

## 不同收入国家 2000-2016 年婴儿死亡率影响因素的岭回归分析

李鸿斌<sup>1\*</sup>, 贲宇<sup>2</sup>, 冯海娟<sup>3</sup>

<sup>1</sup>如皋市妇幼保健计划生育服务中心儿童保健科 江苏如皋

<sup>2</sup>南京大学医学院临床医学专业 2019 级 江苏南京

<sup>3</sup>南通大学附属如皋医院消化内科 江苏如皋

**【摘要】目的** 分析人均 GDP、人均当前卫生支出等经济社会因素对不同收入国家婴儿死亡率的影响程度。**方法** 以高、中、低收入国家 2000-2016 年婴儿死亡率为因变量, 人均 GDP、人均当前卫生支出、总生育率、小学生总毕业率、谷物产量、CO<sub>2</sub> 排放量等经济社会因素为自变量, 进行岭回归分析, 通过比较标准化回归系数绝对值大小, 分析各自变量对婴儿死亡率的影响大小。**结果** 高(k=0.159, F=93.766, P=0.000, R<sup>2</sup>=0.983)、中(k=0.069, F=224.115, p=0.000, R<sup>2</sup>=0.993)、低(k=0.173, F=123.357, P=0.000, R<sup>2</sup>=0.987)收入国家的岭回归方程 F 检验差异均有统计学意义, 不同收入国家除高收入国家总生育率外(t=0.429, P=0.677)其他回归系数的 t 检验差异均有统计学意义(P<0.05)。高收入国家岭回归方程标准化回归系数绝对值前二位是人均当前卫生支出、人均 GDP, 中等收入国家为总生育率、谷物产量, 低收入国家为总生育率、小学生总毕业率。**结论** 人均 GDP、人均当前卫生支出是高收入国家但不是中、低收入国家婴儿死亡率的主要经济社会影响因素。不同经济社会因素对不同收入国家婴儿死亡率的影响程度大小不一、影响顺位各有差别, 防控策略应各有侧重。

**【关键词】** 婴儿死亡率; 人均 GDP; 人均当前卫生支出; 岭回归分析

**【基金项目】** 南通市妇幼健康专科联盟科研项目(TFM202104)

**【收稿日期】** 2023 年 2 月 17 日 **【出刊日期】** 2023 年 3 月 23 日 **【DOI】** 10.12208/j.ijcr.20230115

### Ridge regression analysis of the influencing factors of infant mortality rate in different income countries from 2000 to 2016

Hongbin Li<sup>1</sup>, Yu Ben<sup>2</sup>, Haijuan Feng<sup>3</sup>

<sup>1</sup>Department of Child Healthcare, Maternal and Child Health Care and Family Planning Service Center of Rugao, Rugao, Jiangsu

<sup>2</sup>Clinical medicine, grade 2019, Medical School of Nanjing University, Nanjing, Jiangsu

<sup>3</sup>Department of Gastroenterology, Affiliated Rugao Hospital of Nantong University, Rugao, Jiangsu

**【Abstract】 Objective** To analyze the impact of economic and social factors such as per capita GDP and per capita current health expenditure on infant mortality in different income countries. **Methods** Taking the infant mortality rate in high, middle and low-income countries from 2000 to 2016 as the dependent variable, and the per capita GDP, per capita current health expenditure, total fertility rate, total graduation rate of primary school students, grain production, and CO<sub>2</sub> emissions as the independent variables, ridge regression analysis was carried out. By comparing the absolute values of standardized regression coefficients, the impact of each variable on infant mortality was analyzed. **Results** There were significant differences in ridge regression equation F test among high (k = 0.159, F= 93.766, P = 0.000, R<sup>2</sup> = 0.983), middle (k = 0.069, F = 224.115, P = 0.000, R<sup>2</sup> = 0.993) and low (k = 0.173, F= 123.357, P = 0.000, R<sup>2</sup> = 0.987) income countries. Except for the total fertility rate of high-income countries in different income countries (t=0.429, p=0.677), the t-test differences of other regression coefficients were statistically significant (P < 0.05). According to the absolute value

\*通讯作者: 李鸿斌

of the standardized regression coefficient of the ridge regression equation, the top two were the current per capita health expenditure and per capita GDP in high-income countries, the total fertility rate and grain output in middle-income countries, and the total fertility rate and the total graduation rate of primary school students in low-income countries. **Conclusions** Per capita GDP and per capita current health expenditure are the main economic and social influencing factors of infant mortality in high-income countries, but not in middle and low-income countries. Different economic and social factors have different degrees of influence on infant mortality in different income countries, and the order of influence is different, so the prevention and control strategies should have their own focus.

**【Keywords】** Infant mortality rate; GDP per capita; Current health expenditure per capita; Ridge regression

随着经济社会的不断发展,全球婴儿死亡率(IMR)持续下降<sup>[1,2]</sup>。婴儿死亡的直接致死原因为疾病,纷繁复杂的经济社会因素在婴儿死亡中也挥着不同程度的作用,是一个涉及社会各领域的复杂系统<sup>[3]</sup>。众多研究发现人均GDP、人均卫生支出与IMR显著相关<sup>[4,5]</sup>,但不同国家和地区在不同发展阶段却存在着差异性表现,2010-2017年美国医疗补助扩展状况与全国IMR变化没有关联<sup>[6]</sup>,另有研究2014-2016年美国医疗补助扩展州的IMR下降幅度更大<sup>[7]</sup>。南亚区域合作联盟国家2000-2016年<sup>[8]</sup>卫生支出、人均GDP对IMR有显著影响,1995-2010年<sup>[9]</sup>医疗支出与IMR无显著关系。人均GDP、人均卫生支出在IMR下降中发挥怎样的作用,值得关注。本文通过检索世界银行数据库,收集高、中、低收入国家2000-2016年以来相关数据,以IMR为因变量,人均GDP、人均当前卫生支出(人均CHE)等经济社会因素为自变量,进行岭回归分析,通过比较标准化回归系数的绝对值大小,判断在IMR下降中的作用,为后疫情时代调整防控策略提供参考依据。

## 1 对象与方法

### 1.1 资料来源

以高、中、低收入国家2000-2016年IMR和相关经济社会影响因素为研究对象,经济社会因素包括人均GDP(US\$)、人均CHE(US\$)、总生育率(女性人均生育数)、谷物产量(Kg/公顷)、小学总毕业率(%)、CO<sub>2</sub>排放量(公吨/人)等6个指标。从世界银行数据库下载数据,直接读取。IMR、人均GDP、人均CHE的数据更新时间为2020年12月16日,其他的均为2021年3月19日。见附表。

### 1.2 研究方法

选择IMR作为因变量(y),以人均GDP、人均CHE等指标为自变量(X<sub>i,j</sub>;j=1,2,...,6),进行岭回归分析。定义: $\hat{\beta}(k) = (X'X + kI)^{-1} X'y$ ,其中,X、y已标准化,则该定义为β的标准化岭回归估计, $X'X$

是自变量样本相关阵, $kI$ 为正常数矩阵,k称为岭参数(k>0), $\hat{\beta}(k)$ 作为β的估计比最小二乘法估计 $\hat{\beta}$ 稳定,通过方差扩大因子法选择合适的k值,使岭回归分析比最小二乘法估计的回归系数有较小的均方误差。岭回归方程的检验包括F分析、拟合优度、回归系数t检验等3个方面。回归系数t检验差异有统计学意义,提示该自变量对因变量有影响作用,标准化回归系数绝对值大者影响大。在SPSSPRO1.0.5统计软件上进行数据处理,自动选择k值。p<0.05差异有统计学意义。

## 2 结果

### 2.1 高收入国家IMR经济社会影响因素的岭回归分析

当k=0.159时,岭回归方程F检验差异有统计学意义(F=93.766,P=0.000),R<sup>2</sup>=0.983,调整R<sup>2</sup>=0.972,高收入国家总生育率的回归系数t检验差异无统计学意义(t=0.429,P=0.677),其他回归系数t检验差异均有统计学意义(P<0.05)。见表1、图1。总生育率、CO<sub>2</sub>排放量的标准化回归系数为正号,其余均为负号。各自变量标准化回归系数的绝对值从大到小依次为人均CHE、人均GDP、CO<sub>2</sub>排放量、小学生总毕业率、谷物产量、总生育率。

### 2.2 中等收入国家IMR经济社会影响因素的岭回归分析

当k=0.069时,岭回归方程F检验差异有统计学意义(F=224.115,P=0.000),R<sup>2</sup>=0.993,调整R<sup>2</sup>=0.988,各自变量回归系数t检验差异均有统计学意义(P<0.05)。见表1、图2。总生育率的标准化回归系数为正号,其余均为负号。各自变量标准化回归系数的绝对值从大到小依次为总生育率、谷物产量、小学生总毕业率、CO<sub>2</sub>排放量、人均CHE、人均GDP。

### 2.3 低收入国家IMR经济社会影响因素的岭回归分析

当k=0.173时,岭回归方程F检验差异有统计学

意义 ( $F=123.357, P=0.000$ ),  $R^2=0.987$ , 调整  $R^2=0.979$ , 各自变量回归系数  $t$  检验差异均有统计学意义 ( $P < 0.05$ )。见表 1、图 3。总生育率、 $CO_2$  排放量的标准化回归系数为正号, 其余均为负号。各自变量标准化回归系数的绝对值从大到小依次为总生育率、小学生总毕业率、人均 CHE、人均 GDP、 $CO_2$  排放量、谷物产量。

### 3 讨论

IMR 是国际上公认的基础健康指标, 也是衡量经济社会发展的重要指标。全球大部分国家没有能够实现千年发展儿童死亡率目标<sup>[1,2]</sup>。2019-nCoV 疫情肆虐, 人类健康和经济社会发展受到前所未有的冲击, 数十年不断改善卫生健康的努力因疫情爆发而逆转<sup>[2]</sup>。疫情不仅直接侵害儿童生命健康, 而且促使儿童死亡的社会因素向不利方向发展甚至迅速恶化。2030 年全球可持续发展目标要求消除可预防的新生儿和 5 岁以下

儿童死亡率, 任重而道远。在 5 岁以下儿童死亡中婴儿死亡占很大比重<sup>[2]</sup>, 全球平均约占 75%。全球各国需要采取有效的策略措施, 积极防控 IMR 反弹上升。

结果表明, 高、中、低收入国家 IMR 经济社会影响因素岭回归方程的  $F$  检验差异均有统计学意义, 说明回归模型有意义, 相关自变量与 IMR 间存在着回归关系。决定系数  $R^2$  和调整  $R^2$  均接近于 1, 表示回归模型的拟合效果较好。根据岭迹图确定的  $k$  值均较小, 说明各自变量的标准化回归系数趋于稳定, 偏差较小。在不同收入国家中, 除高收入国家的总生育率外其他自变量的回归系数  $t$  检验差异均有统计学意义, 提示除总生育率对高收入国家 IMR 无明显影响外, 各自变量对不同收入国家 IMR 均有不同程度的影响。标准化回归系数为正值自变量对 IMR 有消极影响, 若其数值增加, IMR 将上升; 标准化回归系数均为负值的自变量对 IMR 有积极影响, 若其数值增加, IMR 将下降。

表 1 2000-2016 年不同收入国家 IMR 影响因素岭回归分析的回归系数  $t$  检验

自变量	高收入国家			中等收入国家			低收入国家		
	$\hat{\beta}$	t	P	$\hat{\beta}$	t	P	$\hat{\beta}$	t	P
人均 GDP	-0.262	-8.075	0.000	-0.076	-2.844	0.017	-0.143	-5.207	0.000
人均 CHE	-0.294	-11.058	0.000	-0.094	-3.368	0.007	-0.179	-8.120	0.000
总生育率	0.016	0.429	0.677	0.311	8.496	0.000	0.238	10.495	0.000
谷类产量	-0.130	-2.840	0.018	-0.230	-6.078	0.000	-0.105	-2.777	0.020
小学总毕业率	-0.138	-3.110	0.011	-0.147	-3.298	0.008	-0.235	-9.512	0.000
$CO_2$ 排放量	0.194	4.333	0.001	-0.137	-4.072	0.000	0.109	3.153	0.010

注:  $\hat{\beta}$  为标准化回归系数

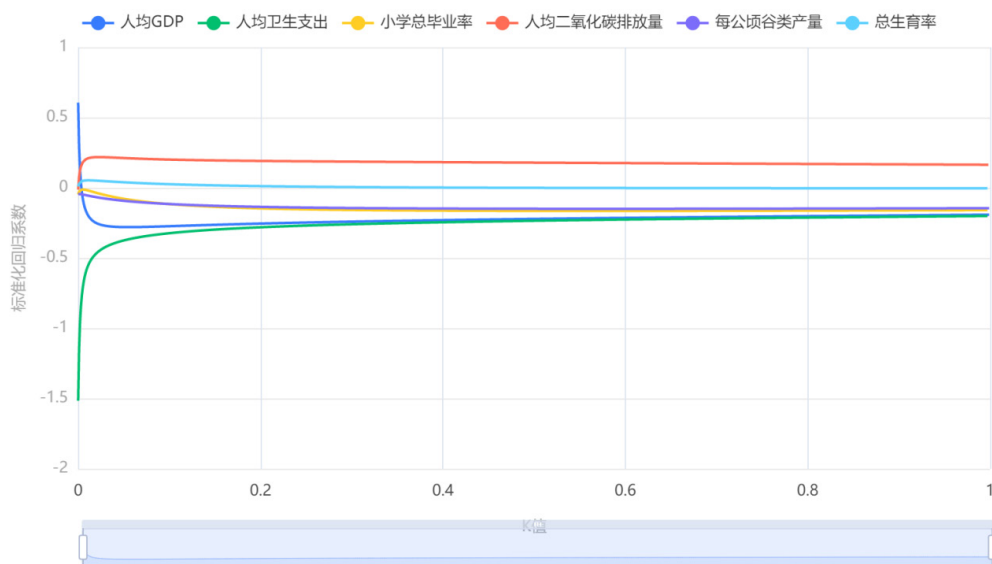


图 1 高收入国家 2000-2016 年 IMR 经济社会影响因素岭回归分析的岭迹图

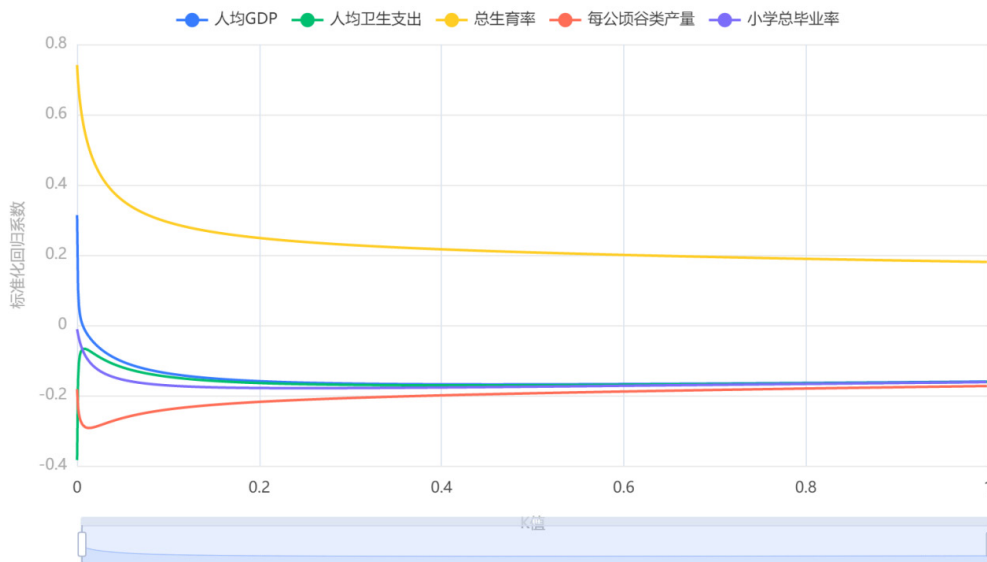


图 2 中等收入国家 2000-2016 年 IMR 经济社会影响因素岭回归分析的岭迹图

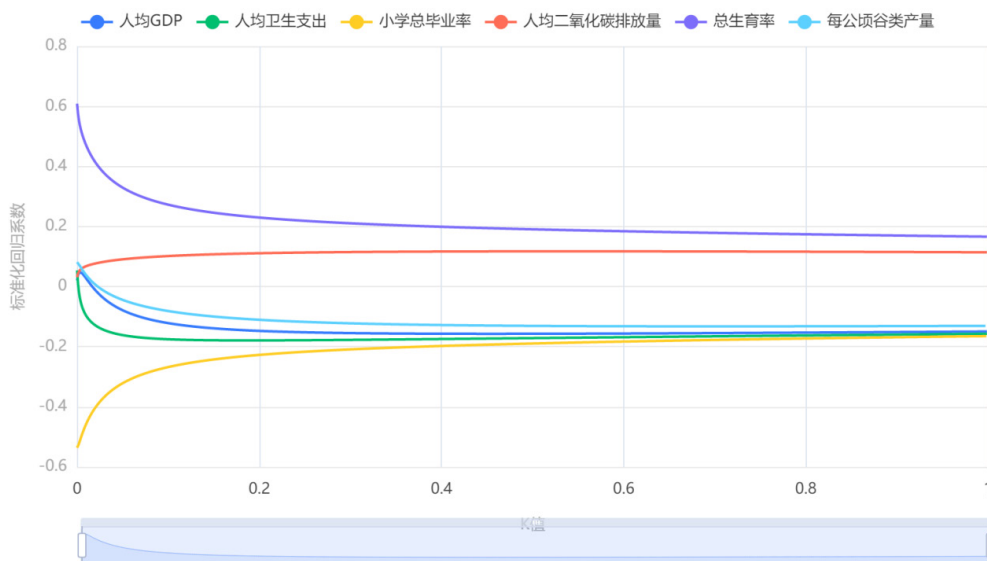


图 3 低收入国家 2000-2016 年 IMR 经济社会影响因素岭回归分析的岭迹图

标准化回归系数消除了计量单位及离散程度的影响,其绝对值可用来比较自变量对因变量的影响大小。通过比较标准化回归系数的绝对值大小,发现各自变量对不同收入国家 IMR 的影响程度大小不一,顺位排序发生了较大变化,不同收入国家的防控重点也要各有侧重。有研究,84 个发展中经济体卫生支出、经济发展会大大降低 IMR<sup>[4]</sup>,结果显示,人均 GDP、人均当前卫生支出的标准化回归系数均为负值,对 IMR 的作用与文献报道一致,但不同收入国家对 IMR 的影响程度存在较大差别,在高收入国家人均 CHE、人均 GDP 位居前二位,在中、低收入国家中总体位次靠后,

加快经济发展、加大卫生投入可作为高收入国家的重要防控策略。不依赖高速增长,通过精心策划的社会扶助项目也可快速降低 IMR<sup>[5]</sup>,中、低收入国家不应过分推崇、过度放大经济和卫生投入的作用,避免减少资源的不必要浪费。总生育率是中、低收入国家 IMR 的首位影响因素,有研究总生育率增加或减少将导致婴儿死亡率的大幅增加或减少<sup>[10]</sup>,实施恰当的人口政策,调整生育率,控制出生人口数量,可作为中、低收入国家的重要防控策略。小学生总毕业率是低收入国家第二位、中等收入国家第三位影响因素,有报道妇女受教育程度低<sup>[11]</sup>与 IMR 密切相关,受教育

程度与保健意识、主动接受保健服务的能力有关,中、低收入国家要重视妇女教育,不断提高国民素质。中等收入国家谷物产量位居第二,在高、低收入国家的位次靠后,有地区<sup>[12]</sup>高 IMR 是由于粮食不安全和营养不良引起,粮食是生存之本,加强农业生产,提高粮食产量,应作为防控的基本举措。高收入国家 CO<sub>2</sub> 排放量仅位居第三,有国家 CO<sub>2</sub> 排放对 IMR 有消极影响<sup>[13, 14]</sup>,尽管在中、低收入国家位次靠后,但应吸取高收入国家 CO<sub>2</sub> 排放量负面效应的教训。另外,中等收入国家 CO<sub>2</sub> 排放量标准化回归系数为负值,对 IMR 的积极影响缺乏充足合理的解释依据,仍需动态观察。认真审视不同收入国家 IMR 的影响因素,分析防控重点,对于调整防控策略、规避无效投入及资源浪费有一定的现实意义。

局限性:IMR 影响因素涉及社会各领域,本文收集资料有限。因纳入变量指标数据的完整性影响,本文仅能收集到 2000-2016 年的数据。以世界银行数据库为基础,其数据尽管具有一定的权威性,但经济社会因素却是动态变化的,研究结果仅供参考,仍需动态观察。

总之,经岭回归分析,2000-2016 年人均 CHE、人均 GDP 是高收入国家但不是中、低收入国家 IMR 的主要经济社会影响因素,总生育率是中、低收入国家的首要经济社会影响因素,中等收入国家谷物产量、低收入国家小学生总毕业率是第二位影响因素,中、低收入国家应吸取高收入国家 CO<sub>2</sub> 排放量负面效应的教训。不同收入国家各自变量对 IMR 的影响程度大小不一、影响顺位各有差别,防控策略应各有侧重。

### 参考文献

- [1] 李鸿斌. 基于《世界卫生统计 2015》资料全球 5 岁以下儿童死亡率的现状分析[J]. 中国循证医学杂志, 2017, 17(3): 269-275.
- [2] 李鸿斌, 贲宇. 基于《世界卫生统计 2015》资料全球婴儿死亡率分析[J]. 国际儿科研究杂志, 2022, 2(1): 12-19.
- [3] 李鸿斌, 贲宇, 冯海娟. 婴儿死亡率影响因素及未来防控策略思考[J]. 临床医学进展, 2022, 12(8): 7992-8000.
- [4] Zhou L, Bandara J S, Paramati S R. Impact of sanitation, safe drinking water and health expenditure on infant mortality rate in developing economies[J]. Australian Economic Papers, 2020, 59(1):13-33.
- [5] Hongbin Li, Haijuang Feng, Jie Wang, et al. Relationships

among gross domestic product per capita, government health expenditure per capita and infant mortality rate in China[J]. Biomedical Research, 2017, 28(6):2859-2864.

- [6] Wiggins A, Karaye I M, Horney J A. Medicaid expansion and infant mortality, revisited: A difference-in-differences analysis[J]. Health Services Research, 2020, 55(3): 93-398.
- [7] Bhatt C B, Beck-Sagué, Consuelo M. Medicaid Expansion and Infant Mortality in the United States[J]. Am J Public Health. 2018, 108(4):565-567.
- [8] Dutta U P, Gupta H, Sarkar A K, et al. Some Determinants of Infant Mortality Rate in SAARC Countries: an Empirical Assessment through Panel Data Analysis[J]. Child Ind Res, 2020, 13(5):2093-2116.
- [9] Hassan S A, Zaman K, Zaman S, et al. Measuring health expenditures and outcomes in saarc region: health is a luxury?[J]. Quality & Quantity, 2014, 48(3):1421-1437.
- [10] Ahmed A H, Muftawu M. THE CAUSAL RELATIONSHIP BETWEEN INFANT MORTALITY AND FERTILITY RATE IN GHANA: ECONOMETRIC AND TIME SERIES ANALYSIS[J]. International Journal of Advanced Research, 2018, 6(10):1192-1200.
- [11] Patel K K, Rai R, Rai A K. Determinants of infant mortality in Pakistan: evidence from Pakistan Demographic and Health Survey 2017-18[J]. J Public Health, 2021, 29(6):693-701
- [12] Sati VP. Nutritional status and infant mortality rate in Saiha district, Mizoram, India[J]. Current science, 2016, 110(12):2280-2285.
- [13] Zayed N M, Sheikh A, Shahi S K, et al. Nexus among Infant Mortality Rate, Literacy Rate and Carbon Emission Rate compared with GDP: An Empirical Analysis in Bangladesh[J]. Int J Arts and Social Science, 2020, 3(5): 20-25.
- [14] 李鸿斌, 冯海娟, 贲宇. 日本阶段性婴儿死亡率经济社会因素的岭回归分析[J]. 国际医学与数据杂志, 2023, 7(1): 4-7.

版权声明: ©2023 作者与开放获取期刊研究中心(OAJRC)所有。本文章按照知识共享署名许可条款发表。

<http://creativecommons.org/licenses/by/4.0/>



OPEN ACCESS

附表 世界银行高、中、低收入国家纳入回归模型指标的数据

年份	IMR (%)			人均 GDP (US\$)			人均 CHE (US\$)			小学总毕业率 (%)			CO <sub>2</sub> 排放量 (公吨/人)			谷类产量 (Kg/公顷)			总生育率 (女性人均生育数)		
	高	中	低	高	中	低	高	中	低	高	中	低	高	中	低	高	中	低	高	中	低
2000	6.8	52.63	88.6	25103.29	1264.83	333.73	2404.24	65.42	15.41	96.81	83.39	44.51	12.04	2.27	0.42	4597.37	2758.28	977.23	1.70	2.68	5.96
2001	6.6	50.84	85.6	24717.60	1263.91	303.88	2492.22	66.01	14.24	97.35	84.22	46.78	11.93	2.29	0.42	4578.62	2876.40	1074.37	1.66	2.64	5.90
2002	6.4	49.04	82.6	25606.47	1271.18	317.06	2695.85	64.70	15.08	96.90	85.01	48.79	11.94	2.36	0.38	4536.8	2858.83	1120.20	1.65	2.62	5.84
2003	6.3	47.27	79.6	28530.66	1413.95	336.8	3060.27	72.04	16.92	97.35	85.75	51.18	12.04	2.54	0.42	4448.75	2886.47	1140.40	1.66	2.59	5.77
2004	6.1	45.61	76.6	31559.04	1658.24	380.12	3370.69	83.39	19.93	97.39	86.39	52.78	12.15	2.72	0.42	5267.77	2981.95	1099.26	1.67	2.56	5.70
2005	6.0	43.83	73.8	33172.96	1952.91	432.32	3549.52	98.26	22.58	96.93	87.25	55.81	12.10	2.87	0.42	4893.77	3015.41	1172.45	1.67	2.54	5.64
2006	5.8	42.23	71.2	34799.05	2294.12	491.13	3738.52	114.63	26.35	97.01	87.79	57.26	12.05	3.02	0.42	4833.66	3091.03	1250.13	1.71	2.51	5.56
2007	5.6	40.63	68.7	37809.43	2804.84	584.30	4058.44	137.70	32.34	97.70	88.92	58.27	12.02	3.06	0.42	4895.96	3151.93	1218.27	1.73	2.50	5.49
2008	5.5	39.08	66.3	39890.22	3326.80	689.48	4349.29	164.47	36.55	97.79	90.71	60.66	11.76	3.29	0.44	5103.43	3330.77	1177.24	1.74	2.48	5.41
2009	5.3	37.56	64.1	37125.80	3214.33	691.50	4413.17	169.95	37.43	97.85	91.33	63.48	11.06	3.25	0.40	5280.55	3293.54	1371.81	1.72	2.46	5.33
2010	5.1	36.11	62.5	38653.90	3867.49	738.05	4556.34	195.93	38.54	97.95	91.08	63.87	11.35	3.52	0.39	5218.11	3353.56	1331.96	1.70	2.44	5.25
2011	5.0	34.79	60.1	41538.63	4518.54	784.36	4896.81	226.11	41.71	98.31	90.66	63.52	11.10	3.74	0.35	5280.77	3503.96	1326.94	1.68	2.42	5.17
2012	4.9	33.53	58.2	41344.81	4771.64	803.90	4913.19	241.63	40.24	98.89	91.58	64.08	10.97	3.83	0.32	4857.81	3526.13	1357.95	1.69	2.41	5.08
2013	4.8	32.37	56.5	41629.15	4992.77	859.27	4951.30	257.92	42.31	98.69	91.94	63.04	10.91	3.81	0.30	5581.15	3638.63	1276.75	1.66	2.39	5.00
2014	4.7	31.30	54.9	42298.55	5119.58	896.71	5097.84	264.21	44.45	98.88	91.60	63.70	10.77	3.83	0.30	5847.39	3680.83	1300.17	1.67	2.38	4.92
2015	4.6	30.28	53.5	39735.36	4804.80	884.79	4936.13	251.66	45.62	98.98	91.21	63.92	10.65	3.79	0.27	5672.15	3731.59	1360.71	1.66	2.37	4.84
2016	4.5	29.40	51.9	40343.82	4780.83	790.16	5099.69	247.17	43.70	98.39	91.48	64.28	10.39	3.75	0.27	6153.83	3770.67	1293.78	1.66	2.36	4.76