Homogenous demand, conduct exercise

# 1 Import & clean data, produce descriptive statistics

## Import data

First, import the data into a data frame, called coffee\_raw:

coffee\_raw <- read.csv("dutch\_coffee.csv")

## Deflate prices using price index

Next, deflate the prices, save a new data frame, coffee:

##   
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':  
##   
## filter, lag

## The following objects are masked from 'package:base':  
##   
## intersect, setdiff, setequal, union

## Explore variations in price over time

Convert year/month variable into R time format, using as.Date and as.yearmon from the zoo package:

library(zoo)

##   
## Attaching package: 'zoo'

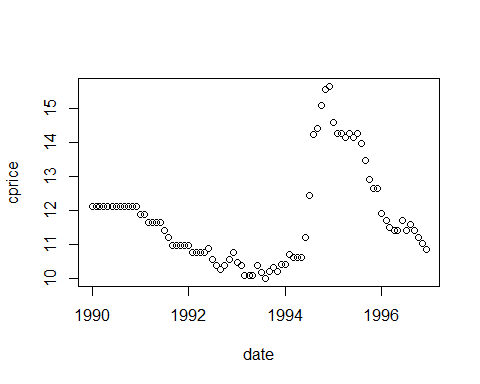
## The following objects are masked from 'package:base':  
##   
## as.Date, as.Date.numeric

coffee <- coffee %>%  
 mutate(date = as.Date(as.yearmon(month)))

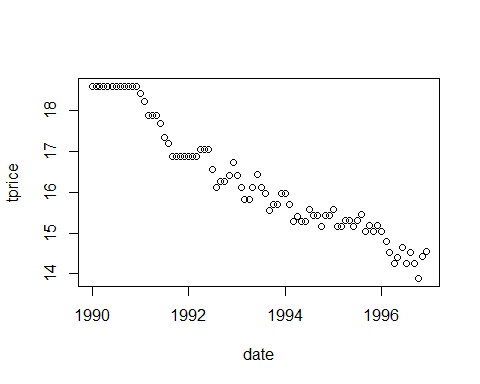
## Make some nice line graphs

Create simple line graphs of key price and income variables over time.

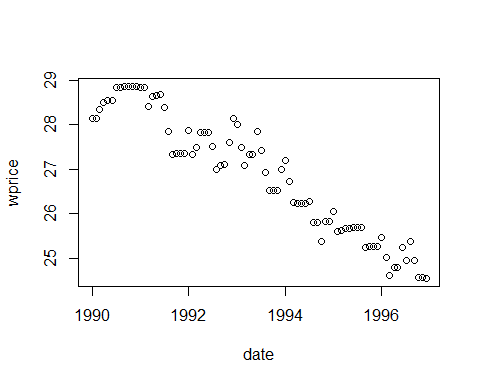
Basic plot of coffee price over time.



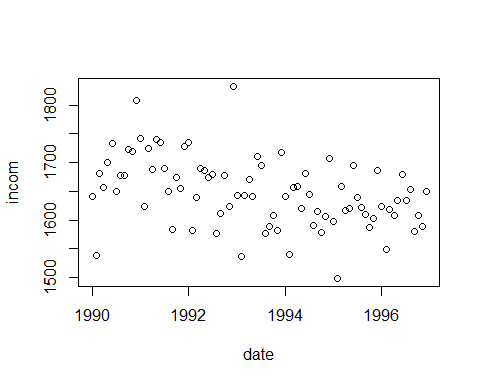
Basic plot of tea price against time:

 Plot of wage prices over time:

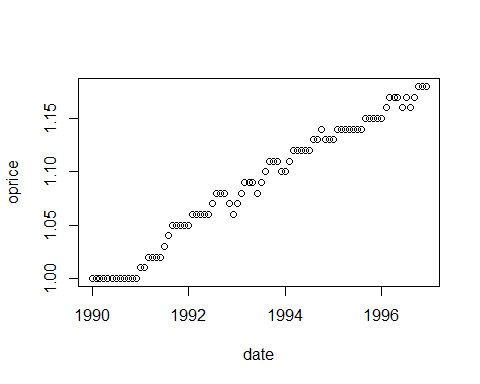
plot(wprice ~ date, coffee)

 Basic plot of income against time:

plot(incom ~ date, coffee)

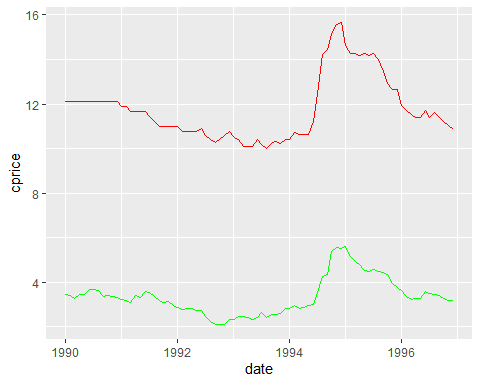
 Plot of other prices over time:

plot(oprice ~ date, coffee)



Now plot coffee retail price against the bean price using ggplot2 and geom\_line:

library(ggplot2)  
  
ggplot(coffee,  
 aes(date)) +   
 geom\_line(aes(y = cprice), colour = "red") +   
 geom\_line(aes(y = bprice), colour = "green")

 ## Summary statistics

Produce a table of summary statistics for qu, cprice, tprice, wprice, bprice, incom using dplyr select:

coffee %>%  
 select(qu, cprice, tprice, wprice, bprice, incom) %>%  
 summary()

## qu cprice tprice wprice   
## Min. :0.5200 Min. :10.02 Min. :13.89 Min. :24.55   
## 1st Qu.:0.6300 1st Qu.:10.70 1st Qu.:15.27 1st Qu.:25.68   
## Median :0.6600 Median :11.42 Median :15.98 Median :27.08   
## Mean :0.6815 Mean :11.76 Mean :16.28 Mean :26.87   
## 3rd Qu.:0.7400 3rd Qu.:12.12 3rd Qu.:17.05 3rd Qu.:27.90   
## Max. :1.0400 Max. :15.66 Max. :18.60 Max. :28.86   
## bprice incom   
## Min. :2.118 Min. :1498   
## 1st Qu.:2.812 1st Qu.:1608   
## Median :3.280 Median :1643   
## Mean :3.363 Mean :1648   
## 3rd Qu.:3.611 3rd Qu.:1686   
## Max. :5.623 Max. :1833

Now consider summary statistics for quarter 1 alone:

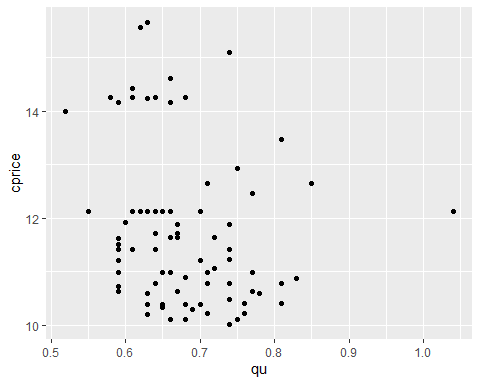
coffee %>%  
 filter(q1 == 1) %>%  
 select(qu, cprice, tprice, wprice, bprice, incom) %>%  
 summary()

## qu cprice tprice wprice   
## Min. :0.5500 Min. :10.10 Min. :14.54 Min. :24.61   
## 1st Qu.:0.5900 1st Qu.:10.72 1st Qu.:15.27 1st Qu.:26.06   
## Median :0.6500 Median :11.65 Median :16.11 Median :27.34   
## Mean :0.6529 Mean :11.68 Mean :16.50 Mean :27.07   
## 3rd Qu.:0.7100 3rd Qu.:12.12 3rd Qu.:17.88 3rd Qu.:28.15   
## Max. :0.7600 Max. :14.61 Max. :18.60 Max. :28.83   
## bprice incom   
## Min. :2.323 Min. :1498   
## 1st Qu.:2.843 1st Qu.:1582   
## Median :3.206 Median :1639   
## Mean :3.333 Mean :1624   
## 3rd Qu.:3.400 3rd Qu.:1656   
## Max. :5.623 Max. :1742

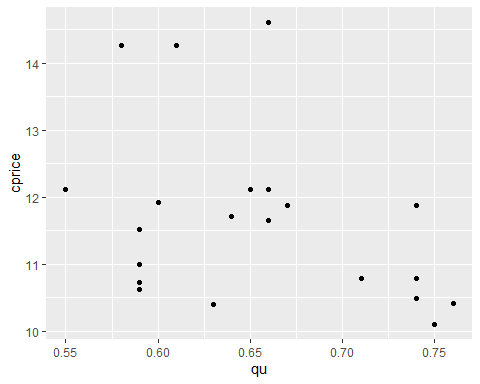
Now do the same for q2, q3, and q4.

## Considering prices and quantities

First, use ggplot to produce a scatter graph of prices and quantities using geom\_point:

 Now do the same thing but only for the first quarter:

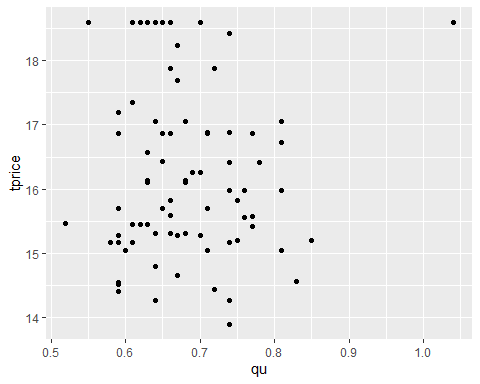
coffee %>%   
 filter(q1 == 1) %>%  
 ggplot(aes(x = qu)) +  
 geom\_point(aes(y = cprice))



Now consider for q2, q3 and q4.

## Assess relationship between tea price and quantity demanded of coffee

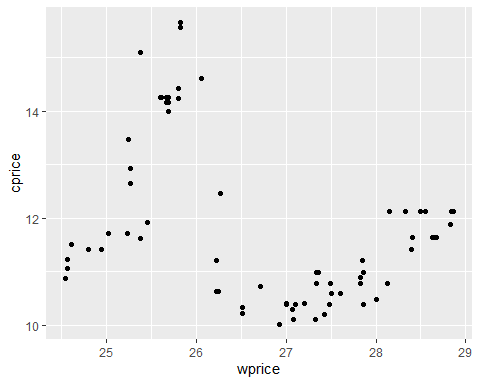
coffee %>%  
 ggplot(aes(x = qu)) +  
 geom\_point(aes(y = tprice))



## Price of coffee and labor costs

Consider the price of coffee and wages:

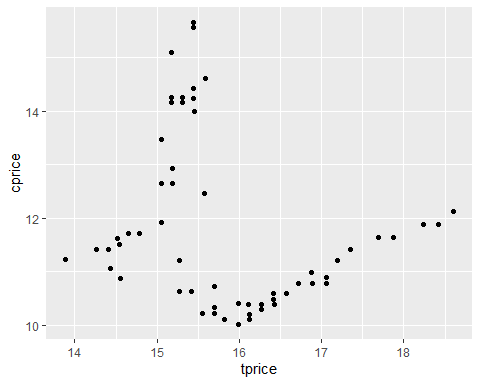
coffee %>%  
 ggplot(aes(x = wprice)) +   
 geom\_point(aes(y = cprice))



## Price of coffee and price of tea

Look at the relationship between the price of coffee and the price of tea:

coffee %>%  
 ggplot(aes(x = tprice)) +   
 geom\_point(aes(y = cprice))



## Overall correlations

Produce a correlation matrix for qu, cprice, tprice, wprice, bprice, incom:

coffee %>%  
 select(qu, cprice, tprice, wprice, bprice, incom) %>%  
 cor()

## qu cprice tprice wprice bprice incom  
## qu 1.00000000 -0.2243179 -0.01592424 0.01687637 -0.2367654 0.2352513  
## cprice -0.22431789 1.0000000 -0.11878101 -0.30051704 0.9595010 -0.1068931  
## tprice -0.01592424 -0.1187810 1.00000000 0.94795657 -0.1978217 0.4703265  
## wprice 0.01687637 -0.3005170 0.94795657 1.00000000 -0.3774048 0.5145200  
## bprice -0.23676540 0.9595010 -0.19782168 -0.37740483 1.0000000 -0.1657889  
## incom 0.23525135 -0.1068931 0.47032651 0.51452003 -0.1657889 1.0000000

## create log variables

We need our variables to be in logs:

# Considering the basic relationships

## Basic demand regressions

Regress log quantity on log price and seasonal dummies and store as q\_p1:

q\_p1 <- lm(ln\_qu ~ ln\_cprice + q1 + q2 + q3, data = coffee)  
summary(q\_p1)

##   
## Call:  
## lm(formula = ln\_qu ~ ln\_cprice + q1 + q2 + q3, data = coffee)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -0.18478 -0.06511 -0.00992 0.06617 0.35404   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 0.31947 0.23357 1.368 0.175270   
## ln\_cprice -0.25424 0.09427 -2.697 0.008552 \*\*   
## q1 -0.12698 0.03032 -4.188 7.26e-05 \*\*\*  
## q2 -0.09192 0.03035 -3.029 0.003315 \*\*   
## q3 -0.11786 0.03029 -3.891 0.000207 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.09815 on 79 degrees of freedom  
## Multiple R-squared: 0.2648, Adjusted R-squared: 0.2275   
## F-statistic: 7.112 on 4 and 79 DF, p-value: 6.07e-05

Including tea price, save as q\_p2:

q\_p2 <- lm(ln\_qu ~ ln\_cprice + q1 + q2 + q3 + ln\_tprice, data = coffee)  
summary(q\_p2)

##   
## Call:  
## lm(formula = ln\_qu ~ ln\_cprice + q1 + q2 + q3 + ln\_tprice, data = coffee)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -0.18522 -0.06472 -0.01006 0.06628 0.35637   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 0.36503 0.45793 0.797 0.427794   
## ln\_cprice -0.25534 0.09534 -2.678 0.009022 \*\*   
## q1 -0.12659 0.03070 -4.124 9.23e-05 \*\*\*  
## q2 -0.09166 0.03062 -2.994 0.003692 \*\*   
## q3 -0.11779 0.03049 -3.864 0.000229 \*\*\*  
## ln\_tprice -0.01545 0.13324 -0.116 0.908002   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.09876 on 78 degrees of freedom  
## Multiple R-squared: 0.2649, Adjusted R-squared: 0.2178   
## F-statistic: 5.621 on 5 and 78 DF, p-value: 0.000178

Include income, save as q\_p3:

q\_p3 <- lm(formula = ln\_qu ~ ln\_cprice + q1 + q2 + q3 + ln\_tprice + ln\_incom, data = coffee)  
summary(q\_p3)

##   
## Call:  
## lm(formula = ln\_qu ~ ln\_cprice + q1 + q2 + q3 + ln\_tprice + ln\_incom,   
## data = coffee)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -0.18762 -0.06862 -0.00937 0.05964 0.33514   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -2.44664 2.62445 -0.932 0.354122   
## ln\_cprice -0.24988 0.09536 -2.620 0.010577 \*   
## q1 -0.11409 0.03274 -3.485 0.000816 \*\*\*  
## q2 -0.09227 0.03059 -3.017 0.003462 \*\*   
## q3 -0.10914 0.03147 -3.468 0.000861 \*\*\*  
## ln\_tprice -0.10242 0.15525 -0.660 0.511407   
## ln\_incom 0.40981 0.37667 1.088 0.279994   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.09865 on 77 degrees of freedom  
## Multiple R-squared: 0.276, Adjusted R-squared: 0.2196   
## F-statistic: 4.893 on 6 and 77 DF, p-value: 0.0002767

# Estimating theta

## Instrumental variable regressions

Run an instrumental variable regression, using prices, quantities, seasonal dummies and instrument, first without team and income, and secondly with tea and income:

## Loading required package: car

## Loading required package: carData

##   
## Attaching package: 'car'

## The following object is masked from 'package:dplyr':  
##   
## recode

## Loading required package: lmtest

## Loading required package: sandwich

## Loading required package: survival

##   
## Call:  
## ivreg(formula = ln\_qu ~ q1 + q2 + q3 + ln\_cprice | q1 + q2 +   
## q3 + ln\_bprice, data = coffee)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -0.181011 -0.067238 -0.008444 0.067356 0.354644   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 0.37349 0.24693 1.513 0.134386   
## q1 -0.12730 0.03033 -4.197 7.04e-05 \*\*\*  
## q2 -0.09235 0.03036 -3.042 0.003194 \*\*   
## q3 -0.11789 0.03030 -3.891 0.000207 \*\*\*  
## ln\_cprice -0.27613 0.09970 -2.770 0.006994 \*\*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.09818 on 79 degrees of freedom  
## Multiple R-Squared: 0.2643, Adjusted R-squared: 0.227   
## Wald test: 7.207 on 4 and 79 DF, p-value: 5.322e-05

## Now lets show the adjusted learner index

Compute the adjusted lerner index using our cost information and our estimate of :

## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## 0.07434 0.08372 0.08748 0.08855 0.09292 0.10597

## Ln\_adj   
## Min. :0.07434   
## 1st Qu.:0.08445   
## Median :0.08834   
## Mean :0.08837   
## 3rd Qu.:0.09311   
## Max. :0.09822

Now, consider for q2, q3, and q4.

## Estimating conduct parameter

conduct1 <- lm(cprice ~ cost + q1 + q2 + q3 - 1, data = coffee) # omits intercept  
summary(conduct1)

##   
## Call:  
## lm(formula = cprice ~ cost + q1 + q2 + q3 - 1, data = coffee)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -1.06569 -0.28829 -0.03361 0.28889 1.00238   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## cost 1.466267 0.010910 134.398 <2e-16 \*\*\*  
## q1 0.002807 0.125644 0.022 0.982   
## q2 -0.007612 0.125362 -0.061 0.952   
## q3 0.084503 0.126116 0.670 0.505   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.4158 on 80 degrees of freedom  
## Multiple R-squared: 0.9988, Adjusted R-squared: 0.9988   
## F-statistic: 1.703e+04 on 4 and 80 DF, p-value: < 2.2e-16

conduct2 <- lm(cprice ~ cost -1, data = coffee) # omits intercept  
summary(conduct2)

##   
## Call:  
## lm(formula = cprice ~ cost - 1, data = coffee)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -1.08922 -0.28022 -0.05497 0.29080 0.98628   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## cost 1.468731 0.005548 264.7 <2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.4099 on 83 degrees of freedom  
## Multiple R-squared: 0.9988, Adjusted R-squared: 0.9988   
## F-statistic: 7.007e+04 on 1 and 83 DF, p-value: < 2.2e-16

From the FOC we have: This is the estimated coefficient for cost.

b=1.47  
theta=gamma\*(1-b)/b  
summary(theta)

## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## 0.08856 0.08856 0.08856 0.08856 0.08856 0.08856

n=1/theta  
summary(n)

## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## 11.29 11.29 11.29 11.29 11.29 11.29

## Linear demand

linear1 <- ivreg(qu ~ incom + q1 + q2 + q3 + cprice | bprice + incom + q1 + q2 + q3, data = coffee)  
summary(linear1)

##   
## Call:  
## ivreg(formula = qu ~ incom + q1 + q2 + q3 + cprice | bprice +   
## incom + q1 + q2 + q3, data = coffee)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -0.116823 -0.049428 -0.009419 0.041008 0.279560   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 0.6353799 0.2493831 2.548 0.012806 \*   
## incom 0.0001617 0.0001392 1.162 0.248701   
## q1 -0.0839335 0.0224717 -3.735 0.000355 \*\*\*  
## q2 -0.0699218 0.0217508 -3.215 0.001901 \*\*   
## q3 -0.0783453 0.0221914 -3.530 0.000700 \*\*\*  
## cprice -0.0138031 0.0057188 -2.414 0.018139 \*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.07023 on 78 degrees of freedom  
## Multiple R-Squared: 0.2691, Adjusted R-squared: 0.2222   
## Wald test: 5.781 on 5 and 78 DF, p-value: 0.0001366

linear2 <- ivreg(qu ~ cprice + q1 + q2 + q3 | bprice + q1 + q2 + q3, data = coffee)  
summary(linear2)

##   
## Call:  
## ivreg(formula = qu ~ cprice + q1 + q2 + q3 | bprice + q1 + q2 +   
## q3, data = coffee)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -0.113895 -0.046497 -0.009371 0.044152 0.302896   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 0.918705 0.069596 13.201 < 2e-16 \*\*\*  
## cprice -0.014984 0.005707 -2.626 0.01038 \*   
## q1 -0.090782 0.021763 -4.171 7.71e-05 \*\*\*  
## q2 -0.068903 0.021786 -3.163 0.00222 \*\*   
## q3 -0.083902 0.021730 -3.861 0.00023 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.07041 on 79 degrees of freedom  
## Multiple R-Squared: 0.2558, Adjusted R-squared: 0.2181   
## Wald test: 6.939 on 4 and 79 DF, p-value: 7.7e-05

Compute adjusted learner index:

coffee$gamma\_l=-0.015\*coffee$cprice/coffee$qu  
coffee$Ln\_adj\_l=-coffee$gamma\_l\*coffee$Ln

Now summarise the lerner indices:

coffee %>%  
 select(Ln\_adj\_l) %>%  
 summary()

## Ln\_adj\_l   
## Min. :0.05623   
## 1st Qu.:0.07298   
## Median :0.08197   
## Mean :0.08407   
## 3rd Qu.:0.09099   
## Max. :0.13442

Now summarise by quarters.

## Demand rotaters

Create seasonal demand rotators:

coffee <- coffee %>%  
 mutate(cprice\_q1=cprice\*q1,  
 bprice\_q1=bprice\*q1,  
 cprice\_q2=cprice\*q2,  
 bprice\_q2=bprice\*q2,  
 cprice\_q3=cprice\*q3,  
 bprice\_q3=bprice\*q3,  
 cprice\_q4=cprice\*q4,  
 bprice\_q4=bprice\*q4  
 )

Can also use use demand rotators in regression:

rotator <- ivreg(qu ~ q1 + q2 + q3 + (cprice + cprice\_q1) |   
 (bprice + bprice\_q1) + q1 + q2 + q3 , data = coffee)  
summary(rotator)

##   
## Call:  
## ivreg(formula = qu ~ q1 + q2 + q3 + (cprice + cprice\_q1) | (bprice +   
## bprice\_q1) + q1 + q2 + q3, data = coffee)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -0.111909 -0.047215 -0.009324 0.041327 0.302610   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 0.903591 0.078116 11.567 < 2e-16 \*\*\*  
## q1 -0.028064 0.166265 -0.169 0.86640   
## q2 -0.068552 0.021913 -3.128 0.00247 \*\*   
## q3 -0.083854 0.021842 -3.839 0.00025 \*\*\*  
## cprice -0.013713 0.006438 -2.130 0.03631 \*   
## cprice\_q1 -0.005345 0.014053 -0.380 0.70473   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.07077 on 78 degrees of freedom  
## Multiple R-Squared: 0.2577, Adjusted R-squared: 0.2101   
## Wald test: 5.503 on 5 and 78 DF, p-value: 0.0002165