

1996-2013年中国制造业企业TFP测算

王贵东

内容提要:在系统性整理中国工业企业数据库的基础上,本文利用OP、LP、GMM、FE等方法,首次较为科学规范地测算了1996-2013年中国制造业企业的全要素生产率。结果显示:30个行业的资本、劳动产出弹性分别处于0.2-0.4、0.3-0.6区间,TFP增速处于3.90-18.94%区间,且呈现一定的收敛趋势;中国的行业存在上下游TFP集聚现象,且与下游行业更为密切;中国各省的TFP在全国层面不具有显性特点,但在局部小区域具有较强的集聚现象;中国企业的TFP整体随时间波动上升,且企业间存在收敛。

关键词:全要素生产率;中国工业企业数据库;交叉识别

DOI: 10.19365/j.issn1000-4181.2018.04.08

一、问题的提出

全要素生产率(Total factor productivity, TFP)原属宏观经济研究范畴,由Solow(1956)提出,是指各传统生产要素之外的可影响产出的部分,其主要来源包括技术进步、组织创新、制度环境和生产创新等。TFP的测算通常由生产函数开始,因为生产函数的估计方法存在一定差异,所以TFP的测算也有所不同,具体可包括3个维度的分类(Del Gatto et al., 2011):(1)前沿分析、非前沿分析;(2)确定性方法、计量参数法、计量半参数法;(3)宏观方法、微观方法、宏微观方法。

近几年,随着微观数据的日益丰富、计量方法的创新突破、数据处理软件的高效化,以及微观TFP既可研究微观又可“加总”研究宏观,所以学者们开始转向微观方法,利用微观数据测算企业TFP。在微观方法领域,OP法(Olley & Pakes, 1996)和LP法(Levinsohn & Petrin, 2003)为最成熟的方法,两者均采用半参数估计。前者旨在解决微观数据的同时性偏差和样本选择偏差问题,后者旨在解决微观数据的同时性偏差问题。另外,Il Kim et al.(2016)在OP法、LP法基础上,进一步解决了资本、劳动和中间投入品的观测误差问题。

本文主要利用OP法,同时也兼顾LP法、GMM法、FE法等,对1996-2013年中国制造业企业TFP进行测算。因为企业TFP是很多实证研究的基础,所以测算过程是否科学规范就直接决定了TFP数据的质量。但是,目前学者们对中国制造业企业的TFP测算都存在两类问题:一是数据处理过程较为粗糙,二是行业异质性的忽略。因此,本文的主要边际贡献为:在系统性处理中国工业企业数据的基

收稿日期:2017-02-20

基金项目:本文获得北京大学光华思想力平台、国家社会科学基金重大项目(15ZDA019)、教育部人文社会科学重点研究基地重大项目(14JJD790004)、国家社会科学基金重点项目(14AZD110)的资助。

作者简介:王贵东,北京大学光华管理学院,博士后。

致谢:感谢北京大学光华管理学院刘俏教授、国务院参事室徐宪平参事、南开大学经济学院周京奎教授提供的帮助,以及匿名审稿专家提出的建设性意见。诚然,文责自负。

基础上,考虑行业中类的要素产出弹性异质性,首次较为科学规范地测算了1996-2013年中国制造业企业TFP。

二、两类问题的解决

第一类问题 数据处理方面。本文数据来源为中国工业企业数据库(1996-2013年)、中国经济普查数据(2004、2008和2013年)、《中国投入产出表》(1997、2002、2007和2012年)、历年《中国统计年鉴》。其中,中国工业企业数据和中国经济普查数据均为微观大数据,不可避免地具有样本匹配混乱、指标缺失严重、指标异常值、测度误差、录入误差和指标定义模糊等问题(聂辉华等,2012)。为更精细化处理数据,本文主要借鉴王贵东(2017a)科学规范地对程序可重复性、变量代码修复、企业ID交叉识别、极端值判断及处理、行业调整、价格调整、规模以上调整等进行了处理,尤其是企业ID交叉识别。

第二类问题 行业异质性方面。在一些经典文献中,鲁晓东和连玉君(2012)忽略了行业大类的要素产出弹性异质性,杨汝岱(2015)忽略了行业中类的要素产出弹性异质性。但是,资本密集型行业与劳动密集型行业对资本和劳动的需求是存在差异的,行业分类越粗糙则意味着这种异质性越为明显。为充分考虑行业中类的要素产出异质性,本文借鉴王贵东(2017b)、王贵东和周京奎(2017)科学巧妙地对企业的行业中类进行划分(也包括行业大类),分别估计不同行业大类(2位代码)、行业中类(3位代码)的资本、劳动产出弹性。对于行业大类,本文以GB/T4754-2002标准为基准统一整理,并对“采矿业”行业门类,“电力、燃气及水的生产和供应业”行业门类,以及其他非“制造业”行业进行剔除,剔除后得到30个制造业行业大类(王贵东,2017a)。对于行业中类,本文将参考GB/T4754-1994、GB/T4754-2002标准对企业的行业中类进行重新梳理,主要采用“自下而上的自选择”处理方式(王贵东,2017b)。以GB/T4754-2002为例,首先需要分别汇总2003-2012年每个行业中类下的所有企业ID,再利用2003-2012年每个行业中类下的企业ID向其他年份扩散,进而得到GB/T4754-2002行业中类的“自下而上的自选择统一”。很明显,这种处理方式可以挖掘出企业更换行业的信息、企业从事多种行业的信息;同时,这种处理方式将样本量扩散至其他年份,大大增加了参与估计的样本量,一定程度上解决了OP法对最低样本量的要求。

三、研究设计

(一) 模型设定

测算TFP需要先设定企业的生产函数,具体包括线性、Leontief、CD、CES、VES、Trans-log等形式。由于CD生产函数结构简单,且经济学含义更为清晰,所以本文采用CD形式

$$Y_{it} = A_{it} L_{it}^{\alpha} K_{it}^{\beta} \quad (1)$$

$$A_{it} \equiv \gamma^t A_{i0} \quad (2)$$

其中, Y_{it} 为产出, L_{it} 为劳动, K_{it} 为资本; A_{it} 为TFP,并分解为 A_{i0} 、 γ^t 两部分: A_{i0} 为基期TFP, γ^t 为 t 期TFP的增长情况。将式(2)代入式(1),并取对数可线性化为:

$$\ln Y_{it} = \alpha \cdot \ln L_{it} + \beta \cdot \ln K_{it} + \underbrace{(\ln \gamma) \cdot t + \ln A_{i0}}_{\ln TFP} \quad (3)$$

(二) 估计方法

通过对式(3)估计,可得到系数 α 、 β 和 $\ln \gamma$ 的估计值,进而得到当期TFP和基期TFP。另外,由 $A_{it} \equiv \gamma^t A_{i0} = e^{(\ln \gamma) \cdot t} A_{i0}$ 可知,系数 $\ln \gamma$ 恰好为TFP增速。但是,仅对式(3)进行OLS估计,将会面临同时性偏差和样本选择偏差问题。其中,同时性偏差是指企业在实际生产过程中,可以观测到部分TFP,

并根据最优原则及时调整各生产要素,这意味着计量估计模型中的误差项与解释变量具有相关性,进而使得估计系数有偏;样本选择偏差是指当企业面临 TFP 下行冲击时,资本存量较大的企业因抗风险能力强而具有较低的退出市场概率,资本存量较小的企业因抗风险能力弱而具有较高的退出市场概率,即 TFP 冲击与企业退出市场的概率存在相关性,进而产生样本选择性偏差。

Olley & Pakes(1996)针对这两类问题提出了 OP 法,其特点为:利用投资作为代理变量解决同时性偏差问题,利用二值选择模型解决样本选择偏差问题。当然,还有其它方法可解决同时性偏差问题,包括 LP 法(Levinsohn & Petrin,2003)、GMM 法(Blundell & Bond,1998;Blundell & Bond,2000)等。其中,LP 法与 OP 法最大的不同在于,前者采用中间投入品作代理变量,而后者采用投资作代理变量。GMM 法与 OP 法最大的不同在于,前者使用变量自身滞后项作工具变量,后者采用当期投资作代理变量。鉴于 OP 法在解决样本选择偏差方面的优势,本文将主要采用 OP 法测算 TFP。

(三) 关键变量处理

在测算 TFP 时,产出 Y 、劳动 L 、资本 K 三个基本变量分别用工业增加值(y_add)、职工数($labor$)、固定资产($capital$)表示。此外,OP 法还需要计算投资($invest$)和企业退出($exit$)两个变量。对于投资($invest$),计算公式为:投资=当年固定资产-上年固定资产估计值+当年折旧,这里的上年固定资产估计值采用“几何”插值法(王贵东,2017a)^①。对于企业退出($exit$),若企业不出现在观察期末年^②且企业时间序列不中断,则在企业最后出现的年份赋值 1;否则,赋值 0。同时,本文还设定 TFP 的基期为 1996 年,将时间($time$)加入到 OP 法估计中, $time$ 的估计系数可以挖掘 TFP 的增长速度。

表 1 变量统计信息

变量	解释	样本量	均值	标准差	最小值	最大值
$\ln(y_add)$	工业增加值对数	2230363	8.6454	1.4827	0.0000	17.5499
$\ln(capital)$	固定资产对数	3776311	8.6431	1.7618	-0.1194	18.9587
$\ln(labor)$	职工数对数	3793467	4.8673	1.1612	0.0000	12.3159
$\ln(invest)$	投资对数	2032950	6.4583	1.8162	-5.6512	16.2614
$exit$	企业退出	1372606	0.0875	0.2826	0.0000	1.0000
$time$	时间	3844966	10.5452	4.4059	0.0000	17.0000
$\ln(y_total)$	工业总产值对数	3795361	10.4526	1.4305	-0.1050	19.7465
$\ln(middle)$	中间投入品对数	1999102	9.7544	1.4539	0.0777	19.0302

注:所有价格变量均用地区价格指数作平减,基期为 2010 年;LP 法将用到工业总产值(y_total)和中间投入品($middle$)。

(四) 企业 TFP 测算

通过对各个关键变量的时间跨度进行梳理,本文发现:固定资产($capital$)、职工数($labor$)的时间跨度为 1996-2013 年,工业增加值(y_add)的时间跨度仅为 1996-2007、2010 年,这对 OP 法估计提出了挑战。如果转用 LP 法估计,虽然工业总产值(y_total)的时间跨度达 1996-2013 年,但中间投入品($middle$)的时间跨度仅为 1996-2007 年,这意味着 LP 法将比 OP 法损失 2010 年的观测值,也就是 OP 法在有效观测值方面优于 LP 法。

为了额外测算 2008-2009、2011-2013 年的 TFP,本文参考中国工业企业数据库中的其他变量进行辅助生成。具体过程为:先用 OP 法估计资本、劳动产出弹性,直接测算 1996-2007、2010 年 TFP;然

^① 当企业上年固定资产为空值时,其估计值为真实插值;当企业上年观测值不存在时,其估计值为虚拟插值。

^② 由于工业增加值 Y_add 在中国工业企业数据中缺失 2008、2009、2011、2012、2013 年数据,所以实际观察期末期为 2010 年。为了估计出缺失年份的 TFP,本文还将结合其他方法,详见后文。

后,参考其他变量估测未知年份的工业增加值;最后,利用资本、劳动产出弹性,以及估测的工业增加值,间接测算 1998-1999、2011-2013 年 TFP。其中,工业增加值的估算主要有以下三种:(1)要素报酬占工业增加值的比重估算。可以预见,若企业完全按照最优化原则进行生产,则其要素报酬占工业增加值的比重必然为要素产出弹性,也就是可以利用劳动报酬和资本报酬去测算未知年份的工业增加值。庆幸的是,中国工业企业数据库中恰好有劳动报酬数据(即应付工资总额,时间跨度为 1996-2008、2011-2013 年),利用应付工资总额可以直接估测出 2008、2011-2013 年的工业增加值(=应付工资总额÷劳动产出弹性)。但是,仍缺少 2009 年的工业增加值,可以再利用资本报酬去补充测算。较为棘手的是,资本报酬数据是无法从中国工业企业数据库中直接获得的,所以还需要额外测算资本报酬,具体公式为“资本报酬=利息支出÷负债总额×固定资产”^①,进而估测出 2009 年的工业增加值(=资本报酬÷资本产出弹性)。需要说明的是,劳动报酬估算工业增加值的准确性可能高于资本报酬,所以建议先用劳动报酬去更新工业增加值缺失值,再用资本报酬去更新。当然,若存在要素扭曲,则该方法的功效将有所降低。(2)增值税与工业增加值的关系估算。因为增值税是针对产品增值而征收的一种税,所以可以利用增值税占工业增加值的比重来估算工业增加值。具体操作为:计算每个企业每年增值税占工业增加值的比重,利用其中位数估算未知年份的工业增加值(=增值税÷增值税占工业增加值比重中位数)。庆幸的是,中国工业企业数据库中恰好有 1996-2013 年增值税数据。当然,若企业每年的税收结构有较大变化或存在增值税改革,则该方法的功效将有所降低。(3)工业总产值与工业增加值的关系估算。当每个企业的工业增加值占工业总产值的比重较为稳定时,可以利用工业总产值估算工业增加值。庆幸的是,中国工业企业数据库中恰好有 1996-2013 年工业总产值数据。当然,若该比重每年都有较大波动,则该方法的功效将有所降低。

表 2 为 OP 法分行业大类的估计结果,参考 GB/T4754-2002 标准对 30 个行业大类进行估计。其中,资本、劳动产出弹性是以年份作因子变量的估计结果,使用该方法能够消除年份固定效应的影响;TFP 增速是以时间作控制变量的估计结果,使用该方法用于捕捉不同行业大类的 TFP 增速。同时,附录为 OP 法分行业中类估计结果,分别参考 GB/T4754-1994、GB/T4754-2002 标准对 172、169 个行业中类进行估计,且不再考虑时间因素。在得到 TFP 的初步测算值后,本文还利用中位数法对其进行了极端值处理。

表 2 OP 法分行业大类估计结果(GB/T4754-2002)

代码	行业名称	资本产出弹性	劳动产出弹性	TFP 增速	OP 样本	总样本
13	农副食品加工业	0.2919***	0.5219***	0.1507***	75293	241185
14	食品制造业	0.3420***	0.5509***	0.1429***	31033	90910
15	饮料制造业	0.3716***	0.5689***	0.1284***	22192	63716
16	烟草制品业	0.2279***	0.5377***	0.0390***	2049	3546
17	纺织业	0.2597***	0.4715***	0.0942***	117291	320184
18	纺织服装_鞋_帽制造业	0.2260***	0.5420***	0.0696***	64440	176510
19	皮革_毛皮_羽毛(绒)及其制品业	0.2290***	0.4656***	0.0845***	32368	87770
20	木材加工及木_竹_藤_棕_草制品业	0.3071***	0.5211***	0.1195***	24213	86407
21	家具制造业	0.2536***	0.5921***	0.0879***	15154	48751
22	造纸及纸制品业	0.2704***	0.5023***	0.0946***	41030	104319
23	印刷业和记录媒介的复制	0.3661***	0.4946***	0.0985***	29130	71528
24	文教体育用品制造业	0.2499***	0.5318***	0.0586***	18241	47318
25	石油加工_炼焦及核燃料加工业	0.3102***	0.3236***	0.1269***	9457	25860

① 充分体现机会成本的概念;也可采用“利息支出÷负债总额×总资产”的形式。

续表

代码	行业名称	资本产出弹性	劳动产出弹性	TFP 增速	OP 样本	总样本
26	化学原料及化学制品制造业	0.3327***	0.3616***	0.1209***	101238	282154
27	医药制造业	0.3565***	0.4995***	0.0846***	29672	76595
28	化学纤维制造业	0.3862***	0.4423***	0.1044***	7206	20902
29	橡胶制品业	0.3634***	0.4292***	0.1012***	16411	44748
30	塑料制品业	0.2911***	0.4437***	0.0856***	62579	179419
31	非金属矿物制品业	0.2802***	0.4227***	0.1213***	115585	324636
32	黑色金属冶炼及压延加工业	0.2503***	0.4970***	0.1622***	30528	86352
33	有色金属冶炼及压延加工业	0.3134***	0.4152***	0.1521***	25542	69723
34	金属制品业	0.3254***	0.4450***	0.0963***	71115	213707
35	通用设备制造业	0.3627***	0.4493***	0.1228***	102865	312527
36	专用设备制造业	0.3611***	0.4489***	0.1347***	56868	170641
37	交通运输设备制造业	0.3415***	0.5255***	0.1176***	61727	180665
39	电气机械及器材制造业	0.3569***	0.4364***	0.0976***	81999	240521
40	通信设备_计算机及其他电子设备制造业	0.3531***	0.4699***	0.0700***	45832	134859
41	仪器仪表及文化_办公用机械制造业	0.3664***	0.4116***	0.1019***	19407	53366
42	工艺品及其他制造业	0.2132***	0.4871***	0.0724***	29888	79290
43	废弃资源和废旧材料回收加工业	0.0577	0.3971***	0.1894***	1291	7718

注: ***, **, * 分别表示 1%、5%、10% 的显著性水平。

在表 2 中,除废弃资源和废旧材料回收加工业(当然,大部分文献都直接剔除了该行业)外,其他 29 个行业的估计系数均在 1% 的显著性水平上显著为正。与其他文献比较,本文的资本、劳动产出弹性要普遍小于杨汝岱(2015)的估计^①,并且该文献的烟草制品业资本产出弹性并不显著;而本文与鲁晓东和连玉君(2012)的估计较为接近。另外,废弃资源和废旧材料回收加工业的资本产出弹性在以年份作因子变量的 OP 法中不显著,而在以时间作控制变量的 OP 法中非常显著(p 值 0.000);这说明该行业的资本几乎与年份固定效应完全共线,而这也恰好符合该行业的特性。

横向比较表 2 中的产出弹性,可以发现 30 个行业的资本产出弹性均低于劳动产出弹性。进一步观察,可以发现:不同行业间的资本产出弹性波动区间为 [0.2132, 0.3862]^②,劳动产出弹性波动区间为 [0.3236, 0.5921]。由于其波动范围较大,所以直接将所有行业统一进行估计是不合理的。通过利用 30 个行业的要素产出弹性,可以计算整个制造业的要素产出弹性,具体思路为:在每个行业中,要素产出弹性都可表示为要素报酬占工业增加值的比重,那么整个制造业的要素报酬占工业增加值的比重相当于 30 个行业的要素报酬占工业增加值的比重的加权平均,而整个制造业的要素报酬占工业增加值的比重又可解释为整个制造业的要素产出弹性。该思路下,利用表 2 的估计结果可计算出整个制造业的资本产出弹性为 0.3112,劳动产出弹性为 0.4665;利用附录的估计结果可计算出整个制造业的资本产出弹性为 0.3159,劳动产出弹性为 0.4071。为了对比,本文还直接对整个制造业进行

① 至少有两个原因导致了此差异。一是杨汝岱(2015)采用了序贯识别法,从而使 OP 法的关键变量 *exit* 数据质量不高。二是杨汝岱(2015)使用的中国工业企业数据不全,仅含有固定资产原值变量,没有固定资产变量(从而只能去估算真实的固定资产,但估算的假设又过于牵强);而本文所用的中国工业企业数据,不仅含有固定资产原值,还有固定资产,甚至是折旧数据。也就是说,对于 OP 法最关键的两个变量,该文献都质量欠佳。当然,该文献也承认了其估计结果与多数文献有所出入,因为大部分文献的资本产出弹性与劳动产出弹性之和都是明显小于 1 的,而他们的结果接近于 1,甚至有些行业大于 1。另外,由于杨汝岱的固定资产变量是额外估算的(精确度较低),而本文的固定资产变量是工企原始数据(缺失值较多),所以在同一时间区间下,杨汝岱的 OP 法估计样本量可能会大于本文。

② 这里,不包括废弃资源和废旧材料回收加工业。

了 OP 法估计,结果显示资本产出弹性为 0.3041,劳动产出弹性为 0.4524。可以发现,不同方法计算的制造业产出弹性有所差异,本文更支持附录的加权平均算法。

最后,本文还分行业采用 LP、GMM、FE 等方法估计了资本、劳动产出弹性,以之对比 OP 法估计结果,感兴趣的读者可向作者索要。

四、TFP 分类比较

利用表 2 和附录的估计结果,本部分将从不同维度分类比较 TFP。其中,不同行业的 TFP 增速由表 2 得出,上中下游行业的 TFP、不同地区的 TFP、不同年份的 TFP 由附录得出。

(一) 不同行业 TFP 增速比较

观察表 2 可以发现,技术进步最快的行业为废弃资源和废旧材料回收加工业(增速 18.94%)、黑色金属冶炼及压延加工业(增速 16.22%)、有色金属冶炼及压延加工业(增速 15.21%)、农副食品加工业(增速 15.07%)和食品制造业(增速 14.29%)。其中,废弃资源和废旧材料回收加工业 TFP 增速最快,反映了中国对资源循环利用的高度重视;金属冶炼及压延加工业 TFP 增速较快,反映了此期间中国用于再生产的物质资料投入较高,并促使该行业技术升级的现象;食品业 TFP 增速较快,折射出此期间中国居民食品种类的丰富化趋势。另一方面,技术进步最慢的行业分别为烟草制品业(增速 3.90%)、文教体育用品制造业(增速 5.86%)、纺织服装_鞋帽制造业(增速 6.96%)、通信设备_计算机及其他电子设备制造业(增速 7.00%)、工艺品及其他制造业(增速 7.24%)。其中,烟草制造业 TFP 增速最慢,反映了中国目前的禁烟成果显著;文教体育用品制造业、纺织服装、鞋帽制造业,以及工艺品及其他制造业 TFP 增速较慢,说明这 3 个行业已经构成夕阳产业。让本文感到意外的是,通信设备_计算机及其他电子设备制造业 TFP 增速较慢,本文猜测其原因可能为:(1)地方政府更乐于建高新技术产业园,扩大园区规模,而很少对企业研发创新进行激励。(2)对高新技术知识产权的保护不够,一定程度上削弱了企业的创新动力。不妨举例,北京小米科技的市场更多局限于国内,而华为技术公司在海外似乎比在国内更具竞争力。

需要说明的是,本文的 TFP 增速要高于部分文献(鲁晓东和连玉君,2012;杨汝岱,2015)。因为这些文献在表述 TFP 水平值和 TFP 增速时,可能存在理解偏误^①。通过分析 OP 法的函数编程^②,本文发现 OP 法对 TFP 的估计值实际为对数 TFP,而非 TFP 水平值。而对这些文献提到的“TFP 水平值”进行对数还原后,正好与本文的 TFP 处于同一数量级。当然,这种理解偏误的后果是,TFP 增速的巨大差异^③。

(二) 上中下游行业 TFP 比较

本文不仅测算了每个企业的 TFP,而且还测算了每个企业所在行业(3 位代码)的 TFP,以及其上游行业(3 位代码)的 TFP 和下游行业(3 位代码)的 TFP。其中,每个行业的上游行业和下游行业权重由 1997、2002、2007、2012 年《中国投入产出表》计算,并对《中国投入产出表》中的行业分类和中国工业企业数据库的 3 位行业代码进行匹配。最终,计算得到 1997、2002、2007、2012 年四个年份的企业上游行业 TFP、企业本行业 TFP、企业下游行业 TFP。限于篇幅,表 3 仅列出 2007 年的测算结果^④。需

^① 不妨举例,杨汝岱(2015)文献的图 3 左图显示全国加权 TFP 仅为个位数字;鲁晓东和连玉君(2012)文献的表 4 显示各省份 TFP 也仅为个位数字。

^② 比如 STATA 下的 opreg 命令。

^③ 不妨举例,杨汝岱(2015)计算出 1998-2007 年全国 TFP 增速为 3.35%;鲁晓东和连玉君(2012)计算出 1999-2007 年金属冶炼行业的 TFP 增速最快,为 4.7%。

^④ 由于大部分文献估计的是 1998-2007 年 TFP,所以本文列出 2007 年的上、中、下游行业 TFP,以便于对比。

要说明的是,本文的 TFP 测算结果考虑了权重因素(以工业增加值为权重^①),并控制了规模以上^②。

表 3 2007 年上、中、下游行业 TFP 测算结果

行业中类	上游行业	本行业	下游行业
谷物磨制业	348.411	339.71	695.868
饲料加工业	373.973	294.128	556.588
植物油加工业	723.186	719.709	756.89
制糖业	310.78	249.129	586.068
屠宰及肉类加工业	509.229	358.761	562.847
水产品加工业	365.761	386.295	598.226
其他食品加工业	396.589	639.635	639.53
其他食品制造业	307.246	266.813	568.168
方便食品制造业	437.154	437.154	473.239
液体乳及乳制品制造业	225.66	156.41	467.056
调味品_发酵制品制造业	278.076	64.39	496.49
酒精及酒的制造业	921.162	221.198	495.079
软饮料及精制茶加工业	787.976	201.435	496.008
烟草制品业	3335.083	10609.77	3831.977
棉_化纤纺织及印染精加工业	448.382	394.645	458.092
毛纺织和染整精加工业	470.544	430.208	546.663
麻纺织_丝绢纺织及精加工业	512.41	586.167	611.726
纺织制成品制造业	462.717	587.833	494.652
针织品_编织品及其制品制造业	476.683	479.497	454.431
纺织服装_鞋帽制造业	726.931	427.554	502.148
皮革_毛皮_羽毛(绒)及其制品业	467.834	495.953	550.822
木材加工及木_竹_藤_棕_草制品业	342.111	283.449	562.833
家具制造业	550.368	208.269	512.623
造纸及纸制品业	923.389	532.078	972.938
印刷业和记录媒介的复制业	1497.063	162.677	598.119
文教体育用品制造业	525.831	282.793	581.925
石油及核燃料加工业	1119.132	3185.631	2169.354
炼焦业	936.032	389.48	1111.097
基础化学原料制造业	816.075	1333.043	1435.342
肥料制造业	568.769	505.593	882.64
农药制造业	651.355	333.401	1032.805
涂料_油墨_颜料及类似产品制造业	574.378	796.332	1178.035
合成材料制造业	483.446	614.318	1837.255
专用化学产品制造业	741.621	863.587	1201.576
日用化学产品制造业	2132.442	3920.241	1463.31
医药制造业	527.556	474.736	665.492
化学纤维制造业	402.774	164.421	847.488
橡胶制品业	502.556	217.634	696.406

① 考虑到 OP 法的被解释变量为工业增加值,所以本文优先选择工业增加值作权重。当然,也有一部分文献采用职工数或工业总产值加权。但本文认为职工数加权相当于给予劳动密集型企业更高权重,容易造成有偏;而工业总产值加权相当于给予下游行业的企业更高权重,也会造成有偏。

② 规模以上已调整为“年主营业务收入达到 2000 万元及以上的工业法人企业”。后文不再一一说明。

续表

行业中类	上游行业	本行业	下游行业
塑料制品业	656.76	488.91	693.442
水泥_石灰和石膏制造业	475.636	920.446	913.781
水泥及石膏制品制造业	455.091	221.557	926.461
砖瓦_石材及其他建筑材料制造业	478.047	329.067	899.509
玻璃及玻璃制品制造业	523.933	364.811	1060.347
陶瓷制品制造业	732.059	693.939	1005.274
耐火材料制品制造业	974.542	441.561	794.31
石墨及其他非金属矿物制品制造业	711.442	569.326	935.868
炼铁业	1101.136	533.625	1049.493
炼钢业	909.344	1750.187	1075.651
钢压延加工业	620.37	956.929	1265.874
铁合金冶炼业	856.57	658.628	879.845
有色金属冶炼及合金制造业	971.969	1244.277	1522.176
有色金属压延加工业	591.626	1413.151	1269.818
金属制品业	645.827	538.801	903.931
锅炉及原动机制造业	423.533	390.369	681.279
金属加工机械制造业	568.636	280.819	823.602
起重运输设备制造业	804.119	799.546	758.342
泵_阀门_压缩机及类似机械的制造业	602.784	510.715	853.512
其他通用设备制造业	670.13	549.853	836.908
矿山_冶金_建筑专用设备制造业	864.348	936.566	803.93
化工_木材_非金属加工专用设备制造业	789.645	268.71	775.977
其他专用设备制造业	509.526	324.358	746.254
农林牧渔专用机械制造业	135.393	128.299	565.731
铁路运输设备制造业	834.009	918.662	835.605
汽车制造业	458.661	400.221	574.501
其他交通运输设备制造业	355.653	251.414	582.477
船舶及浮动装置制造业	285.021	279.941	689.48
电机制造业	564.657	526.967	768.338
输配电及控制设备制造业	730.294	917.272	877.161
电线_电缆_光缆及电工器材制造业	771.031	327.117	1177.452
其他电气机械及器材制造业	679.828	306.888	867.07
家用电力和非电力器具制造业	562.01	562.681	667.382
通信设备制造业	1776.25	2019.548	964.491
雷达及广播设备制造业	211.856	103.725	698.14
电子计算机制造业	863.707	887.211	762.661
电子元器件制造业	863.373	692.255	770.298
家用视听设备制造业	512.088	551.63	691.191
其他电子设备制造业	740.371	858.357	712.717
仪器仪表制造业	724.537	769.525	706.272
文化_办公用机械制造业	555.696	290.588	621.194
工艺品及其他制造业	660.464	719.419	755.021
废品废料	1036.062	2453.865	2105.966

横向观察表 3 不难发现 TFP 在产业链中具有明显的集聚现象。具体而言,1997、2002、2007 年^①本行业与上游行业的相关系数分别为 0.7338、0.7655、0.8793,而本行业与下游行业的相关系数分别为 0.7941、0.8571、0.9068。进一步挖掘,可以发现本行业与下游行业的关系更为密切,这说明本行业的 TFP 更容易使下游行业的 TFP 联动,这也一定程度上支持了市场是以需求为导向的基本观点。

纵向观察表 3 在本行业 TFP 一列,可以发现 2007 年 TFP 最高的行业为烟草制品业、日用化学产品制造业、石油及核燃料加工业、废品废料、通信设备制造业、炼钢业、有色金属压延加工业、基础化学原料制造业、有色金属冶炼及合金制造业、钢压延加工业; TFP 最低的行业为调味品_发酵制品制造业、雷达及广播设备制造业、农林牧渔专用机械制造业、液体乳及乳制品制造业、印刷业和记录媒介的复制业、化学纤维制造业、软饮料及精制茶加工业、家具制造业、橡胶制造业、酒精及酒的制造业。很明显,与第一产业相关的制造业普遍具有较低的 TFP,这也一定程度上肯定了政府产业升级的正确性。但另一方面,虽然与第一产业相关的制造业的 TFP 绝对值较低,但其增速却较高,这说明 TFP 在行业间存在一定收敛。

(三) 不同地区 TFP 比较

这部分将测算不同地区的 TFP,限于篇幅不再给出城市级、甚至是县级的测算结果^②,仅给出省级比较结果。需要说明的是,本文将重庆直辖之前原数四川的企业归属重庆。表 4 为不同省份的 TFP 测算结果。同样,该测算结果考虑了权重因素(以工业增加值为权重),并控制了规模以上。

表 4 中国各省 TFP 测算

行政区划代码	省份	均值	标准差	行政区划代码	省份	均值	标准差
11	北京	751.2336	934.0471	41	河南	605.6753	1554.2350
12	天津	1182.8620	1795.1940	42	湖北	357.2276	751.8540
13	河北	514.6903	723.6859	43	湖南	576.2693	939.2037
14	山西	380.4614	414.5183	44	广东	817.0693	1949.2360
15	内蒙古	585.4835	984.7889	45	广西	709.1753	1822.5540
21	辽宁	764.1803	1284.2850	46	海南	1155.2770	2006.3210
22	吉林	517.4935	572.0616	50	重庆	568.2391	1486.8790
23	黑龙江	496.3882	611.8380	51	四川	443.7576	920.2334
31	上海	660.7057	836.3466	52	贵州	1734.2830	4072.5090
32	江苏	509.2438	758.5412	53	云南	5097.4560	8892.9700
33	浙江	319.6247	549.5650	54	西藏	376.8614	308.3636
34	安徽	701.8198	2118.2600	61	陕西	648.9189	1596.6860
35	福建	590.1463	1311.8100	62	甘肃	875.1017	1817.3080
36	江西	577.4651	1278.0080	63	青海	443.4370	408.5138
37	山东	569.8630	841.8297	64	宁夏	280.1779	302.6444
				65	新疆	418.7306	541.0109

观察表 4 可以发现:京津冀城市群中,北京、天津的 TFP 要高于河北;长三角城市群中,上海的 TFP 要高于江苏、浙江;东北地区,辽宁的 TFP 要高于吉林、黑龙江。然而,从全国层面来,似乎 TFP 的地区分布并没有太明显的特点^③。本文猜测,可能是 TFP 具有局部集聚、全局分散的特点,也就是集

① 1997 年、2002 年的上、中、下游行业 TFP 测算结果可向作者索要。

② 本文实际还对中国县级行政区划代码进行了历史数据调整,具体用到爬虫技术,并结合 Google 地图等辅助。当然,这里不详细介绍,该内容在另一篇文章中有所交代。

③ 这一点,本文与鲁晓东和连玉君(2012)的观点略有不同。

聚效应只发生在一定的距离以内。

(四) 不同年份 TFP 比较

正如前文所述,由于工业增加值缺失 2008-2009、2011-2013 年数据,所以不再报告这些年份估测的 TFP^①,仅报告 1996-2007、2010 年测算的 TFP。目前,大部分文献测算 TFP 的年份为 1998-2007 年,正好也方便与本文进行比较。

表 5 中国分年度 TFP 均值测算

年份	A	B	C	D	E
1996	72.1158	305.6748	310.0993	300.9519	4.8622
1997	73.7730	303.1380	307.8834	279.3797	4.8570
1998	91.0942	252.3754	280.2920	187.0724	4.8117
1999	98.2339	242.9254	261.4132	212.7794	4.8708
2000	105.6690	277.3472	299.2892	246.5319	5.0074
2001	113.0367	300.4504	324.0538	253.9857	5.0950
2002	122.6855	370.1136	396.5334	423.4246	5.2514
2003	160.2735	667.6421	711.7152	1012.0050	5.7593
2004	170.1832	623.3575	668.3539	1266.8030	5.8901
2005	189.3941	758.2707	798.2191	1198.7750	5.9403
2006	212.8162	799.5546	835.0013	1226.2360	6.0167
2007	235.6506	1006.4520	1045.5360	1351.4780	6.0947
2010	151.8831	1253.0380	1296.6850	1472.4250	5.8958
极端值	中位数法	中位数法	中位数法	前后各 0.5%法	中位数法
规模以上	修复	无修复	修复	修复	修复
权重	无	工业增加值	工业增加值	工业增加值	工业增加值
TFP	是	是	是	是	否
对数 TFP	否	否	否	否	是
1998-2007 年增速	11.14%	16.61%	15.75%	24.57%	2.66%

注:极端值一行中,中位数法的判断对象为 TFP,并以“年份+行业小类”、企业 ID 为分组;前后各 0.5%法的判断对象为工业增加值,并以行业大类为分组^②。由于 2010 年中国工业企业数据库中的固定资产极端值较多,故 2010 年的 TFP 不适合与其他年份直接作比较。

在表 5 中,参照组为 C 列。对比 A 列和 C 列,可以发现:TFP 简单均值要小于 TFP 工业增加值加权平均。对比 B 列和 C 列,可以发现:对规模以上修复后,TFP 均值有所增加。对比 D 列和 C 列,可以发现:1996-2001 年,前后各 0.5%极端值法下的 TFP 均值小于中位数极端值法下的 TFP 均值;2002-2010 年,前后各 0.5%极端值法下的 TFP 均值大于中位数极端值法下的 TFP 均值。这意味着工业增加值可能与 TFP 存在非线性关系,具体表现为:一般情况下,随着工业增加值的增加,企业的 TFP 也增加;但是一旦企业增加值达到行业最大(比如行业前 0.5%),企业的 TFP 既可能降低(1996-2001 年)也可能上升(2002-2010 年)。这意味着 1996-2001 年中国的企业创新主要由大中型企业贡献,而特大型企业则更倾向于寻租来保持现有的行业地位;2002-2010 年中国的企业创新不仅由大中型企业贡献,还由特大型企业贡献。对于 1996-2010 年特大型企业的市场行为发生转变,其原因可能为:(1) 中国国有企业改革的巨大成功,尤其是针对特大型国有企业;(2) 2001 年底中国加入 WTO,特大

① 感兴趣的读者,可向作者索要结果。需要说明的是,由于工业增加值采用了估测,所以,最终测算的 TFP 不可与 1996-2007、2010 年直接比较,但在 2008-2009、2011-2013 年的内部可进行比较。

② 前后各 0.5%法参照杨汝岱(2015)。

型企业面临国际市场竞争,而不得不提高自身创新能力。需要说明的是,这正好暗示了前后各0.5%法处理极端值是非常不合理的,因为它忽略了特大型企业对中国TFP的影响,相当于删除了权重最大的企业,这不符合统计学基本要求。

对比E列和C列,可以发现对数TFP和TFP的不同是全方位的,对数TFP水平值基本是个位数,而TFP水平值基本是三位数。E列显示,1998-2007年,中国对数TFP增速为2.66%,同类文献中,李玉红、王皓和郑玉歆(2008)估计出2000-2005年中国“TFP”增速为2.5%;杨汝岱(2015)估计出1998-2007年中国“TFP”增速为3.35%;鲁晓东和连玉君(2012)估计出1999-2007年中国各省“TFP”增速为1.01-5.05%。将E列的对数TFP增速转为TFP增速,则为15.75%(即C列)。

五、结语

利用中国工业企业数据库,本文首次测算了中国制造业企业在1996-2013年的全要素生产率(TFP)。结果显示:中国不同行业大类的资本产出弹性大约处在0.2-0.4之间,劳动产出弹性大约处在0.3-0.6之间,而整个制造业的资本产出弹性为0.3159,劳动产出弹性为0.4071;所有行业大类的劳动产出弹性均大于其资本产出弹性;所有行业大类的TFP增速均大于零,最快的行业达18.94%(废弃资源和废旧材料回收加工业),而最慢的行业仅为3.90%(烟草制品业);比较意外的是,通信设备、计算机及其他电子设备制造业的TFP增速相对较低,但是其TFP绝对值相对较高。中国的各行业间的TFP差距有所缩小,呈现一定的收敛趋势。从产业链视角来看,中国的行业存在上下游TFP集聚现象,尤其是与下游行业关系更为密切,整体呈现出需求为导向的TFP联动。从地区分布来看,中国各省的TFP在全国层面不具有显性特点,但在局部小区域具有较强的集聚现象。从时间维度来看,中国企业的TFP整体呈波动增长趋势,且企业间的TFP差距呈缩小收敛趋势。

参考文献:

- [1] Bai C. E., J. Lu and Z. Tao 2009, "How Does Privatization Work in China?" *Journal of Comparative Economics*, 37(3): 453-470.
- [2] Blundell R. and S. Bond, 2000, "GMM Estimation with Persistent Panel Data: An Application to Production Functions," *Econometric Reviews*, 19(3): 321-340.
- [3] Blundell R. and S. Bond, 1998, "Initial Conditions and Moment Restrictions in Dynamic Panel Data Models," *Journal of Econometrics*, 87(1): 115-143.
- [4] Brandt L., J. Van Biesebroeck and Y. Zhang, 2012, "Creative Accounting or Creative Destruction? Firm-level Productivity Growth in Chinese Manufacturing," *Journal of Development Economics*, 97(2): 339-351.
- [5] Cai H. and Q. Liu, 2009, "Competition and Corporate Tax Avoidance: Evidence from Chinese Industrial Firms," *Economic Journal*, 119(537): 764-795.
- [6] il Kim K., A. Petrin and S. Song, 2016, "Estimating Production Functions with Control Functions when Capital is Measured with Error," *Journal of Econometrics*, 190(2): 267-279.
- [7] Levinsohn J. and A. Petrin, 2003, "Estimating Production Function Using Inputs to Control for Observables," *Review of Economic Studies*, 70(2): 317-341.
- [8] Del Gatto M., A. Di Liberto and C. Petraglia, 2011, "Measuring Productivity," *Journal of Economic Surveys*, 25(5): 952-1008.
- [9] Olley G. S. and A. Pakes, 1996, "The Dynamics of Productivity in the Telecommunications Equipment Industry," *Econometrica: Journal of the Econometric Society*, 64(6): 1263-1297.
- [10] Solow R. M., 1956, "A Contribution to the Theory of Economic Growth," *The Quarterly Journal of Economics*, 70(1): 65-94.
- [11] 盖庆恩、朱喜、程名望、史清华, 2015, "要素市场扭曲、垄断势力与全要素生产率", 《经济研究》, 第5期, 第61-75页。
- [12] 国务院发展研究中心《进一步化解产能过剩的政策研究》课题组、赵昌文、许召元、袁东、廖博, 2015, "当前我国

- 产能过剩的特征、风险及对策研究——基于实地调研及微观数据的分析’，《管理世界》第4期，第1-10页。
- [13] 李玉红、王皓、郑玉歆 2008，“企业演化：中国工业生产率增长的重要途径”，《经济研究》第6期，第12-24页。
- [14] 刘啟仁、黄建忠 2015，“异质出口倾向、学习效应与‘低加成率陷阱’”，《经济研究》第12期，第143-157页。
- [15] 鲁晓东、连玉君 2012，“中国工业企业全要素生产率估计：1999-2007”，《经济学季刊》第2期，第541-558页。
- [16] 罗来军、蒋承、王亚章 2016，“融资歧视、市场扭曲与利润迷失——兼议虚拟经济对实体经济的影响”，《经济研究》第4期，第74-88页。
- [17] 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，第142-158页。
- [18] 钱学锋、王胜、陈勇兵 2013，“中国的多产品出口企业及其产品范围事实与解释”，《管理世界》第1期，第9-27页。
- [19] 王贵东 2017a，“中国制造业企业的垄断行为：寻租型还是创新型”，《中国工业经济》第3期，第83-100页。
- [20] 王贵东 2017b，“中国城镇化发展的理论与实证研究——以生产效率、土地配置为视角”，天津：南开大学经济学院。
- [21] 王贵东、周京奎 2017，“中国制造业企业垄断势力测度——兼论市场边界”，《经济评论》第4期，第30-44页。
- [22] 王小龙、方金金 2015，“财政‘省直管县’改革与基层政府税收竞争”，《经济研究》第11期，第79-93页。
- [23] 谢千里、罗斯基、张轶凡 2008，“中国工业生产率的增长与收敛”，《经济学季刊》第3期，第809-826页。
- [24] 杨汝岱 2015，“中国制造业企业全要素生产率研究”，《经济研究》第2期，第61-74页。
- [25] 张杰、周晓艳、李勇 2011，“要素市场扭曲抑制了中国企业 R&D”，《经济研究》第8期，第78-91页。

Calculation on the TFP of Manufacturing Enterprises in China during 1996-2013

WANG Guidong

Peking University, Beijing, 100871

Abstract: Through systematically arranging the Chinese Industrial Enterprises Database, the article calculates the TFP of Manufacturing Enterprises in China during 1996-2013 in term of OP, LP, GMM, and FE method. The result shows that 30 industries' capital-output elasticities are located in the 0.2-0.4 range, labor-output elasticities are located in the 0.3-0.6 range, the TFP's growing rates range from 3.90% to 18.94%, and the industrial TFP gaps present a certain trend of convergence. From the point of industry chain, TFP exists agglomeration phenomenon, especially between the industry itself and its downstream industry. From the point of regional distribution, the provincial TFP has no dominant characteristics in the whole national level, but has strong agglomeration phenomenon in local area. From the point of time dimension, the enterprises' TFP presents a volatility rising trend, and the enterprises' TFP gaps present a certain trend of convergence.

Key Words: TFP; the Chinese industrial enterprises database; cross-identification

(责任编辑:靳涛) [校对:张靖]