

计算中国分省资本存量：一个 R 包实现 *

陈普

(华东交通大学经管学院, 江西 南昌, 330013)

内容提要: 资本存量计算是宏观经济研究中一个重要组成部分。本文依据张军等 (2004) 的算法提供了计算资本存量的一个 R 软件包。同时, 基于张军等 (2004) 的思路, 通过修正折旧率进一步改进了该算法。不同于张军等 (2004) 对各省各年统一的 9.6% 折旧率的设置, 本文使用的折旧数据源自各省公布数据, 计算结果应更接近现实情况。

关键词: 资本存量; R 包; 永续盘存法

中图分类号: **文献标志码:**

一 引言

资本存量在宏观经济研究中一直是非常有用的变量, 凡涉及到经济增长的实证研究, 一般都绕不开该变量。但遗憾的是, 该变量官方一直未有公布, 多位学者为估算该变量提供了多种方式 (张军等, 2004; 单豪杰, 2008; 叶宗裕, 2010; 陈昌兵, 2014, 2020)。估算资本存量, 如果使用永续盘存法, 一般要用到四个数据, 即初期资本存量、折旧率、投资和投资价格指数。后两个变量一般统计局都有公布或可通过常用的方法进行估算, 而初期资本存量和折旧率一般都是主观设定。但对于初期资本存量而言, 只要其设置的时期足够远, 比如从 1952 年开始, 那么在 1978 年以后, 它的影响就非常小了 (李宾, 2011)。不过, 若从 1952 年开始计算资本存量, 其中要用到的中国各省的相关数据多有缺失 (见文后的表 1), 补充这些数据, 不仅繁琐还容易出错。因此, 在中国经济的宏观研究中, 尽管研究者在文中提到的算法一致, 但很难确保研究者在实证中最终使用的资本存量数据是一致的。如张健华和王鹏 (2012) 资本存量的估算与张军等 (2004) 方法类似, 但最终估算结果存在些许差异, 其文中的表 2 是对该差异的一个比较。

基于此, 考虑到张军等 (2004) 在中国分省资本存量计算中的广泛引用, 本文依据张军等 (2004) 的方法, 对中国各省的资本存量提供了一个 R 包实现, 该包已在 R 的官网收录。此外, 本文也拟对折旧率做出进一步修正。因为折旧率不同于初期资本存量, 它的变化往往极大影响后续资本存量计算 (李宾, 2011)。然而相关文献对该变量往往采取主观认定 (张军等, 2004; 单豪杰, 2008) 或模型估算 (陈昌兵, 2020; 贾润松和张四灿, 2014), 实际上, 固定资产折旧数据各省在 1993 年以后均已公布。因此, 本文尝试对资本存量估算提供另外一种思路。即基于张军等 (2004) 的算法得到各省 1992 年的资本存量, 然后以此作为初值, 利用各省公布固定资产折旧数据计算 1993 年之后的资本存量。该算法, 一般而言, 折旧率要小于张军等 (2004) 建议的 9.6%。整体上, 本文的研究应能为后续宏观经济研究的可复制性提供便利的实证工具。

本文后续结构安排如下, 第二部分是对资本存量算法的细节描述, 第三部分是对 R 语言软件包 `CHNCapitalStock` 的语法阐述, 以及如何使用的一个案例展示, 第四部分是用本文提出的思路对资本存量的重新估算, 第五部分是总结。

二 分省资本存量的计算细节

计算资本存量，通常利用 Goldsmith(1951) 的永续盘存法，该算法公式为，

$$K_t = (1 - \delta)K_{t-1} + I_t \tag{1}$$

其中， K_t 为 t 期的资本存量， δ 是折旧率， I_t 是投资。如果给定初期资本存量，每期的投资数据，以及对应的当年折旧率，那么资本存量序列可以从(1)式递归计算得到。利用(1)式计算中国分省资本存量的主要困难有三：第一，初期资本存量未知。第二，折旧率未知。第三，投资 I_t 是真实变量，因此，不仅需要当年计价的投资额，还需要投资的价格指数平减来得到真实变量的值。而这两个数据，部分省份存在不同程度缺失。特别是投资的价格指数，多数省份在 1995 年之前是没有的。

为解决上述问题，张军等 (2004) 提供了相应的处理方式。他们从 1952 年开始计算资本存量，初期的资本存量假设为当年投资数据的 10 倍，折旧率设为 9.6%。对于投资数据，多数文献 (单豪杰,2008; 张健华与王鹏, 2012) 与张军等 (2004) 类似，均选用固定资本形成总额。而投资的价格指数，张军等 (2004) 在文中的公式 (3) 和 (4) 提供了一个计算方法，从而可以得到大部分省份 1995 年之前的投资价格指数。对于 1995 年之后的价格指数则可直接采用已经公布的固定资产投资价格指数 (以下称 IPI)。对于缺失数据，张军等 (2004) 做出如下处理：

表 1: 张军等 (2004) 缺失数据的处理方式

年份	缺失的数据	处理方式
天津	1952-1989 年的 IPI	1952-1978 年之前用 $RPI \times 0.985$ (RPI 为商品零售价格指数), 1979-1988 年直接用 RPI 。
江西、广东	1952-1978 年的固定资本形成总额	用 1979-2000 年的固定资本形成总额对基本建设投资回归，利用该回归 1952-1978 年固定资本形成总额的样本外预测值作为缺失值填补。
广东	1952-1977 年的 IPI	用广东对应年份的 RPI 填补。
广东	1996-2000 年的 IPI	用福建对应年份的 IPI 填补。
海南	1952-1977 年的固定资本形成总额	用 1952-1977 年基本建设投资填补。
	1952-1977 年的 IPI	直接取 1。
	1978-2000 年的 IPI	用海南的 RPI 填补。
西藏	1952-1958 年的固定资本形成总额	用西藏 1959 年的基本建设投资填补。
	1959-1977 年的固定资本形成总额	用西藏基本建设投资填补。
	1978-1991 年的固定资本形成总额	用全社会固定资产投资填补。

张军等 (2004) 利用上述处理，将资本存量计算至 2000 年。我们在上述缺失数据处理过程中，进行了下述补充处理：第一，西藏的 IPI 到 2019 年一直没有公布，张军等 (2004) 也未在文中提到如何填补该指数，我们在 1990 年前用 1，1991 年以后用 RPI ，最终得到的结果与张军等 (2004) 表 2 的资本存量计算结果基本一致。第二，海南基本建设投资数据在 1952-1970 年，多数年份是缺失的，张军等 (2004) 也未提到如何处理该情况。但“一五”、“二五”、...、“六五”时期的基本建设投资加总数据在《海南统计年鉴 (1987)》有公布，我们用此加总数据对缺失年份进行了平均，最终资本存量的计算结果与张军等 (2004) 表 2 相差不大。第三，张军等 (2004) 将重庆和四川是合并处理的，为今后研究的便利，我们对这两个区域进行了分开处理。重庆的大部分数据在 1996 年以后才开始记录，因此，在计算重庆时，1996 年的初始投资价格指数用四川的替代，之后用重庆的固定资产价格指数计算。1996 年的初值资本存量用 1995 年重庆和四川的固定资本形成比例乘以 1995 年四川的资本存量，来获得。第四，对于 2000 年以后的数据，我们直接利用固定资本形成总额作为投资，用 IPI 进行平减，但西藏用 RPI 进行平减。

三 资本存量估算的 CHNCapitalStock 包应用

基于上述算法，我们提供了一个 R 包予以实现，该 R 包已被 R 软件官网收录。在联网情况下，在 R 语言控制台直接输入`install.packages('CHNCapitalStock')`进行安装。但我们建议安装本包的开发版本，该版本新增了本文第四部分阐述的算法。后续我们也会在 R 的官网中更新该包。安装开发版本的命令为

```
devtools::install_github('common2016/CapitalStock')
```

该包包含的主要函数为CompK，其语法如下，

```
CompK(prv, method = 'ZJ', startyr = 1996, yr = NULL, invest = NULL, InvestPrice = NULL,
      depr = NULL, delta = 0.096, bt = 1952)
```

从中可以看到该函数包含多个参数，每个参数的含义我们列表如下：

表 2: CompK 函数参数的含义	
参数	含义
prv	标量字符串，为要计算的省份的名字，用汉语拼音标记。如北京,prv = beijing，但陕西和山西有着同样的拼音，因此我们用分别用shan3xi和shan1xi标记这两个省份。
method	设置为'ZJ'或'CP'，前者表示使用张军等 (2004) 的方法计算资本存量，后者表示使用本文第四部分的算法计算资本存量。缺省值为'ZJ'。
startyr	如果method = 'CP'，意味着自startyr年之后使用'CP'算法。如果method = 'ZJ'，startyr参数是没有用的。
yr	是一个数值向量。如果你仅需计算 2017 年之前的资本存量，那么该参数可用缺省值NULL。如果你打算计算 2017 年以后的资本存量，如 2018 年和 2019 年，那么该参数可设置为yr = c(2018,2019)。
invest	当年价的投资数据，单位亿元。与yr一致，计算 2017 年前的资本存量，该参数无须提供。计算 2017 年以后的数据时提供该参数值，该参数的向量元素数目与yr的一致。
InvestPrice	投资价格指数。与yr一致，计算 2017 年前的资本存量，该参数无须提供。计算 2017 年以后的数据时提供该参数值，该参数的向量元素数目与yr的一致。
depr	当年折旧数额，单位亿元。与yr一致，计算 2017 年前的资本存量，该参数无须提供。计算 2017 年以后的数据时提供该参数值，该参数的向量元素数目与yr的一致。
delta	折旧率，缺省值为张军等 (2004) 建议的 0.096。
bt	数值标量，为 IPI 计算的基年，缺省值我们设为 1952 年。

该函数的最终输出为一个数据框类型的变量：其第一列为省份的汉语拼音，第二列为对应年份，第三列为资本存量，第四列为投资的价格指数。

如果我们要计算安徽 1952-2017 年的资本存量，相关代码如下，

```
library(CHNCapitalStock)
CompK(prv = 'anhui')
```

其结果输出为，

#	prv	yr	K	InvestPrice
# 1	anhui	1952	23.80000	1.000000
# 2	anhui	1953	24.44022	1.148710
# 3	anhui	1954	25.20224	1.142111
# ...				
# 64	anhui	2015	4470.31591	12.623461
# 65	anhui	2016	5026.90503	12.523735
# 66	anhui	2017	5553.14400	13.450492

上述资本存量的价值是基于 1952 年等于 1 的投资价格指数而计算，如果价格指数设置为 2000 年等于 1，资本存量可用如下命令重新计算，

```
CompK(prv = 'anhui', bt = 2000)
```

```
#   prv   yr      K      InvestPrice
# 1  anhui 1952   23.80000   0.1120472
# 2  anhui 1953   47.62046   0.1287097
# 3  anhui 1954   70.78971   0.1279703
# ...
# 64 anhui 2015 39896.40695   1.4144230
# 65 anhui 2016 44863.89217   1.4032490
# 66 anhui 2017 49560.50491   1.5070895
```

类似的，如果要更换不同的折旧率比如 0.1，可如下重新计算，

```
CompK(prv = 'anhui', bt = 2000, delta = 0.1)
```

```
#   prv   yr      K      InvestPrice
# 1  anhui 1952   23.80000   0.1120472
# 2  anhui 1953   47.52526   0.1287097
# 3  anhui 1954   70.51355   0.1279703
# ...
# 64 anhui 2015 39246.80386   1.4144230
# 65 anhui 2016 44119.66376   1.4032490
# 66 anhui 2017 48711.24377   1.5070895
```

如果要进一步计算 2018 年和 2019 年北京的资本存量，使用者只需提供对应年份的固定资本形成总额和 *IPI*，则利用 `CompK` 函数可以方便得到。但要注意投资价格指数的基期计算。假设以当年价格度量的北京固定资本形成总额在 2018 年和 2019 年分别为 10801 和 11100，其 *IPI* 在 Wind 数据库可以查到上年等于 100 时，*IPI* 分别为 103 和 102.1。那么，首先计算以 1952 年等于 1 时的 *IPI*，可以通过如下命令得到以 1952 年为 1 的 2017 年的 *IPI* 为 1.86。

```
CompK(prv = 'beijing', bt = 1952)
```

```
#   prv   yr      K      InvestPrice
#   ...
#  beijing 2016  29813   1.7845
#  beijing 2017  32506   1.8678
```

那么，对应的以 1952 年为 1 的北京 2018 和 2019 年的 *IPI* 为 $1.86 \times 1.03 = 1.91$ 和 $1.86 \times 1.03 \times 1.021 = 1.95$ 。然后，通过如下命令即可计算北京 2018 和 2019 年的资本存量。

```
CompK(yr = c(2018,2019), invest = c(10801.2,11100),
      InvestPrice = c(1.91,1.95),
      prv = 'beijing',delta = 0.096)
```

```
#   prv   yr      K      InvestPrice
# 67  beijing 1952   7.90000   1.0000000
# 68  beijing 1953  12.50886   0.9986474
# 69  beijing 1954  19.50900   0.9986599
# ...
# 132 beijing 2017 32506.39871   1.8678465
# 133 beijing 2018 35040.86296   1.9100000
# 134 beijing 2019 37369.24781   1.9500000
```

四 分省资本存量的另一个算法

在资本存量的估算中,如果初期足够久远,比如从 1952 年开始计算,那么其对 1978 年以后后资本存量影响非常小,1990 之后的则可忽略不计(贾润松和张四灿, 2014)。实际上,计算资本存量影响最大的参数是折旧率,其他条件相同,折旧率相差 1% 会导致资本存量 25 年后相差 10%(李宾, 2011)。张军等(2004) 直接设定了 9.6% 的折旧率,单豪杰(2008) 取 10.96%,陈昌兵(2020) 则对各省折旧率进行了估算,各省范围大约在 5% 至 8% 之间。而我国在 1993 年以后,各省都提供了固定资产折旧数据(见 Wind 数据库),那么只要有该年的资本存量,折旧数据可用固定资产投资价格指数平减以后直接使用,而无须再次估算折旧率,从而消除折旧率主观设置或模型估计带来的偏误。

基于该想法,我们先利用张军等(2004) 的算法,计算得到各省直至 1992 年的资本存量,然后再利用已经公布的固定资产折旧数据,利用(1)式重新计算资本存量。因此,利用该算法,我们仅能得到 1993 年以后的资本存量数据,对于重庆而言,只能得到 1996 年以后的资本存量数据。在CHNCapitalStock包中,要利用该算法计算安徽的资本存量,可用如下命令实现该算法:

```
CompK(prv = 'anhui', method = 'CP', startyr = 1993, bt = 1992)
#      prv   yr      K InvestPrice
# 1  anhui 1952      NA    0.1632730
# 2  anhui 1953      NA    0.1875533
# ...
# 41 anhui 1992      NA    1.0000000
# 42 anhui 1993 1748.124    1.1900042
# 43 anhui 1994 1979.990    1.2637930
# ...
# 65 anhui 2016 36367.587    2.0447877
# 66 anhui 2017 40891.434    2.1961020
```

该命令中的startyr参数取值意味着我们从 1993 年以后利用新算法来得到资本存量。同样的,如果数据可行,我们也可以基于新算法来得到 2018 年和 2019 年安徽的资本存量,

```
CompK(prv = 'anhui', method = 'CP', yr = 2018:2019, invest = c(10801.2, 11100),
      InvestPrice = c(1.86*1.03, 1.86*1.03*1.021),
      depr = c(3800, 3850),
      delta = 0.096, startyr = 1993, bt = 1992)
#      prv   yr      K InvestPrice
# ...
# 66 anhui 2017 40891.434    2.1961020
# 67 anhui 2018 63273.902    0.3127984
# 68 anhui 2019 85975.046    0.3193672
```

基于新的算法,我们将计算得到的 1993 年至 2017 年的资本存量绘制在图 1 中。

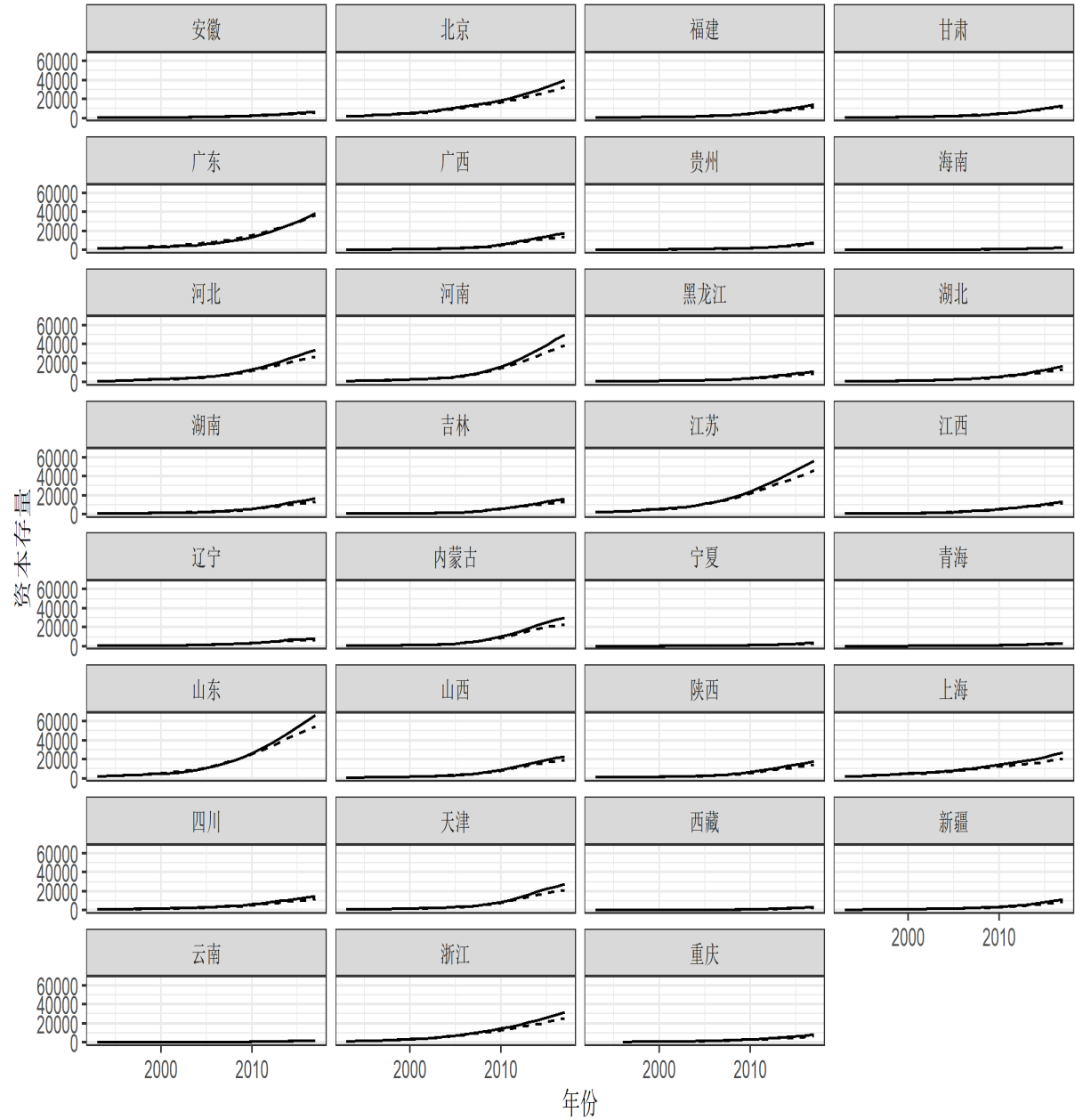


图 1: 1993-2017 年各省资本存量走势图

图中的虚线是利用张军等 (2004) 的方法得到的, 实线是利用我们前述的方法得到的。总体上, 我们估算的资本存量比张军等 (2004) 估算的要高一些。另外, 也注意到各省差异的大小也十分不同。有些省份不同算法间差异较小, 如安徽、广东、甘肃、贵州、青海等, 但有些省份在后期出现较大的分歧, 如河南、山东、北京等。实际上这些分歧的出现, 主要原因在于折旧率与张军等 (2004) 中默认的 9.6% 不同。利用各省自 1993 年公布的固定资产折旧数据以及张军等 (2004) 测算得到的 1993 年的初始资本存量, 我们可以为各省各年估计出一个折旧率, 我们将这些折旧率按年平均后, 可以与张军等 (2004) 9.6% 的折旧率做一个比较。我们将折旧率数据绘制在图 2 中。图中的粗黑线是 0.096, 除了广东以外, 其他各省的折旧率都低于该折旧率。一般地, 可以看到, 往往经济发展水平低一些, 折旧就慢一些, 折旧率就低一些。

可以看到, 我们的算法因不再对各省折旧率统一设定为某个数值, 如 9.6%, 从而有着更少的约束。特别是我们的折旧数据源自各省统计公布, 相对而言, 我们算法估计的资本存量应该更加接近现实。

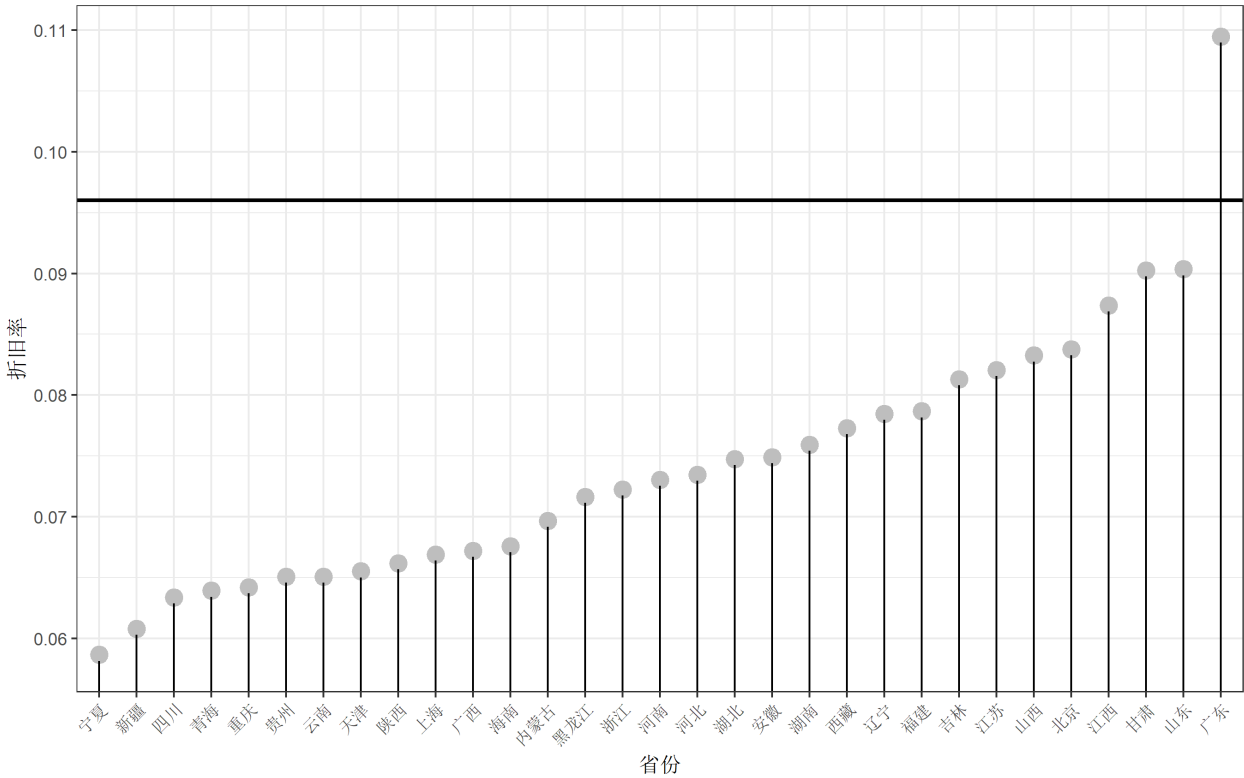


图 2: 1993 年-2017 年各省平均折旧率

五 结论

本文重新梳理了张军等 (2004) 对资本存量的算法，并为该算法提供了一个 R 软件包。为方便该包的使用，本文详细阐述了该包的使用语法，并举例加以说明。资本存量计算中，折旧率是一个关键参数。考虑到张军等 (2004) 对各省各年折旧率统一设定为 9.6%，可能并不准确。本文利用各省公布的固定资产折旧数据，重新对资本存量进行了估算。总体上看，我们的算法，资本存量要略高于张军等 (2004)。但因我们放松了固定的折旧率的约束，基于各省公开的折旧数据，从而使得折旧率各省各年各不相同，我们估算的资本存量应该更加接近现实。

参考文献

[1] 陈昌兵, 可变折旧率估计及资本存量测算. 经济研究, 2014(12): 第 72-85 页.

[2] 陈昌兵, 可变折旧率的另一种估计方法——基于中国各省份资本折旧的极大似然估计. 经济研究, 2020. 55(01): 第 49-64 页.

[3] 贾润崧与张四灿, 中国省际资本存量与资本回报率. 统计研究, 2014(11): 第 35-42 页.

[4] 李宾, 我国资本存量估算的比较分析. 数量经济技术经济研究, 2011(12): 第 21-36+54 页.

[5] 单豪杰, 中国资本存量 K 的再估算:1952~2006 年. 数量经济技术经济研究, 2008(10): 第 17-31 页.

[6] 叶宗裕, 中国省际资本存量估算. 统计研究, 2010(12): 第 65-71 页.

[7] 张健华与王鹏, 中国全要素生产率: 基于分省份资本折旧率的再估计. 管理世界, 2012(10): 第 18-30+187 页.

[8] 张军, 吴桂英与张吉鹏, 中国省际物质资本存量估算:1952—2000. 经济研究, 2004(10): 第 35-44 页.

[9] Goldsmith, R.W., A perpetual inventory of national wealth, in Studies in Income and Wealth. 1951, NBER. p. 5-73.