

中国工业企业规模与生产率的异质性

高凌云 屈小博 贾鹏*

内容提要 本文利用中国经济普查数据库中的全样本工业企业以及分布参数的极大似然估计和非参数检验,从总体和行业层面,具体估计、检验和比较了中国工业企业规模和生产率的异质性特征。结果显示:(1)中国企业整体上仍面临扩张约束,较大规模企业所占比重略小于目标状态;(2)部分规模较大企业的生产率水平实际不高,存在明显的资源错配问题;(3)总体和细分行业层面的生产率分布存在较大差异;(4)在细分行业层面,企业规模和生产率的分布并不能简单类推。

关键词 企业规模 生产率异质性 帕累托分布 替代性分布检验

一 引言

大量基于微观企业生产和贸易数据的研究发现,即使定义在很狭窄的行业内,企业之间在生产规模和效率等方面仍然存在显著的差异(Bartelsman 和 Doms, 2000),而这显然有悖于新古典贸易理论和新贸易理论中同质代表性企业的经典假定。在这一背景下,以 Melitz (2003) 等为代表,通过引入企业异质性假设,建立并拓展了异质性贸

* 高凌云:中国社会科学院世界经济与政治研究所国际贸易研究室 北京建国门内大街5号 100732 电子信箱:gaoly@cass.org.cn;屈小博、贾鹏:中国社会科学院人口与劳动经济研究所 北京市朝阳区曙光西里28号中冶大厦 100028 电子信箱:quxb@cass.org.cn(屈小博);电子信箱:jiapeng@jiapeng.org(贾鹏)。

本文是中国社会科学院哲学社会科学创新项目“国际视角下的中国贸易结构转型研究”、国家社科基金重点项目(10AZD014、12AZD054)、国家自然科学基金项目(71173234)、中国社会科学院世界经济与政治研究所所级重点课题的阶段性成果。作者感谢佟家栋、盛斌、宋泓、姚枝仲、余森杰、包群、孙楚仁、陈勇兵、施炳展、钱震杰、刘青、朱希伟以及2013年《世界经济》编辑部贸易主题论文研讨会和匿名审稿人提出的建设性意见。文责自负。

中国工业企业规模与生产率的异质性

易理论,将分析进一步深入到企业的生产、贸易和投资行为层面,突破了以产业为对象的研究范畴,提出了国际贸易研究中一些新的发展方向,如多产品企业、外包、公司内贸易等,构成了目前国际贸易研究的前沿领域。

虽然异质性贸易理论的发展,如 Melitz (2003) 在分析出口企业的自选择效应时,最初仅是直接引入企业异质性假设,并没有给出生产率分布的具体函数形式,但是考虑到贸易中的企业行为特征等因素 (Bernard 等, 2007), 后续基于异质性贸易理论的研究,在分析和检验特定问题,尤其是在需要给出模型解析解或具体定量结果时,又必须假设分布函数恰当的具体形式 (Chaney, 2008)。

现有文献中假设的企业生产率可能服从的分布较多,常见的有帕累托分布 (Chaney, 2008) 和对数正态分布 (Capuano 和 Schmerer, 2013) 以及指数分布 (Caliendo 和 Rossi-Hansberg, 2012) 等。至于为什么选择这些分布,上述研究普遍认为,除数理上方便处理外,主要是因为企业规模服从类似分布。但问题是,即使这一逻辑合理,现有假设及检验企业规模分布的文献,得到的企业规模分布类型也不一致,常见的有齐普夫定律 (Axtell, 2001), 对数正态分布 (Cefis 等, 2009), 幂律分布 (Gaffeo 等, 2003) 和帕累托分布 (Konno, 2013)。

相比于异质性贸易理论及其在不同领域的快速发展,现有研究对一些基础性的问题,如对企业规模和生产率分布所做的假设是否合理? 为什么不同研究之间具有如此大的差异? 这些分布之间具有什么样的内在联系? 企业规模的分布和生产率的分布能否直接类比等问题缺乏科学合理的解答。不仅如此,由于某经济单元某一时点有大量不同规模或生产率水平的新企业进入,新企业进入后会与该时点仍在位的、不同规模或生产率水平的企业在产品、要素等市场展开竞争,企业之间的竞争使得该经济单元同时出现获利和亏损企业 (也称获胜企业和失败企业),进而新进入企业和原在位企业在对其生产和经营环境认识加深的基础上,做出扩张、收缩或退出的决策;与此同时,退出企业的生产要素和市场份额在该进入单元仍在位的企业间重新配置,并形成某种均衡的分布状态。因此,企业规模或生产率的分布可以在很大程度上反映出该经济单元的生态特征和健康状态,对它们的正确认识,还可以为经济政策的制定提供参考 (杨其静等, 2010)。

本文利用中国经济普查的全样本工业企业数据,依据分布参数的极大似然估计和分布的非参数检验,从总体和行业层面,具体估计、检验和比较了中国工业企业规模和生产率的异质性,试图对上述疑问提供严格的分析,以期对异质性贸易理论的发展及其在中国的应用提供坚实的基础。本文的后续安排为:第二部分是相关文献的综述,

以判断本文在现有文献中的地位;第三部分以中国为例,具体分析了齐普夫定律、帕累托分布和幂律之间的内在一致性,以及帕累托分布、对数正态分布与指数分布之间的区别和联系;以此为基础,第四部分提供了针对帕累托分布尺度参数估计和检验,以及替代性分布检验的具体思路和方法;第五部分对经济普查数据库和指标进行了详细的说明;第六和第七部分则分别提供企业规模与生产率分布的估计检验结果;最后是结论与启示。

二 文献综述

本文的研究主要与两类文献相关。一类是探讨企业规模分布特征、动态变化及原因的文献。首先,对企业规模分布特征的研究,迄今已有近一个世纪的历史。Gibrat (1931)认为同一行业中的企业,无论其规模大小,同一时期内其规模成长的概率是相同的,即企业的成长率独立于其初始的规模(通常称为 Gibrat 法则或比例效应法则),因而企业规模应近似服从对数正态分布。但之后, Mansfield (1962)从理论上提出,对那些特定行业的新进入企业来说,生存的前提是,必须迅速达到能以最低长期平均成本生产的有效规模;在 Mansfield (1962)的基础上, Gabaix (1999)证明,如果企业规模服从有界的随机游走过程,那么其分布呈现的是幂律形态,而不是对数正态。与此对应,检验企业规模服从何种分布的经验研究,其结论也存在类似差异。如 Bottazzi 和 Secchi (2003)利用 3000 家美国企业的销售额数据,发现美国企业规模服从对数正态分布;但 Axtell (2001)利用美国 500 万家纳税企业的数据,发现美国企业规模近似服从的是幂律分布。

其次,对企业规模分布动态变化和影响因素的研究。Cabral 和 Mata (2003)发现,因为融资约束的原因,企业规模分布呈现明显的右偏,而且不同的年份几乎没有什么变化。但不同经营年限的队列 (cohort) 分析显示,随着企业经营年限的增加,企业规模的分布逐渐逼近对数正态分布。Angelini 和 Generale (2008)认为,放松融资约束对企业规模分布的影响并不明显,因为融资约束的放松虽然可以促进在位企业更快扩张,但也会引致更多规模较小的新企业进入。另外, Luttmer (2007) 等还认为,负向的外部冲击导致部分企业退出也是企业规模分布变化的重要原因; Rossi-Hansberg 和 Wright (2007) 还发现,不同行业内企业规模的分布存在非常大的差异,人力资本存量越少的行业会越快经历规模报酬递减,因此特定行业人力资本的密集度与大规模企业的比重正相关。

中国工业企业规模与生产率的异质性

另一类是研究企业生产率分布及其与规模分布关系的文献。首先,尽管对生产率分布而非生产率水平的关注,是近些年来经济增长理论的核心主题之一,但现有研究多侧重于对生产率分布中不同分位企业生产率相对差异(也称生产率离散化)的研究。Syverson(2011)等发现,即使在市场经济最为发达的美国,生产率离散化现象也是长期存在的。其次,更为关键的是,尽管不同分位企业的生产率差异很难刻画企业生产率的整体分布,但现有针对企业生产率分布估计和检验的研究很少,以至于异质性贸易理论引入生产率异质性更多是通过假设的方式。^①再次,在封闭经济的条件下,di Giovanni等(2011)证明,企业生产率如果服从帕累托分布,企业规模同样会服从帕累托分布,尽管分布参数不同。最后,章韬和孙楚仁(2012)在非出口和出口企业生产率分布参数不同但都为帕累托分布的假设下发现,中国非出口企业和出口企业生产率分布与规模分布形态均存在显著差异,但它们的规模分布形态与生产率分布形态高度一致。

整体来看,针对本文的主题,上述企业规模和生产率分布方面的研究还存在一些不足。第一,认为企业规模服从某种分布的研究,如 Fujiwara 等(2004)等,除了依据的样本不具有代表性、多数没有深入到行业层次^②等问题之外,通常是先简单假设某一具体的分布,然后利用 OLS 或修正的 OLS 估计出该分布的参数,并直接采用拟合优度值代替分布检验,没有考虑存在替代性分布拟合程度更好的可能性。第二,Axtell(2001)等认为偏离目标(target)分布是由竞争排斥所引致的推断,忽略了企业规模或生产率分布并不服从目标分布的可能性。第三,结合 Chaney(2008)和 di Giovanni 等(2011)研究还可以发现,现有文献在企业规模异质性和生产率异质性的关系上,存在循环论证的问题,缺乏对企业生产率分布的科学估计和严格检验,在生产率分布的选择上较为随意。

本文的边际贡献主要体现在两个方面:第一,相比 Axtell(2001)、di Giovanni 等(2011)的研究,本文细致阐释了传统估计方法可能存在的问题,并在参数极大似然估计的基础上,将分布检验和替代性分布检验引入企业规模和生产率分布研究中;第二,利用中国经济普查数据库中的全样本工业企业,分别从总体和行业层次对中国企业规

^① Arkolakis(2010)在产品差异化、企业生产率异质性和市场边际覆盖成本递增的基础上,发展了一个动态贸易理论模型,该模型在进入企业生产率变动具有几何布朗运动形态等条件下,内生地决定了企业生产率服从右拖尾的帕累托分布,但这一研究并没有提供经验估计和检验。Combes 等(2012)虽利用法国企业的数据估计了不同类城市企业生产率对数正态和帕累托混合分布的分布参数,但也没有提供检验。

^② 章韬和孙楚仁(2012)在帕累托分布假设下,分 GBT 2 位数行业对出口和非出口企业规模及生产率的帕累托分布参数进行了估计。

模和生产率的分布进行了严格估计和检验,是对 Combes 等(2012)研究的有益补充。

三 不同分布的联系与区别

如引言部分所归纳的,现有的异质性贸易理论及经验文献,经常假设或验证企业规模和生产率服从的分布包括:幂律分布、齐普夫定律、帕累托分布、对数正态分布和指数分布等。这几类分布的共性是,大多数事件的规模、程度或频次很小,而只有少数事件的规模、程度或频次相当大。由于齐普夫定律、帕累托分布和幂律分布本质上具有一致性,所以本部分先论证齐普夫定律、帕累托分布、幂律的内在联系,以简化后文的分析。事实上,现有文献在齐普夫定律、帕累托分布或幂律分布相互之间关系的认识上有些模糊,有些甚至将这些分布截然对立。然后,再以帕累托分布为代表,比较它们与对数正态分布和指数分布之间的联系。

(一) 齐普夫定律、帕累托分布与幂律分布的一致性

齐普夫定律,是由 Zipf(1936)在对英语文献中单词出现频次进行大量统计、以检验前人的定量化公式时提出的词频分布定律,具体可以表述为:如果把一篇较长文章中每个词出现的频次统计起来,按照高频词在前、低频词在后的递减顺序排列,那么不同词的序号与其出现频次的积,将近似地为一个常数。若用 x 表示频次, r 表示等级序号,即有 $x \times r = C$, C 为常数,这一关系常用 *rank-size* 图刻画。帕累托分布可能是目前文献中最常见的生产率分布假设;除此以外,还被广泛用于描述自然和社会中的众多现象,如城市规模大小的分布、地震规模大小的分布等。与齐普夫定律不同的是,帕累托分布是累积分布的概念,指的是规模或频次大于 x 事件发生的概率与 x 之间的反向关系。如果 X 是一个随机变量,则 X 的概率分布为: $P(X > x) = \left(\frac{x}{x_{\min}}\right)^{-k}$ 。其中 x 是任何大于 x_{\min} 的数, x_{\min} 是 X 最小的可能值,^① k 为正,通常被称为尺度参数(*scaling parameter*)。而幂律一般是指概率密度函数近似服从幂函数,即 $p_X(x) dx = \Pr(x \leq X \leq x + dx) = Cx^{-a} dx$ 的随机变量 X 的分布;其中 $C = (a - 1)x_{\min}^{a-1}$ 。^② 与帕累托分布不同的是,幂律是概率密度的概念,指的是规模、程度等于 x 事件发生的概率与

① 因为 $k > 0$, 分布 $P(X > x) = \left(\frac{x}{x_{\min}}\right)^{-k}$ 在 $x \rightarrow 0$ 的时候发散,所以帕累托分布必然存在下限 x_{\min} 。

② 由 $\int_{x_{\min}}^{\infty} p(x) dx = C \int_{x_{\min}}^{\infty} x^{-a} dx = 1$ 得到。

中国工业企业规模与生产率的异质性

x 之间的反向关系。

由上述定义可以推知,对同一随机变量分布的描述,无论齐普夫定律、帕累托分布,还是幂律,其内在分布关系是一致的,只是表述方式有所不同。首先,尽管齐普夫定律是规模与其排序的关系,而帕累托分布是规模与其频数的关系,但齐普夫定律和帕累托分布本质上是可以等同的。以企业规模为例说明,按定义,齐普夫定律描述的是排在第 r 位的企业,其规模为 x ,这等价于:有 r 个企业,其规模大于等于 x ,而这又恰恰是帕累托分布的定义。唯一的差别在于,对齐普夫定律来说, r 在横轴上, x 在纵轴上;而帕累托分布是 r 在纵轴上, x 在横轴上。其次,由累积分布函数 $P(X <= x) = 1 - P(X > x)$,求导就可以得到服从幂律分布的随机变量 X 的密度分布函数: $p_X(x) = \frac{k}{x_{\min}^{k+1}}$

$$\left(\frac{x}{x_{\min}}\right)^{-k-1}, k = a - 1。$$

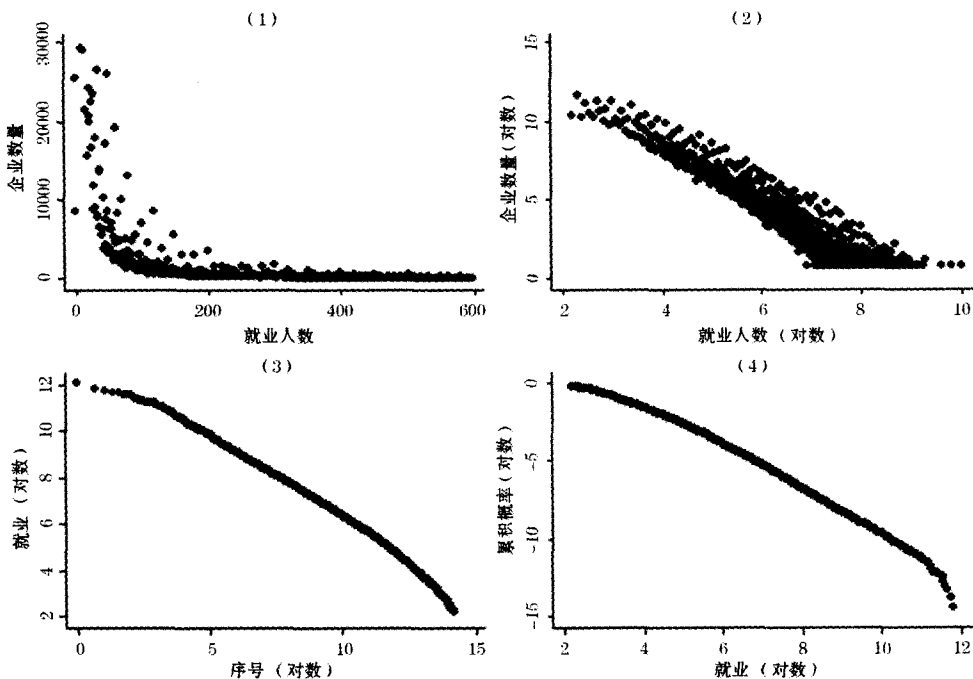


图1 中国全样本工业企业从业人员合计的分布情况

为了明确上述分析,本处利用中国工业企业就业规模的总体分布情况,提供齐普夫定律、帕累托分布和幂律分布之间关系的例证。与 Axtell(2001)类似,此处暂且采

用从业人员合计测度企业规模。^① 首先,图 1(1)刻画了企业从业人员合计与其频次的关系,散点图显示绝大多数企业的规模很小,只有少数企业的规模比较大。^② 其次,图 1(2)描述对数刻度下的企业规模与其频次的对应,散点图显示了企业规模等于 x 事件发生的概率与 x 之间的反向关系。而且,在等分间隔的情况下,呈现出越往尾部噪声越大的典型帕累托分布特征,即越往右相同间隔包括的样本企业数量越少、统计波动越大。再次,图 1(3)、(4)分别是双对数刻度的企业规模与其排序的 rank-size、双对数刻度的企业规模与累积概率。正如前文所分析的,两图之间的差别仅在于,对齐普夫定律来说,排序在横轴上,规模在纵轴上;而帕累托分布是累积概率在纵轴上,企业规模在横轴上。因此,后文不再区分齐普夫定律、帕累托分布和幂律分布,而是仅从帕累托分布出发,估计、检验和比较中国工业企业规模和生产率的异质性。

(二) 帕累托分布、对数正态分布和指数分布的联系

在双对数坐标图上,幂律分布、齐普夫定律和帕累托分布一般都表现为一条斜率为幂指数的负数直线,特别是在大样本的情况下。但是,与幂律分布、齐普夫定律和帕累托分布相比,对数正态分布和指数分布存在明显不同。下面以帕累托分布作为联系纽带,分析和比较三类分布之间的联系和差异。

首先,对数正态分布是指:若随机变量 X 的对数 $\ln X$ 服从正态分布 $N(\mu, \sigma)$, 则 X 服从对数正态分布。它的概率密度函数为: $C_1 \frac{1}{x} \exp\left[-\frac{(\ln x - \mu)^2}{2\sigma^2}\right]$, 根据 $\int_{x_{\min}}^{\infty} Cf(x) dx = 1$ 可得: $C_1 = \sqrt{\frac{2}{\pi\sigma^2}} \left[\operatorname{erfc}\left(\frac{\ln x_{\min} - \mu}{\sqrt{2}\sigma}\right) \right]^{-1}$, erfc 为误差函数, μ 和 σ 分别为对数正态分布的对数均值和对数标准差。对数正态分布与帕累托分布的主要差别在于,对对数正态分布的密度函数两边取对数只能得到: $\ln f(x) = -\ln x - \ln\sqrt{2\pi}\sigma - \frac{(\ln x - \mu)^2}{2\sigma^2}$, 在双对数坐标图上,仅在对数标准差足够大的情况下才能近似表现为一条直线。而对帕累托分布的两边取对数,可以直接得到线性形式。

其次,相比于帕累托分布,指数分布在取对数的情况下,同样不能得到线性形式的表达式。因为指数分布的密度函数为: $C_2 e^{-\lambda x}$, 分布的区间是 $[0, \infty)$ 。根据 $\int_{x_{\min}}^{\infty} Cf(x) dx = 1$ 可得: $C_2 = \lambda e^{\lambda x_{\min}}$ 。其中, $\lambda > 0$ 是分布的一个参数,常被称为率参

① 数据来自 2008 年中国第二次经济普查工业企业数据库,论文第五部分提供了详细的处理和说明。

② 为了清晰表述这一关系,我们仅保留了就业人数小于 600 人的企业,如果包括所有的企业,规模和频次的散点图将简化为与坐标轴重合的“L”型折线。

中国工业企业规模与生产率的异质性

数。指数分布和帕累托分布的联系在于,如果随机变量 X 服从参数为 x_{min} 和 k 帕累托分布,可以推得 $Y = \ln\left(\frac{X}{x_{min}}\right)$ 服从率参数为 k 的指数分布。反过来,如果 Y 服从率参数为 k 的指数分布,则 $x_{min}e^Y$ 服从参数为 x_{min} 和 k 的帕累托分布。因为由 $\Pr\left(\ln\left(\frac{X}{x_{min}}\right) < y\right)$, 可得 $\Pr(X < x_{min}e^y) = 1 - \left(\frac{x_{min}}{x_{min}e^y}\right)^k$, 即 $\Pr(Y < y) = 1 - e^{-ky}$ 。

为了对帕累托分布、对数正态分布和指数分布之间的差异有更直观的认识,本文仍然利用2008年中国工业企业就业规模的总体分布情况,提供帕累托分布与对数正态分布和指数分布之间关系的例证。图2显示,在就业规模分布的右尾部位,帕累托分布和对数正态分布对就业规模分布的拟合效果明显好于指数分布(其中,实直线为帕累托分布拟合线,虚线为对数正态分布拟合线,点线为指数分布拟合线),而在右尾底部,对数正态分布的偏离程度要大于帕累托分布。但是,在就业规模分布的左部,指数分布的拟合效果要好于对数正态分布和帕累托分布。

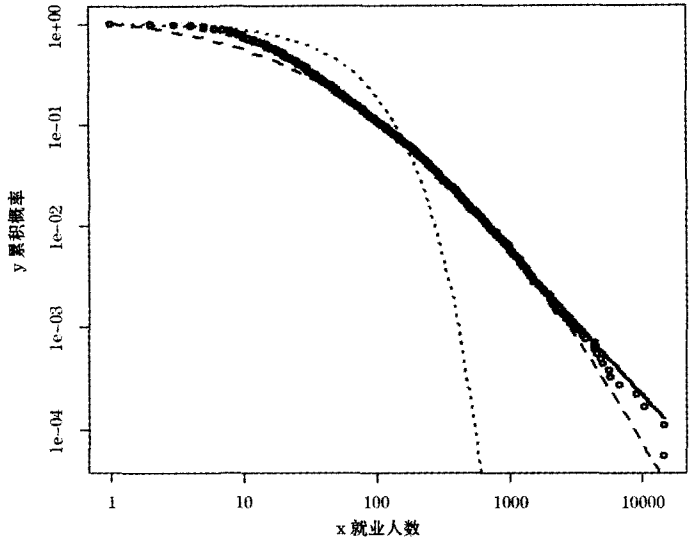


图2 帕累托、对数正态和指数分布对企业就业规模拟合效果

四 分布估计与检验的方法体系

判断随机变量是否服从某种分布 $P(x)$, 需要通过三个阶段的工作。第一,假设关注的随机变量确实服从分布 $P(x)$, 进而利用合适的方法估计该分布的所有参数; 第二,在参数估计的基础上,设定检验统计量,判断假设的分布是否合理; 第三,利用替代性分布检验,判断是否有其他分布拟合效果更好。

(一) 帕累托分布的估计与检验

针对帕累托分布,从估计的角度,最常用的方法是对帕累托分布函数的两边直接取对数,得到线性形式的 $\log P(x) = k \times \log(x_{min}) - k \times \log(x)$, 然后采用最小二乘法估计尺度参数(Axtell 等,2001),而且还可以考虑直接利用最小二乘估计的 R^2 值作为分布拟合优度检验的依据(di Giovanni 等,2011)。但是,这一简单的估计方法可能存在较多问题:首先,最小二乘估计要求误差项服从正态分布,即被解释变量是高斯白噪声的。但是,对 $P(x)$ 取对数后,累积概率变量就不再为白噪声了,因为一旦假定 $\log P(x)$ 服从正态分布,那么 $P(x)$ 只能是服从对数正态分布,这导致估计及检验随机变量服从何种分布的工作变得毫无意义。其次,由于帕累托分布是累积分布的概念,即 $P(x) = P(x-1) - p(x)$,这使得被解释变量存在严重的自相关问题。再次,不管随机变量总体服从何种分布,最小二乘回归的拟合线通常无法满足 $\int_{x_{min}}^{\infty} p(x) dx = 1$ 的约束条件(Weisberg,1985)。最后,最小二乘估计并没有考虑 x_{min} 值对尺度参数估计的重要性。如前文所分析的,帕累托分布只在 $[x_{min}, \infty)$ 区间成立,估计时如果 x_{min} 取值过小,会导致最小二乘估计纳入那些不服从分布的样本,从而影响尺度参数的正确估计,而取值过大又会丢掉有价值的样本点信息,增大尺度参数估计的误差。另外,从分布拟合评价标准的角度,最小二乘估计的 R^2 值可能只是必要条件,其他的常见分布,如对数正态分布、指数分布等,一定区间内采用最小二乘估计也可能得到较大的 R^2 值。

基于此,Wasserman(2003)等证明了在大样本的情况下,极大似然估计能提供尺度参数的准确估计。不妨假设企业规模或生产率确实服从参数为 k 和 x_{min} 的帕累托分布,由此可构造帕累托分布的对数似然函数: $L = \ln p(x | k) = \ln \prod_{i=1}^n \frac{k-1}{x_{min}} \left(\frac{x_i}{x_{min}}\right)^{-k}$, 其中, n 表示样本数量。由 $\partial L / \partial k = 0$,可以得到具有渐近正态、一致等特性,同时标准差等于 $\frac{k-1}{\sqrt{n}} + O(1/n)$, ①尺度参数的极大似然估计为:② $\hat{k} = 1 + n \left[\sum_{i=1}^n \ln \frac{x_i}{x_{min}} \right]^{-1}$ 。

但是,上面尺度参数的极大似然估计是建立在真实的 x_{min} 值已知的假定下,正如在评价最小二乘估计缺点时所提到的,选择不同的 x_{min} 值对尺度参数的估计结果影响

① 因为均值 $(\int_{x_{min}}^{\infty} xp(x) dx = C \int_{x_{min}}^{\infty} x^{-k+1} dx)$ 在 $k \leq 2$ 的情况下发散,此处隐含假设 $k > 1$ 。

② 由于本文估计采用的是2008年第二次全国经济普查的企业数据,文中仅提供了连续型变量极大似然估计结果,离散型变量帕累托分布参数极大似然估计结果可见 Clauset 等(2009)的研究。

中国工业企业规模与生产率的异质性

非常大。那么实际估计过程中,该如何确定 x_{min} 值? 是否可以假定其为已知的值,如随机变量序列中的最小值,又或者从累积概率分布图、概率密度分布图等图形中通过直观的观测得到呢? Goldstein 等(2004)利用非参数 Kolmogorov-Smirnov 统计量(后简称 K-S 统计量),较好地解决了这个问题。K-S 统计量的基本算法在于,计算和寻找实际累积分布和理论累积分布之间的最大差值(取绝对值),即:

$$D = \max_{x \geq x_{min}} |S(x) - P(x)|$$

其中, $S(x)$ 表示观测值的实际累积分布函数, $P(x)$ 表示具有估计参数的理论累积分布函数,然后选择 $\operatorname{argmin}_{x_{min}}(D)$ 作为所求的 x_{min} 估计量。不过,针对验证企业生产率是否服从帕累托分布的意图,仅准确地估计出企业生产率分布的尺度参数和 x_{min} 值还远远不够。上文提到,验证随机变量是否服从帕累托分布之前,需要先假定该随机变量服从帕累托分布。可事实上,无论随机变量是否服从帕累托分布,通过极大似然估计和 K-S 统计量,可以将尺度参数和 x_{min} 值都估计出来,但这并不能说明企业的生产率就服从帕累托分布。因此,与 Clauset 等(2007)及 Clauset 等(2009)一致,我们还需采用较为常用的、基于 K-S 统计量的 K-S 拟合优度检验法,对企业生产率是否服从帕累托分布进行验证。

(二) 替代性分布检验

针对本文的研究,利用极大似然估计和 K-S 统计量准确估计出帕累托分布的尺度参数和 x_{min} 值,进而利用 K-S 分布拟合优度检验,即使发现企业生产率服从帕累托分布的原假设不能拒绝,也还不足以说明帕累托分布就是拟合程度最好的,因为仍然存在其他分布,如常用的对数正态分布(Capuano 和 Schmerer, 2013)以及指数分布(Caliendo 和 Rossi-Hansberg, 2012)假设,具有更好拟合的可能。这就要求我们在进行帕累托分布参数的估计和分布拟合检验后,继续进行替代性分布检验。

K-S 分布拟合优度检验需要对不同的替代性分布重复本文第三部分的过程,然后比较不同替代性分布 p 值的大小。而现在的问题是,在通过帕累托分布检验之后,我们只需要知道是否有其他分布比帕累托分布拟合得更好。针对随机变量的实际观测值数据集,常用的直接检验何种分布拟合程度更好的方法,包括 χ^2 检验、似然比检验(Vuong, 1989)和 Bayesian 检验(Stouffer 等, 2005)。但 χ^2 检验更适用于离散分布的检验,对于连续分布,由于需要对资料分组而会失去较多信息, Bayesian 检验会在一定程度上减弱抽样波动的影响,从而不能识别偶然性对检验结果的影响,因此本文采用似然比检验对替代性分布进行判断。

利用似然比检验对替代性分布进行选择的基本思路是,对相同的数据集,比较不

同分布假设下的似然值,越大说明拟合得越好,比较似然比对数的正负号与此等价。因此,针对企业规模和生产率数据,假设存在两个候选的分布,其概率密度函数分别为

$$p_1(x), p_2(x), \text{由此可以得到两个分布的似然值为: } L_1 = \prod_{i=1}^n p_1(x_i), L_2 = \prod_{i=1}^n p_2(x_i)。$$

进而,这两个分布的对数似然比就可以表示为: $R = \sum_{i=1}^n [\ln p_1(x_i) - \ln p_2(x_i)]。$

Vuong(1989)证明,利用实际数据集估计出的 \hat{R} 值大于等于真实 R 值绝对值的概率可以表示为: $p(\hat{R} \geq |R|) = \frac{1}{\sqrt{2\pi n\sigma^2}} \left[\int_{-\infty}^{-|R|} e^{-t^2/2n\sigma^2} dt + \int_{|R|}^{+\infty} e^{-t^2/2n\sigma^2} dt \right]。$

$$\text{其中, } \sigma^2 = \frac{1}{n} \sum_{i=1}^n \left[(\ln p_1(x_i) - \ln p_2(x_i)) - \left(\frac{1}{n} \sum_{i=1}^n \ln p_1(x_i) - \frac{1}{n} \sum_{i=1}^n \ln p_2(x_i) \right) \right]^2。$$

考虑真实的 R 值为零或接近于零的情况,这表示似然比对数不能区分哪种分布更合适, $p(\hat{R} \geq |R|)$ 等价于得到给定 R 的概率,此时,如果 $p(R)$ 很小(如 $p < 0.1$),就可以拒绝真实的 R 值为零或接近于零的原假设,说明 \hat{R} 值的符号可以作为判断何种分布拟合程度更好的依据。

五 数据库、指标的选择与说明

根据帕累托分布的概率分布函数,上文分析了 x_{min} 的选择,对于尺度参数估计以及后续的帕累托分布检验和替代性分布检验为什么至关重要,是因为选择的 x_{min} 值过小,不符合帕累托分布的样本点进入会导致错误的估计结果,而选择的 x_{min} 值过大又会损失大量符合要求的样本信息,从而降低估计和检验的准确性。这也是本文选择 K-S 统计量确定 x_{min} 值的根本原因。更进一步,上述逻辑对估计企业规模、生产率分布所用数据库的选择也提出了更高的要求。因为,如果估计和检验依据的数据库,不是全样本或者来自全样本随机抽样产生的具有代表性的子样本,就无法估计出准确的 x_{min} 值,自然也无法得到企业相关变量的真实分布(Combes 等,2012)。

目前,对中国经济问题的研究,有相当多的国内外学者使用了中国工业企业数据库,但是,这一数据库除了样本匹配混乱、指标大小异常、测度误差明显和变量定义模糊等问题之外(聂辉华等,2012),针对本文的主题,最大的不足在于,它仅仅提供了全部国有及规模以上非国有工业企业的数据。对具有典型右拖尾特征的分布来说(Bernard 等,2007),这一数据库实际上截除了大部分处于分布左边的、规模较小和生产率

中国工业企业规模与生产率的异质性

较低的样本企业,这必然导致实际估计得到的 x_{min} 值过大,并增加服从帕累托分布的可能性。帕累托分布虽然具有标度无关性(scale-free),但应该是在证明某一随机变量服从这一分布之后,再利用任意子样本进行相关研究,而不是相反。

因此,本文后续估计和检验基本建立在 2004、2008 年两次经济普查数据库中全样本工业企业的基础上。这两次全国经济普查的对象是在中国境内从事第二产业和第三产业的全部法人单位、产业活动单位和个体经营户,其优势在于即包括规模以上企业,又包括规模以下企业。如 2008 年第二产业的具体范围包括:采矿业,制造业,电力、热力、燃气及水生产和供应业,合计有工业企业 1979 135 家。而中国工业企业数据库中,在不做任何剔除的情况下,2008 年的工业企业只有 399 594 家,只占全部工业企业数的 20.1%。不过,经济普查数据库也有一些劣势,如对生产率分布的研究只能限于单要素生产率分析;同时,经济普查数据库仅有规模以上企业提供了财务数据,而规模以下企业除了企业代码、名称和国民经济行业代码等定性指标外,只提供了年末从业人员合计、全年营业收入合计等少数几个相关的数量指标。

本文与 Axtell (2011)、Fujiwara 等(2004)一致,同样利用年末从业人员合计、全年营业收入合计作为企业规模的代理变量,而用全年营业收入除以年末从业人员合计表示企业生产率。需特别强调的是,对劳动生产率或全要素生产率的衡量,标准的做法是采用更能反映企业生产效率的工业增加值,而不是营业收入。只是由于中国两次经济普查对规模以

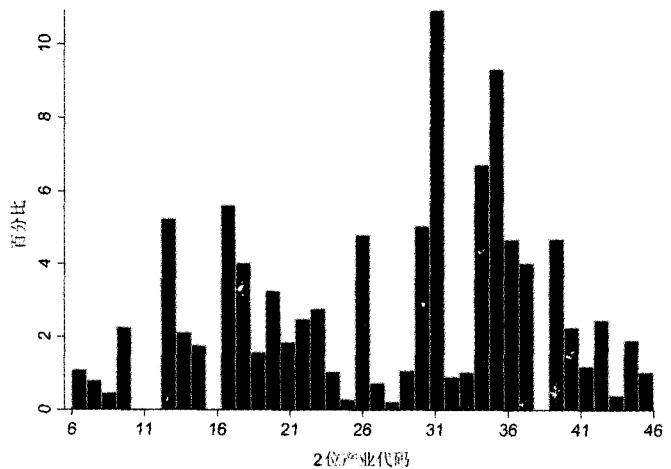


图3 不同2位数工业行业内企业数量占总体的比重

下企业仅提供了少数指标,没有工业增加值数据,也无法通过会计恒等式计算获得,因而此处利用人均收入来代替企业的生产效率。虽然全年营业收入合计与工业增加值具有较高的相关系数,但全年营业收入合计是指企业在销售商品、提供劳务及让渡资产使用权等活动中形成的收入,核算的基础是工业产品的销售总量,不管是否为本期生产;而工业增加值的核算前提是工业产品的生产总量,只要是本期生产的,不论是否

销售,都需计入工业总产值,并反映在减去中间投入后的工业增加值中,两者之间还存在一定的出入。

另外,不同工业行业内部的企业在很多方面均存在较大差异。图3^①显示了2008年不同2位数工业行业内企业数量占样本总体的比重,其中,烟草制品业(代码06)所含的企业数量最少,为206个,约占样本总量的0.01%;而非金属矿物制品业(代码31)所含的企业数量最多,为199256个,是烟草制品业的967倍,占样本总量的比重超过10.9%。这说明对企业异质性的研究,除了总体层次的分析外,还需深入到工业行业内部。

六 中国工业企业规模的分布特征

针对工业企业规模分布的估计与相关检验,^②本文使用了年末从业人员合计和全年营业收入合计两个指标,^③并分行业总体和2位数行业两个层次进行判断,其中,每一层次采用的处理顺序为:(1)利用极大似然方法和K-S统计量分别估计 \hat{x}_{\min} 和 \hat{k} ,并绘制累积分布图与帕累托分布拟合线;^④(2)对这一结果进行非参数K-S分布检验,并计算检验p值;(3)对应现有研究中常用的对数正态分布和指数分布假设,进行两次替代性分布检验,计算帕累托分布相对它们的对数似然比和检验p值;(4)根据检验p值的选择原则,判断合适的分布类型。

(一)企业就业规模的分布特征

考虑以年末从业人员合计测度的企业规模,附表1中的检验结果显示,整体而言,相比于对数正态分布和指数分布,帕累托分布对企业规模的拟合效果更好。

首先,在全样本的情况下,估计的 \hat{x}_{\min} 和 \hat{k} 分别为201和2.26,K-S分布检验的p值为1.00;同时,对数正态分布的替代性分布检验p值为0.95,超过接受原假设的临界值,而指数分布的替代性分布检验p值为0.07,尽管符合接受原假设的临界值,但是R值为1.80,说明帕累托分布的似然值更大,拟合程度更好。因此,全样本估计与

① 剔除了年末从业人员合计、全年营业收入合计、资产总计等指标缺失的企业135614家,剔除企业生产率两端各0.5%分位的异常企业18434家,本处和后文用于估计和检验的企业为1825087家。

② 本文的计算和检验过程由R软件的软件包(poweRlaw)完成;由于计算设备的限制,在进行全样本分布的参数估计和检验时,仅从中随机抽取50%的样本;本文所有过程数据、结果及程序代码备索。

③ 单位分别为人、千元人民币。

④ 由于篇幅限制,我们略去了所有的累积分布图与帕累托分布拟合线,感兴趣的读者可函索。

中国工业企业规模与生产率的异质性

检验结果支持企业就业规模服从帕累托分布的判断,这一结果与 Axtell(2001)、Fujiwara 等(2004)的结果一致。

其次,如 Rossi-Hansberg 等(2007)和本文第五部分所分析的,分行业估计得到的 x_{\min} 和 k ,与全样本相比差异较大。如烟草制品业的 x_{\min} 值等于 811,而电力、热力的生产和供应业的 x_{\min} 值只有 11;有色金属矿采选业的 k 为 2.74,而石油与天然气开采业的 k 只有 1.55。这说明总体层次的估计并不能体现行业层次就业规模分布的差异。

再次,除食品制造业、工艺品及其他制造业外,分行业工业企业就业规模均服从帕累托分布;值得注意的是,食品制造业、工艺品及其他制造业同时服从帕累托分布、对数正态分布和指数分布;其中,食品制造业对数正态分布和指数分布的替代性分布检验,得到的 p 值分别为 0.08、0.00, R 值分别为 -1.78、8.04,而工艺品及其他制造业对数正态分布和指数分布的替代性分布检验,得到的 p 值分别为 0.10、0.00, R 值分别为 -1.67、6.14,说明对这两个行业,对数正态分布的拟合程度比帕累托分布更高,而帕累托分布又优于指数分布。

(二)企业收入规模的分布特征

考虑到与 Axtell(2001)、Fujiwara 等(2004)的可比性,本处对以全年营业收入合计测度的企业规模分布做类似分析,另一层含义是为企业就业规模分布提供稳健性检验。附表 2 的结果显示,在总体特征上,基于收入规模的估计和检验结果与就业规模大体类似。

首先,在全样本的情况下,估计的 x_{\min} 和 k 分别为 86 500 和 2.18, $K-S$ 分布检验的 p 值为 1.00;同时,对数正态分布的替代性分布检验的 p 值为 0.89,同样超过接受原假设的临界值,而指数分布的替代性分布检验的 p 值虽为 0.00,符合接受原假设的临界值,但 R 值为 4.88,说明帕累托分布的似然值更大,拟合程度更好。因此,全样本估计与检验的结果支持企业收入规模服从帕累托分布的判断。而分行业估计得到的参数值与全样本相比,差异同样较大。这再次说明总体层次的分布估计,并不能体现行业内部企业收入规模分布的显著差异。

其次,除饮料制造业、纺织业、仪器仪表及文化办公机械制造业、废弃资源和废旧材料回收加工业外,其余分行业工业企业收入规模全部服从帕累托分布;其中,饮料制造业对数正态分布和指数分布的替代性分布检验,得到的 p 值分别为 0.07、0.00, R 值分别为 -1.82、7.96;纺织业对数正态分布和指数分布的替代性分布检验,得到的 p

值分别为 0.07、0.00, R 值分别为 -1.79、9.79; 仪器仪表及文化办公机械制造业对数正态分布和指数分布的替代性分布检验, 得到的 p 值分别为 0.04、0.00, R 值分别为 -2.01、4.44; 废弃资源和废旧材料回收加工业对数正态分布和指数分布的替代性分布检验, 得到的 p 值分别为 0.02、0.00, R 值分别为 -2.26、7.01; 说明这四个行业对数正态分布比帕累托分布具有更好的拟合性, 而帕累托分布又优于指数分布。

(三) 企业规模分布估计和检验结果的经济含义

Luttmer(2007)等强调, 只有在政府管制较少、较自由的竞争市场条件下, 企业在规模上才会呈现帕累托指数大致等于 1 的分布状态,^①当存在扭曲竞争的因素时, 该经济体的帕累托指数就会偏离 1。正向偏离(大于 1)说明企业之间的异质性较小, 较大规模企业所占比重明显小于目标状态, 企业存在扩张约束, 如融资和准入限制; 而负向偏离表示企业之间的异质性较大, 较大规模企业所占比重明显大于目标状态, 存在垄断等竞争不足问题。但是, 由于对任何随机分布都可以估计出帕累托分布参数, 因此检验企业规模的帕累托指数是否偏离 1, 其前提需要验证, 表征企业规模的变量确实服从帕累托分布, 之前的研究大多没有注意帕累托指数因分布函数误设产生偏离的情况。而本节的估计和检验结果表明, 中国全样本工业企业的就业和收入规模, 整体服从帕累托分布, 从而排除了因分布函数误设产生偏差的可能。

在此基础上, 本文利用 2008 年中国全样本工业企业以就业和收入衡量的企业规模, 得到的帕累托指数分别为 1.26 和 1.18, 说明中国较大规模企业所占比重略小于目标状态, 整体来看企业发展面临扩张约束; 从细分行业的结果也可以看出, 帕累托指数大于 1 的行业所占比重大于帕累托指数小于 1 的行业, 如从就业和收入规模分布的角度, 帕累托指数大于 1 的行业分别为 29 和 25 个。但是, 帕累托指数大于和小于 1 的行业同时存在, 说明那些旨在消除企业扩张约束或促进竞争的单一目标政策, 其效应在不同的行业会截然不同。因此, 新形势下中国深化企业改革的一系列政策, 其指向应该力求明确。

七 中国工业企业生产率的分布特征

对中国工业企业生产率分布特征的分析, 同样包括所有行业和 2 位数行业个体两

^① R 软件软件包 (poweRlaw) 采用幂律分布的函数形式, 因而得到的帕累托指数应该是 $k = a - 1$ 。

中国工业企业规模与生产率的异质性

个层次,其中每一层次的处理顺序,与对企业规模分布的分析完全相同。

(一) 工业企业生产率的分布特征

附表3的估计检验结果显示,首先,在全样本条件下,中国工业企业生产率分布的 x_{\min} 值为666.67,^① k 为3.21,表明企业规模分布的上尾比生产率分布的上尾更厚;同时,帕累托分布检验的 p 值为1.00,而对数正态分布替代性检验的 p 值和指数分布替代性检验的 p 值均大于0.10,说明帕累托分布对全样本工业企业生产率的拟合效果更好。另外,对那些帕累托分布提供了更好拟合的2位数工业行业来说,其 x_{\min} 值和 k 与全样本相比,差异依然较大。如其他采矿业的 x_{\min} 值和 k 仅为186.73和2.65,这些都与企业就业、收入规模分布的结论类似。

其次,企业就业规模只在两个行业、企业收入规模只在四个行业存在对数正态分布拟合程度更高的情况。但是,企业生产率分布,在所有39个2位数工业行业内,帕累托分布拟合优于对数正态分布和指数分布的只有10个行业,而对数正态分布拟合优于帕累托分布和指数分布的占到23个,指数分布优于对数正态分布和帕累托分布的也有2个,分别是石油加工、炼焦及核燃料加工业以及黑色金属冶炼及压延加工业。也就是说,Chaney(2008)、Caliendo和Rossi-Hansberg(2012)以及Capuano和Schmerer(2013)等研究所做的生产率分布假设,只有在特定行业才有其合理性。这意味着,仅仅依据企业规模的分布,简单类推企业生产率的分布是不合理的。

再次,附表1、2的结果显示,针对企业就业规模分布和收入规模分布,如果仅仅是进行帕累托分布的参数估计和分布拟合检验,那么总体样本和所有行业样本都不能拒绝服从帕累托分布的假设;但是,在估计和检验工业企业生产率分布的时候,附表3第4列中,煤炭开采和洗选业、烟草制品业、化学纤维制造业、有色金属冶炼及压延加工业等四个行业,K-S分布检验的 p 值仅分别为0.02、0.00、0.01、0.00,这验证了无论企业生产率是否服从帕累托分布,通过极大似然估计和K-S统计量可以将尺度参数和 x_{\min} 值都估计出来,但这并不能说明企业的生产率就服从帕累托分布。

最后,煤炭开采和洗选业、烟草制品业、化学纤维制造业、有色金属冶炼及压延加工业等四个行业,没有通过帕累托分布检验,虽然通过对数正态分布或指数分布的替代性分布检验,但替代性检验只是判断何种分布拟合程度更好。在没有通过初始分布检验的情况下,并不能得到这四个行业内企业生产率服从对数正态分布或指数分布

① 单位为千元人民币。

的结论。而且,对这四个行业分别利用对数正态分布和指数分布替换帕累托分布,重复本文第三部分的过程,得到的分布检验 p 值仍都小于 0.10。

(二) 稳健性分析

上节的结果表明,与企业就业、收入规模分布不同,大部分细分工业行业中,对数正态分布对企业生产率的拟合效果,要优于帕累托分布和指数分布。但这一结论是普遍性规律还是会随不同年份的经济普查样本、相同年份但不同营业年限等发生改变?本节利用相同估计和检验方法,依据不同类别的中国全样本工业企业,为企业生产率分布特征的稳健性提供证据。

首先,我们利用中国 2004 年第一次经济普查中的全样本工业数据库,比较不同年份工业企业生产率的分布是否存在显著差异。相比于 2008 年的第二次经济普查,2004 年中国的第一次经济普查,包括工业企业 1355 278 家,剔除年末从业人员合计、全年营业收入合计、资产总计等指标缺失的企业 142 085 家、企业生产率两端各 0.5% 分位的异常企业 13 088 家,最后可用于估计和检验的工业企业为 1295 562 家。附表 4 的估计和检验结果显示,中国工业企业生产率的分布特征非常稳定。从总体样本的角度,帕累托分布对 2004 年工业企业生产率分布的拟合效果最好;从细分行业角度,相比于 2008 年的 23 个工业行业,2004 年有 22 个工业行业,对数正态分布的拟合效果优于帕累托分布和指数分布;而且,除煤炭开采和洗选业、烟草制品业等少数行业外,2004 和 2008 年生产率分布保持一致的工业行业有 31 个。

其次,按 Cabral 和 Mata(2003)、Angelini 和 Generale(2008)等的逻辑,伴随企业经营年限的增加,其内、外部经营环境会逐渐改善。因此,我们推测,经营年限较长的企业样本,其生产率分布的右偏程度会更小,从而更可能服从对数正态分布。我们按经营年限是否超过 10 年为标准,^①将 2008 年的全样本工业企业划分为两个子样本后,附表 5 的结果显示,虽然不同子样本的细分行业中,有 28 个行业生产率服从的最优分布没有发生变化,但正如我们所预计的,有 8 个行业,如食品制造业、黑色金属冶炼及压延加工业等,在经营年限小于等于 10 年的子样本里服从帕累托分布或无法判断分布,而在经营年限大于 10 年的子样本里服从对数正态分布。

(三) 企业生产率分布估计和检验结果的经济含义

本文针对企业生产率分布的估计和检验结果具有明确的经济含义。首先,从异质

^① 经营年限依据 2008 年经济普查数据库中的“开业时间”计算,实际估计时剔除了经营年限超过 60 年的企业;另外,采用其他年数(如 5、15 年)作为划分标准,不会改变总体特征;对其他分类样本估计和检验结果感兴趣的读者可向作者函索。

中国工业企业规模与生产率的异质性

性贸易理论发展的角度,由于生产率分布形态在不同行业并不相同,因此,后续非总体层面的理论或经验研究应该尽量与 Melitz(2003)、Combes 等(2012)的研究一致,采用生产率分布的一般形式,这是异质性贸易文献应该遵循的基本原则。其次,2004、2008 年中国工业企业总体样本生产率分布的 k 由 4.24 下降为 3.21,同时 x_{\min} 值由 373.17 上升为 666.67,说明伴随着对外开放和社会主义市场经济体制改革的逐渐深入,中国工业企业整体的生产率水平有显著提升,且异质性程度加大,高生产率企业所占的比重明显增加。再次,如章韬和孙楚仁(2012)等所指出的,企业生产规模分布是否和生产效率分布一致,在一定程度上反映了现实经济中的扭曲情况。而 2008 年中国工业企业就业规模和收入规模的 k 分别为 2.26 和 2.18,低于同年生产率分布的 k ,企业规模分布的右尾,比生产率分布的右尾更厚,规模分布的差异程度明显高于生产率分布,说明中国仍然存在明显的资源错配问题,企业间生产要素的流动受到限制,部分较大规模企业的生产率水平实际不高。

八 结论

企业规模和生产率异质性的假设,是异质性贸易理论发展的基石。相比新古典贸易理论和新贸易理论中的同质性企业假设,虽然涉及企业异质性的贸易文献都认为企业异质性假设更接近现实,但至今鲜有文献对企业异质性假设,特别是企业生产率的异质性进行科学的估计和严格的检验。针对企业规模分布的研究,也存在样本不具有代表性、以估计替代检验等问题,有些将齐普夫定律、帕累托分布和幂律分布对立,有些甚至将企业规模异质性与企业生产率异质性等同。

基于此,本文首先分析了企业规模和生产率异质性文献中常见的分布,然后引入了针对帕累托分布尺度参数估计和非参数检验的方法;而且,考虑到企业规模或生产率通过了帕累托分布的检验,仍然存在其他分布拟合程度更好的可能性,论文进一步引入了替代性分布检验的具体思路和方法;进而,根据帕累托分布的特征,利用中国两次经济普查的全样本数据库,从总体和细分行业两个层次具体估计检验和比较了中国工业企业规模和企业生产率的异质性,并提供了相应的稳健性分析。

本文的基本结论和政策启示有:(1)中国较大规模企业所占比重略小于目标状态,整体来看企业发展面临扩张约束;(2)总体层次的企业规模和生产率分布估计,并不能体现行业之间在企业规模分布和生产率分布上的显著差异;(3)对中国工业企业

来说,企业规模的异质性和生产率异质性存在明显的区别,规模分布的尾部比生产率分布的尾部更厚,显示中国仍然存在明显的资源配置问题;(4)本文利用经济普查全样本数据得到的企业规模和生产率的 x_{\min} 值、 k 等,还可用于那些以异质性贸易理论为基础的数值模拟研究之中。

本文还存在一些不足之处,如与侧重于研究中国工业企业规模、生产率分布差别相比,探究企业规模或生产率分布变动的影响因素可能更加具有实践价值。

参考文献:

- 聂辉华、江艇、杨汝岱(2012):《中国工业企业数据库的使用现状和潜在问题》,《世界经济》第5期。
- 杨其静、李小斌、方明月(2010):《市场、政府与企业规模分布——一个经验研究》,《世界经济文汇》第1期。
- 章韬、孙楚仁(2012):《贸易开放、生产率形态与企业规模》,《世界经济》第8期。
- Angelini, P. and Generale, A. "On the Evolution of Firm Size Distributions." *American Economic Review*, 2008, 98, pp. 426-438.
- Arkolakis, C. "Market Penetration Costs and the New Consumers Margin in International Trade." *Journal of Political Economy*, 2010, 118(6), pp. 1151-1199.
- Axtell, R. L. "Zipf Distribution of U. S. Firm Sizes." *Science*, 2001, 293(5536), pp. 1818-1820.
- Bartelsman, E. and Doms, M. "Understanding Productivity: Lessons from Longitudinal Microdata." *Journal of Economic Literature*, 2000, 38(3), pp. 569-595.
- Bernard, A. B.; Jensen, J.; Redding, S. J. and Schott, P. K. "Firms in International Trade." *Journal of Economic Perspectives*, 2007, 21(3), pp. 105-130.
- Bottazzi, G. and Secchi, A. "Common Properties and Sectoral Specificities in the Dynamics of U. S. Manufacturing Companies." *Review of Industrial Organization*, 2003, 23, pp. 217-232.
- Cabral, M. B. and Mata, J. "On the Evolution of the Firm Size Distribution: Facts and Theory." *American Economic Review*, 2003, 93(4), pp. 1075-1090.
- Caliendo, L. and Rossi-Hansberg, E. "The Impact of Trade on Organization and Productivity." *The Quarterly Journal of Economics*, 2012, 127(3), pp. 1393-1467.
- Capuano, S. and Schmerer, H. J. "Trade and Unemployment Revisited: Do Institutions Matter?" *World Economy*, 2014, forthcoming.
- Cefis, E.; Marsili, O. and Schenk, H. "The Effects of Mergers and Acquisitions on the Firm Size Distribution." *Journal of Evolutionary Economics*, 2009, 19, pp. 1-20.
- Chaney, T. "Distorted Gravity: the Intensive and Extensive Margins of International Trade." *American Economic Review*, 2008, 98(4), pp. 1707-1721.
- Clauset, A.; Young, M. and Gleditsch, K. S. "On the Frequency of Severe Terrorist Attacks." *Journal of Conflict Resolution*, 2007, 51(1), pp. 58-88.
- Clauset, A.; Shalizi, C. R. and Newman, M. "Power-law Distributions in Empirical Data." *SIAM Review*, 2009, 51

中国工业企业规模与生产率的异质性

(4), pp. 661-703.

Combes, Pierre-Philippe; Gilles, D.; Laurent, G.; Diego, P. and Sébastien, R. "The Productivity Advantages of Large Cities: Distinguishing Agglomeration from Firm Selection." *Econometrica*, 2012, 80(6), pp. 2543-2594.

Fujiwara, Y.; di Guilmi, C.; Aoyama, H.; Gallegati, M. and Souma, W. "Do Pareto-Zipf and Gibrat Laws Hold True? An Analysis with European Firms." *Physica A: Statistical Mechanics and its Applications*, 2004, 335(1-2), pp. 197-216.

Gabaix, X. "Zipf's Law For Cities: An Explanation." *Quarterly Journal of Economics*, 1999, 114(3), pp. 739-767.

Caffeo, E.; Gallegati, M. and Palestrini, A. "On the Size Distribution of Firms: Additional Evidence from the G7 Countries." *Physica A: Statistical Mechanics and its Applications*, 2003, 324(1-2), pp. 117-123.

Gibrat, R. *Les Inegalite's Economiques. Applications: Aux ine'galite's Des Richesses, A la Concentration Des Entreprises, aux Populations des Villes, aux Statistiques des Familles, etc., D'une Loi Nouvelle: La loi De l'effect Proportionnel*. Paris: Sirey, 1931.

di Giovanni, J.; Levchenko, A. and Ranciére, R. "Power Laws in Firm Size and Openness to Trade: Measurement and Implications." *Journal of International Economics*, 2011, 85(1), pp. 42-52.

Goldstein, M.; Morris, S. and Yen, G. "Problems with Fitting to the Power-law Distribution." *The European Physical Journal*, 2004, 41(2), pp. 255-258.

Konno, T. "Pareto Distribution of Firm Size and Knowledge Spillover Process as A Network." *Economics Bulletin*, 2013, 33(3), pp. 2407-2413.

Luttmer, E. "Selection, Growth, and the Size Distribution of Firms." *The Quarterly Journal of Economics*, 2007, 122(3), pp. 1103-1144.

Mansfield, E. "Entry, Gibrat's Law, Innovation and the Growth of Firms." *American Economic Review*, 1962, 52(5), pp. 1023-1051.

Melitz, M. J. "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity." *Econometrica*, 2003, 71(6), pp. 1695-1725.

Rossi-Hansberg, E. and Wright, M. "Establishment Size Dynamics in the Aggregate Economy." *American Economic Review*, 2007, 97(5), pp. 1639-1666.

Stouffer, D. B.; Malmgren, R. D. and Amaral, L. "Comment on Barabasi." *Nature*, 2005, 435, pp. 207-211.

Syversen, C. "What Determines Productivity?" *Journal of Economic Literature*, 2011, 49(2), pp. 326-365.

Vuong, Q. "Likelihood Ratio Tests for Model Selection and Non-nested Hypotheses." *Econometrica*, 1989, 57(2), pp. 307-333.

Wasserman, L. A. *All of Statistics*. New York: Springer-Verlag, 2003.

Weisberg, S. *Applied Linear Regression*. New York: Wiley and Sons, 1985.

Zipf, G. *The Psycho-Biology of Language: An Introduction to Dynamic Philology*. Boston: Psychology Press, 1936.

附表1 中国工业企业就业规模分布的估计及检验结果

样本行业	帕累托分布			对数正态分布		指数分布		支持分布
	x_{min}	k	p值	R	p值	R	p值	
所有行业	201	2.26	1.00	0.07	0.95	1.80	0.07	P
煤炭开采和洗选业	452	1.93	1.00	0.07	0.94	3.11	0.00	P
石油和天然气开采业	43	1.55	0.99	0.12	0.91	7.13	0.00	P
黑色金属矿采选业	52	2.24	1.00	-0.32	0.75	5.91	0.00	P
有色金属矿采选业	420	2.74	1.00	0.12	0.91	1.90	0.06	P
非金属矿采选业	51	2.31	1.00	-0.22	0.82	2.46	0.01	P
其他采矿业	28	2.68	1.00	0.51	0.61	2.64	0.01	P
农副食品加工业	34	2.21	1.00	-0.91	0.36	5.89	0.00	P
食品制造业	23	1.99	1.00	-1.78	0.08	8.04	0.00	L
饮料制造业	31	2.08	1.00	-0.68	0.50	5.96	0.00	P
烟草制品业	811	2.11	0.81	-0.21	0.83	1.74	0.08	P
纺织业	216	2.58	1.00	-0.29	0.77	3.35	0.00	P
纺织服装、鞋帽制造业	198	2.60	1.00	-1.13	0.26	1.64	0.10	P
皮革、毛皮、羽毛(绒)及其制品业	287	2.32	1.00	-0.59	0.56	3.77	0.00	P
木材加工及木、竹、藤、草制品业	34	2.39	1.00	-0.89	0.37	4.41	0.00	P
家具制造业	44	2.21	1.00	-0.51	0.61	5.43	0.00	P
造纸及纸制品业	79	2.34	1.00	-0.30	0.77	3.50	0.00	P
印刷业和记录媒介的复制	33	2.37	1.00	-0.20	0.84	6.28	0.00	P
文教体育用品制造业	130	2.23	1.00	-0.65	0.52	2.80	0.01	P
石油加工、炼焦及核燃料加工业	14	1.73	0.93	-1.26	0.21	9.77	0.00	P
化学原料及化学制品制造业	81	2.33	1.00	-0.27	0.78	4.22	0.00	P
医药制造业	243	2.41	1.00	-0.81	0.42	3.03	0.00	P
化学纤维制造业	174	2.09	1.00	-0.45	0.65	5.21	0.00	P
橡胶制品业	23	1.94	1.00	-1.48	0.14	6.13	0.00	P
塑料制品业	45	2.27	1.00	-0.28	0.78	4.80	0.00	P
非金属矿物制品业	34	2.34	1.00	-0.44	0.66	2.79	0.01	P
黑色金属冶炼及压延加工业	117	1.99	1.00	0.18	0.86	3.27	0.00	P
有色金属冶炼及压延加工业	117	2.14	1.00	-0.33	0.74	3.43	0.00	P
金属制品业	46	2.26	1.00	-0.98	0.32	4.53	0.00	P
通用设备制造业	127	2.39	1.00	-0.19	0.85	2.54	0.01	P
专用设备制造业	158	2.54	1.00	0.20	0.84	2.68	0.01	P
交通运输设备制造业	28	1.91	1.00	-1.60	0.11	8.40	0.00	P
电气机械及器材制造业	263	2.51	1.00	-0.31	0.76	3.39	0.00	P
通信、计算机及其他电子设备制造业	80	1.94	1.00	-0.90	0.37	7.58	0.00	P
仪器仪表及文化、办公用机械制造业	35	2.02	1.00	-0.11	0.91	7.10	0.00	P
工艺品及其他制造业	34	2.10	1.00	-1.67	0.10	6.14	0.00	L
废弃资源和废旧材料回收加工业	27	2.40	1.00	-0.17	0.87	6.52	0.00	P
电力、热力的生产和供应业	11	1.71	1.00	-0.62	0.53	5.06	0.00	P
燃气生产和供应业	27	1.91	1.00	-0.51	0.61	10.71	0.00	P
水的生产和供应业	181	2.39	1.00	-0.04	0.97	2.88	0.00	P

中国工业企业规模与生产率的异质性

附表 2 中国工业企业收入规模分布的估计及检验结果

样本行业	帕累托分布			对数正态分布		指数分布		支持分布
	x_{min}	k	p 值	R	p 值	R	p 值	
所有行业	86 500	2.18	1.00	0.14	0.89	4.88	0.00	P
煤炭开采和洗选业	41892	2.04	1.00	-0.95	0.34	7.40	0.00	P
石油和天然气开采业	4905	1.37	0.99	-1.17	0.24	125.7	0.00	P
黑色金属矿采选业	65 110	2.27	1.00	-0.28	0.78	5.05	0.00	P
有色金属矿采选业	276 914	2.73	1.00	0.08	0.93	4.12	0.00	P
非金属矿采选业	71 850	3.70	0.99	-0.25	0.81	4.06	0.00	P
其他采矿业	6600	1.82	0.93	-0.47	0.63	2.47	0.01	P
农副食品加工业	35 512	2.02	1.00	-0.80	0.42	4.85	0.00	P
食品制造业	65 312	2.22	1.00	-0.03	0.98	3.95	0.00	P
饮料制造业	4936	1.64	0.59	-1.82	0.07	7.96	0.00	L
烟草制品业	32 621	1.46	1.00	-1.11	0.27	3.98	0.00	P
纺织业	54 970	2.28	1.00	-1.79	0.07	9.79	0.00	L
纺织服装、鞋、帽制造业	65 725	2.62	1.00	-0.07	0.95	4.31	0.00	P
皮革、毛皮、羽毛(绒)及其制品业	45 962	2.20	1.00	-0.89	0.37	6.50	0.00	P
木材加工及木、竹、藤、草制品业	69 078	2.68	1.00	-0.56	0.58	5.87	0.00	P
家具制造业	65 024	2.41	1.00	0.08	0.93	3.93	0.00	P
造纸及纸制品业	53 420	2.27	0.98	-1.52	0.13	6.57	0.00	P
印刷业和记录媒介的复制	4910	1.82	0.63	-1.08	0.28	6.54	0.00	P
文教体育用品制造业	56 077	2.46	1.00	-0.50	0.61	5.80	0.00	P
石油加工、炼焦及核燃料加工业	556 908	2.19	1.00	0.09	0.93	2.63	0.01	P
化学原料及化学制品制造业	139 122	2.22	1.00	-0.04	0.97	117.3	0.00	P
医药制造业	57 381	2.12	1.00	-0.98	0.33	7.25	0.00	P
化学纤维制造业	87 035	1.93	1.00	-0.85	0.40	6.10	0.00	P
橡胶制品业	37 429	2.10	1.00	0.07	0.95	2.55	0.01	P
塑料制品业	40 511	2.32	1.00	-0.85	0.40	6.76	0.00	P
非金属矿物制品业	99 127	2.53	1.00	-0.20	0.84	4.08	0.00	P
黑色金属冶炼及压延加工业	141 252	1.82	1.00	-0.09	0.93	3.32	0.00	P
有色金属冶炼及压延加工业	120 823	2.15	1.00	-0.12	0.91	4.11	0.00	P
金属制品业	28 006	2.11	1.00	0.07	0.94	2.64	0.01	P
通用设备制造业	28 905	2.15	1.00	-0.31	0.75	4.66	0.00	P
专用设备制造业	56 556	2.17	1.00	-0.56	0.57	4.80	0.00	P
交通运输设备制造业	140 817	1.98	1.00	0.07	0.94	1.45	0.15	P
电气机械及器材制造业	46 551	1.91	1.00	-0.81	0.42	4.34	0.00	P
通信、计算机及其他电子设备制造业	53 400	1.79	1.00	-1.24	0.22	5.21	0.00	P
仪器仪表及文化、办公机械制造业	2217	1.71	0.15	-2.01	0.04	4.44	0.00	L
工艺品及其他制造业	96 540	3.10	1.00	0.14	0.89	3.10	0.00	P
废弃资源和废旧材料回收加工业	4951	1.69	1.00	-2.26	0.02	7.01	0.00	L
电力、热力的生产和供应业	90 285	1.82	1.00	-0.46	0.65	2.72	0.01	P
燃气生产和供应业	191 324	2.21	1.00	-0.51	0.61	4.36	0.00	P
水的生产和供应业	11 773	1.97	1.00	-0.77	0.44	4.01	0.00	P

附表3 2008年中国工业企业生产率分布的估计及检验结果

样本行业	帕累托分布			对数正态分布		指数分布		支持分布
	x_{\min}	k	p值	R	p值	R	p值	
所有行业	666.67	3.21	1.00	-1.50	0.13	0.46	0.65	P
煤炭开采和洗选业	89.89	1.85	0.02	-8.65	0.00	6.50	0.00	N
石油和天然气开采业	180.00	2.21	0.13	-2.98	0.00	4.03	0.00	L
黑色金属矿采选业	216.28	2.20	0.33	-5.76	0.00	2.92	0.00	L
有色金属矿采选业	127.62	2.03	0.68	-5.74	0.00	7.39	0.00	L
非金属矿采选业	460.00	3.06	1.00	-0.41	0.68	3.44	0.00	P
其他采矿业	186.73	2.65	0.99	-0.16	0.87	3.45	0.00	P
农副食品加工业	265.80	2.29	0.60	-5.20	0.00	1.49	0.14	L
食品制造业	341.67	2.70	1.00	-2.15	0.03	2.07	0.04	L
饮料制造业	360.00	2.68	1.00	-1.78	0.07	2.99	0.00	L
烟草制品业	80.94	1.60	0.00	-3.78	0.00	-0.52	0.60	N
纺织业	428.51	2.94	1.00	-2.09	0.04	1.11	0.27	L
纺织服装、鞋、帽制造业	269.43	2.78	1.00	-0.48	0.63	4.42	0.00	P
皮革、毛皮、羽毛(绒)及其制品业	190.32	2.51	1.00	-2.10	0.04	5.17	0.00	L
木材加工及木、竹、藤、草制品业	247.50	2.83	1.00	-2.13	0.03	3.00	0.00	L
家具制造业	356.37	3.06	1.00	-1.45	0.15	1.83	0.07	P
造纸及纸制品业	287.60	2.85	1.00	-2.14	0.03	2.56	0.01	L
印刷业和记录媒介的复制	292.69	3.04	1.00	-0.79	0.43	4.20	0.00	P
文教体育用品制造业	246.73	2.90	1.00	-1.70	0.09	2.67	0.01	L
石油加工、炼焦及核燃料加工业	1914.66	5.07	0.30	-2.74	0.01	-6.95	0.00	E
化学原料及化学制品制造业	357.51	2.43	0.40	-5.05	0.00	-1.17	0.24	L
医药制造业	851.76	3.38	1.00	-2.32	0.02	-1.63	0.10	L
化学纤维制造业	1033.20	3.11	0.01	-5.54	0.00	-9.65	0.00	N
橡胶制品业	260.00	2.63	1.00	-2.06	0.04	3.86	0.00	L
塑料制品业	360.00	2.89	1.00	-1.61	0.11	2.73	0.01	P
非金属矿物制品业	286.53	2.69	1.00	-1.43	0.15	4.18	0.00	P
黑色金属冶炼及压延加工业	2311.88	6.56	0.34	-2.44	0.01	-4.46	0.00	E
有色金属冶炼及压延加工业	194.44	1.95	0.00	-9.32	0.00	-0.53	0.60	N
金属制品业	246.16	2.62	1.00	-2.44	0.01	4.54	0.00	L
通用设备制造业	266.67	2.82	1.00	-2.31	0.02	3.25	0.00	L
专用设备制造业	480.00	2.99	1.00	-0.96	0.34	2.74	0.01	P
交通运输设备制造业	286.67	2.72	1.00	-2.14	0.03	3.41	0.00	L
电气机械及器材制造业	219.29	2.40	0.97	-3.91	0.00	4.75	0.00	L
通信、计算机及其他电子制造业	298.43	2.49	1.00	-3.19	0.00	2.54	0.01	L
仪器仪表及文化、办公机械制造业	318.38	2.77	1.00	-1.91	0.06	2.69	0.01	L
工艺品及其他制造业	283.38	2.76	1.00	-1.15	0.25	4.34	0.00	P
废弃资源和废旧材料回收加工业	348.33	2.40	0.92	-3.92	0.00	2.18	0.03	L
电力、热力的生产和供应业	246.67	2.40	1.00	-2.99	0.00	3.81	0.00	L
燃气生产和供应业	729.60	2.90	0.96	-3.06	0.00	-1.86	0.06	L
水的生产和供应业	449.50	3.32	1.00	-1.17	0.24	1.02	0.31	P

中国工业企业规模与生产率的异质性

附表 4 2004 年中国工业企业生产率分布的估计及检验结果

样本行业	帕累托分布			对数正态分布		指数分布		支持分布
	x_{\min}	k	p 值	R	p 值	R	p 值	
所有行业	373.17	4.24	1.00	0.03	0.97	1.69	0.09	<i>P</i>
煤炭开采和洗选业	350.00	3.80	0.97	0.35	0.73	1.28	0.20	<i>P</i>
石油和天然气开采业	191.67	2.77	1.00	-3.44	0.00	2.57	0.01	<i>L</i>
黑色金属矿采选业	237.89	2.60	0.86	-5.44	0.00	-2.81	0.00	<i>L</i>
有色金属矿采选业	148.08	2.60	1.00	-3.14	0.00	1.78	0.07	<i>L</i>
非金属矿采选业	342.08	3.24	1.00	-1.48	0.14	0.56	0.58	<i>P</i>
其他采矿业	1180.39	8.34	0.53	-0.23	0.82	-0.57	0.57	<i>P</i>
农副食品加工业	408.00	3.56	1.00	-1.64	0.10	-0.76	0.45	<i>L</i>
食品制造业	366.00	2.97	0.97	-4.28	0.00	-3.00	0.00	<i>L</i>
饮料制造业	133.05	2.52	1.00	-3.30	0.00	1.97	0.05	<i>L</i>
烟草制品业	288.89	3.72	1.00	0.03	0.97	2.76	0.01	<i>P</i>
纺织业	276.67	3.87	1.00	-0.90	0.37	0.90	0.37	<i>P</i>
纺织服装、鞋、帽制造业	318.65	3.62	1.00	-1.10	0.27	0.80	0.42	<i>P</i>
皮革、毛皮、羽毛(绒)及其制品业	224.56	3.32	1.00	-1.15	0.25	1.74	0.08	<i>P</i>
木材加工及木、竹、藤、草制品业	221.87	2.66	1.00	-4.21	0.00	-1.54	0.12	<i>L</i>
家具制造业	161.26	3.15	1.00	-0.32	0.75	3.29	0.00	<i>P</i>
造纸及纸制品业	329.25	3.09	1.00	-2.80	0.01	-1.10	0.27	<i>L</i>
印刷业和记录媒介的复制	198.29	3.08	1.00	-1.26	0.21	2.19	0.03	<i>P</i>
文教体育用品制造业	176.31	2.70	1.00	-3.44	0.00	0.70	0.48	<i>L</i>
石油加工、炼焦及核燃料加工业	816.67	4.63	1.00	-2.78	0.01	-4.32	0.00	<i>E</i>
化学原料及化学制品制造业	232.50	3.21	1.00	-1.91	0.06	0.90	0.37	<i>L</i>
医药制造业	215.75	2.87	1.00	-1.71	0.09	0.97	0.33	<i>L</i>
化学纤维制造业	150.00	2.24	0.40	-6.83	0.00	-0.99	0.32	<i>L</i>
橡胶制品业	499.50	3.17	0.97	-4.01	0.00	-4.77	0.00	<i>E</i>
塑料制品业	437.50	3.81	1.00	-0.59	0.55	0.32	0.75	<i>P</i>
非金属矿物制品业	320.00	3.36	1.00	-1.29	0.20	1.21	0.23	<i>P</i>
黑色金属冶炼及压延加工业	129.99	2.51	1.00	-4.25	0.00	1.50	0.13	<i>L</i>
有色金属冶炼及压延加工业	187.50	2.84	1.00	-1.84	0.07	2.74	0.01	<i>L</i>
金属制品业	233.84	2.89	1.00	-2.66	0.01	0.85	0.39	<i>L</i>
通用设备制造业	154.77	2.49	1.00	-3.94	0.00	2.14	0.03	<i>L</i>
专用设备制造业	190.22	3.03	1.00	-1.54	0.12	2.10	0.04	<i>P</i>
交通运输设备制造业	162.42	2.79	1.00	-1.79	0.07	3.00	0.00	<i>L</i>
电气机械及器材制造业	568.75	3.62	1.00	-2.01	0.04	-2.04	0.04	<i>L</i>
通信、计算机及其他电子制造业	82.40	2.10	0.89	-6.06	0.00	2.00	0.05	<i>L</i>
仪器仪表及文化、办公机械制造业	397.57	3.11	0.97	-2.07	0.04	-1.65	0.10	<i>L</i>
工艺品及其他制造业	314.18	3.14	1.00	-1.26	0.21	1.32	0.19	<i>P</i>
废弃资源和废旧材料回收加工业	95.80	2.53	1.00	-2.08	0.04	4.72	0.00	<i>L</i>
电力、热力的生产和供应业	275.47	2.48	0.00	-2.49	0.01	-2.14	0.03	<i>N</i>
燃气生产和供应业	179.82	2.57	1.00	-3.10	0.00	1.77	0.08	<i>L</i>
水的生产和供应业	282.35	3.49	1.00	-0.29	0.77	3.93	0.00	<i>P</i>

附表5 2008年不同经营年限企业生产率分布的估计及检验结果

样本行业	经营年限 ≤ 10							经营年限 > 10								
	帕累托分布		对数正态分布		指数分布			支持分布	帕累托分布		对数正态分布		指数分布			支持分布
	k	p值	R	p值	R	p值	k		p值	R	p值	R	p值			
0	2.54	1.00	-2.50	0.01	4.82	0.00	L	2.41	1.00	-4.41	0.00	5.88	0.00	L		
1	1.78	0.00	-9.90	0.00	2.10	0.04	N	2.06	0.98	-4.14	0.00	7.71	0.00	L		
2	2.49	0.59	-1.20	0.23	2.52	0.01	P	2.08	0.00	-2.41	0.02	2.81	0.00	N		
3	2.20	0.65	-5.18	0.00	4.45	0.00	L	2.31	0.91	-4.08	0.00	3.21	0.00	L		
4	2.10	0.95	-4.11	0.00	8.44	0.00	L	1.92	0.38	-6.15	0.00	9.35	0.00	L		
5	2.92	1.00	-0.38	0.71	4.34	0.00	P	2.94	1.00	-0.83	0.41	4.33	0.00	P		
6	2.63	0.58	-0.06	0.96	3.81	0.00	P	2.91	0.82	0.08	0.94	1.85	0.06	P		
7	2.46	0.90	-3.25	0.00	1.54	0.12	L	2.46	0.79	-4.12	0.00	0.89	0.37	L		
8	2.75	1.00	-0.97	0.33	3.91	0.00	P	2.52	1.00	-2.30	0.02	3.48	0.00	L		
9	2.43	0.98	-3.12	0.00	3.63	0.00	L	2.45	1.00	-2.72	0.01	4.13	0.00	L		
10	1.69	0.02	-1.47	0.14	0.26	0.80	N	2.18	0.01	-1.06	0.29	-0.06	0.95	N		
11	2.54	0.98	-3.37	0.00	2.91	0.00	L	2.76	1.00	-2.01	0.04	3.36	0.00	L		
12	2.81	1.00	-0.18	0.85	5.08	0.00	P	2.80	1.00	-0.99	0.32	3.34	0.00	P		
13	2.67	1.00	-1.25	0.21	3.96	0.00	P	2.37	1.00	-3.06	0.00	5.86	0.00	L		
14	2.76	1.00	-2.29	0.02	2.47	0.01	L	2.78	1.00	-2.00	0.05	2.98	0.00	L		
15	2.97	1.00	-0.41	0.68	4.51	0.00	P	3.09	1.00	-0.67	0.50	2.97	0.00	P		
16	3.03	1.00	-1.68	0.09	1.15	0.25	L	2.65	1.00	-2.33	0.02	4.11	0.00	L		
17	2.98	1.00	-0.56	0.57	3.33	0.00	P	3.22	1.00	-1.50	0.13	1.51	0.13	P		
18	3.14	1.00	-1.34	0.18	1.23	0.22	P	2.75	1.00	-1.75	0.08	4.08	0.00	L		
19	2.79	0.01	-6.09	0.00	-9.05	0.00	N	2.15	0.00	-8.15	0.00	-7.08	0.00	N		
20	2.46	0.67	-4.49	0.00	0.25	0.81	L	2.69	0.84	-3.74	0.00	-1.89	0.06	L		
21	3.05	1.00	-2.17	0.03	-0.76	0.45	L	2.41	0.65	-4.63	0.00	-0.06	0.95	L		
22	2.96	0.01	-5.12	0.00	-7.03	0.00	E	2.06	0.00	-6.42	0.00	-2.31	0.02	N		
23	2.66	1.00	-2.01	0.04	4.43	0.00	L	2.85	1.00	-1.16	0.25	3.30	0.00	P		
24	2.79	1.00	-1.47	0.14	4.23	0.00	P	2.91	1.00	-1.38	0.17	1.99	0.05	P		
25	2.66	1.00	-1.80	0.07	3.54	0.00	L	2.67	1.00	-2.37	0.02	2.70	0.01	L		
26	1.86	0.00	-11.9	0.00	-4.56	0.00	N	5.02	0.21	-2.14	0.03	-3.97	0.00	L		
27	1.91	0.00	-10.2	0.00	-1.05	0.29	N	2.46	0.01	-6.02	0.00	-5.36	0.00	N		
28	2.84	1.00	-1.64	0.10	2.62	0.01	P	2.74	1.00	-1.99	0.05	2.40	0.02	L		
29	2.84	1.00	-0.76	0.44	5.52	0.00	P	2.52	0.99	-3.09	0.00	4.29	0.00	L		
30	2.77	1.00	-1.91	0.06	2.04	0.04	L	2.70	1.00	-2.20	0.03	2.50	0.01	L		
31	2.64	1.00	-2.05	0.04	4.07	0.00	L	2.55	1.00	-2.52	0.01	4.96	0.00	L		
32	2.62	1.00	-2.60	0.01	3.15	0.00	L	2.57	1.00	-2.80	0.01	1.34	0.18	L		
33	2.73	1.00	-1.76	0.08	2.22	0.03	L	2.56	1.00	-3.29	0.00	1.77	0.08	L		
34	2.66	1.00	-1.93	0.05	2.39	0.02	L	2.59	1.00	-1.71	0.09	4.53	0.00	L		
35	2.84	1.00	-0.36	0.72	4.59	0.00	P	2.94	1.00	-0.58	0.56	4.67	0.00	P		
36	2.35	0.94	-3.85	0.00	3.08	0.00	L	2.47	0.83	-2.94	0.00	1.83	0.07	L		
37	2.73	1.00	-2.14	0.03	1.51	0.13	L	2.39	1.00	-3.09	0.00	3.15	0.00	L		
38	3.01	0.90	-1.45	0.15	-0.20	0.84	P	2.64	0.35	-2.26	0.02	0.46	0.65	L		
39	3.15	1.00	-1.39	0.17	1.86	0.06	P	3.92	1.00	-0.31	0.76	0.73	0.46	P		

说明:0~39与附表1、2、3、4的样本行业一一对应。

(截稿:2014年4月 责任编辑:李元玉)