|  |
| --- |
| **データの説明** |

藪友良『入門　実践する計量経済学』(東洋経済新報社、2023年)の実証例を再現するためのデータです。

1. 同性愛者の割合 gay\_data.csv

2章で用いた同性愛者に関する米国州別データ

出所：同性愛を自己申告する割合(%)はギャラップ調査のadmit LGBT。

同性婚を支持する割合(%)はネイト・シルバーによる推定値support for same-sex marriage。データは全て2012年の値となる。サンプルサイズは50。

state: 州の名前

support: 各州の同性婚を支持する割合(%表示)

gay:各州の同性愛を自己申告する割合(%表示)

注)コロンビア自治区は州ではないため、データに含まれていない

2. 親子の身長 height\_data.csv

2章、4章、5章で用いた親子の身長データ

出所：授業で収集した計164人(男性113人、女性51人)の学生の身長とその両親の身長データになります。

student:学生の身長(cm)

father: 学生の父親の身長(cm)

mother: 学生の母親の身長(cm)

female: 学生が女性なら1としたダミー変数

3. 成績 timss\_data.csv

4章と5章で用いた中学2年生の成績データ

出所：国際数学・理科教育動向調査(TIMSS)の2003年の中学2年生のデータ

URL: https://timssandpirls.bc.edu/timss2003i/userguide.html

サンプルサイズ: 4508(男2259人、女2249人)

変数の定義

school: 学校のid番号(idschool)

math: 数学の成績(BSMNRSC)

science:理科の成績(BSSNRSC)

female: 女性なら1となるダミー変数(itsex, bsbgsex)

birth\_q1: 4月から6月生まれのダミー変数(itbirthm,bsbgbrtm)

birth\_q2: 7月から9月生まれのダミー変数(itbirthm, bsbgbrtm)

birth\_q3: 10月から12月生まれのダミー変数(itbirthm, bsbgbrtm)

birth\_q4: 1月から3月生まれのダミー変数(itbirthm, bsbgbrtm)

mother: 母親の教育年数(bsbgmfed)、中学校9、高校12、短大・高専14、大学16、大学院18、ただし、不明は0とした

father: 父親の教育年数(bsbgfmed)、中学校9、高校12、短大・高専14、大学16、大学院18、ただし、不明は0とした

birth\_m1: 1月生まれのダミー変数(itbirthm,bsbgbrtm)

birth\_m2: 2月生まれのダミー変数(itbirthm,bsbgbrtm)

birth\_m3: 3月生まれのダミー変数(itbirthm,bsbgbrtm)

birth\_m4: 4月生まれのダミー変数(itbirthm,bsbgbrtm)

birth\_m5: 5月生まれのダミー変数(itbirthm,bsbgbrtm)

birth\_m6: 6月生まれのダミー変数(itbirthm,bsbgbrtm)

birth\_m7: 7月生まれのダミー変数(itbirthm,bsbgbrtm)

birth\_m8: 8月生まれのダミー変数(itbirthm,bsbgbrtm)

birth\_m9: 9月生まれのダミー変数(itbirthm,bsbgbrtm)

birth\_m10: 10月生まれのダミー変数(itbirthm,bsbgbrtm)

birth\_m11: 11月生まれのダミー変数(itbirthm,bsbgbrtm)

birth\_m12: 12月生まれのダミー変数(itbirthm,bsbgbrtm)

データ作成上の注意点

全てのデータから未回答/分からないと回答されたものを除く

誕生月は99を除く

誕生年は99を除く

誕生年月で88年４月～89年3月、以外を除く

4. 賃料 rent\_data.csv

3章、5章、6章、9章で用いた賃貸物件のデータ

出所：H駅周辺にある計724の賃貸物件に関するデータ(2021年に収集)

rent:　家賃(万円)

space:　専有面積(㎡)

distance:　最寄り駅までの時間(分)

age:　築年数(年)

floor:　階数(階)

5. 高齢化とGDP demography\_data.csv

5章で用いた高齢化と経済成長率のデータ

出所

このデータは、以下の研究で用いられたデータとなります。

Daron Acemoglu and Pascual Restrepo "Secular Stagnation? The Effect of Aging on Economic Growth

in the Age of Automation", American Economic Review: Papers & Proceedings 2017, 107(5): 174–179

URL: https://economics.mit.edu/faculty/acemoglu/data/secular

データの解説

pop1990: 1990年の人口

pop2015: 2015年の人口

rgdp1990: 1990年の実質GDP

rgdp2015: 2015年の実質GDP

depend1990: 1990年の高齢人口(50歳以上)÷若者人口(20-49歳)

depend2015: 2015年の高齢人口(50歳以上)÷若者人口(20-49歳)

depend1990\_ave: 1990年の20歳以上の人口の平均年齢

depend2015\_ave: 2015年の20歳以上の人口の平均年齢

6. 貨幣需要関数 money\_data.csv

6章で用いた米国の貨幣需要データ

出所

このデータは、以下の研究で用いられたデータとなります。詳しくは論文を参照してください。

Watanabe,Tsutomu and Yabu, Tomoyoshi (2018)."The Demand for Money at the Zero Interest Rate Bound," CARF Working Paper Series, CARF-F-444.

標本期間：1980Q1-2013Q4

データの解説

M1: Sweep Adjusted M1

NGDP: 名目GDP

interest: 名目金利

7. ガソリン価格 oil\_data.csv

6章と10章で用いたガソリン価格と原油価格のデータ

出所：原油価格はCrude Oil Prices: Brent - Europe, Dollars per Barrel(average)、

ドル円レートはJapanese Yen to U.S. Dollar Spot Exchange Rate(average)とし、FREDから入手した。また、ガソリン価格は、国内企業物価指数 ガソリン(税込み)を日経NEEDSから入手した。

期間: 1987年6月～2017年12月

変数の定義

brent: 北海ブレント原油の価格(ドル建て)

spot: ドル円レート

price: ガソリン価格

8. 効率市場仮説 stockprice\_data.csv

7章で用いた市場効率仮説検証のためのデータ

期間：1991月1月4日-2018年2月28日

ds: 日経平均の変化率(%)

ds\_1: 1日前の日経平均変化率(%)

ds\_2: 2日前の日経平均変化率(%)

ds\_3: 3日前の日経平均変化率(%)

s\_ma: 1日前の日経平均の移動平均(過去1か月）からの乖離率(%)

time: 1991/1/4を1とし、1づつ増加させた変数。2018/2/28は6677となる。サンプルサイズは6677となる。

9. 日本の実質GDP成長率 growth\_data.csv

 7章で用いた日本の実質GDP成長率のデータ

期間：1981Q1-2007Q4

データの定義

growth: 実質GDPの成長率(季節性を除くため前年同期比とし、100を掛けることで%表示とした)

10. 日次為替介入 dailyint\_data.csv

7章で用いた為替介入の日次データ

出所：このデータは以下の研究で用いられたものです。詳細は論文を参照してください。

Ito, Takatoshi and Yabu, Tomoyoshi.“What prompts Japan to intervene in the Forex market: A new approach to a reaction function,” Journal of International Money and Finance 26, March 2007, Pages 193-212.

標本期間：1991年4月1日-2002年12月31日

データの定義

intervention: 為替介入額(プラスなら円買いドル売り介入、マイナスなら円売りドル買い介入,単位：億円)

spot:　ドル円レート

11. 月次為替介入 monthlyint\_data.csv

7章の補足で用いた為替介入の月次データ

出所

このデータは次の研究で用いられたものです。詳細は論文を参照してください。

伊藤隆敏、藪友良 (2017)「為替介入と外貨準備―運用損益の長期推計」、『日本経済研究』

Ito, T. and Yabu, T (2020) “Japanese Foreign Exchange Interventions, 1971-2018:

Estimating a Reaction Function Using the Best Proxy” Journal of the Japanese and International Economies 58, Article 101106.

標本期間：1991年4-1999年3月

変数の定義

intj:　実際の介入額(億円)

intj\_gaika: 外貨準備から推定した介入額(億円)

intj\_taimin: 対民収支から推定した介入額(億円)

12. 差別 black\_data.csv

9章で用いた差別に関するデータ

出所

このデータは、以下の研究で用いられたデータとなります。詳細は論文を参照してください。

Bertrand, Marianne and Sendhil Mullainathan (2004) "Are Emily and Greg More Employable Than Lakisha and Jamal? A Field Experiment On Labor Market Discrimination",

American Economic Review 94, 991-1013.

URL: https://www.aeaweb.org/articles?id=10.1257/0002828042002561

データの定義

name: first name

call: 企業から連絡があれば1となるダミー変数

black: 黒人風の名前であれば1となるダミー変数

experience: 勤続年数

female: 女性なら1となるダミー変数

13. 消費 cons\_data.csv

9章で用いた都道府県別の消費データ

出所

全国家計構造調査（旧全国消費実態調査） 平成26年全国消費実態調査 都道府県別 家計収支に関する結果［二人以上の世帯］地域編第42表　地域別１世帯当たり１か月間の収入と支出

URL: https://www.e-stat.go.jp/dbview?sid=0003122134

注）世帯区分\_2014は、「二人以上の世帯のうち勤労者世帯」を使用した。

変数の定義

number: 集計世帯数

age: 世帯主の年齢（歳）

income 可処分所得(万円)

cons: 消費支出(万円)

house: 持ち家率(現住所、%表示)

14. 自殺国際比較 suicide\_data.csv

11章で用いた自殺率の国際比較

出所

世界銀行データベース

データは64カ国、2000年～2019年の20年分

データの定義

suicide:自殺率(10万人当たり)

unemployment:失業率(%)

15. 自殺都道府県別 suicide\_data.csv

11章で用いた都道府県別の自殺率

出所

自殺率は「社会・人口統計体系」、完全失業率は労働力調査から入手した。

分析期間は2009年から2019年、47都道府県のデータとなる。

\* データは年度。完全失業率は、四半期単位から年度単位に変換した値とする。

変数の定義

suicide:自殺率(10万人当たり)

unemployment:都道府県別完全失業率(%)

16. 緊急事態宣言とステイホーム stayhome\_data.csv

11章で用いた都道府県別ステーホーム指標

出所

このデータは、以下の研究で用いられたデータとなります。詳細は論文を参照して

ください。

Watanabe, T. and Yabu, T. (2021) Japan's Voluntary Lockdown, PLoS ONE 16(6).

日本語版『日本の自発的ロックダウンに関する考察』

標本期間 2020年1月6日から2020年6月28日まで

データの解説

stay: ステイホーム指標(%)

infection: 新規感染者数

rain: 県庁所在地の降水量

emerg\_start:緊急事態の開始日から1となるダミー変数

emerg\_end:緊急事態の終了日から1となるダミー変数

close\_start:学校閉鎖の開始日から1となるダミー変数

close\_end:学校閉鎖の終了日から1となるダミー変数

mobilephones:調査に用いた携帯電話数

17. 結婚 marriage\_data.csv

12章で用いた結婚に関するデータ

出所

国際成人力調査(PIAAC)

16～65歳までの個人を対象

2011年8月から2012年2月に実施された調査結果

male:男性なら1となるダミー変数

age:年齢

partner:パートナーがいたら1となるダミー変数

undergrad:大学学部卒なら1となるダミー変数

grad:大学院卒なら1となるダミー変数

18. 労働参加 participation\_data.csv

12章で用いた労働参加に関するデータ

出所

国際成人力調査(PIAAC)

16～65歳までの個人を対象

2011年8月から2012年2月に実施された調査結果

male:男性なら1となるダミー変数

age:年齢

partner:パートナーがいたら1となるダミー変数

children:子供の数

yongest:一番下の子供の年齢

experience: 過去の就業年数

undergrad:大学学部卒なら1となるダミー変数

grad:大学院卒なら1となるダミー変数

job:1ならフルタイム、2ならパートタイム、0なら未就業

19. 制度とGDP protection\_data.csv

13章で用いた制度とGDPに関するデータ

出所

以下の研究で用いられたデータとなります。

Acemoglu, D., S. Johnson and J.A. Robinson (2001) ”The Colonial Origins of Comparative Development:

An Empirical Investigation” American Economic Review 91(5): 1369-1401.

データの解説

protection:搾取からの保護指数

gdp: 1995年のGDP(1人当たり)の対数

mortality:植民者死亡率の対数

latitude：緯度の絶対値

europop: 1900年時の欧州系人口割合

20. 女性の労働時間 laborsupplyj\_data.csv

13章で用いた女性労働時間に関するデータ(21～35歳の2人以上の子供がいる母親)

出所

このデータは、J.Stock and M.Watson "Introduction to Econometrics"

で提供されているデータセットをもとに作成しました。元データは、以下の研究となります。

Joshua Angrist and William Evans“Children and Their Parents’ Labor Supply: Evidence from Exogenous

Variation in Family Size,” American Economic Review, June 1998, Vol. 88, No. 3, 450-477.

URL: https://economics.mit.edu/faculty/angrist/data1/data/angev98

変数の定義

morekids:子供が3人以上なら1となるダミー変数

samesex:最初の2人の子供が同性なら1となるダミー変数

age: 母親の年齢

black:　母親が黒人なら1となるダミー変数

hispan:　母親がヒスパニック系なら1となるダミー変数

othrace:　母親が黒人、ヒスパニック、白人でもないなら1となるダミー変数

weeks:　母親の労働時間(1年で何週働いたか)

21. STAR実験 star\_data.csv

14章で用いたSTAR実験のデータ

出所

このデータは、J.Stock and M.Watson "Introduction to Econometrics"

で提供されているデータセットをもとに作成しました。元データは以下から

ダウンロードできます。

https://ideas.repec.org/p/boc/bocins/webstar.html

データの解説

total\_score:リーディングと算数の合計点

total\_read:リーディングの点数

total\_math:算数の点数

small:少人数なら1となるダミー変数

regular\_aid:標準/補助クラスならば、1となるダミー変数

total\_experience:担当教員の勤続年数

school:学校のid番号

black:黒人なら1となるダミー変数

boy:男なら1となるダミー変数

freelunch:フリーランチの資格があるなら1となるダミー変数(所得の代理変数となり、1なら貧しい家庭出身)

22. 最低賃金 mwage\_data.csv

14章で用いた最低賃金のデータ

出所

このデータは、以下の研究で用いられたデータとなります。

Card, D., and Krueger, A. B. (1994), “Minimum Wages and Employment:

A Case Study of the Fast-Food Industry in New Jersey and Pennsylvania,”

American Economic Review, 84, 772–793

データの解説

store:店のid番号

time:介入後なら1となるダミー変数

state:ニュージャージー州に属するなら1となるダミー変数

fulltime:フルタイム労働者数

hours:店舗の営業時間

register:レジ台数

23. 選挙 election\_data.csv

14章で用いた選挙のデータ

出所

このデータは、以下の研究で用いられたデータとなります。

Lee, D. (2008), “Randomized experiments from non-random selection in U.S. House elections,”Journal of Econometrics 142 (2008) 675–697

データの解説

diff\_share:t期の選挙の得票差(0より大きいなら選挙に勝利したことを意味する)

myoutcomenext:t+1期の選挙の結果(選挙に勝ったら1となるダミー変数)

24. 金利スプレッド spread\_data.csv

15章と16章で用いた金利スプレッドのデータ

出所

以下の論文で用いたデータ。

Watanabe, T. and Yabu, T. (2019)“How large is the demand for Money at the ZLB? Evidence from Japan,” CARF Working Paper Series, CARF-F-465.

期間：1980年第1四半期-2017年第4四半期

変数の定義

r\_long: 長期金利(5年物国債金利)

r\_short: 短期金利(90-180日物国債金利)

25. 株価VAR var\_stockprice\_data.csv

15章で用いた日米の株価データ

出所

日経平均はFRED、ダウ工業平均は日経NEEDSより入手した。

期間：2010月1月5日-2018年12月28日(土日祝日を除く、

日米どちらかで休みならデータから除去している)

Dow: ＮＹダウ工業株３０種平均株価(終値)の変化率(%)

Nikkei: 日経平均225株価(終値)の変化率(%)

Time: 2010年1月5日を1とし、1づつ増加させた値である。2018年12月28日が2132となる。サンプルサイズは2132。

26. 見せかけの回帰 spurious\_data.csv

16章で用いた見せかけの回帰のデータ

PCを用いて生成された仮想データ

Time: 1から200まで、1づつ増加する値

Y: ランダムウォークから生成したデータ

X: ランダムウォークから生成したデータ

27. 原油の需給ショックoil2.csv

15章　新しく追加する予定の練習問題のデータ

出所

データは、次の研究で用いられたものです。

詳細は論文を参照してください。

Kilian, L. (2009) Not All Oil Price Shocks Are Alike: Disentangling Demand and Supply Shocks in the Crude Oil Market, American Economic Review 99(3), 1053-1069.

分析期間：1973年3月から2007年12月までの月次データ

oil: 世界原油生産量の変化率(%)

output: 世界経済活動指数

price: 実質原油価格