

关于收入和自由的因果关系的分析

董天锐

加拿大多伦多大学, 加拿大·安省多伦多 M4Y1M7

【摘要】没有足够的证据表明, 随着收入的增长, 公民会要求更多的民主和自由。虽然这两个变量呈正相关, 通常高收入的国家公民会要求更多的民主与自由, 但它们之间并不存在因果关系。

【关键词】收入; 自由; 线性回归模型; 计量经济学

引言

经济学家们曾经相信“各种生活水平衡量指标的提高预示着民主会逐渐兴起”(罗伯特·保罗, 1999, 160), 但阿西莫格鲁认为“虽然一个国家公民的收入和民主呈正相关, 但没有证据表明两者之间存在必然的因果关系”(2008, 836)。

1 数据概括

我们有147个国家和地区从1950年到2000年(每五年一个时间段)的数据, 总共有1617个观测值。记录的实际人均GDP对数平均值为8.165034, 标准差为1.03723。权利指数(Augmented freedom house political rights index, fhpolrigaug)是阿西莫格鲁使用的用以衡量民主的指标, 它的范围是0到1, 均值为0.5683026, 标准差为0.3577456。

殖民地的人均实际GDP对数平均值为7.963443, 非殖民地为8.748208。殖民地的权利指数的平均值为0.5353503, 而非殖民地为0.6553049。

社会主义国家人均实际GDP对数平均值为8.65161, 非社会主义国家为8.137597。

社会主义国家的权利指数的平均值为0.4183521, 而非社会主义国家为0.5803801。

在高收入国家(人均GDP高于平均水平的国家), 权利指数的平均值为0.7109018, 而在低收入国家仅为0.3865905。

从这些数据中, 我们不难发现, 民主与收入呈高度正相关。通常来说, 收入高的国家权利指数也更高。

详细结果见表1。

2 回归分析

我们做了6种不同的线性回归模型:

2.1 最基础的模型: 权利指数在实际人均GDP对数的一个周期滞后上的SLR, 具有稳健的标准误差

权利指数 $i = \beta_0 + \beta_1$ 实际人均GDP对数 $i + \varepsilon_i$

在其他条件不变的情况下, 实际人均GDP对数值每增加1, 权利指数平均增加0.2254。这是一个很大的影响, 因为权利指数的区间是[0, 1]。因为 $t = 30.91$, p 值等于0.000, 说明我们的结果具有很强的统计显著性, 足以否定原假设($\beta_1 = 0$)。95%置信区间为(0.2111657, 0.2397974), β_1 的真实值有95%的概率在这个区间内。

2.2 权利指数在实际人均GDP对数的一个周期滞后上的SLR, 标准误差聚集在国家层面

与(1)相比, 我们的结果没有发生变化。95%置信区间为(0.2008087, 0.251544), β_1 的真实值有95%的概率在这个区间内。如果FE6运行失败, 则OLS SE存在错误。一个时不变因素是国家的意识形态(社会主义或非社会主义)。社会主义国家的人均实际GDP对数值较高, 但权利指数低于平均水平, 因此可能存在负的遗漏变量偏误。

2.3 权利指数对实际人均GDP对数的一个时期滞后的回归, 包括国家固定效应

权利指数 $i = \beta_0 + \beta_1$ 实际人均GDP对数 $i + \sum$ 国家固定效应 $i + \varepsilon_i$

在其他条件不变的情况下, 实际人均GDP对数值每增加1, 权利指数平均增加0.0550704。影响仍然很大, 由(2)变形而来。 p 值为0.097, 大于0.05, 所以, 这结果在95%置信水平下不具有统计显著性。95%置信区间为(-0.0100479, 0.1201887), β_1 的真实值有95%的概率在这个区间内。一个时变因素可能是一年中发生的重大事件。例如, 在2020年, 新冠疫情爆发, 它对经济造成了巨大的影响, 使人们感到压力、焦虑和两极分化。

2.4 权利指数对实际人均GDP对数的一个时期滞后的回归, 包括国家和年份固定效应

在其他条件不变的情况下, 实际人均GDP对数值每增加1, 权利指数平均增加0.0587276。影响仍然很大。95%置信区间为(-0.0310453, 0.1485006), β_1 的真实值有95%的概率在这个区间内。 F 统计量等于5.78, p 值等于0.0000。我们应该在我们的模型中考虑年固定效应。

2.5 在我们的数据库里一些关于人口的数据是缺失的, 我们将去掉这些缺少部分数据的样本并且重新做一次(4)

估计系数变为-0.0228115, p 值高, 所以我们不能推断出收入与民主指数之间有因果关系。

2.6 使用(5)的数据, 运行权利指数对实际人均GDP对数的一个时期滞后的回归, 包括国家和年份固定效应, 以及滞后的人口变量: 年龄、教育、对数人口。

表1 描述性统计

	(1)	(2)	(3)	(4)	(5)
变量	数量	平均值	标准偏差	最小值	最大值
年	1,617	1,975	15.82	1,950	2,000
年龄_中年	1,551	0.175	0.0298	0.0897	0.298
年龄_老年	1,551	0.118	0.0374	0.056	0.221
年龄_高龄	1,551	0.0883	0.0493	0.0316	0.241
年龄_低龄	1,551	0.363	0.0947	0.143	0.516
年龄_年轻	1,551	0.255	0.0256	0.169	0.347
代码_数字	1,617	101.3	58.37	3	211
教育	722	4.458	2.874	0.042	12.18
权利指数	1,194	0.568	0.358	0	1
劳动力占比	504	0.37	0.131	0.0763	0.738
总人口对数	1,308	8.682	1.907	3.714	14.05
实际人均GDP对数	1,199	8.165	1.037	5.774	10.69
年龄中位数	1,529	23.12	6.664	14.4	41.3
名义存储率	1,206	0.165	0.136	-0.766	0.74
第四制度指数	1,092	0.556	0.379	0	1
社会主义	1,617	0.15	0.357	0	1
年_数字	1,617	36	3.163	31	41
样本	1,617	0.649	0.477	0	1
世界收入	1,331	11.63	8.623	0.589	81.02
世界民主	1,331	0.349	0.251	0.0147	2.27
殖民地	1,617	0.701	0.458	0	1

在其他条件不变的情况下,实际人均GDP对数值每增加1个单位,权利指数平均增加0.0159635。与(4)相比,这种影响相对较小。这是因为我们控制了更多的变量。 p 值为0.765,大于0.05,所以这个结果在95%的置信水平上不具有统计显著性。年固定效应的F统计量等于5.57, p 值等于0.0000,所以我们应该在我们的模型中考虑年固定效应。年固定效应的F统计量等于2.79, p 值等于0.0448。我们应该在我们的模型中考虑年龄固定效应。排除所有的人口控制变量,F统计量为1.48, p 值为0.1929,我们应该添加这些控制变量。在其他条件不变的情况下,处于更早的年份会降低权利指数。例如,1970年,权利指数下降0.1140042,在2000年则上升0.1639049。年龄组的权重也可能对民主有一定的影响,例如中年组的权重每增加1%,增广自由之家权利指数就会下降2.61807。随着教育和人口的增加,增广权利指数也随之下降。详细结果见表2。

3 延伸

第四政策指数(Polity IV Index)是衡量民主的另一个指标。我们可以在设定6中将权利指数替换为第四制度指数,看看是否得到类似的结论。系数为-0.0084042, p 值远大于0.05(0.893),因此我们没有证据表明收入与民主之间没有因果关系。此外,人均GDP并不完全意味着收入水平,因此我们可以在我们的模型中增加一个控制变量:名义储蓄率。将此变量添加到(6)

表 2

	(1)	(2)	(3)	(4)	(5)	(6)
利指数	利指数	利指数	利指数	利指数	利指数	利指数
L. 实际人均GDP对数	0.225***	0.225***	0.0551	0.0587	-0.0228	0.016
	-0.0073	-0.0125	-0.0329	-0.0454	-0.0538	-0.0532
	[0.211, 0.240]	[0.201, 0.250]	[-0.0100, 0.120]	[-0.0310, 0.149]	[-0.130, 0.0840]	[-0.0896, 0.122]
L. 年龄_年轻						-0.509
						-0.63
L. 年龄_中年						-2.618**
						-0.949
L. 年龄_老年						0.452
						-1.413
L. 年龄_高龄						0.848
						-1.39
L. 教育						-0.00351
						-0.0237
L. 总人口对数						-0.104
						-0.133
F统计量和P值				5.78	5	5.57
年固定效应年				0	-0.0001	0
龄控制变量						2.79
人口控制变量						-0.00448
						1.48
						-0.1929
年固定效应国家固定效应	No	No	No	Yes	Yes	Yes
聚类调整标准误	No	No	Yes	Yes	Yes	Yes
人口样本	No	Yes	Yes	Yes	Yes	Yes
	No	No	No	No	Yes	Yes
N	1023	1023	1023	1023	685	685
r ²	0.403	0.403	0.727	0.757	0.726	0.735

参考文献:

[1]Acemoglu, Daron, Simon Johnson, James A. Robinson and Pierre Yared. 2008. “收入与民主”. 美国经济评论, 98(3):

后,我们再次得到高 p 值(0.800),这表明收入与民主之间没有因果关系,这再次验证了我们的结论。

4 结果讨论

beta估计值的大小在开始时相当大,随着控制变量的增加而变小。也就是说,权利指数和收入呈正相关,但并没有直接的因果关系。所有控制变量(国家、年份、人口控制变量)显著,它们的 p 值都等于0,它们的F统计量都高(正如我们在前一部分中讨论的那样)。这些控制变量是收入和民主之间的混杂因素,因此有必要控制它们。

但是我们的结果存在以下两点限制条件:第一,如果我们按地理区域对这些国家进行分组,我们可能会得到更精确的结果,因为地理区域可能是一个同时影响收入和民主的遗漏变量。第二,民主和自由很难衡量。增广自由之家权利指数的大多数数值为1/6的倍数或两位数,而第四制度指数的值为0.05的倍数;这些数据不够精确,可能不准确,因此无法提供足够有价值的信息。

不过我们仍然可以得出结论:正如我们观察和分析的那样,收入与民主之间存在很强的正相关关系,但是随着我们添加越来越多的控制变量,包括国家、年份和人口变量,这种相关性消失了,因此我们不能将其视为因果关系。

回归分析

括号中为标准误差* $p<0.05$, ** $p<0.01$, *** $p<0.001$

808-42. DOI: 10.1257/aer.98.3.808.

[2]Barro, Robert J. 1999. “民主的决定因素”. 政治经济杂志, 107(6): S158-83.