Econometrics Report Assignment 2

Matthew Gitu (Student)

Introduction:

House prices vary across the UK, potentially due to differences in regional working populations and demographics. In this econometrics report there will be an analysis on the real-world factors that affect the house prices. In England, house prices were defined as affordable until 2001, but they have since increased twice as fast as household income" (Valentina Romei, 2024). This essay will aim to explain and critique the variation in house prices through econometrics methodology and economics theory.

Methodology:

This report will use econometric regression analysis to analyse those factors which affect house prices. The four models were used in this report: a simple linear regression to analyse one variable, adding more variables in multiple regression, and a logarithmic model to consider nonlinear relationships and elasticity. The significance and strength of the models are tested with different statistical tests, such as the p-value, R2, and the F-statistic.

Image 1-Regression coefficients

	Dependent variable:			
	House_price			log(House_price)
	(1)	(2)	(3)	(4)
Working_population	0.104^{***}	0.114^{***}	0.098^{**}	
	(0.021)	(0.018)	(0.041)	
Working_population_squared			0.00001	
			(0.00002)	
log(Working_population)				0.138^{***}
				(0.023)
Employment_rate		7.920^{***}	7.954^{***}	0.029^{***}
		(1.768)	(1.775)	(0.007)
Claimant_count		-20.891^{***}	-21.023^{***}	-0.088^{***}
		(6.975)	(7.002)	(0.026)
Constant	225.914^{***}	-303.237^{**}	-303.108^{**}	2.885^{***}
	(8.617)	(148.950)	(149.376)	(0.560)
Observations	146	146	146	146
R^2	0.147	0.444	0.444	0.456
Adjusted R^2	0.141	0.432	0.429	0.444
Residual Std. Error	87.567 (df = 144)	71.217 (df = 142)	71.421 (df = 141)	0.272 (df = 142)
F Statistic	24.836^{***} (df = 1; 144)	37.753^{***} (df = 3; 142)	28.202^{***} (df = 4; 141)	39.606^{***} (df = 3; 142)
Note:				*p<0.1; **p<0.05; ***p<0.01

Results and analysis:

Regression 1:

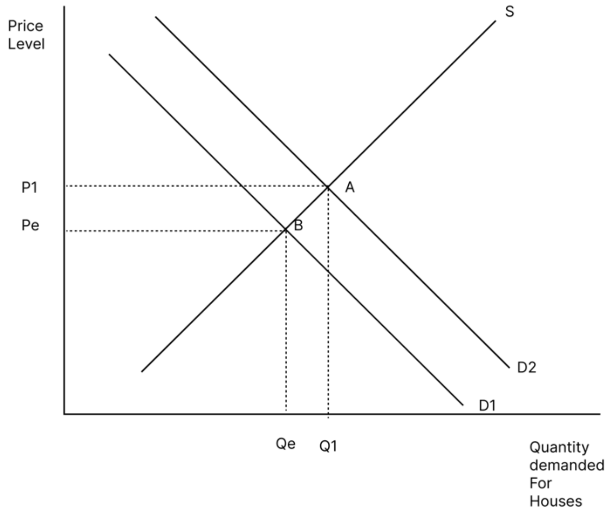
House pricei = β! + β"Working populationi + εi House pricei = 225.91 + β"0.10i + εi

According to the UK government there is a need for new houses midst a growth in population (Cassie Barton, 2023). The intercept shows that the predicted value of house prices while the working population is zero would be £225.91 controlling for the effects of other variables. The coefficient has a p-value of 2e-16, which is less than 0.05 meaning the coefficient is statistically significant. A 1% increase in the working population is associated with a £0.10 increase in house prices, holding other variables constant. The coefficient is statistically significant at the 1% significant level with a p-value of 1.77e-06.

Image 2 shows, as the demand (D1 to D2) for houses goes up from Qe to Q1, the price increases from Pe to P1. Which presents the theory of inelastic supply. Which theorises that the quantity supplied does not change autonomously with changes in price (Jon Guest, Dean Garrat, John Sloman, 2021).

Image 2-Supply and Demand

On the other hand, the R-squared value of 0.1412, shows that the working population explains only 14.12%ofthevariationsinthehouseprices.Furthermore,thehypothesistestingof,($H_{\#}$:$R^{''}$ = 0 $H_{!}$: $R^{''} \neq 0$) the variable working population can explain the variations in House prices. (F = 24.84 > $F_{\%\&'()\%}$ =3.91).Thissuggeststhatwhiletheworkingpopulationhasastatisticallysignificanteffect,it has a limited explanatory power in accounting for the overall variability in house prices.



600

500

400

300

200

100

Marginal Effect of Working Population on House Price

0 1000 2000

Working Population

House Price

Furthermore, image 5 shows the model is heteroscedastic, possibly a result of omitted variables: levels of income, rates of interest, and rates of inflation. These exogenous factors could influence house price and make the linear model produce biased estimates. Furthermore, the residual range—running from -143.25 to 302.41—is very large, which would indicate that the model hasn't accounted for the variability in house prices and, therefore, it might be suggesting missing critical factors.

Image 5-Marginal effect of Working population on House Price

Regression 2:

House price = β1 + β2Working populationi + β3Employment ratei+ β4Claimant counti + εi

House pricei = −303.23746 + β"0.11396i + β+37.92043 − β,20.89076 + εi

In regression analysis two, two new variables have been added, employment rate and claimant count. Thus, in the addition of the two variables there should be an increase in the explanatory power of the of the model. When the independent variables are 0, the predicted average house price is £-303.24 however this statistic is not realistic, the coefficient has a p-value which significant at the 5% level (0.04363). Secondly an additional unit increase in the working population, causes house prices to increase by £0.11 controlling for the effects of other variables. Along with a p-value which is significant at the 1% level (1.34e-09). For every additional unit increase in the employment rate, house prices increase by £7.92, on average controlling for the effects of other variables. Along with a p- value of 1.52e-05 ($p < 0.001$) it is statistically significant at the 1% level. On that note, for every additional unit in the claimant count, house prices decrease by £20.89 controlling for the effects of other variables, with a significant p-value at the 5% level (0.00324). Furthermore, we find the expected relationship between employment and prices: job creation increases housing prices and job destruction decreases housing prices (Kerri Agnew, Ronan C. Lyons, 2018). Through conducting the F-Test (H# : β! = β" = β+ = β, =

0, H!:Atleastoneβ' ≠0,F−statistic37.75>Criticalvalue5.13212)Thisimpliesthatthe inclusion of employment rate and claimant count significantly enhances the model's ability to explain house price variations. Furthermore, the R squared on the second regression has

increased by 2.01% (44.37-14.71= "-.//) explaining 44.37% of the variation in house prices. !,.0!

Regression 3:

House pricei = β1 + β2Working populationi +β3Working_populationi" + β4Employment ratei+β5Claimant counti + εi

House pricei = -303.1 + 0.00978 β2 i + 0.0000008264 β3 i+ 7.954 β4 i -21.02 β4 i+εi

The coefficients show that the intercept is -303.1. This means that when the independent variables are 0 house prices are -£303.10, with a significant p-value of 0.04432 which is significant at the 5% level. The intercept value here is mathematically valid but doesn't necessarily provide a meaningful interpretation in this case.

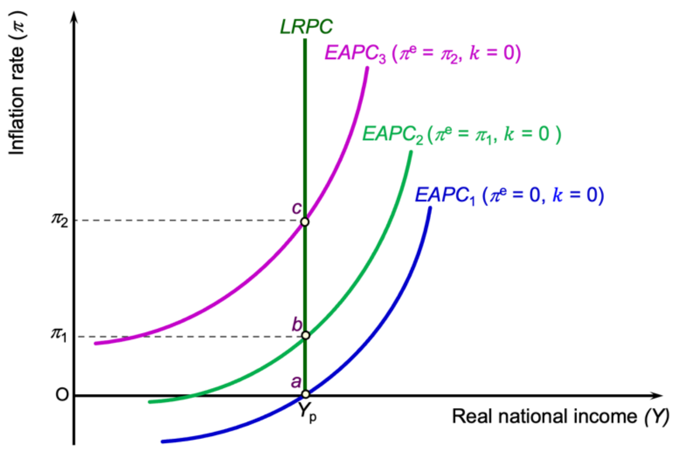
The coefficient Working population is 9.786e-02, which means for a one-unit increase, there is a £0.01 increase in house prices. The coefficient is also significant at the 5% level, 0.01774. Working_population_squared has a coefficient of 8.264e-06. Yes, that number is small. But what's even more important is the p-value of 0.66237, meaning this term is not statistically significant (at the 5% level). That means this variable isn't explaining a meaningful amount of variation in house prices. However, there was a moderate collinearity between the

Working population and Working population squared, with the VIFs 5.77 and 5.74, correspondingly. This may further indicate the presence of a collinearity effect between the two variables, which would, in turn, inflate the standard error, probably at the cost of reducing their ability to know their separate contributions to the model. The very high correlation coefficient 0.9078 of Working_population and working_population_squared is indicative of a very strong positive relation. Though expected—the squared term is derived directly from the original—the high correlation does also confirm that inclusion of both terms will render the model redundant.

Image 6- Philips Curve-Inflation and Real national income

The coefficients of employment rate denote each 1% increase in the employment rate, causes a £7.95 in house prices, holding all other variables constant. Image 6 (Philips curve) explains how in theory the increase in the employment rate could affect house prices. The EAPC is plotted as the relationship between the rate of inflation and national output (real national income) to reflect the effect of employment rate.

The curve denotes that as the employment rate increases from $EAPC_1$to $EAPC_2$ to $EAPC_3$, the inflation rate increases from A, B and to C additionally the vertical position of the curve continues to visualise inflationary pressures due to an increase in disposable income. Thus, the movement along the *EAPC* reflect demand-pull inflation ($\pi_1\ to\ \pi_2$). Thus, when the employment rate increases, inflation rates increase simultaneously. Causing a spill-over effect onto the housing market (Jon Guest, Dean Garrat, John Sloman, 2021).



The claimant coefficient shows that for every additional person that claims benefits the house prices are expected to decrease by £21.02, furthermore with a p-value of 0.00317 the coefficients have a negative relationship with house prices.

Through the R^2, the coefficient of the model explains 44.4% of the variations in house prices. However, the $R^2$of model 2 is identical, meaning the addition of the variable (Working_population_squared) did not improve the significance of the model. Furthermore, the adjusted R^2 of model 2 shows, 43.2% of the variations explain changes in house prices, while model (3) shows 42.9% explanation in house prices. This further reinstates the fact he inclusion of (Working_population_squared) does not substantially enhance the model's explanatory power, as the change in Adjusted R2 is minimal.

Regression 4:

log (House pricei) = β1 + β2log (Working populationi) + β3Employment ratei + β4Claimant counti + εi

House pricei = 2.885403 + 0.138483 β2 i + 0.138483 β3 i+ 0.029086 β4 i -0.087548 β4 i+εi

The logarithmic transformations are used to linearise relationships that may otherwise be non- linear. In theory it should also reduce skewness, stabilise variance, and interpret coefficients in terms of elasticity or percentage changes. The model explored will be the log-log model, where the dependent variable and independent are logged.

The intercept of the coefficients is 2.89, however with the coefficients fully expressed the work coefficientisexpandedto7.85(($e^{".1-}$ =7.88)Meaningthatwhenallthecoefficientsare0the house prices would be £7.888. To further prove this the p value of the coefficient is 8.37e-08, which is statistically significant at the 1% level. It proves that the elasticity of working population is 0.138. According to this, a 1% rise in working population increases the average house price by 14%, showing relative inelasticity, ceteris paribus. The p-value = 1.22e-08, which is statistically significant at 1%. It's observed that the coefficient of the employment rate is 0.029; therefore, with every 1% increase in employment, house prices will increase by 2.7% ceteris paribus. The p- value stands at 0.00000323 and hence the coefficients obtained are statistically significant at the 1% level, suggesting high semi elasticity due to the inelasticity in the supply of demand (Christopher Dougherty, 2016). Claimants count coefficient -0.088, so if the claimant count increased by 1% then house prices would fall by 8.8%, ceteris paribus. The p value is 0.00105 so the coefficient is significant at the 5% level of significance. The F-statistic proves that the coefficients statistically significant F-test 39.61>2.67 with a p-value of 2.2e-16 making the model significant at the 1% level. The R^2 is 0.4556, meaning that 45.56% of the coefficients explain the variation in house prices. Furthermore the R^2 is larger than the three different models, as the logarithm helps reduce the dispersion in the data.

I believe that in the end the logarithmic model is the best as it statistically explains the change in the house prices. Furthermore, it could be argued that increasing the variables (Model 2+3) in the dataset increases multicollinearity as the coefficients could be highly related, decreasing the statistical significance of an independent variable. This is proven with model 4 having the largest adjusted R^2.

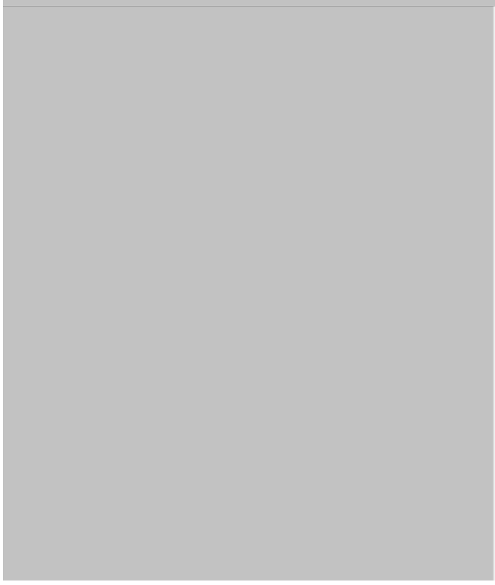
Conclusion:

In conclusion house prices of a country are not just affected by the six endogenously related variables. But are shaped by the macro-environment of a country. The models used fail to account for the remaining 55.6% of variation in house prices, as indicated by Model 4's adjusted R2.

References:

Jon Guest, Dean Garrat, John Sloman, 2021. Economics. s.l.:Pearson Education Limited.
Anon., 2020. UK's population density. [Online]
Available at: https://ukmap360.com/united-kingdom-%28uk%29-population-map Cassie Barton, 2023. Tackling the under-supply of housing in England. [Online]

Available at: https://commonslibrary.parliament.uk/research-briefings/cbp-7671/
HM LAND REGISTRY, 2023. UK House Price Index England: June 2023. [Online]
Available at: https://www.gov.uk/government/statistics/uk-house-price-index-for-june-2023/uk- house-price-index-england-june-2023
Kerri Agnew, Ronan C. Lyons, 2018. The impact of employment on housing prices: detailed evidence from FDI in Ireland. [Online]
Available at:
https://eprints.lse.ac.uk/87223/1/Lyons_Impact%20of%20employment_2018.pdf?utm_source =chatgpt.com



Appendix:

Stargazer:
install. packages("stargazer")
library(stargazer)
stargazer (House_price_1, House_price_2, House_price_3, House_price_4, type = "text")

Marginal effect of Working population on House Price:

ggplot (prediction_data, aes (x = Employment_rate, y = Predicted_House_price)) +
geom_line (color = "red", size = 1.2) +
labs (title = "Marginal Effect of Employment Rate on House Price",

x = "Employment Rate",

y = "Predicted House Price") + theme_minimal ()

Critical values:

```
qf (0.975, 1, 144)
qf (0.975,1,142)
qf (0.975,1,142)
qf (0.95, df1 = 3, df2 = 142)
```

Regressions:

```
House_price_1 <- lm (House_price~Working_population, data=Coursework_Data)

House_price_2 <- lm (House_price
~Working_population+Employment_rate+Claimant_count, data=Coursework_Data)

House_price_3 <- lm
(House_price~Working_population+(Working_population_squared)
+Employment_rate+Claimant_count, data=Coursework_Data)

House_price_4 <- lm (log (House_price) ~log (Working_population) + Employment_rate
+ Claimant_count, data = Coursework_Data)
```

Summary:

```
summary (House_price_1) summary (House_price_2) summary (House_price_3)
summary (House_price_4)
```

Variance inflation factor:

```
vif (House_price_3) vif (House_price_4)
```

Correlation coefficient:

```
cor (Coursework_Data$Working_population,
Coursework_Data$Working_population_squared)
```