Effect of Education and Residence on Life Satisfaction – WVS China 2018

Alex Shen

2025-07-11

Data and Methods

World Values Survey, Wave 7 (China, 2018). Logistic regression predicting high life satisfaction (score ≥ 7) by education, urban/rural status, age, and sex, with education × urban interaction.

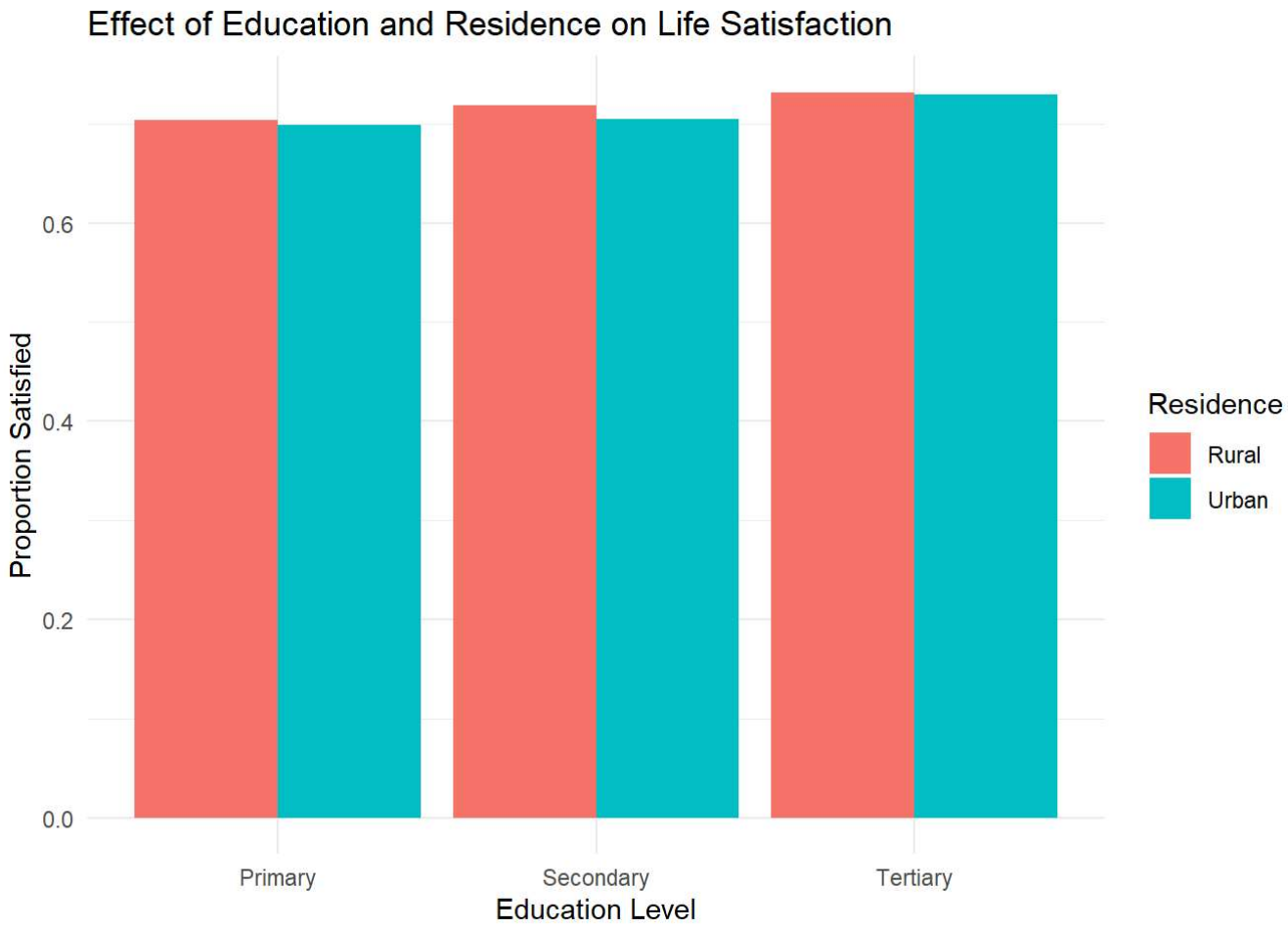
```
# Adjust the path as needed
df_raw <- read_excel("data/F00013183-WVS_Wave_7_China_Excel_v5.1.xlsx", sheet = 1)
df <- df_raw %>%
  transmute(
    life_satisfaction = as.numeric(`Q49: Satisfaction with your life`),
    education = factor(
      as.numeric(`Q275R: Highest educational level: Respondent (recoded into 3 groups)`),
      levels = 1:3,
      labels = c("Primary", "Secondary", "Tertiary")
    ),
    urban = as.numeric(`H_URBRURAL: Urban-Rural`),
    age = as.numeric(`Q262: Age`),
    sex = as.numeric(`Q260: Sex`)
  ) %>%
  filter(
    !is.na(life_satisfaction), !is.na(education), !is.na(urban), !is.na(age), !is.na(sex)
  ) %>%
  mutate(
    urban_bin = if_else(urban == 1, 1, 0),
    satisfied = if_else(life_satisfaction >= 7, 1, 0)
  )
nrow(df)
```

```
## [1] 3006
```

```
model <- glm(satisfied ~ education * urban_bin + age + sex, data = df, family = "binomial")
summary(model)
```

```
##
## Call:
## glm(formula = satisfied ~ education * urban_bin + age + sex,
##     family = "binomial", data = df)
##
## Coefficients:
##                               Estimate Std. Error z value Pr(>|z|)
## (Intercept)                  -0.135423   0.222617  -0.608   0.5430
## educationSecondary            0.295802   0.188457   1.570   0.1165
## educationTertiary             0.496872   0.250355   1.985   0.0472 *
## urban_bin                    -0.022624   0.108451  -0.209   0.8347
## age                           0.017344   0.003169   5.473 4.43e-08 ***
## sex                           0.087271   0.081581   1.070   0.2847
## educationSecondary:urban_bin -0.122208   0.222885  -0.548   0.5835
## educationTertiary:urban_bin  -0.076170   0.271535  -0.281   0.7791
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 3624.2  on 3005  degrees of freedom
## Residual deviance: 3591.3  on 2998  degrees of freedom
## AIC: 3607.3
##
## Number of Fisher Scoring iterations: 4
```

```
df %>%
  group_by(education, urban_bin) %>%
  summarise(mean_satisfied = mean(satisfied, na.rm = TRUE)) %>%
  ggplot(aes(x = education, y = mean_satisfied, fill = factor(urban_bin, labels = c("Rural", "U
rban")))) +
  geom_bar(stat = "identity", position = "dodge") +
  labs(
    x = "Education Level",
    y = "Proportion Satisfied",
    fill = "Residence",
    title = "Effect of Education and Residence on Life Satisfaction"
  ) +
  theme_minimal()
```


Effect of Education and Residence on Life Satisfaction
Proportion Satisfied
0.0
0.2
0.4
0.6
Primary
Secondary
Tertiary
Education Level
Residence
Rural
Urban