# Package ‘camerondata’

July 22, 2025

**Title** Datasets from ``Microeconometrics: Methods and Applications" by Cameron and Trivedi

**Version** 1.0.0

**Description** Quick and easy access to datasets that let you replicate the empirical examples in Cameron and Trivedi (2005) ``Microeconometrics: Methods and Applications" (ISBN: 9780521848053).The data are available as soon as you install and load the package (lazy-loading) as data frames. The documentation includes reference to chapter sections and page numbers where the datasets are used.

**License** CC BY 4.0

**Depends** R (>= 3.5.0)

**URL** https://github.com/juvlac/camerondata

**BugReports** https://github.com/juvlac/camerondata/issues

**Encoding** UTF-8

**LazyData** true

**RoxygenNote** 7.1.2

**NeedsCompilation** no

**Author** Juliana Vega-Lacorte [aut, cre]

**Maintainer** Juliana Vega-Lacorte <jv@jv-lacorte.de>

**Repository** CRAN

**Date/Publication** 2022-03-21 17:50:02 UTC

## Contents

fishing . . . . . 2
incpanel . . . . . 3
jobless . . . . . 4
laborpanel . . . . . 6
laborpanelprec . . . . . 7
nswproject . . . . . 8
patentsrd . . . . . 9

randhealth . . . . . 11
schooling . . . . . 13
strikes . . . . . 16
vietnamlss . . . . . 17
vietnam_hh . . . . . 18
vietnam_ind . . . . . 19

**Index** **21**

---

`fishing` *Fishing mode choice*

---

## Description

Data sample of 1,182 people from a survey conducted by Thomson and Crooke (1991) and analyzed by Herriges and Kling (1999). Cameron and Trivedi (2005).

## Usage

```
fishing
```

## Format

A data frame with 1182 observations and 16 variables:

**mode** fishing mode choice, = 1 beach, = 2 pier, = 3 private boat, = 4 charter boat

**price** price for chosen alternative, usd

**crate** catch rate for chosen alternative, sum of per-hour catch rates of targeted species.

**dbeach** = 1 if beach mode chosen, = 0 otherwise

**dpier** = 1 if pier mode chosen, = 0 otherwise

**dprivate** = 1 if private boat mode chosen, = 0 otherwise

**dcharter** = 1 if charter boat mode chosen, = 0 otherwise

**pbeach** price for beach mode, usd

**ppier** price for pier mode, usd

**pprivate** price for private boat mode, usd

**pcharter** price for charter boat mode, usd

**qbeach** catch rate for beach mode

**qpier** catch rate for pier mode

**qprivate** catch rate for private boat mode

**qcharter** catch rate for charter boat mode

**income** monthly income, usd

### Section in Text

14.2 Binary Outcome Example: Fishing Mode Choice, pp. 464-6, 486

15.2 Choice of Fishing Mode, pp. 491-5

### Source

http://cameron.econ.ucdavis.edu/mmabook/mmadata.html

### References

Cameron, A. and Trivedi, P. (2005), "Microeconometrics: Methods and Applications," Cambridge University Press, New York.

Herriges, J. and Kling, C. (1999), "Nonlinear Income Effects in Random Utility Models," Review of Economics and Statistics, 81, 62-72.

Thomson, C., and Crooke, S. (1991), "Results of the Southern California Sportfish Economic Survey," NOAA Technical Memorandum, National Marine Fisheries Service, Southwest Fisheries Science Center.

### Examples

```
summary(fishing)
```

---

## incpanel *Hourly wages*

---

### Description

Data from the Michigan Panel Survey of Income Dynamics, Individual Level Final Release 1993. Sample of 4856 women, extracted by Cameron and Trivedi (2005).

### Usage

```
incpanel
```

### Format

A data frame with 4856 observations and 9 variables:

**intnum** interview number 1968

**persnum** person number

**age** age of individual in 1993

**educatn** highest grade/year of school completed 1993

**earnings** total labor income of individual received in 1992, dollars

**hours** total annual work hours in 1992

**sex** sex of individual,= 2 if female

**kids** total number of children born to this individual

**married** last known marital status: 1 = married, 2 = never married, 3 = widowed, 4 = divorced, 5 = separated, 8 = NA, 9 = no histories 85-93

## Section in Text

9.2.1 Nonparametric density estimation, pp. 295 9.2.2 Nonparametric Regression, pp. 297

## Source

http://cameron.econ.ucdavis.edu/mmabook/mmadata.html

## References

Cameron, A. and Trivedi, P. (2005), "Microeconometrics: Methods and Applications," Cambridge University Press, New York.

Michigan Panel Study of Income Dynamics (PSID), https://psidonline.isr.umich.edu/

## Examples

```
summary(incpanel)
```

---

# jobless *Unemployment duration*

---

## Description

Data from the January Current Population Survey's Displaced Workers Supplements (DWS) for the years 1986, 1988, 1990, and 1992. Only individuals between 20 and 61 years old who were displaced from nonagricultural jobs due to plant closure, slack work, or abolished positions are included in the sample (McCall, 1996). Cameron and Trivedi (2005).

## Usage

```
jobless
```

## Format

A data frame with 3343 observations and 43 variables:

**spell** length of spell (joblessness duration) in number of two-week intervals

**censor1** = 1 if re-employed at full-time job

**censor2** = 1 if re-employed at part-time job

**censor3** = 1 if re-employed but left job: pt–ft status unknown

**censor4** = 1 if still jobless

**ui** = 1 if filed unemployment insurance claim

**reprate** eligible replacement rate, weekly benefit amount divided by weekly earnings in the lost job,

**logwage** log weekly earnings in lost job, 1985 prices

**tenure** years tenure in lost job

**disrate** eligible disregard rate

**slack** = 1 if lost job due to slack work

**abolpos** = 1 if lost job due to abolished position

**explose** = 1 if expected to lose job

**stateur** state unemployment rate, percent

**houshead** = 1 if household head

**married** = 1 if married

**female** = 1 if female

**child** = 1 if has children

**ychild** = 1 if has children five age and under

**nonwhite** = 1 if nonwhite

**age** age

**schlt12** = 1 if less than 12 years schooling

**schgt12** = 1 if more than 12 years schooling

**smsa** = 1 if resides in SMSA (standard metropolitan statistical area)

**bluecoll** = 1 if los job blue collar

**mining** = 1 if lost job in mining

**constr** = 1 if lost job in construction

**transp** = 1 if lost job in transportation

**trade** = 1 if lost job in trade

**fire** = 1 if lost job in finance, insurance and real estate sector

**services** = 1 if lost job in services sector

**pubadmin** = 1 if lost job in the public administration

**year85** = 1 if year of job loss is 1985

**year87** = 1 if year of job loss is 1987

**year89** = 1 if year of job loss is 1989

**midatl** = 1 if residence in Middle Atlantic

**encen** = 1 if residence in East North Central

**wncen** = 1 if residence in West North Central

**southatl** = 1 if residence in South Atlantic

**escen** = 1 if residence in East South Central

**wscen** = 1 if residence in West South Central

**mountain** = 1 if residence in Mountain region

**pacific** = 1 if residence in Pacific region

### Section in Text

17.11 Duration Example: Unemployment Duration, pp. 603-8, 632-6, 658-62

### Source

http://cameron.econ.ucdavis.edu/mmabook/mmadata.html

### References

Cameron, A. and Trivedi, P. (2005), "Microeconometrics: Methods and Applications," Cambridge University Press, New York.

McCall, B. (1996), Unemployment Insurance Rules, Joblessness, and Part-time Work," Econometrica, 64, 647-682.

### Examples

```
summary(jobless)
```

---

## laborpanel *Hours worked and wages*

---

### Description

Data on 532 males over 10 years (1979-1988) from Ziliak (1997), collected from the Panel Study of Income Dynamics.

### Usage

```
laborpanel
```

### Format

A data frame with 5320 observations and 8 variables:

**lnhr** log of annual hours worked

**lnwg** log of of hourly wage

**kids** number of children

**ageh** age

**agesq** quadratic age

**disab** = 1 if bad health

**id** identification code

**year** interview year

### Section in Text

21.3 Linear Panel Example: Hours and Wages, pp. 708-15

## Source

http://cameron.econ.ucdavis.edu/mmabook/mmadata.html

## References

Cameron, A. and Trivedi, P. (2005), "Microeconometrics: Methods and Applications," Cambridge University Press, New York.

Ziliak, J. (1997), "Efficient Estimation With Panel Data when Instruments are Predetermined: An Empirical Comparison of Moment-Condition Estimators," Journal of Business and Economic Statistics, 15, 419-431. https://amstat.tandfonline.com/doi/abs/10.1080/07350015.1997.10524720

Panel Study of Income Dynamics (PSID), https://psidonline.isr.umich.edu

## Examples

```
summary(laborpanel)
```

---

# laborpanelprec *Hours worked and wages (more precision)*

---

## Description

Data on 532 males over 10 years (1979-1988) from Ziliak (1997), with more significant digits (seven decimals) than the data originally posted on JBES website with two decimal places (Cameron and Trivedi, 2005).

## Usage

```
laborpanelprec
```

## Format

A data frame with 5320 observations and 8 variables:

**lnhr** log of annual hours worked

**lnwg** log of of hourly wage

**kids** number of children

**ageh** age

**agesq** quadratic age

**disab** = 1 if bad health

**id** identification code

**year** interview year ...

## Section in Text

22.3 Panel GMM Example: Hours and Wages, pp. 754-6

## Source

http://cameron.econ.ucdavis.edu/mmabook/mmadata.html

## References

Cameron, A. and Trivedi, P. (2005), "Microeconometrics: Methods and Applications," Cambridge University Press, New York.

Ziliak, J. (1997), "Efficient Estimation With Panel Data when Instruments are Predetermined: An Empirical Comparison of Moment-Condition Estimators," Journal of Business and Economic Statistics, 15, 419-431. https://amstat.tandfonline.com/doi/abs/10.1080/07350015.1997.10524720

Panel Study of Income Dynamics (PSID), https://psidonline.isr.umich.edu

## Examples

```
summary(laborpanelprec)
```

---

# nswproject *Training and earnings*

---

## Description

Data from the National Supported Work (NSW) demonstration project used by Lalonde (1986), and Dehejia and Wahba (1999, 2002). This sample has 185 observations in the treatment group and 2490 in the control group. The treatment sample consists of males who received training during 1976-1977. THe control group consists of male household heads under the age of 55 who are not retired, drawn from the PSID (Cameron and Trivedi, 2005).

## Usage

```
nswproject
```

## Format

A data frame with 2675 observations and 18 variables:

**treat** = 1 if individual is in treatment group, = 0 if in control group

**age** age in years

**educ** education in years

**black** = 1 if black

**hisp** = 1 if hispanic

**marr** = 1 if married

**re74** real annual earnings in 1974 (pre-treatment), in 1982 usd

**re75** real annual earnings in 1975 (pre-treatment), in 1982 usd

**re78** real annual earnings in 1978 (post-treatment), in 1982 usd

**u74** = 1 if unemployed in 1974

**u75** = 1 if unemployed in 1975

**agesq** age squared

**educsq** educ squared

**nodegree** = 1 if years of education < 12

**re74sq** re74 squared

**re75sq** re75 squared

**u74black** interaction term u74 x black

**u74hisp** interaction term u74 x hisp

## Section in Text

25.8 Treatment Evaluation Example: The Effect of Training on Earnings, pp. 889-95

## Source

http://cameron.econ.ucdavis.edu/mmabook/mmadata.html

## References

Cameron, A. and Trivedi, P. (2005), "Microeconometrics: Methods and Applications," Cambridge University Press, New York.

Dehejia R. and Wahba S. (1999), "Causal Effects in Nonexperimental Studies: Reevaluating the Evaluation of Training Programs," JASA, 1053-1062.

Dehejia R. and Wahba S. (2002), "Propensity-score Matching Methods for Nonexperimental Causal Studies", ReStat, 151-161

Lalonde, R. (1986), "Evaluating the Econometric Evaluations of Training Programs with Experimental Data," AER, 604-620.

## Examples

```
summary(nswproject)
```

---

# patentsrd *Patents and R&D*

---

## Description

Panel data on patents and R&D expenditures. The sample includes 346 firms with five years of data from 1975 to 1979 used by Hall, Griliches, and Hausman (1986).

## Usage

```
patentsrd
```

## Format

A data frame with 346 observations and 25 variables:

**cusip** Compustat's identifying number for the firm (Committee on Uniform Security Identification Procedures number).

**ardssic** A two-digit code for the applied R&D industrial classification.

**scisect** = 1 if firm is in the scientific sector.

**logk** log of the book value of capital in 1972.

**sumpat** sum of patents applied for between 1972-1979.

**logr70** log of R&D spending in 1970, in 1972 dollars.

**logr71** log of R&D spending in 1971, in 1972 dollars.

**logr72** log of R&D spending in 1972, in 1972 dollars.

**logr73** log of R&D spending in 1973, in 1972 dollars.

**logr74** log of R&D spending in 1974, in 1972 dollars.

**logr75** log of R&D spending in 1975, in 1972 dollars.

**logr76** log of R&D spending in 1976, in 1972 dollars.

**logr77** log of R&D spending in 1977, in 1972 dollars.

**logr78** log of R&D spending in 1978, in 1972 dollars.

**logr79** log of R&D spending in 1979, in 1972 dollars.

**pat70** number of patents applied in the year that were eventually granted (1970).

**pat71** number of patents applied in the year that were eventually granted (1971).

**pat72** number of patents applied in the year that were eventually granted (1972).

**pat73** number of patents applied in the year that were eventually granted (1973).

**pat74** number of patents applied in the year that were eventually granted (1974).

**pat75** number of patents applied in the year that were eventually granted (1975).

**pat76** number of patents applied in the year that were eventually granted (1976).

**pat77** number of patents applied in the year that were eventually granted (1977).

**pat78** number of patents applied in the year that were eventually granted (1978).

**pat79** number of patents applied in the year that were eventually granted (1979).

## Section in Text

23.3 Nonlinear Panel Example: Patents and R&D, pp. 792-5

## Source

http://cameron.econ.ucdavis.edu/mmabook/mmadata.html

## References

Cameron, A. and Trivedi, P. (2005), "Microeconometrics: Methods and Applications," Cambridge University Press, New York.

Hall, B., Griliches, Z. and Hausman J. (1986), "Patents and R and D: Is There a Lag?," International Economic Review, 27, issue 2, p. 265-83.

## Examples

```
summary(patentsrd)
```

---

`randhealth` *Health expenditures and insurance plans*

---

## Description

Data from the RAND Health Insurance Experiment. The data comes from Deb and Trivedi (2002). It includes variables on the number of contacts with a medical doctor, medical expenditures, demographics, health status, and insurance status. Cameron and Trivedi (2005).

## Usage

```
randhealth
```

## Format

A data frame with 20,190 observations and 45 variables:

**plan** health insurance plan number

**site** one of six sites where experiment was conducted

**coins** medical coinsurance

**tookphys** took baseline physical

**year** study year

**zper** person id, leading digit is sit

**black** = 1 if race of household head is black

**income** income based on annual income

**xage** age that year

**female** = 1 if person is female

**educdec** years of schooling of decision maker

**time** time eligible during the year

**outpdol** outpatient exp. excl. ment and

**drugdol** drugs purchased, outpatient

**suppdol** supplies purchased, outpatient

**mentdol** psychotherapy exp., outpatient

**inpdol** inpatient exp., facilities and md

**meddol** annual medical expenditures in constant dollars, excluding dental and outpatient mental

**totadm** number of hospital admissions

**inpmis** missing any inpatient charges

**mentvis** number psychotehrapy visits

**mdvis** number face-to-face md visits

**notmdvis** number face-to-face, not md visits

**num** family size

**mhi** mental health index, baseline

**disea** number of chronic diseases

**physlm** = 1 if person has physical limitation

**ghindx** general health index, baseline

**mdeoff** maximum expenditure offer

**pioff** participation incentive

**child** = 1 if age is less than 18

**fchild** = 1 if female child

**lfam** log of family size

**lpi** log of annual participation incentive payment or 0 if no payment

**idp** = 1 if individual deductible plan

**logc** log(coinsurance + 1) where coinsurance rate is 0 to 100

**fmde** log(max(medical deductible expenditure)) if idp=1 and mde>1, 0 otherwise

**hlthg** = 1 if self-rated health is good

**hlthf** = 1 if self-rated health is fair

**hlthp** = 1 if self-rated health is poor, (omitted is excellent)

**xghindx** ghi with imputation

**linc** log of annual family income, usd

**lnum** log of family size

**lnmeddol** log of medical expenditures given meddol > 0; missing otherwise

**binexp** = 1 if medical expenditures > 0

## Section in Text

16.6 Selection Models, pp. 553-6, 565 20.3 Count Example: Contacts with Medical Doctor, p.671

## Source

http://cameron.econ.ucdavis.edu/mmabook/mmadata.html

## References

Cameron, A. and Trivedi, P. (2005), "Microeconometrics: Methods and Applications," Cambridge University Press, New York.

Deb, P. and Trivedi, P.K. (2002), "The Structure of Demand for Health Care: Latent Class versus Two-Part Models," Journal of Health Economics, 21, 601-625.

RAND Corporation. "RAND's Health Insurance Experiment ." https://www.rand.org/health-care/projects/hie.html

## Examples

```
summary(randhealth)
```

schooling *Returns to schooling*

## Description

Data from the National Longitudinal Survey of Young Men. Cohort includes 3,010 males aged 24 to 34 years old in 1976, who were ages 14-24 when first interviewed in 1966. Cameron and Trivedi (2005)

## Usage

```
schooling
```

## Format

A data frame with 5226 observations and 101 variables:

**id** identification code

**black** = 1 if black race

**imigrnt** = 1 if born in the US

**hhead** person lived with at age 14 (in 1966)

**mag_14** = 1 if magazines available at age 14

**news_14** = 1 if newspapers available at age 14

**lib_14** = 1 if library card available at age 14

**num_sib** total number of siblings

**fgrade** highest grade completed by father (1966)

**mgrade** highest grade completed by mother (1966)

**iq** IQ score in 1968

**bdate** date of birth

**gfill76** highest grade completed 1976, some values filled from prevs reports

**wt76** sampling weights 1976

**grade76** highest grade completed in 1976

**grade66** highest grade completed in 1966

**age76** age in 1976

**age66** age in 1966

**smsa76** current residence, = 1 if lived in central city in 1976

**smsa66** current residence, = 1 if lived in central city in 1966

**region** census region in 1966

**col4** = 1 if there is a 4-year college nearby

**mcol4** = 1 if male 4-year college nearby

**col4pub** = 1 if public 4-year college nearby

**south76** = 1 if lived in South in 1976

**wage76** hourly wage in 1976, ln

**exp76** work experience in 1976, years calculated as (10 + age66) - grade76 - 6

**expsq76** experience 1976 squared, exp76^2/100

**agesq76** age squared (1976)

**reg1** region, = 1 if lived in region NE

**reg2** region, = 1 if lived in region MidAtl

**reg3** region, = 1 if lived in region ENC

**reg4** region, = 1 if lived in region WNC

**reg5** region, = 1 if lived in region SA

**reg6** region, = 1 if lived in region ESC

**reg7** region, = 1 if lived in region WSC

**reg8** region, = 1 if lived in region M

**reg9** region, = 1 if lived in region P

**momdad14** = 1 if lived with both parents at age 14

**sinmom14** = 1 if lived with mother only at age 14

**nodaded** = 1 if father has no formal education

**nomomed** = 1 if mother has no formal education

**daded** mean grade level of father

**momed** mean grade level of mother

**famed** father's and mother's education

**famed1** = 1 if mgrade> 12 & fgrade> 12

**famed2** = 1 if mgrade>=12 & fgrade>=12

**famed3** = 1 if mgrade==12 & fgrade==12

**famed4** = 1 if mgrade>=12 & fgrade==-1

**famed5** = 1 if fgrade>=12

**famed6** = 1 if mgrade>=12 & fgrade> -1

**famed7** = 1 if mgrade>=9 & fgrade>=9

**famed8** = 1 if mgrade> -1 & fgrade> -1

**famed9** = 1 if famed not in range 1-8

**int76** = 1 if wt76 not missing

**age1415** = 1 if in age group 14-15

**age1617** = 1 if in age group 16-17

**age1819** = 1 if in age group 18-19

**age2021** = 1 if in age group 20-21

**age2224** = 1 if in age group 22-24

**cage1415** = 1 if in age group 14-15 and lived near college
**cage1617** = 1 = 1 if in age group 16-17 and lived near college
**cage1819** = 1 if in age group 18-19 and lived near college
**cage2021** = 1 if in age group 20-21 and lived near college
**cage2224** = 1 if in age group 22-24 and lived near college
**cage66** age in 1966 and lived near college
**a1** = 1 if age in 1966 is 14
**a2** = 1 if age in 1966 is 15
**a3** = 1 if age in 1966 is 16
**a4** = 1 if age in 1966 is 17
**a5** = 1 if age in 1966 is 18
**a6** = 1 if age in 1966 is 19
**a7** = 1 if age in 1966 is 20
**a8** = 1 if age in 1966 is 21
**a9** = 1 if age in 1966 is 22
**a10** = 1 if age in 1966 is 23
**a11** = 1 if age in 1966 is 24
**ca1** = 1 if did not live near college in 1966
**ca2** = 1 if lived near college and age in 1966 = 14
**ca3** = 1 if lived near college and age in 1966 = 15
**ca4** = 1 if lived near college and age in 1966 = 16
**ca5** = 1 if lived near college and age in 1966 = 17
**ca6** = 1 if lived near college and age in 1966 = 18
**ca7** = 1 if lived near college and age in 1966 = 19
**ca8** = 1 if lived near college and age in 1966 = 20
**ca9** = 1 if lived near college and age in 1966 = 21
**ca10** = 1 if lived near college and age in 1966 = 22
**ca11** = 1 if lived near college and age in 1966 = 23
**ca12** = 1 if lived near college and age in 1966 = 24
**g25** grade level when 25 years old
**g25i** = 1 if =g25 and intrvwed in year used for determining g25
**intmo66** interview month in 1966, used to identify cases incl by Card
**nlsflt** flag to identify if the case was used by Card
**nsib** number of siblings
**ns1** = 1 if the person has no siblings
**ns2** = 1 if number of siblings is 2
**ns3** = 1 if number of siblings is 3
**ns4** = 1 if number of siblings is 4
**ns5** = 1 if number of siblings is 6
**ns6** = 1 if number of siblings is 9
**ns7** = 1 if number of siblings is 18

## Section in Text

4.9.6 Instrumental Variables Application, pp. 110-2

## Source

http://cameron.econ.ucdavis.edu/mmabook/mmadata.html

## References

Cameron, A. and Trivedi, P. (2005), "Microeconometrics: Methods and Applications," Cambridge University Press, New York.

Card, D. (1995), "Using Geographic Variation in College Proximity to Estimate the Returns to Schooling", in Aspects of Labor Market Behavior: Essays in Honor of John Vanderkamp, eds. L.N. Christofides et al., Toronto: University of Toronto Press, pp.201-221.

Kling, J.R. (2001) "Interpreting Instrumental Variables Estimates of the Return to Schooling," Journal of Business and Economic Statistics, 19, 358-364.

https://www.nlsinfo.org/content/cohorts/older-and-young-men

## Examples

```
summary(schooling)
```

---

# strikes *Strikes duration*

---

## Description

Data set on 566 contract strikes in U.S. manufacturing for the period 1968-76. The data has been used by Kennan (1985), Jaggia (1991), and others, and was originally published by the U.S. Department of Labor. Cameron and Trivedi (2005).

## Usage

```
strikes
```

## Format

A data frame with 566 observations and 2 variables:

**dur** duration of the strike, number of days from the start of the strike.

**gdp** measure of business cycle stage, deviation of monthly log industrial production in manufacturing.

## Section in Text

17.2 Duration Models, pp. 574-5, 582

## Source

http://cameron.econ.ucdavis.edu/mmabook/mmadata.html

## References

Cameron, A. and Trivedi, P. (2005), "Microeconometrics: Methods and Applications," Cambridge University Press, New York.

Kennan, J. (1985), "The Duration of Contract strikes in U.S. Manufacturing," Journal of Econometrics, 28, 5-28.

Jaggia, S. (1991), "Specification Tests Based on the Heterogeneous Generalized Gamma Model of Duration: With an Application to Kennan's Strike Data," Journal of Applied Econometrics, 6, 169–180.

## Examples

```
summary(strikes)
```

---

# vietnamlss *Household medical expenditure*

---

## Description

Data from the World Bank's 1997 Vietnam Living Standards Survey 1997-98 at the household level. Cameron and Trivedi (2005)

## Usage

```
vietnamlss
```

## Format

A data frame with 5999 observations and 9 variables:

**sex** gender of household head, 1 = male; 2 = female

**age** age of household head

**educyr98** schooling year of household head

**farm** type of household, = 1 if farm

**urban98** = 1 if urban area, = 0 if rural area

**hhsize** household size

**lhhexp1** household total expenditure, ln

**lhhex12m** household medical expenditure, ln

**lnrlfood** household food expenditure, ln

### Section in Text

4.6.4 Quantile Regression Example, pp. 88-90

### Source

http://cameron.econ.ucdavis.edu/mmabook/mmadata.html

### References

Cameron, A. and Trivedi, P. (2005), "Microeconometrics: Methods and Applications," Cambridge University Press, New York.

World Bank Living Standards Survey 1997-1998 Vietnam. https://microdata.worldbank.org/index.php/catalog/2694

### Examples

```
summary(vietnamlss)
```

---

## vietnam_hh *Vietnam health care use (household level)*

---

### Description

Data from the World Bank's Vietnam Living Standards Survey of 1997-1998 at the household level. Sample extract by Cameron and Trivedi (2005).

### Usage

```
vietnam_hh
```

### Format

A data frame with 5999 observations and 8 variables:

**sex** = 1 if head of household is female

**age** age of head of household

**educ** Highest education obtained by head of household

**farm** = 1 for agricultural household

**hhsize** household size

**commune** commune code

**lnhhexp** total household expenditure, ln

**lnexp12m** household healthcare expenditure in the past 12 months, ln

### Section in Text

24.7 Clustering Example: Vietnam Health Care Use, pp 848-53

## Source

http://cameron.econ.ucdavis.edu/mmabook/mmadata.html

## References

Cameron, A. and Trivedi, P. (2005), "Microeconometrics: Methods and Applications," Cambridge University Press, New York.

World Bank Living Standards Survey 1997-1998 Vietnam. https://microdata.worldbank.org/index.php/catalog/2694

## Examples

```
summary(vietnam_hh)
```

---

# vietnam_ind *Vietnam health care use (individual level)*

---

## Description

Data from the World Bank's Vietnam Living Standards Survey of 1997-1998 at the individual level. Sample extract by Cameron and Trivedi (2005).

## Usage

```
vietnam_ind
```

## Format

A data frame with 27766 observations and 12 variables:

**educ** Completed diploma level

**sex** = 1 if respondent is male

**age** age in years

**married** = 1 for married person

**illness** number of illnesses experienced in past 12 months

**injury** = 1 if injured during survey period

**illdays** number of illness days

**actdays** number od days of limited activity

**pharvis** number of direct pharmacy visits

**insurance** = 1 if respondent has health insurance coverage

**lnhhexp** total household expenditure, ln

**commune** commune code

## Section in Text

Section

## Source

http://cameron.econ.ucdavis.edu/mmabook/mmadata.html

## References

Cameron, A. and Trivedi, P. (2005), "Microeconometrics: Methods and Applications," Cambridge University Press, New York.

World Bank Living Standards Survey 1997-1998 Vietnam. https://microdata.worldbank.org/index.php/catalog/2694

## Examples

```
summary(vietnam_ind)
```

# Index

∗ **datasets**
fishing, 2
incpanel, 3
jobless, 4
laborpanel, 6
laborpanelprec, 7
nswproject, 8
patentsrd, 9
randhealth, 11
schooling, 13
strikes, 16
vietnam_hh, 18
vietnam_ind, 19
vietnamlss, 17

fishing, 2

incpanel, 3

jobless, 4

laborpanel, 6
laborpanelprec, 7

nswproject, 8

patentsrd, 9

randhealth, 11

schooling, 13
strikes, 16

vietnam_hh, 18
vietnam_ind, 19
vietnamlss, 17