

附录 1

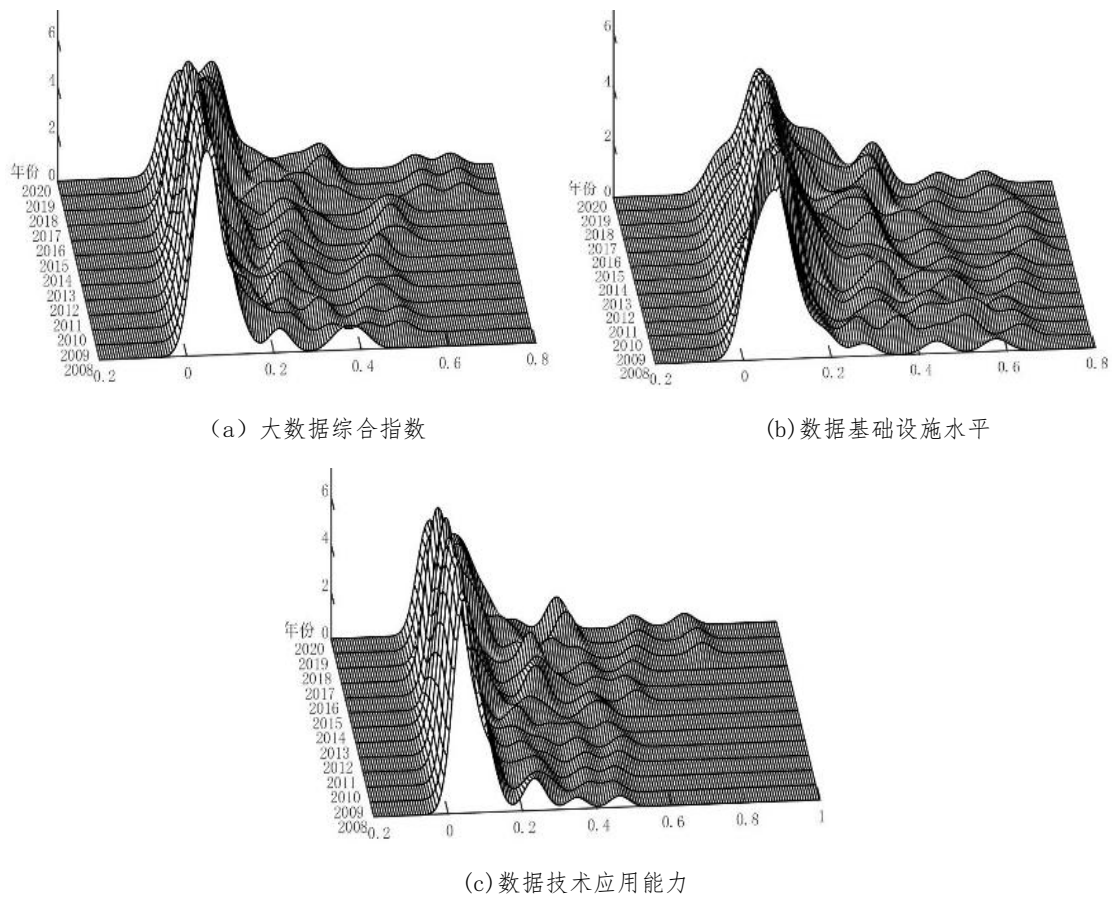
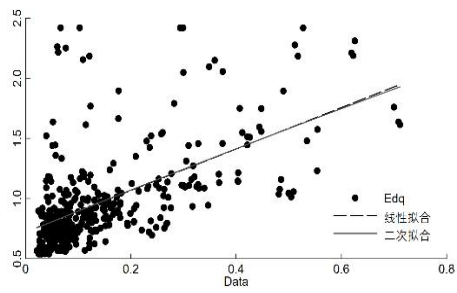
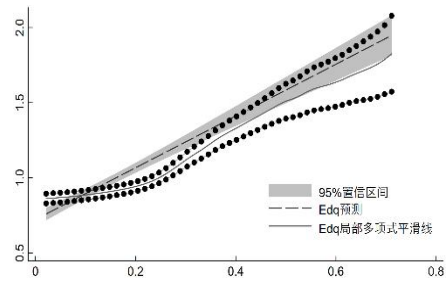
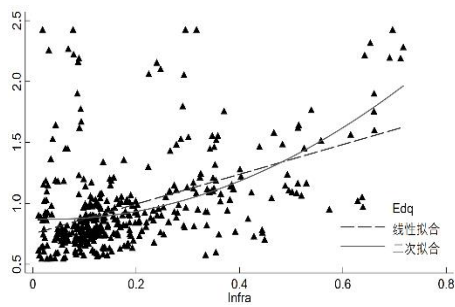
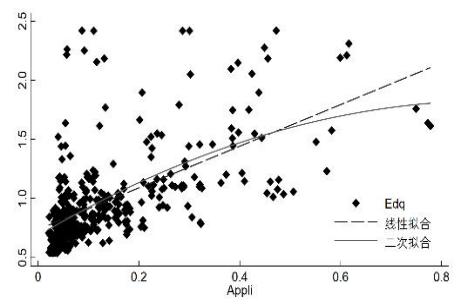


图 1 各地区大数据发展水平立体图

附录 2

图2 D^c 与 Edq 散点拟合图图3 D^c 与 Edq 多项式拟合图图4 D^I 与 Edq 散点拟合图图5 D^A 与 Edq 散点拟合图

附录 3

借鉴 Lesage 和 Pace(2008)的方法,对大数据影响高质量发展的空间效应估计结果进行分解,在此,展示 SDM 模型的分解过程,如式(1)一(6)所示。将 SDM 模型的回归方程进行偏微分,转化为式(1)的形式:

$$(I_N - \rho W) Edq_i = \tau_N \beta'_0 + \theta W X_M + \beta_M X_M + \varepsilon_i \quad (1)$$

式(1)中 I_N 为 N 阶单位矩阵; τ_N 为 N 行 1 列所有元素为 1 的矩阵; W 为空间权重矩阵,代表 W^G 、 W^E 和 W^T ; ρ 为被解释变量空间滞后项系数; θ 为解释变量空间交互项系数; β_0 为常数项, β_M 为解释变量系数; X_M 为第 M 个解释变量。

$$\text{令 } T(W) = (I_N - \rho W)^{-1}, \quad C(W) = T(W) \cdot (I_N \beta_M + \theta W)$$

则式(1)可以转化为式(2):

$$Edq_i = \sum_{M=1}^K C_M(W) Data_M + T(W) \tau_N \beta'_0 + T(W) \varepsilon \quad (2)$$

将式(2)转化为矩阵形式,如式(3)所示:

$$\begin{bmatrix} Edq_1 \\ Edq_2 \\ \vdots \\ Edq_{N-1} \\ Edq_N \end{bmatrix} = \sum_{M=1}^K \begin{bmatrix} C_M(W)_{11} & C_M(W)_{12} & \cdots & \cdots & C_M(W)_{1N} \\ C_M(W)_{21} & C_M(W)_{22} & \cdots & \cdots & C_M(W)_{2N} \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ C_M(W)_{(N-1)1} & C_M(W)_{(N-1)2} & \cdots & \cdots & C_M(W)_{(N-1)N} \\ C_M(W)_{N1} & C_M(W)_{N2} & \cdots & \cdots & C_M(W)_{NN} \end{bmatrix} \begin{bmatrix} Data_{1M} \\ Data_{2M} \\ \vdots \\ Data_{(N-1)M} \\ Data_{NM} \end{bmatrix} + T(W) (\tau_N \beta'_0 + \varepsilon) \quad (3)$$

式(3)中等式右边第一个矩阵为 Lesage 和 Pace(2008)提出的偏微分矩阵,其对角线上的元素表示了特定空间单元 X_{im} 对本单元内被解释变量的直接影响效应,可表示为公式(4)的形式:

$$Direct = \frac{\partial Edq_i}{\partial X_{im}} = C_M(W)_{ii} \quad (4)$$

非对角线上的元素反映了特定空间单元里 X_{im} 对其他空间单元被解释变量的平均影响,即空间效应,也可表示为式(5):

$$Indirect = \frac{\partial Edq_i}{\partial X_{jm}} = C_M(W)_{ij} \quad (5)$$

右侧第一个矩阵所有元素之和即为解释变量 X_{im} 变动所造成的本单元内直接影响和其他空间单元间接影响的总和,即总效应,可表示为式(6):

$$Total = C_M(W)_{ii} + C_M(W)_{ij} \quad (6)$$

SAR 模型的分解过程,如式(7)所示。将 SAR 模型的回归方程进行偏微分,转化为式(7)的形式:

$$(I_N - \rho W)Edq_i = \tau_N \beta'_0 + \beta_M X_M + \varepsilon_i \quad (7)$$

其中各符号含义与式（1）相同，但不同的是，

$$C(W) = T(W) \cdot I_N \beta_M$$

后续分解过程如式（2）—（6）所示。

附录 4

表 1 大数据影响经济发展质量的基准回归稳健性检验结果

变量	(1)	(2)	(3)	(4)
D^C	1.4759* (0.7651)			
D^I		1.2152** (0.4763)		0.8021*** (0.3103)
D^A			1.2652* (0.7279)	1.1237* (0.6228)
控制变量	是	是	是	是
C	0.8629*** (0.1908)	0.9932*** (0.1465)	0.9405*** (0.1745)	0.8878*** (0.1489)
R^2	0.1101	0.0627	0.1007	0.0705
F	4.3523 [0.0007]	9.2917 [0.0000]	4.0776 [0.0013]	4.6312 [0.0001]

注：***、**、*分别表示在 1%、5%和 10%的显著性水平上显著，括号内为标准误，中括号内为 P 值，下同。

表 2 大数据影响经济发展质量的非线性效应稳健性检验

变量		(1) D^C	(2) D^I	(3) D^A	(4) D^I/D^A
门槛值	$q1$	0.0453**	0.0215**	0.0475***	0.0214**
	$q2$	0.2529**	0.0808*	0.2080*	0.1240***
$D \leq q1$		-2.8828 (1.8642)	-2.4166 (2.3462)	0.1453*** (0.0336)	0.2603* (0.1370)
$q1 < D \leq q2$		1.5703** (0.5972)	2.3028*** (0.8195)	0.8566** (0.3220)	0.9001*** (0.1582)
$D > q2$		2.7005*** (0.6452)	1.6164*** (0.5857)	2.7466*** (0.4679)	3.4994*** (1.1822)
R^2		0.0952	0.1045	0.1085	0.1125

注： D^C 、 D^I 、 D^A 表示以大数据综合发展水平、大数据基础设施水平以及大数据技术应用能力为门槛变量， D^I/D^A 表示以大数据基础设施为门槛变量时，大数据技术应用能力的作用估计。

表 3 PSTR 模型下大数据影响经济发展质量的非线性效应稳健性检验

指标	(1) D^C	(2) D^I	(3) D^A	(4) D^I/D^A
转换函数个数 r	1	1	2	2

位置参数个数 m	1	1	2	2
β_0	3.9149*** (1.2770)	0.9653*** (0.1245)	1.3224*** (0.3580)	0.6734* (0.3742)
β_1	-2.2877* (1.3064)	-0.5556** (0.2825)	1.3174*** (0.3930)	2.5434*** (0.4203)
β_2			1.4552*** (0.5039)	7.3498*** (1.6025)
位置参数 c_1	0.1119	0.0938	0.0421	0.0908
位置参数 c_2			0.3232	0.1522
平滑参数 γ_1	3.0905	5.0672	1.0267	1.8027
平滑参数 γ_2			4.4442	7.0601
RSS	15.9992	15.4274	16.2039	12.4523
AIC	-3.1337	-3.1548	-3.1138	-3.2864
BIC	-3.0151	-3.0172	-2.9853	-3.1376

注： D^C 、 D^I 、 D^A 表示以大数据综合发展水平、大数据基础设施水平以及大数据技术应用能力为门槛变量， D^I/D^A 表示以大数据基础设施为门槛变量时，大数据技术应用能力的作用估计。 β_0 为线性部分的参数， β_1 与 β_2 为非线性部分的参数。

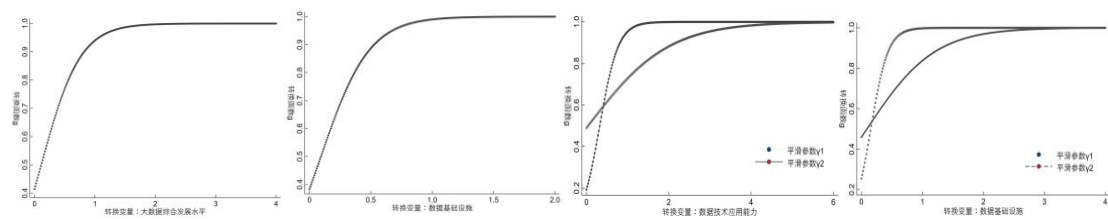


图 6 PSTR 模型估计结果的转换函数图

表 4

大数据影响经济发展质量的空间效应稳健性检验

	大数据综合水平			数据基础设施水平			数据技术应用能力		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	地理距离	经济距离	技术距离	地理距离	经济距离	技术距离	地理距离	经济距离	技术距离
D^C	0.4375* (0.2457)	0.8715*** (0.2585)	0.7740*** (0.2513)						
D^I				0.7991*** (0.1262)	0.1451 (0.1444)	0.2028 (0.1477)			
D^A							0.3726* (0.2119)	0.7214*** (0.2241)	0.6245*** (0.2175)
控制变量	是	是	是	是	是	是	是	是	是
ρ	3.8967 (4.5384)	0.0017 (0.0015)	0.0431*** (0.0085)	5.6855*** (1.7313)	-0.0015 (0.0016)	0.0414*** (0.0123)	3.6744 (4.5360)	0.0021 (0.0006)	0.0436*** (0.0085)
$W \times D^C$	0.3558 (0.7606)								
$W \times D^I$					1.1999** (0.4955)	0.8879** (0.3597)			
$W \times D^A$							0.6348 (0.6490)		
$W \times$ 控制变量	是	是	是	是	是	是	是	是	是
C	1.1058*** (0.3185)	0.6962*** (0.1262)	0.8514*** (0.1166)	0.9100*** (0.0665)	1.2190*** (0.1407)	0.8908*** (0.2388)	1.1305*** (0.3055)	0.7988*** (0.0979)	0.9443*** (0.0994)
R^2	0.9126	0.8962	0.9025	0.9061	0.9388	0.9030	0.9129	0.8959	0.9022
Sargan	1.968 [0.5790]	2.593 [0.4587]	2.263 [0.5196]	3.402 [0.3336]	1.467 [0.6899]	4.387 [0.2226]	2.926 [0.4032]	3.244 [0.3555]	3.183 [0.3643]

附录 5

表 5 大数据综合试验区设立对经济高质量发展的冲击效应估计结果^①

变量	不含空间因素		含空间因素		
			地理距离	经济距离	技术距离
<i>DID</i>	0.3100*** (0.0369)	0.0867** (0.0336)	0.1078*** (0.0305)	0.1100*** (0.0369)	0.1526*** (0.0335)
控制变量	否	是	是	是	是
个体(时间)效应	是	是	是	是	是
常数项	0.9477*** (0.0092)	0.4634*** (0.0672)			
ρ	否	否	0.1700 (0.1341)	0.1840** (0.0747)	0.4300*** (0.0631)
<i>SDID</i>	否	否	0.3111 (0.2092)	-0.0083 (0.0820)	
<i>W</i> *控制变量	否	否	是	是	否
总效应	否	否	0.5146* (0.2655)	0.1286 (0.0878)	0.2727*** (0.0665)
直接效应	否	否	0.1114*** (0.0317)	0.1110*** (0.0351)	0.1593*** (0.0348)
间接效应	否	否	0.4032 (0.2587)	0.0175 (0.0960)	0.1134*** (0.0386)
R^2	0.1597	0.4911	0.9027	0.9199	0.8782

表 6 “宽带中国”试点政策对经济高质量发展的冲击效应估计结果^②

变量	不含空间因素		含空间因素		
			地理距离	经济距离	技术距离
<i>DID</i>	0.2368*** (0.0147)	0.1201*** (0.0212)	0.0776*** (0.0227)	0.0471 (0.0396)	0.0524 (0.0386)
控制变量	否	是	是	是	是
个体(时间)效应	是	是	是	是	是
常数项	0.8550*** (0.0103)	0.6770*** (0.0769)			
ρ	否	否	0.6480*** (0.0619)	0.1990*** (0.0742)	-0.1600* (0.0903)
<i>SDID</i>	否	否		-0.0667 (0.1079)	-0.1098 (0.0944)
<i>W</i> *控制变量	否	否	否	是	是
总效应	否	否	0.2232*** (0.0591)	-0.0250 (0.1482)	-0.0517 (0.0858)
直接效应	否	否	0.0821*** (0.0230)	0.0442 (0.0397)	0.0560 (0.0382)

① 2016 年我国分两批设立了八大国家级大数据综合试验区，其中，包括贵州、京津冀、珠三角、重庆、河南、上海、内蒙古和沈阳，但由于沈阳为地级市、内蒙古发展目标在于提升大数据基础设施水平，故为保障处理组及对照组设置合理，在样本中剔除掉了辽宁和内蒙古，并将在 2016 年 2 月获批大数据综合试验区资格的贵州政策实施时间设立在 2016 年，而将其余在 2016 年 10 月获批大数据综合试验区资格的其他省份政策处理时间设置在 2017 年。

② 由于“宽带中国”试点地区为地级市，而本文的样本为省级地区，故使用地级市数量加总的方式匹配到省级地区进行检验。

间接效应	否	否	0.1411*** (0.0427)	-0.0692 (0.1328)	-0.1077 (0.0850)
R ²	0.4116	0.5236	0.9134	0.8988	0.9037

表7 “珠峰计划”试点对经济高质量发展的冲击效应估计结果^①

变量	不含空间因素		含空间因素		
			地理距离	经济距离	技术距离
<i>DID</i>	0.2720*** (0.0246)	0.1573*** (0.0251)	0.2041*** (0.0421)	0.0467 (0.0289)	0.1377*** (0.0246)
控制变量	否	是	是	是	是
个体(时间)效应	是	是	是	是	是
常数项	0.9069*** (0.0102)	0.1885*** (0.0923)			
ρ	否	否	0.0200 (0.1471)	0.1920*** (0.0735)	0.3910*** (0.0662)
<i>SDID</i>	否	否	0.2146 (0.2855)	0.2696*** (0.0790)	
<i>W</i> *控制变量	否	否	是	是	否
总效应	否	否	0.4348 (0.3075)	0.3930*** (0.0944)	0.2265*** (0.0407)
直接效应	否	否	0.2039*** (0.0413)	0.0581* (0.0289)	0.1411*** (0.0243)
间接效应	否	否	0.2309 (0.3033)	0.3349*** (0.0943)	0.0854*** (0.0241)
R ²	0.2476	0.4084	0.4797	0.9039	0.8812

附录6

为分析要素配置 $Faal$ 的中介作用，借鉴并修正白俊红和刘宇英(2018)的做法，测算资本配置指数 F_u^C 和劳动力配置指数 F_u^L ，具体测算过程如式(8)—(11)所示：

$$\gamma_u^C = \frac{I}{I + \tau_u^C}, \gamma_u^L = \frac{I}{I + \tau_u^L} \quad (8)$$

式(8)中， γ_u^C 和 γ_u^L 分别为 i 省份 t 时期资本和劳动要素价格绝对扭曲系数， τ_u^C 和 τ_u^L 分别为资本与劳动的错配系数，在此，使用资本与劳动的相对价格扭曲系数 $\hat{\gamma}_u^C$ 和 $\hat{\gamma}_u^L$ 替代绝对扭曲系数，其测算公式如式(9)所示：

$$\hat{\gamma}_u^C = \left(\frac{C_u}{C_t} \right) / \left(\frac{s_u \beta_u^C}{\beta_t^C} \right), \hat{\gamma}_u^L = \left(\frac{L_u}{L_t} \right) / \left(\frac{s_u \beta_u^L}{\beta_t^L} \right) \quad (9)$$

式(9)中， $s_u = \frac{y_u}{Y_t}$ 表示省份 i 在 t 期的产出 y_u 占 t 期总产出 Y_t 的份额， $\beta_t^C = \sum_{i=1}^N s_{it} \beta_i^C$ ， $\beta_t^L = \sum_{i=1}^N s_{it} \beta_i^L$ 分别表示产出加权的资本、劳动贡献值，其中， β_u^C 和 β_u^L 分别为资本与劳动的

① 第一批的试点高校包括：北京大学、清华大学、北京师范大学、南开大学、吉林大学、复旦大学、上海交通大学、南京大学、中国科学技术大学、浙江大学、厦门大学、山东大学、武汉大学、中山大学、四川大学、西安交通大学、兰州大学。

产出弹性。 $\frac{C_{it}}{C_t}$ 表示 i 省份 t 期使用的资本占 t 期资本总量的实际比例， $\frac{L_{it}}{L_t}$ 表示 i 省份 t 期投入的劳动力占 t 期劳动力总量的实际比例，而 $\frac{s_{it}\beta_{it}^c}{\beta_{it}^c}$ 和 $\frac{s_{it}\beta_{it}^L}{\beta_{it}^L}$ 表示资本与劳动有效配置时的理论比例。

进一步，令资本的要素产出弹性 β_{it}^c 为 0.45、劳动的要素产出弹性 β_{it}^L 为 0.55。以 2011 年为基期平减计算产出变量 y_{it} ；用各省份的年末就业人数表征劳动力投入量 L_{it} ；用各省份的固定资本存量表征资本投入量 C_{it} ，并使用永续盘存法计算，如式(10)所示：

$$C_{it} = C_{it-1}(1-\delta) + I_{it} \quad (10)$$

式(10)中，资本存量以 2008 年为基期， δ 代表折旧率，令其为 9.6%， I_{it} 表示各省份固定资产投资， C_{it-1} 表示上一期固定资本存量。而后根据式(8)、(9)计算出各省份的资本与劳动力错配系数 τ_{it}^c 和 τ_{it}^L ，代入式(8)中求得资本配置指数(F_{it}^c)与劳动力配置指数(F_{it}^L)：

$$F_{it}^c = \frac{1}{|\tau_{it}^c|}, \quad F_{it}^L = \frac{1}{|\tau_{it}^L|} \quad (11)$$