，以工商注册企业数据库企业类型划分

exitsoe：该省份该行业当年注销的国有企业数，以工商注册企业数据库企业类型划分

existsoe：该省份该行业年末存续的国有企业数，以工商注册企业数据库企业类型划分

entrypri：该省份该行业当年新注册的私营企业数，以工商注册企业数据库企业类型划分

exitpri：该省份该行业当年注销的私营企业数，以工商注册企业数据库企业类型划分

existpri：该省份该行业年末存续的私营企业数，以工商注册企业数据库企业类型划分

upentry：该省份该行业对应上游行业的注册企业数，计算方法见正文变量说明，基于2005年投入产出表的行业间投入产出占比加权计算得到

upexit：该省份该行业对应上游行业的注销企业数，计算方法见正文变量说明，基于2005年投入产出表的行业间投入产出占比加权计算得到

upexist：该省份该行业对应上游行业的年末存续企业数，计算方法见正文变量说明，基于2005年投入产出表的行业间投入产出占比加权计算得到

downentry：该省份该行业对应下游行业的注册企业数，计算方法见正文变量说明，基于2005年投入产出表的行业间投入产出占比加权计算得到

downexit：该省份该行业对应下游行业的注销企业数，计算方法见正文变量说明，基于2005年投入产出表的行业间投入产出占比加权计算得到

downexist：该省份该行业对应下游行业的年末存续企业数，计算方法见正文变量说明，基于2005年投入产出表的行业间投入产出占比加权计算得到

val：行业类型的分类变量，1代表行业属于竞争性行业，2代表行业属于垄断性行业，分类基准是基于工商注册企业注册资本规模计算得行业HHI指标（赫芬达尔赫希曼指数）

REF：该省份该行业是否已经发生首次国企混改的0-1变量，若是则取1

从pre10到pre1的10个变量：0-1变量，pre10代表是否是首次国企混改发生前的第10年，pre1代表是否是首次国企混改发生前1年，以此类推

current：0-1变量，是否是首次国企混改发生的当年

post1、post2和post3plus：0-1变量，post1代表是否是首次国企混改发生后的第1年，post2代表混改发生后的第2年，post3plus发表首次国企混改发生后的第3年及之后

financenum\_pri：该省份该行业当年对应注册时间不大于1年的私营企业获得融资支持的事件数，基于RESSET投融资事件库和工商注册企业数据库匹配后计算得到

financenum\_soe：该省份该行业当年对应注册时间不大于1年的国有企业获得融资支持的事件数，基于RESSET投融资事件库和工商注册企业数据库匹配后计算得到