Estimating the Demand of Fish in Singapore Wet Markets

Naveen Mathew Verghese - A0262734A
Yumo Yao - A0262687N
Samin Batra - A0262665X

Wet markets are very popular in Singapore. These are the markets where people go and bargain with vendors to buy fresh food products like poultry, pork, vegetables, fruits and spices. The reason they're referred to as "wet markets" is because of the regular practices of hosing down and washing the market floors to maintain cleanliness.

In this report, we will try to estimate the demand curve of fish at these wet markets using the parameters that are given to us. The raw data consists of 111 rows of data from which we can formulate the demand function of fish in the wet market. Here are the variables which we'll use:

Mon	1 if Monday, 0 otherwise
Tue	1 if Tuesday, 0 otherwise
Wed	1 if Wednesday, 0 otherwise
Thu	1 if Thursday, 0 otherwise
Date	Date in yymmdd format
Stormy	1 if wave height greater than 1.5m, and wind speed greater than 18 knots. Based on moving averages of the last three days' wind speed and wave height.
Mixed	1 if wave height greater than 1m, and wind speed greater than 13 knots, excluding Stormy days. Based on moving averages of the last three days' wind speed and wave height.
p	**Log** average daily prices per kg of fish
q	**Log** average daily quantity sold of fish in kg
Rainy	1 if rainy weather on shore
Cold	1 if cold weather on shore
Wind	Wind speed in knots

Selecting variables for estimating Demand Function

As known in basic economic theorems, the demand for fish is highly dependent on its price. We could expect a negative coefficient for the price within the model due to the negative correlation between the price and the demand. Also, considering a wet market that is open to the local neighborhood, its demand could be highly dependent on the public lifestyle. People tend to cook when they are less busy, while outside dining could be a better choice for busy people. Thus, we can expect that the specific day of the week would matter.

Based on these assumptions, the demand function is taken to be :

$$Q_d = \beta_0 + \beta_1 p + \sum_{i=1}^{n} \gamma_i \, Days_i + u_1$$

Where:

$$p \rightarrow \log \quad price\ of\ fish$$

$$Days_i \rightarrow Days\ (Mon,\ Tue,\ Wed,\ Thurs)$$

$$u_1 \rightarrow error\ term$$

Endogeneity

From this demand function, we can hypothesize that the price (p) is endogenous. The demand for fish can vary with other parameters that are not included in the current list and we can assume that price is correlated to the error that comes from not including them. Consequently, the supply function could consist of these additional parameters which would influence the change in price. However, this influence is not accounted for in the demand function and so these additional parameters can be considered as supply shocks. These supply shocks are being considered in the u_1 error term. Since p and u_1 are correlated, any supply shocks will affect the price.

Hence, we need to consider some instrumental variables that will impact the price (p), thereby taking care of the correlation with the error term. These variables will be coming from the supply function. From the raw data, we can logically assume that the supply of fish is dependent on certain weather conditions. The weather conditions available to us are “Stormy”, “Mixed”, “Rainy”, “Cold” and “Wind”.

Extreme wave height and wind speed at the sea would significantly influence the feasibility of fishing, which would further impact the price of fish caught. “Stormy” and “Mixed” variables take these factors into account. Hence, we can consider them as our instrumental variables. Even though “Rainy”, “Cold”, and “Wind” may have certain correlations with the fish price, we believe that these variables may not be as significant since fishermen could also work in such weather conditions.

Considering the instruments that we have selected and the exogenous variables of days, we can express the price p as:

$$p = \delta_0 + \delta_1 Mon + \delta_2 Tue + \delta_3 Wed + \delta_4 Thurs + \delta_5 Stormy + \delta_6 Mixed + v_i$$

To test these instruments and the endogeneity of price, we will perform three tests.

Weak Instruments Test

To statistically test the instrumental variables, we could conduct a weak instrument test by running the first stage regression of price p on exogenous variables (Mon, Tue, Wed, and Thu) and instrumental variables (Stormy and Mixed). Below is the result of the regression:

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                      p   R-squared:                       0.245
Model:                            OLS   Adj. R-squared:                  0.201
Method:                 Least Squares   F-statistic:                     5.624
Date:                Wed, 21 Sep 2022   Prob (F-statistic):           4.35e-05
Time:                        09:17:26   Log-Likelihood:                -34.566
No. Observations:                 111   AIC:                             83.13
Df Residuals:                     104   BIC:                             102.1
Df Model:                           6
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const         -0.3596      0.079     -4.543      0.000      -0.517      -0.203
Mon           -0.1082      0.103     -1.047      0.298      -0.313       0.097
Tue           -0.0661      0.101     -0.654      0.514      -0.266       0.134
Wed           -0.0493      0.104     -0.475      0.636      -0.255       0.157
Thu            0.0393      0.101      0.391      0.697      -0.160       0.239
Stormy         0.4463      0.079      5.635      0.000       0.289       0.603
Mixed          0.2369      0.079      3.017      0.003       0.081       0.393
==============================================================================
Omnibus:                        2.130   Durbin-Watson:                   0.697
Prob(Omnibus):                  0.345   Jarque-Bera (JB):                1.994
Skew:                          -0.326   Prob(JB):                        0.369
Kurtosis:                       2.924   Cond. No.                         6.20
==============================================================================
```

Given that the null hypothesis is that $\delta_i = 0$, both p-values for Stormy and Mixed are smaller than 0.05, indicating that we strongly reject the null hypothesis. Hence, the coefficients between price and these 2 variables are statistically significant. The high F_test value of 16.14 also indicates that these instruments are significant. We may conclude that “Stormy” and “Mixed” are appropriate instrumental variables for estimating the demand function.

From the regression result, the equation for price is given below:

$$p = -0.36 - 0.1 Mon - 0.07 Tue - 0.05 Wed + 0.04 Thu + 0.45 Stormy + 0.24 Mixed + v$$

Hausman Test

We use Hausman test to check whether price (p) is endogenous. On running the Hausman test for the residuals of the first stage regression values ($\dot{v}$) obtained from the reduced form, we notice the p-value for the coefficient of the residual term is 0.225, which is greater than 0.05. This means that we cannot reject the null hypothesis that price variable is exogenous.

Considering that the Hausman test failed to prove that price is indeed endogenous, we can use OLS to estimate the demand function. By running the regression of demand quantity (q) on price (p) and days of the week (Mon, Tue, Wed, Thu), the result is as below:

OLS Regression Results

Dep. Variable:	q	R-squared:	0.220
Model:	OLS	Adj. R-squared:	0.183
Method:	Least Squares	F-statistic:	5.940
Date:	Wed, 21 Sep 2022	Prob (F-statistic):	7.08e-05
Time:	10:58:03	Log-Likelihood:	-110.00
No. Observations:	111	AIC:	232.0
Df Residuals:	105	BIC:	248.3
Df Model:	5		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
const	8.6069	0.143	60.170	0.000	8.323	8.891
p	-0.5625	0.168	-3.344	0.001	-0.896	-0.229
Mon	0.0143	0.203	0.071	0.944	-0.387	0.416
Tue	-0.5162	0.198	-2.611	0.010	-0.908	-0.124
Wed	-0.5554	0.202	-2.745	0.007	-0.957	-0.154
Thu	0.0816	0.198	0.413	0.681	-0.311	0.474

Omnibus:	14.325	Durbin-Watson:	1.487
Prob(Omnibus):	0.001	Jarque-Bera (JB):	15.876
Skew:	-0.804	Prob(JB):	0.000357
Kurtosis:	3.920	Cond. No.	5.85

The resulting equation of the demand function from OLS is:

$$Q_d = 8.61 - 0.56p + 0.01Mon - 0.52Tue - 0.56Wed + 0.08Thu + u_1$$

However, as the result of the Hausman test is contrary to intuition, we will consider using 2SLS to estimate the demand function and to see the difference between its result with that of OLS.

Sargan's Test

Assuming that the price is endogenous, we can conduct the Sargan's test to validate if we should indeed use Stormy and Mixed as our instrumental variables.

We see that the $R^2$is coming to 0.007, which is low. The p-value from the χ^2 test gives us 0.378, which is greater than 0.05, thus signifying that we can use these weather conditions as our instrumental variables.

Given the reduced form of 2SLS, we can perform the second stage by running a regression of demand quantity (q) on the estimated price ($\dot{y}$) and the exogenous variables (Mon, Tue, Wed, and Thu):

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                      q   R-squared:                       0.191
Model:                            OLS   Adj. R-squared:                  0.153
Method:                 Least Squares   F-statistic:                     4.965
Date:                Wed, 21 Sep 2022   Prob (F-statistic):           0.000403
Time:                        09:18:16   Log-Likelihood:                -112.05
No. Observations:                 111   AIC:                             236.1
Df Residuals:                     105   BIC:                             252.4
Df Model:                           5
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const          8.5402      0.156     54.741      0.000       8.231       8.850
hat_y1        -0.9301      0.352     -2.642      0.010      -1.628      -0.232
Mon           -0.0119      0.208     -0.057      0.954      -0.423       0.400
Tue           -0.5258      0.202     -2.609      0.010      -0.925      -0.126
Wed           -0.5626      0.206     -2.729      0.007      -0.971      -0.154
Thu            0.0999      0.202      0.494      0.622      -0.301       0.501
==============================================================================
Omnibus:                       15.211   Durbin-Watson:                   1.526
Prob(Omnibus):                  0.000   Jarque-Bera (JB):               17.790
Skew:                          -0.796   Prob(JB):                     0.000137
Kurtosis:                       4.146   Cond. No.                         6.13
==============================================================================
```

Using this price equation, we can use linear regression to determine the equation for the demand function, which is given below:

$$Q_d = 8.54 - 0.93\overline{p} - 0.01Mon - 0.53Tue - 0.56Wed + 0.09Thu + \dot{u}$$

From this equation, we can see that given other variables unchanged, a unit increase in the price of fish decreases the demand of fish by 0.93**.** This negative correlation is logically aligned with the concept of demand function.

Conclusion

We assumed that the price was endogenous and hence expressed it as a function of exogenous and instrumental variables. We considered the exogenous variables to be "Mon", Tue", "Wed" and "Thu", while we considered our instrument variables to be "Stormy" and "Mixed".

The results from the 3 tests are given below:

Test	**Result**	**Inference**
Weak instrument test	Pass	Stormy and mixed were the best instrument variables
Hausman Test	Fail	Price cannot be considered as endogenous, and so, the demand equation using OLS is: $Q_d = 8.61 - 0.56p + 0.01Mon - 0.52Tue - 0.56Wed + 0.08Thu + u_1$
Sargan's Test	Pass	Strengthened our assumption that weather conditions could be used as instrument variables and so, the demand equation from the 2SLS is: $Q_d = 8.54 - 0.93\overline{p} - 0.01Mon - 0.53Tue - 0.56Wed + 0.09Thu + \dot{u}$