

1. 泰尔指数的测算

$$\begin{aligned}
 T &= \frac{1}{N_i} \sum_{k=1}^N \frac{Y_{ijk}}{\bar{Y}_i} \log\left(\frac{Y_{ijk}}{\bar{Y}_i}\right) \\
 &= T_b + T_w \\
 &= \sum_{j=1}^J Y_{ij} \log\left(\frac{Y_{ij}}{N_{ij} / N_i}\right) + \sum_{j=1}^J Y_{ij} \left(\sum_{K \in j} \frac{Y_{ijk}}{Y_{ij}} \log\left(\frac{Y_{ijk} / Y_{ij}}{1 / N_{ij}}\right) \right)
 \end{aligned} \quad (1)$$

其中, T 为省内总泰尔指数, T_b 为省内地级市之间泰尔指数 (四个直辖市该值取 0), T_w 为省内地级市内部泰尔指数。 Y_{ijk} 为 i 省 j 市 k 县的人均实际 GDP 水平, 相应地, Y_{ij} 和 $\bar{Y}_i$ 分别代表 i 省 j 市的人均实际 GDP 和 i 省的人均实际 GDP 的平均水平; N_{ij} 代表 i 省 j 市中县级行政单位的个数, N_i 表示 i 省县级行政单位的个数。

2. 广义熵指数的测算

$$GE(\bar{\alpha}) = \frac{1}{N\bar{\alpha}(\bar{\alpha}-1)} \sum_{i=1}^N \left(\left(\frac{y_i}{\bar{y}} \right)^{\bar{\alpha}} - 1 \right) \quad (2)$$

其中, N 为组内包含的地区数, y_i 为地区经济发展水平, $\bar{y}$ 为平均值。该指数的大小与 α 的取值有关, α 越大, GE 指数对收入分布的上尾部越敏感, 反之, 则对下尾部较敏感。通常 α 的取值为 0 或者 1, 泰尔指数是广义熵指数的特例, 常用的泰尔指数又名泰尔 T 指数 (如公式 (1) 所示), 是 α 取值为 1 的情况, 本部分计算了 α 取值为 0 的 GE 指数, 又称为泰尔 L 指数。如公式 (3) 所示。

$$GE(0) = \frac{1}{N} \sum_{i=1}^N \ln\left(\frac{\bar{y}}{y_i}\right) \quad (3)$$

3. 阿特金森指数的测算

$$A_{\check{r}} = 1 - \left(\frac{1}{n} \sum_{i=1}^n \left(\frac{y_i^{1-\check{r}}}{\mu} \right)^{\frac{1}{1-\check{r}}} \right) \quad (4)$$

其中, n 为分组内包含的地区数量, y_i 为第 i 组的实际经济发展水平, μ 为平均经济发展情况, ε 为不平等厌恶参数, 反映社会对不平等的厌恶 (偏好) 程度, 其值越大, 表示社会对收入相对较低的人群给予更大权重, 其中 ε 取值范围为 $\check{r} \in (0, +\infty)$ 。本文分别计算了 $\check{r}=1$ 和 $\check{r}=2$ 两种情况下的阿特金森指数。并以地级市为组别对该指数进行分解, 其中总指数、组内指数和组间指数的关系如式 (5) 所示。

$$A_{\check{r}}^T = A_{\check{r}}^{between} + A_{\check{r}}^{within} + Residual \quad (5)$$

其中, $A_{\check{r}}^T$ 表示总阿特金森指数, $A_{\check{r}}^{between}$ 表示组间阿特金森指数, $A_{\check{r}}^{within}$ 表示组内指数。

4. 全球夜间灯光数据的处理过程。

全球夜间灯光影像包括 DMSP-OLS 和 SNPP-VIIRS 两套卫星数据, 其中 DMSP-OLS 为年度数据, 时间跨度为 1992~2013 年, 共包含 F10 (1992~1994)、F12 (1994~1999)、F14 (1997~2003)、F15 (2000~2007)、F16 (2004~2009) 和 F18 (2010~2013) 6 台卫星观测数据, 空间分辨率为 1km*1km, 虽然数据时间跨度较大, 但不得不面对不同卫星数据不可比

的问题。SNPP-VIIRS 有日度、月度和年度数据^①, 时间跨度为 2012 年 4 月-至今, 数据时效性好, 空间分辨率达到 0.5km*0.5km, 优于 DMSP-OLS 卫星数据, 并克服了 DMSP-OLS 数据的过饱和问题。基于上述考虑并结合研究需要, 本文接下来提到的灯光数据均指 SNPP-VIIRS 卫星灯光数据。

虽然 SNPP-VIIRS 数据克服了 DMSP-OLS 分辨率低、过饱和的问题, 并充分考虑了原始数据中云层、闪电、月光等的影响, 但其采集的地球影像数据不可避免地包含了异常值、负值、背景噪声和除人工照明外的其他微光现象, 如闪电、杂散光、生物质燃烧和气体燃烧等, 因此需要对原始数据进行处理以消除低质量光源和无关特征的影响 (Elvidge 等, 2017)。本文所用数据为 2012~2020 年 VIIRS 年度夜间灯光数据^②。虽然 NOAA/NGDC 地球观察小组 (下文简称 EOG) 已经充分考虑到了背景噪声的影响并进行了剔除, 但对于气体燃烧并未进行处理。基于此, 本文需要对 VIIRS 年度全球夜间灯光数据做进一步的处理, 具体处理过程如下。

(1) 数据裁剪与重采样。本文研究区域范围为中国大陆境内, 故需要从美国国家地球物理数据中心 (NOAA) 发布的全球夜间灯光影像中裁剪出中国区域夜间灯光影像。具体操作如下, 首先借助 WGS84 地理坐标系下的中国行政区划矢量地图裁剪出全国区域的夜间灯光影像栅格数据; 接着, 将其重投影到 Asia Lambert Conformal Conic 平面投影坐标系以避免影像栅格形变所造成的数据不准确问题; 最后将栅格影像数据重采样到空间分辨率为 500m*500m 的栅格范围内, 得到中国区域省、市、县三级行政单位的年度夜间灯光栅格数据。

(2) 背景噪声和异常值的校正。虽然 VNL_V2 年度数据在月度无云灯光影像的基础上充分过滤了生物质燃烧、极光等的影响, 但对于气体燃烧并未进行处理。Elvidge 等 (2017) 指出气体燃烧在所有地球表面光源中亮度最高, 并且会产生大量耀斑围绕在其周围, 大大影响了灯光数据的稳定性。针对该问题, 本文参考潘竞虎和胡艳兴 (2016) 的做法, 将原始灯光影像中的负值全部赋值为 0, 采用不变目标区域法, 提取出 2013 年 DMSP-OLS 数据灰度值大于 0 的区域作为掩膜, 通过将这部分像元值赋值为 1, 构造掩膜数据与 VIIRS 灯光数据的交互项, 剔除 NPP-VIIRS 中的异常值。

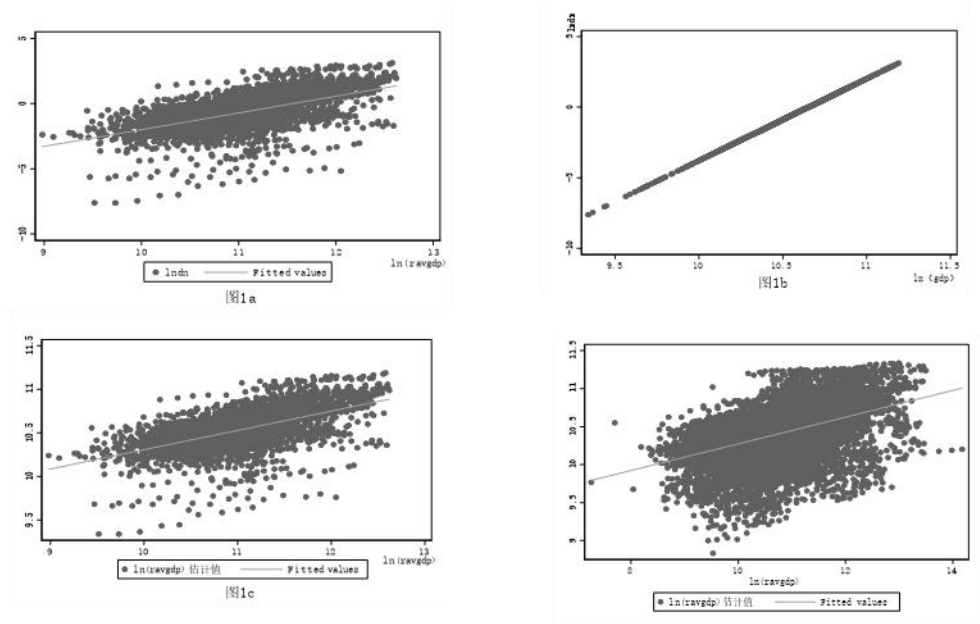
5. 灯光数据与 GDP 指标间的拟合图像

附图 1a 显示人均 GDP 与灯光亮度之间存在对数线性相关关系, 附图 1b 是模型的拟合效果, 从图像可以清晰的看出, 二者均存在明显的线性相关关系, 说明了上述模型设定的合理性; 附图 1c 和附图 1d 分别是市级和县级人均实际 GDP 对数值与灯光数据拟合值之间的

^①2021 年 EOG 刚公布 2012-2020 年的年度 VIIRS 夜间灯光数据。

^②VIIRS 年度夜间灯光数据一共包括 VNL_V1 和 VNL_V2 两种类型数据, 其中 VNL_V1 年度数据是基于 VIIRS-DNB 日度数据合成, 时间截止到 2016 年; VNL_V2 年度数据是在 VIIRS 月度数据的基础上, 过滤了生物质燃烧、极光和背景噪声等影响后得到的, 数据范围为 2012-2020 年。本文采用的数据为年度 VNL_V2 夜间灯光数据。

散点图，可以看出，无论是市级数据还是县级数据，二者均呈现线性相关关系，进一步验证了采用灯光数据作为 GDP 替代指标的合理性。



附图 1 灯光数据与人均实际 GDP 的散点图

6. 数字经济发展水平对地区发展不平衡的异质性检验

(1) 南北差异

附表 1 数字经济发展水平与地区发展不平衡：南北差异						
变量	省总指数		省内地级市间		省内地级市内	
	T_{i_nan}	T_{i_bei}	Tb_{i_nan}	Tb_{i_bei}	Tw_{i_nan}	Tw_{i_bei}
DEI	-0.022	-1.196***	0.074	-0.436*	-0.142	-0.872***
	(0.183)	(0.228)	(0.104)	(0.229)	(0.125)	(0.146)
常数项	5.352**	2.583	2.098	-2.880	3.194**	3.970***
	(2.337)	(3.075)	(1.509)	(5.053)	(1.371)	(0.973)
控制变量	是	是	是	是	是	是
地区固定效应	是	是	是	是	是	是
样本量	135	135	135	135	131	135
R^2	0.671	0.608	0.519	0.058	0.591	0.462
经验 p 值	0.003		0.098		0.005	

注：同表 2，fisher 检验的原假设为 H_0 ：南北区域不存在差异。

(2) 资源禀赋差异

附表 2 数字经济发展水平与地区发展不平衡：资源禀赋差异						
变量	省内总指数		省内地级市间		省内地级市内	
	T_{i_YES}	T_{i_NO}	Tb_{i_YES}	Tb_{i_NO}	Tw_{i_YES}	Tw_{i_NO}
DEI	-0.503	-0.727***	-0.030	-0.349	-0.481**	-0.536***
	(0.437)	(0.217)	(0.231)	(0.213)	(0.182)	(0.151)
常数项	2.467	5.202	1.383	-0.467	4.509**	3.243***
	(3.289)	(3.078)	(1.869)	(3.892)	(1.534)	(1.077)
控制变量	是	是	是	是	是	是
地区固定效应	是	是	是	是	是	是
样本量	63	207	63	207	63	203
R^2	0.565	0.62	0.481	0.062	0.332	0.576

经验 p 值	0.312	0.254	0.487
----------	-------	-------	-------

注：同表 2，fisher 检验的原假设为 H_0 ：资源禀赋是否富裕不存在差异。

7. 数字经济发展水平对地区发展不平衡的异质性检验

(1) 替换地区发展不平衡指标。

附表 3 数字经济发展水平与地区不平衡：替换不平衡指标

变量	广义熵指数 $GE(0)$			阿特金森指数 ($\varepsilon=1$)			阿特金森指数 ($\varepsilon=2$)		
	GE	WGE	BGE	A_ε	WA_ε	BA_ε	A_ε	WA_ε	BA_ε
DEI	-0.007*** (0.002)	-0.005*** (0.002)	-0.002 (0.001)	-0.007*** (0.002)	-0.005*** (0.001)	-0.001 (0.001)	-0.012*** (0.003)	-0.011*** (0.003)	-0.001 (0.002)
常数项	0.049* (0.025)	0.029*** (0.008)	0.020 (0.022)	0.046* (0.024)	0.024*** (0.009)	0.023 (0.020)	0.087* (0.045)	0.043** (0.017)	0.047 (0.039)
控制变量	是	是	是	是	是	是	是	是	是
地区固定	是	是	是	是	是	是	是	是	是
样本量	270	270	270	270	270	270	270	270	270
R^2	0.544	0.537	0.334	0.552	0.547	0.309	0.515	0.545	0.250

注：同表 2。

(2) 内生性问题处理：工具变量回归。

附表 4 数字经济发展水平与区域不平衡：工具变量回归

变量	第一阶段	第二阶段		
	DEI	省内总指数 (T_i)	省内地级市间 (Tb_i)	省内地级市内 (Tw_i)
IV	0.778*** (0.071)			
DEI		-5.126*** (1.206)	-3.719*** (1.393)	-1.741* (0.931)
常数项	-0.656 (0.663)	17.184 (12.543)	21.468* (11.927)	-5.929 (4.425)
控制变量	是	是	是	是
地区固定效应	是	是	是	是
样本量	270	270	270	270
R^2		0.511	0.390	0.315

注：同表 2。

(3) 删除直辖市。

附表 5 数字经济与地区不平衡：删除直辖市

变量	省内总指数	省内地级市间	省内地级市内
	T_i	Tb_i	Tw_i
DEI	-0.814*** (-3.56)	-0.514** (-2.21)	-0.428*** (-3.32)
常数项	3.851 (1.52)	-0.102 (-0.03)	3.177*** (2.91)
控制变量	控制	控制	控制
地区固定效应	控制	控制	控制
样本量	234	234	230
R^2	0.593	0.066	0.432

注：同表 2

8. 数字经济发展速度对地区经济发展不平衡的异质性检验

(1) 南北差异

附表 6 数字经济发展速度与地区经济发展不平衡：南北差异

变量	省总指数		省内地级市间		省内地级市内	
	<i>T_{i_nan}</i>	<i>T_{i_bei}</i>	<i>Tb_{i_nan}</i>	<i>Tb_{i_bei}</i>	<i>Tw_{i_nan}</i>	<i>Tw_{i_bei}</i>
数字经济发展速度	-1.180** (0.468)	-3.132*** (0.503)	-0.419** (0.171)	-0.204** (0.087)	-0.638* (0.321)	-2.731*** (0.813)
常数项	5.369** (2.148)	5.272** (2.407)	1.353 (1.002)	-0.396 (0.535)	2.865* (1.345)	3.956*** (1.236)
控制变量	是	是	是	是	是	是
地区固定效应	是	是	是	是	是	是
样本量	120	120	104	104	120	120
R ²	0.649	0.634	0.434	0.332	0.566	0.51
经验 <i>P</i> 值	0.01		0.126		0.019	

注：同表 2。

(2) 资源禀赋差异

附表 7 数字经济发展速度与地区发展不平衡：资源禀赋差异

变量	省总指数		省内地级市间		省内地级市内	
	<i>T_{i_YES}</i>	<i>T_{i_NO}</i>	<i>Tb_{i_YES}</i>	<i>Tb_{i_NO}</i>	<i>Tw_{i_YES}</i>	<i>Tw_{i_NO}</i>
数字经济发展速度	-2.669*** (0.572)	-1.979*** (0.459)	-0.192 (0.209)	-0.324** (0.138)	-3.805* (1.829)	-1.106*** (0.263)
常数项	2.765 (0.950)	6.658** (2.280)	-0.023 (0.767)	0.668 (0.534)	3.370 (2.340)	3.127** (1.154)
控制变量	是	是	是	是	是	是
地区固定效应	是	是	是	是	是	是
样本量	56	184	56	152	56	184
R ²	0.626	0.612	0.540	0.379	0.472	0.538
经验 <i>P</i> 值	0.229		0.265		0.011	

注：同表 2。