

# The Demand for Coffee in Ruritania, 1987-2016

Z0983692

May 6, 2019

## 1 Introduction

In this report I estimate the determinants of the demand for coffee in Ruritania from a multivariate time series over 1987-2016. I find empirical evidence for a double-log static model of demand.

### 1.1 Summary Statistics and Hypotheses

The dependent variable is *QCOFFE10*: the total consumption of coffee. The standard deviation ( $s_{\text{QCOFF}} = 21648.4$ ) of coffee consumption in Ruritania is far greater than a comparable time series from Germany (Deutscher Kaffeverband n.d.). The prices of a bundle of consumer goods are listed in the dataset; there is also an index of all other goods in the consumption bundle. The table below lists the variables along with hypotheses regarding their effect on coffee demand. I hypothesise that coffee demand is independent to other food items like alcoholic beverages and food groceries, as coffee is commonly bought in either a cafe or a grocery setting. While grocery-bought demand for coffee may be affected by other grocery prices, as they are usually purchased under the same budget constraint, I believe that this split in consumption method along with the addictive effect provides a strong basis for its aggregative independence.

The income data is assumed to be disposable income. We assume that the coffee supply price to the consumer price transmission mechanism is constant in Ruritania. Additionally I hypothesise

Table 1: List of variables and their hypothesised relationship to demand for coffee.

Variable	Definition	Proposed Relationship
QCOFFE10 ( $Q_C$ )	Quantity demanded coffee (total)	n/a
PCOFF ( $P_C$ )	Price of coffee	Strong Own-price elastic ( $\eta_C < 0$ )
PTEA ( $P_T$ )	Price of tea	Strong Substitute ( $\eta_T > 0$ )
INCOME ( $Y$ )	Household income	Strong Necessity Good ( $0 < \beta_Y < 1$ )
PBEER ( $P_B$ )	Price of beer	None
PWINE ( $P_W$ )	Price of wine	None
PMTFH ( $P_M$ )	Price of meat & fish	Weak complement ( $\eta_{<0}$ )
PFTVG ( $P_F$ )	Price of fruit & vegetables	Weak complement ( $\eta_F < 0$ )
PLEIS ( $P_L$ )	Price of leisure	None
PTRAV ( $P_{TR}$ )	Price of travel	None
PALLOTH ( $P_A$ )	Weighted index of all other prices: electricity bills <i>etc.</i>	Strong complement due to general income effects ( $\eta_A < 0$ )

that coffee demand will obey Engel’s Law where the income share spent on foodstuffs decreases with increased income, implying that  $0 < \eta_Y < 1$ . Coffee should also be relatively inelastic due to habit formation effects. As other grocery items, meat and fish and fruit and vegetables should have a weakly complementary effect, as their increase in price causes the ‘weekly shop’ to become more expensive.

Visual inspection of the time series (Figure 1) yields important observations. *QCOFFE10* has very large outliers until 1998, after which there is a notable decrease in standard deviation:  $s_{87-98} = 23782$ ;  $s_{98-16} = 9074.5$ . This suggests a structural break and that a log transformation should be performed to stabilise the variance. *PTEA* - a potential substitute for coffee - trends downwards. *INCOME* and *PCOFF* both show an upward nominal trend, as would be expected over time.

## 1.2 Interpretation of Preliminary Model: Theoretical Framework

Classical demand theory places a number of constraints on our framework. Assuming the primal form of the demand function, consumers maximise utility through choosing a set bundle of goods to match their expenditure constraint. Demand functions follow from this maximisation procedure,

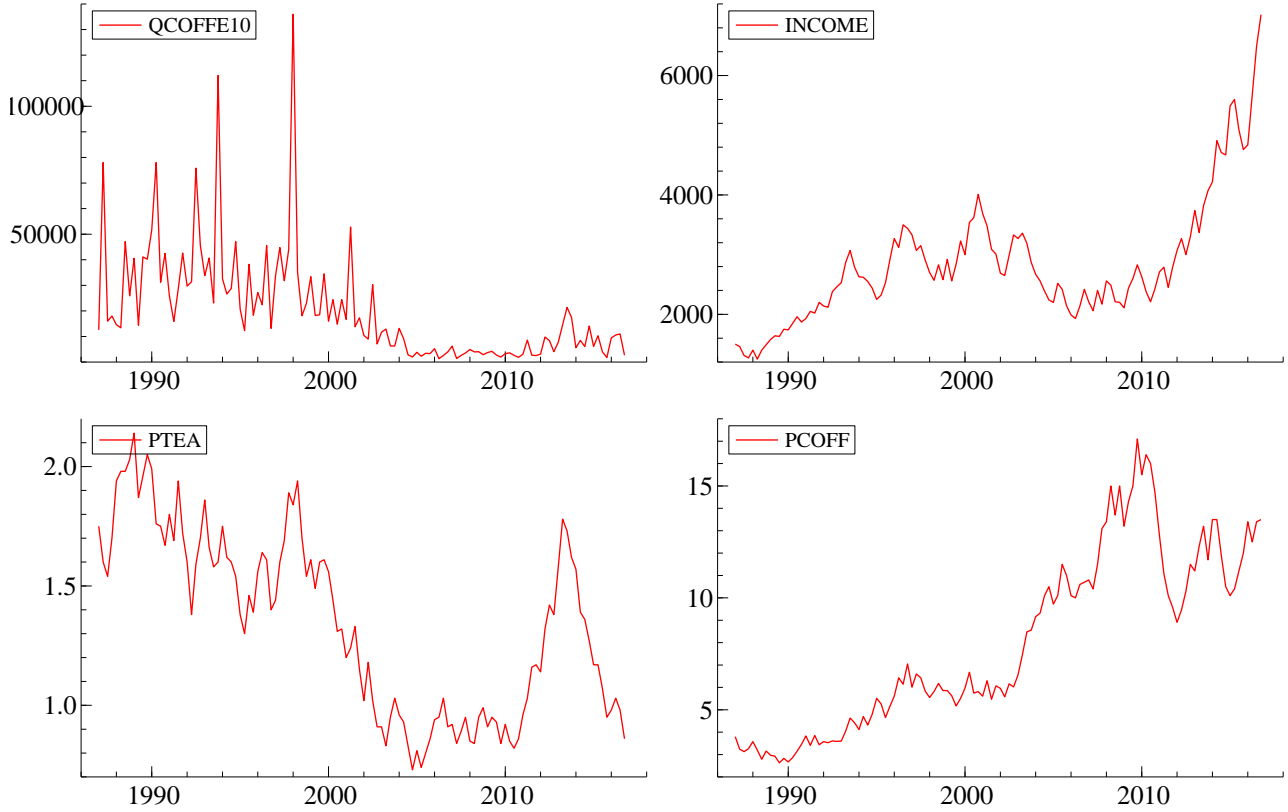


Figure 1: Time series of  $PCOFF$ ,  $QCOFFE10$ ,  $PTEA$ ,  $INCOME$ .

for which we estimate a single equation. The dataset does not include supply-side determinants of demand, thereby precluding a simultaneous equations model. There is no data available on the quantity demanded of other consumer goods, ruling out a demand-system approach in the form of the Almost Ideal Demand System, which can accommodate more of the theoretical properties of a demand function (Deaton and Muellbauer 1980). According to standard microeconomic theory, there are four salient properties of our demand function: additivity (sum of expenditure share equals income), homogeneity (no price effect), symmetry (cross-price compensated elasticities for two goods are equal), and separability (consumers have preferences regarding separate goods) (Snyder and Nicholson 2012). Since we only have quantity data for coffee, we assume additivity, symmetry, and separability. I test for homogeneity in section 4.3.

## 2 Choice of Functional Form

Log-Log formulations have been used widely in estimating single equations of demand for coffee (Okunade 1992). Alternatives are estimating a demand system or a flexible Box-Cox single equation. I compare four functional forms: Linear, Log-Lin, Lin-Log, and Log-Log with diagnostic tests: Jarque-Bera test for normality of the error term, Durbin-Watson and Breusch-Godfrey test for serial correlation, White test for heteroskedasticity, Ramsey-RESET test for functional form misspecification, and F test for overall significance. The fit is compared using the adjusted  $R^2$ , Schwartz, and Akaike Information Criteria. After selecting the most appropriate functional form, I then perform t- and f-tests to drop insignificant variables. I then test whether the model can be restricted to match some characteristics suggested by theory: habit formation, seasonality, and a structural break. The significance level is  $\alpha = 0.05$ .

Table 2: Summary of Diagnostic Tests with decision rules.

Test	Null Hypothesis	Decision Rule (fail to reject null)
F test	Variables are not jointly significant	$F(10, 109) < 1.9186$
Jarque-Bera	Residuals are normally distributed	$\chi^2(2) < 5.99$
Durbin-Watson	No autocorrelation, positive or negative	See Below
Breusch-Godfrey	No serial correlation of any order	$F(5, 104) < 2.3017$
White	No heteroskedasticity	$F(65, 54) < 1.548$
Ramsey RESET	Model correctly specified	$F(2, 107) < 3.081$

The Durbin-Watson test produces values between 0 and 4, with 0 indicating negative and 4 indicating positive autocorrelation, and 2 indicating no autocorrelation. It has a decision rule that falls within a range of critical values, with an inconclusive range either side of the rejection range. For our model this the lower bound is 1.526 and the upper bound is 1.885, giving an inconclusive (regards negative and positive autocorrelation, respectively) range of  $1.526 < D < 1.885$ ;  $2.114 < D < 2.474$  and a conclusive fail-to-reject region of  $1.885 < D < 2.114$ .

The Akaike and Schwartz Criteria are both used to compare models. They minimize the sum of squared errors while penalising too many variables (Hill, Griffiths, and Lim 2011, p. 238). The

criteria are

$$\begin{aligned} \text{AIC} &= \ln\left(\frac{SSE}{N}\right) + \frac{2K}{N} \\ \text{SC} &= \ln\left(\frac{SSE}{N}\right) + \frac{K \ln(N)}{N}. \end{aligned}$$

The model with the smallest criterion is preferred.

The full results of the initial model selection diagnostic tests are in appendix 6.1. The main result is the rejection of a linear *QCOFFE10*; as hypothesised, this is likely due to the large variance and non-normality of the residuals. In no preliminary regression was the own-price coefficient significant at 5%; although the Double-Log and Log-Lin functions were close to this level, the p-values for the linear independent variable functions were 21% and 50%. Due to this disagreement with theory, failing the Jarque-Bera tests, and showing  $R^2$  values well below the two logarithmic models, the Linear and Log-Lin specifications were discarded. A choice remains between Log-Lin and Double-Log specifications: Log-Lin passes the Durbin-Watson test, whereas it is inconclusive in the double-log. However, the  $R^2$ , AIC, and SC criteria are in favour of the Double-Log; in addition, it is readily interpretable in economic theory.

### 3 Finding the Preferred Model

#### 3.1 Dropping Insignificant Variables

Therefore, the current preliminary model is:

As hypothesised, non-grocery items are irrelevant variables in the preliminary specification. There is a difference in significance between Meat & Fish and Fruit & Veg as cross-price variables; Fruit & Veg is perhaps considered a necessary staple, and Meat & Fish more ‘luxury’ food. This might be the case if Ruritania is a developing economy. Most surprisingly, *LPCOFF* is not significant in the preliminary result. It is significant at 10%, however, and I suspect that the insignificant

Table 3: Coefficients of unrestricted model. Significant variables in **bold**.

<b>Variable</b>	<b>Coefficient</b>	<b>T-Value</b>	<b>T-Prob</b>
LPCOFF	-0.650794	-1.92	0.0573
Constant	8.40743	2.46	<b>0.0155</b>
LPBEER	-0.0467172	-0.155	0.8769
LPWINE	0.112133	0.278	0.7818
LPLEIS	-0.125489	-0.563	0.5747
LPTRAV	0.0301442	0.153	0.8789
LPALLOTH	-0.935608	-1.36	0.1756
LPMTFH	-0.262636	-1.14	0.2561
LPFTVG	0.0718850	0.193	0.8471
LPTEA	1.52944	4.60	<b>0.0000</b>
LINCOME	0.910242	3.29	<b>0.0014</b>

result is due to the influence of the other insignificant variables on the model. Its vital theoretical relevance means we must keep it in the model.

Following the logic of a ‘nested model’ (Gujarati 2002, p. 529), I remove variables stepwise, removing the least significant variable in the preceding model, until the model fails diagnostic tests. This process is summarised in appendix 6.2 The summary of the final equation shows that PMTFH is insignificant at 5%. However, removing this variable causes PCOFF to become insignificant. This interaction can signify that the functional form has been misspecified; however, when checking against the other viable functional form, Log-Lin, we find that the same issue exists. It can also signify that an omitted variable is present; however, the model passes the RESET test for omitted variable bias and the F-test below shows that there are no jointly significant variables omitted. Hence, PMTFH is included in the final model despite being individually insignificant at 5%, as it maintains the theoretical validity of a significant own-price effect on coffee demand.

The final model passes all diagnostic tests, with the exception of the Durbin-Watson statistic which shows an unclear result for positive autocorrelation. We will therefore investigate this further in 3.2.3. The full diagnostic summary is in appendix 6.3.

I confirm that none of the variables removed were significant with an f-test. The order of removal is LPTRAV, LPBEER, LPWINE, LPLEIS, LPMTFH, LPALLOTH. The null hypothesis is that all of these variables equal zero and are insignificant to the constrained equation. The

Table 4: Final model specification.

Variable	Coefficient	Std. Error	T-Value	T-Prob	Partial $R^2$
Constant	8.185	1.499	5.46	0.0000	0.2073
LPCOFF	-0.606	0.275	-2.20	0.0297	0.0408
LPALLOTH	-1.128	0.2919	-3.860	0.0002	0.1158
LPMTFH	-0.329	0.1915	-1.72	0.0878	0.0253
LPTEA	1.53497	0.2495	6.15	0	0.2493
LINCOME	0.990306	0.2152	4.6	0	0.1566
$R^2$	0.80724	Adj. $R^2$	0.798786	RSS	26.8070245
N	120	Params	6		

F-statistic is given by

$$F = \frac{(SSR_R - RSS_U)/q}{RSS_U/(N - k)} = \frac{(26.807 - 26.599)/5}{26.599/(120 - 11)} = 0.1707$$

Where q is the number of variables removed (5) and k is the total regressors (11). We compare this to our critical value of  $F(5, 109) = 2.297 > 0.1707$ . Hence we fail to reject the null hypothesis and infer that they are not jointly significant.

## 3.2 Additional Tests

### 3.2.1 Seasonality Test

Demand for hot beverages may increase around the colder season; demand may also be affected by the coffee harvest which happens at a certain point in the year. In this way demand for coffee may show seasonal effects: we can test for this by running a new regression with three quarterly dummies.

The regression shows that none of the dummies are individually significant. The F-statistic is

$$F = \frac{(SSR_R - RSS_U)/q}{RSS_U/(N - k)} = \frac{(26.807 - 26.460)/3}{26.460/(120 - 9)} \\ = 0.4852$$

$$F(3, 111) = 2.69 > 0.4852.$$

Table 5: Model estimated with seasonal dummies included.

<b>Variable</b>	<b>Coefficient</b>	<b>T-prob</b>	<b>Partial <math>R^2</math></b>
Constant	8.03348	0	0.2019
LPCOFF	-0.598697	0.0329	0.0403
LPALLOTH	-1.13626	0.0002	0.1186
LPMTFH	-0.321518	0.0991	0.0243
LPTEA	1.53443	0	0.2512
LINCOME	1.00468	0	0.1617
Seasonal	0.102932	0.4168	0.0059
Seasonal_1	0.136962	0.2807	0.0105
Seasonal_2	0.0363281	0.774	0.0007
Adj. $R^2$	0.796021	RSS	26.4602
N	120	Params	9

Hence, we fail to reject the null and conclude they are not jointly significant. The adjusted  $R^2$  has also decreased slightly, meaning that less of the data is explained.

### 3.2.2 Structural Break: Chow Test

The next hypothesis is that there may be a structural break in the data. This was identified from visual inspection of the linear time series of QCOFFE10, which includes several large residuals. We use the break point of 1998(1) identified earlier to find a structural break with the Chow test. I hypothesise that there will be no change, as applying the log transformation to QCOFFE10 has largely smoothed the noted change in variance.

The chow test is set up as follows; the F statistic is

$$\frac{(SSR_P - SSR_U)/k}{SSR_U/(n_1 + n_2 - 2k)}.$$

The null hypothesis is that there is no evidence of a structural break.

Since the test statistic does not exceed the critical value we fail to reject the null and conclude that the underlying structure does not change over time.



Table 6: Summary of Chow test for final model.

Parameter	Interpretation	Value
SSRP	Sum of sq. residuals pooled regression	26.807
SSRU	Sum of sq. residuals of subsample regressions	24.329
k	Number of parameters	6
n1 + n2	Total number of observations	45+75=120
<b>Test Statistic</b>	1.834	
<b>Critical Value</b> F(6,108)	2.184	

### 3.2.3 Habit Formation: Lagged Dependent Variable Test

Finally we may want to check for some form of habit formation in our model. This follows from the fact that coffee is an addictive substance, since it contains caffeine. The quantity of coffee demanded may depend on the previous time periods, with habit effects as an autoregressive process. We can test for this by running a new regression with lagged dependent variables and considering the outcome. Habit formation effects should occur at least in the immediately preceding period (Pollak 1970); I propose that the habit effect will diminish after a period of 1 year, implying a four-quarter lag to be included in the model. We see that there is no significant relationship on

Table 7: Summary of regression with lagged variables. Significant coefficients in **bold**.

Variable	Coefficient	T-value	T-prob
LQCOFFE10_1	-0.04	-0.43	0.67
LQCOFFE10_2	-0.12	-1.40	0.16
LQCOFFE10_3	-0.10	-1.18	0.24
LQCOFFE10_4	0.11	1.35	0.18
Constant	9.38	4.37	<b>0.00</b>
LPALLOTH	-1.34	-3.66	<b>0.00</b>
LPTEA	1.74	6.07	<b>0.00</b>
LINCOME	1.16	4.30	<b>0.00</b>
LPCOFF	-0.66	-2.26	<b>0.03</b>
LPMTFH	-0.38	-1.86	0.07
Adj. $R^2$	0.0810263	RSS	23.68
N	116	Params	10

the dependent variable on any lag period; the common single-period lag seen in the literature is not significant either. This shows that there is no particular habit formation effect; however, more refined modelling which takes into account addiction asymmetry - the idea that addicted

consumers have smaller elasticities in response to a price rise versus a price fall - may show an effect (Young 1982).

## 4 Preferred Model

We have been unable to improve our model by implementing seasonality, a structural break, or habit formation. However, we have confirmed that Marshallian demand restriction of homogeneity holds. The final model (with brackets indicating t-values) is

$$\begin{aligned} \ln Q_C = & 8.185 - 0.6057 \ln P_C - 1.128 \ln P_A \\ & (1.5) \quad (0.275) \quad (0.292) \\ & - 0.3297 \ln P_M + 1.535 \ln P_T + 0.9903 \ln Y \\ & (0.192) \quad (0.249) \quad (0.215) \end{aligned}$$

### 4.1 Interpretation of Coefficients

Table 8: Economic summary of coefficients in final model.

Variable	Coefficient	Interpretation of Elasticity Coefficient
Constant	8.18	No economic interpretation.
LPCOFF	-0.605	Demand for coffee is relatively price inelastic
LPALLOTH	-1.128	A 1 increase in general prices reflects a 1.12 fall in coffee demand. They are therefore close complements.
LPMTFH	-0.329	Coffee, Meat and Fish are weak complements.
LPTEA	1.53	Coffee and Tea are strong substitutes: a rise in Tea price leads to a 1.53 rise in coffee demand.
LINCOME	0.990	Coffee is very close to unit elastic in income. It is a necessity good but is close to being a luxury good.

These interpretations track with my hypotheses. The inelasticity of coffee reflects its addictive nature; the price effect of other goods is likely due to an income effect on the budget constraint; Tea is a strong substitute, and Meat/Fish is a weak complement. A somewhat surprising result is the income elasticity being 0.990: this implies that it is close to being a luxury good. This would

not obey Engel's Law that foodstuffs do not show increasing expenditure share with income. As it stands, the income share stays stable.

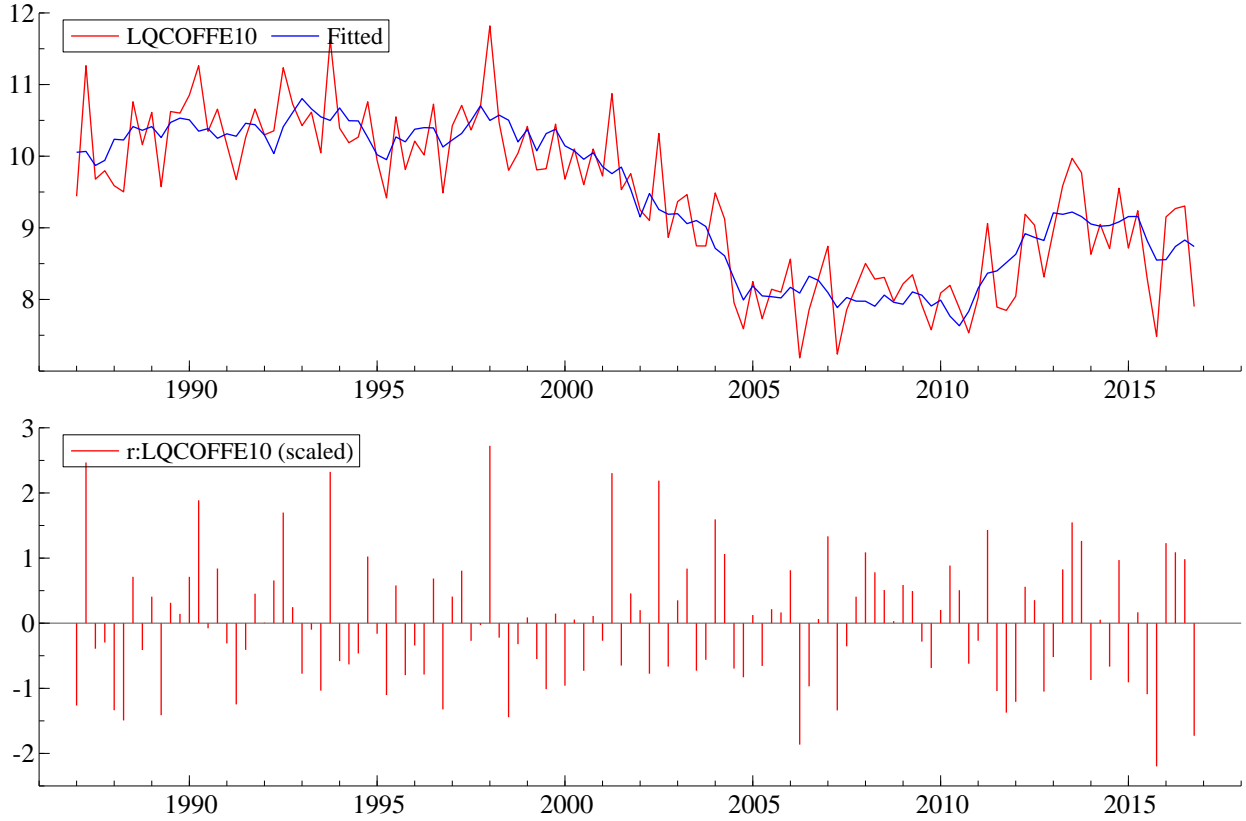


Figure 2: Fitted path and residual plot of the preferred model.

Visual inspection of the fitted residual plot shows large deviations but with a zero mean, where the residuals are normally distributed. An improvement on this model may be a Box-Cox transformation to better deal with the large outlier residuals.

There are some improvements that can be made to the model: for example there could be a problem of identification. Given information on the determinants of supply, we could run a simultaneous equations procedure to better model the market and reduce the chance of simultaneity bias, rather than consumer demand alone. The variables are also non-stationary as figure 3 shows; this may be accounted for using a more nuanced functional form such as ARDL, or GARCH as in Aradhyula and Holt (1988).

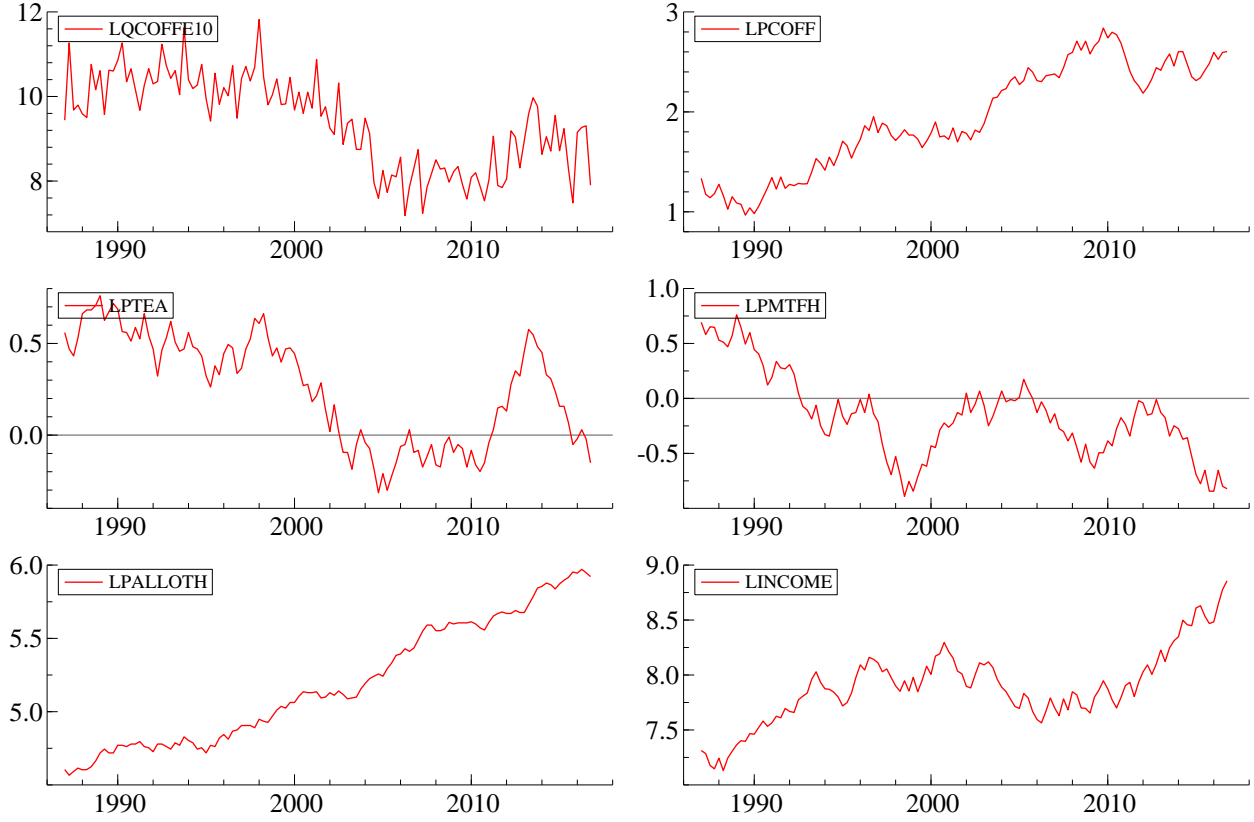


Figure 3: Time series plot of the variables in the final model.

## 4.2 Slutsky Equation

We can use the Slutsky equation to decompose the own-price elasticity into the substitution effect and income effect. This result follows from the equivalence of the compensated and uncompensated demand functions when utility is maximised. The derivation follows

$$\begin{aligned}
 \frac{\partial Q}{\partial P_C} &= \frac{\partial Q}{\partial P_C} - Q \frac{\partial Q}{\partial Y} \\
 \frac{P_C}{Q} \cdot \frac{\partial Q}{\partial P_C} &= \frac{P_C}{Q} \cdot \frac{\partial Q}{\partial P_C} - \frac{P_C}{Q} \cdot Q \frac{\partial Q}{\partial Y} \\
 \eta_{Q,P_C} &= \eta_{Q^*} - s \cdot \eta_Y.
 \end{aligned}$$

We can use this equation to decompose the responses amounting from a price change into a price response - keeping the same utility - and an income response - shifting to a new utility. We find the compensated elasticity by substituting our estimated parameters:

$$\begin{aligned}\eta_{Q^*} &= \eta_{Q,P_C} + s \cdot \eta_Y \\ \eta_{Q^*} &= -0.605 + (0.04 \cdot 0.99) \\ &= -0.5654\end{aligned}$$

This means that, faced with a 1% increase in price, consumers will substitute 0.57% of coffee expenditure to other goods. Coffee is thus relatively inelastic, and the substitution effect is the most important in decomposing the price response. The income compensation required to offset the price change is quite small (Snyder and Nicholson 2012, p. 153).

### 4.3 Homogeneity Condition

Demand theory implies that there should be a restriction amounting to demand homogeneity on the estimated equation: since demand functions are homogenous of degree zero, any instance of pure inflation should scale the budget constraint by the same amount, leaving the values achieved by the maximisation the same Snyder and Nicholson 2012, p. 154. The elasticities should sum to zero:

$$0 = \eta_{P_C} + \eta_{P_T} + \eta_{P_M} + \eta_{P_A} + \eta_Y.$$

Our log-log model makes it easy to find the elasticities: the coefficients simply give values for the elasticity in relation to coffee. In our model  $-0.6057 + -1.128 + -0.3297 + 1.535 + 0.9903 = 0.4619$ ; we check whether this value is statistically significant to 0 using an F test. The null hypothesis is that all the coefficients satisfy  $\beta_1 + \beta_2 + \beta_3 + \beta_4 + \beta_5 = 0$ . We use a parameter-restriction F test,

deriving the constrained equation as follows:

$$Q_C = \alpha + \beta_1 LP_C + \beta_2 LP_A + \beta_3 LP_M + \beta_4 LP_T + \beta_5 LY$$

$$\beta_5 = -(\beta_1 + \beta_2 + \beta_3 + \beta_4)$$

$$Q_C = \alpha + \beta_1 LP_C + \beta_2 LP_A + \beta_3 LP_M + \beta_4 LP_T + (-\beta_1 - \beta_2 - \beta_3 - \beta_4) LY$$

$$Q_C = \alpha + \beta_1(LP_C - LY) + \beta_2(LP_A - LY) + \beta_3(LP_M - LY) + \beta_4(LP_T - LY)$$

We then estimate this restraint and compare it against the unconstrained equation above. The F-statistic is given as

$$F = \frac{(RSS_M - RSS)/M}{RSS/(N - K - 1)} = \frac{(26.9860276 - 26.8070245)/1}{26.8070245/(113)} = 0.75455$$

Where  $RSS_M$  is the sum of squares in the constrained equation, M is the number of constraints imposed (1 in this case), and (N-K-1) is the degrees of freedom of the unconstrained equation. The critical value is  $F(1, 113) = 3.92 > 0.755$ . Therefore we fail to reject the null, and as theory predicts the demand is statistically homogenous of degree zero. There is therefore no price effect.

## 5 Literature Discussion

Double-log single-equation demand estimation has been popular with demand estimators for its ease of estimation and interpretation, given its property that the coefficients equal the equation's elasticities. However, this single-equation approach cannot test for three key restrictions that theory places: additivity, symmetry, and separability; it also assumes constant elasticity. Since Stone (1954)'s Linear Equation System, approaches have therefore usually focussed on specifying *systems* of demand equations. Deaton and Muellbauer (1980) introduces the Almost Ideal Demand System, which seeks to model these characteristics. Other examples of demand estimation techniques are

the Translog model (Pollak and Wales 1978) and the Rotterdam model (Lee, Brown, and Seale 1994).

Ruritania's demand broadly fits a Marshallian demand function with significant elasticities on the own price, general price index, meat and fish, income, and tea. The final estimated model is static and non-seasonal. Homogeneity was found to hold, and the compensated own-price elasticity was calculated. The finding of no significant habit formation effect is surprising: coffee is both addictive and bought irregularly, which implies that response to prices may be 1 quarter or more (Durevall 2007a). My model estimates the income elasticity as 0.99, which is high: while Huang, Siegfried, and Zardoshty (1980), Hughes (1969), Daly (1958), and Abaelu and Mander-scheid (1968) all estimate coffee to be a necessity good, no coefficient goes above 0.51, and Timms (1973), Lawrance (1977), and Heien and Pompelli (1989) all estimate income elasticities of coffee below 0, implying it is an inferior good. Finally, Okunade (1992) and Parikh (1973) also find strong evidence for habit formation effects, while Bettendorf and Verboven (1998) reaches only an inconclusive result requiring a longer sample. Durevall (2007b) also fails to find a significant role for tea in the demand equation, and attributes this to coffee's predominance in the studied market; by extension given tea's importance to Ruritania I surmise strong competition between tea and coffee.

2490 words.

## References

- Abaelu, John Nduka and Lester V Manderscheid (1968). “US Import Demand for Green Coffee by Variety”. In: *American Journal of Agricultural Economics* 50.2, pp. 232–242. ISSN: 1467-8276.
- Aradhyula, Satheesh V. and Matthew T. Holt (1988). *GARCH Time Series Models : An Application to Retail Livestock Prices*. Iowa State University.
- Bettendorf, L and F Verboven (1998). *Competition on the Dutch Coffee Market*. CentER Discussion Paper 9810. Tilburg University.
- Daly, Rex F. (July 1958). “Coffee Consumption and Prices in the United States”. In: *Agricultural Economics Research* 10, pp. 61–71.
- Deaton, Angus and John Muellbauer (1980). “An Almost Ideal Demand System”. In: *The American Economic Review* 70.3, pp. 312–326. ISSN: 0002-8282.
- Deutscher Kaffeverband (n.d.). *Coffee: Per Capita Consumption in Germany 1960-2015*.
- Durevall, Dick (Aug. 2007a). “Competition in the Swedish Coffee Market, 1978–2002”. In: *International Journal of Industrial Organization* 25.4, pp. 721–739. ISSN: 01677187.
- (Oct. 2007b). “Demand for Coffee in Sweden: The Role of Prices, Preferences and Market Power”. In: *Food Policy* 32.5-6, pp. 566–584. ISSN: 03069192.
- Gujarati, Damodar (2002). *Basic Econometrics*. 4th ed. OCLC: 248758244. New York London: McGraw-Hill. 838 pp. ISBN: 978-0-07-247852-5 978-0-07-112343-3.
- Heien, Dale and Greg Pompelli (1989). “The Demand for Alcoholic Beverages: Economic and Demographic Effects”. In: *Southern Economic Journal* 55.3, pp. 759–770. ISSN: 0038-4038.
- Hill, R. Carter, William E. Griffiths, and G. C. Lim (2011). *Principles of Econometrics*. 4th ed. OCLC: ocn656452679. Hoboken, NJ: Wiley. 758 pp. ISBN: 978-0-470-62673-3.
- Huang, C. J., J. J. Siegfried, and F. Zardoshty (1980). “The Demand for Coffee in the United States, 1963-77.” In: *Quarterly Review of Economics and Business* 20.2, pp. 36–50. ISSN: 0033-5797.
- Hughes, John J. (1969). “Note on the U.S. Demand for Coffee”. In: *American Journal of Agricultural Economics* 51.4, pp. 912–914. ISSN: 0002-9092.



- Lawrance, Neil A (1977). *US Coffee Consumption, 1946-76*. Foreign Agricultural Service, US Department of Agriculture.
- Lee, Jonq-Ying, Mark G. Brown, and James L. Seale (Aug. 1994). "Model Choice in Consumer Analysis: Taiwan, 1970-89". In: *American Journal of Agricultural Economics* 76.3, p. 504. ISSN: 00029092.
- Okunade, Albert A. (Nov. 1, 1992). "Functional Forms and Habit Effects in the US Demand for Coffee". In: *Applied Economics* 24.11, pp. 1203–1212. ISSN: 0003-6846.
- Parikh, A. (Aug. 1973). "United States, European, and World Demand Functions for Coffee". In: *American Journal of Agricultural Economics* 55.3, p. 490. ISSN: 00029092.
- Pollak, Robert A. (1970). "Habit Formation and Dynamic Demand Functions". In: *Journal of Political Economy* 78.4, pp. 745–763. ISSN: 0022-3808.
- Pollak, Robert A. and Terence J. Wales (1978). "Estimation of Complete Demand Systems from Household Budget Data: The Linear and Quadratic Expenditure Systems". In: *The American Economic Review* 68.3, pp. 348–359. ISSN: 0002-8282.
- Snyder, Christopher M. and Walter Nicholson (2012). *Microeconomic Theory: Basic Principles and Extensions*. 11. ed., international ed. OCLC: 815570442. Mason: South-Western [u.a.] 738 pp. ISBN: 978-1-111-52551-4.
- Stone, Richard (1954). "Linear Expenditure Systems and Demand Analysis: An Application to the Pattern of British Demand". In: *The Economic Journal* 64.255, pp. 511–527. ISSN: 0013-0133.
- Timms, D (1973). *World Demand Prospects for Coffee in 1980*. Washington, DC: USDA/ERS.
- Young, Trevor (Feb. 1982). "ADDICTION ASYMMETRY AND THE DEMAND FOR COFFEE". In: *Scottish Journal of Political Economy* 29.1, pp. 89–98. ISSN: 0036-9292, 1467-9485.

## 6 Appendices

### 6.1 Diagnostic summary of preliminary functional forms

Table 9: Model comparison with test parameters and critical values listed. Tests rejecting the null in **bold**.

Test	F-test	$R^2$	DW	BG	JB	White	RESET	AIC	SC
<b>Model</b>	F(10,109)			F(5,104)	$\chi^2 < (2)$	F(65,54)	F(2,107)		
$\alpha = 0.05$	<i>1.9186</i>			<i>2.3017</i>	<i>5.99</i>	<i>1.5477</i>	<i>3.081</i>		
Linear	10.03	0.479246	<b>2.17194</b>	0.69389	<b>226.84</b>	1.0437	1.5118	22.326	22.581
Log-Lin	43.04	0.797914	2.06533	1.0684	5.6425	0.57342	1.1268	1.5696	1.8252
Lin-Log	10.47	0.49003	<b>2.16626</b>	0.61544	<b>193.82</b>	1.2125	2.5873	22.305	22.56
Log-Log	46.09	0.808738	<b>2.14233</b>	1.4329	5.4254	0.81353	0.19218	1.5146	1.7701

Note the Durbin-Watson inconclusive bounds of  $1.526 < D < 1.885$ ;  $2.114 < D < 2.474$  and a conclusive fail-to-reject region of  $1.885 < 2.114$ . **Bold** for Durbin-Watson indicates an *inconclusive* result; there were no outright failures.

### 6.2 P-Values for variables removed from preliminary model

Table 10: P-values for insignificant variables removed from preliminary double-log model. Significant variables in **bold**.

Variable	Constant	$P_C$	$P_A$	$P_T$	$Y$	$P_M$	$P_W$	$P_L$	$P_F$	$P_B$	$P_{TR}$
1	0.0573	0.0573	0.1756	<b>0</b>	<b>0.0014</b>	0.2561	0.7818	0.5747	0.8471	0.8769	0.8789
2	<b>0.008</b>	<b>0.0364</b>	0.1572	<b>0</b>	<b>0.0004</b>	0.2541	0.6987	0.572	0.875	0.8874	
3	<b>0.001</b>	<b>0.0353</b>	<b>0.0108</b>	<b>0</b>	<b>0.0002</b>	0.2556	0.502	0.5669	0.904		
4	<b>0.0003</b>	<b>0.0206</b>	<b>0.0048</b>	<b>0</b>	<b>0.0002</b>	0.2543	0.4634	0.5527			
5	<b>0</b>	<b>0.0234</b>	<b>0.0033</b>	<b>0</b>	<b>0</b>	0.2283	0.4966				
6	<b>0</b>	<b>0.0297</b>	<b>0.0002</b>	<b>0</b>	<b>0</b>	0.0878					
7	<b>0</b>	0.1103	<b>0</b>	<b>0</b>	<b>0</b>						

7 consecutive regressions were run, stopping at model 7. Regression **6** is the final specified model.

### 6.3 Diagnostic summary of final model

Table 11: Model comparison with test parameters and critical values listed. Tests rejecting the null in **bold**.

Test	F-test	$R^2$	DW	BG	JB	White	RESET	AIC	SC
<b>Model</b>	F(5,114)			F(5,109)	$\chi^2 < (2)$	F(20,99)	F(2,112)		
$\alpha = 0.05$	<i>2.294</i>			<i>2.298</i>	<i>5.99</i>	<i>1.678</i>	<i>3.077</i>		
Model 7	95.48	0.80724	<b>2.13086</b>	1.3444	4.5854	0.60696	0.07049	1.4390	1.5784