

《科研管理》一论文被指“数据存疑”

近日，一封“读者来信”对发表于《科研管理》的一篇文章 [1] 的数据提出了疑问。该文章采用多元回归模型，考察 93 家汽车制造业上市公司数字化建设（以“数字技术领域投资金额”等指标衡量）对两类创新行文的影响，包括探索式创新（以“对应 IPC 分类号未在该企业五年内出现过的专利申请”等指标衡量）以及利用式创新（以“对应 IPC 分类号已在该企业五年内出现过的专利申请”等指标衡量）。

10.19571/j.cnki.1000-2995.2024.12.015

表 2 企业数字化对探索式创新和利用式创新的影响
Table 2 Impact of digitalization on explorative innovation and exploitative innovation of enterprises

变量	Model 1	Model 3	Model 4	Model 5
Digi		0.277*** (0.293)	0.416** (0.371)	
Digi2				-0.399*** (0.291)
Top10share	-0.068** (0.009)	-0.055** (0.008)	-0.477 (0.007)	-0.463 (0.006)
Bdnum	0.090** (0.010)	0.102** (0.012)	-0.029 (0.011)	-0.030 (0.012)
FirmAge	-0.012 (0.013)	-0.027** (0.017)	-0.120*** (0.016)	-0.119*** (0.017)
Roe	0.036 (0.013)	0.064 (0.015)	-0.038** (0.013)	-0.048** (0.015)
Cash	-0.044 (0.010)	-0.066 (0.008)	-0.056*** (0.007)	-0.033*** (0.009)
Banklev	0.150 (0.015)	-0.141 (0.013)	-0.847*** (0.016)	-0.827*** (0.015)
Size	0.424*** (0.012)	0.078 (0.015)	0.397*** (0.018)	0.423*** (0.017)
Constant	-1.093* (0.456)	-3.529*** (0.449)	0.756 (0.346)	0.920 (0.455)
Observations	957	957	957	957
χ^2	409 831.23	406 001.89	408 001.31	399 073.91
Log likelihood	-588 100.19	-575 991.03	-576 009.33	-556 903.38

注：*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$, 括号内是聚类稳健标准误，回归模型中会时间虚拟变量。

表 3 联盟伙伴多样性的调节作用
Table 3 Moderating effect of the diversity of alliance partners

变量	Model 5	Model 6
Digi	0.231*** (0.218)	0.345*** (0.257)
Digi* Diversity	0.113*** (0.266)	-0.105** (0.176)
Digi2		-0.166*** (0.211)
Digi2* Diversity		0.011** (0.102)
Top10share	-0.010 (0.007)	-0.001 (0.006)
Bdnum	0.004*** (0.010)	-0.001*** (0.011)
FirmAge	0.008 (0.003)	-0.005 (0.002)
Roe	0.065 (0.009)	0.002 (0.005)
Cash	0.325 (0.013)	-0.002 (0.005)
Banklev	0.394 (0.011)	-0.004* (0.014)
Size	-0.052 (0.031)	0.001 (0.020)
Constant	-7.858*** (0.718)	4.218*** (0.326)
Observations	727	727
χ^2	399 018.93	398 077.68
Log likelihood	-556 881.91	-548 780.07

注：*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$, 括号内是聚类稳健标准误，回归模型中会时间虚拟变量。

论文 [1] 指出“企业数字化对应的回归系数在 1% 水平下显著为正”；然而，这个说法遭受了质疑。来函信件指出，模型得到的回归系数 (0.277) 与 0 之间的差值不到一个标准误差 (0.293)，并不能支持“回归系数与 0 间具有统计学意义的显著差异”的结论。论文 [1] 中其它模型的回归系数也存在着类似的问题。例如，模型 5 (Model 5) 中 Digi 变量的回归系数 (0.231) 与 0 之间的差值也仅仅略大于一个标准误差 (0.218)，Digi* Diversity 变量的回归系数 (0.113) 与 0 之间的差值不足一个标准误差 (0.266)。模型 6 (Model 6) 的回归系数也出现了类似情况。

这些数据都表明，论文 [1] 作者所使用的数据并不能支持他们的结论；但期刊以及审稿人在同行评议的过程中并未发现/指出上述不足，反映着当前期刊审稿质量的欠缺。同行评议的质量被认为保证学术记录可靠的重要保障，而这个保障正在面临日益增长的质疑。

目前尚不清楚论文 [1] 作者对回归系数显著性的错误标注是“刻意为之”，还是因为对数据理解的不足而造成。限于作者没有公开所使用的数据集，目前也没法判断论文所使用模型的计算是否准确。但基于上述存疑的数据，读者在理解论文 [1] 结论的时候需要额外的谨慎。

[1] 10.19571/j.cnki.1000-2995.2024.12.015

This article is licensed to the 5GH Foundation under a CC BY-NC-ND 4.0 International License