Department of Economics

Prof. Damian Kozbur

Machine Learning for Economic Analysis

University of Zurich

January 2021

Applying Post Double Selection for Inference

An example on the Melbourne Housing Market

Authors:

Josia Schellenberg	josia.schellenberg@uzh.ch	13-608-351
Kevin Hardegger	kevin.hardegger@student.unisg.ch	12-758-785

Abstract

This paper analyzes the effect of number of rooms on housing prices while taking geographical features as control variables. For comparison reasons two different approaches are conducted this inference problem. The first approach follows the standard OLS method and the second the advanced post double selection which includes a formal and rigorous control variable selection process. This methodology is applied on housing prices data of the city of Melbourne. The results show that the number of rooms have significant effect on housing prices. This effect is even higher when controlled for geographical variables. Furthermore, post double selection provides to be a useful tool for introducing sparsity into a regression problem that contains numerous controls.

1. Introduction

After more than a decade of quantitative easing economists begin to question why the anticipated rise in inflation hasn't manifested yet. It appears as though the real economy has thoroughly been decoupled from markets (Gaba, 2020). However, as economic rules don't vanish suddenly, inflation is inevitable and with it the drop in prices of financial assets due to rising discounting rates for future earnings. With depreciation of financial assets and higher inflation this will ultimately lead to investments into real assets inter alia real estate value. While there are many ways to estimate housing prices this paper focuses on the effect of the number of rooms on the housing prices. It is common sense that the number of rooms have a significant impact on an individual house's price. However, one will realize that housing prices additionally depend on where said accommodations are located. For example, in contrast to the agglomerations and rural areas, densely populated regions will have less accommodations with many rooms. Furthermore, apartments near a newly gentrified district are considerably more expensive than in rather deteriorating quarters or regions in the outskirts of a city.

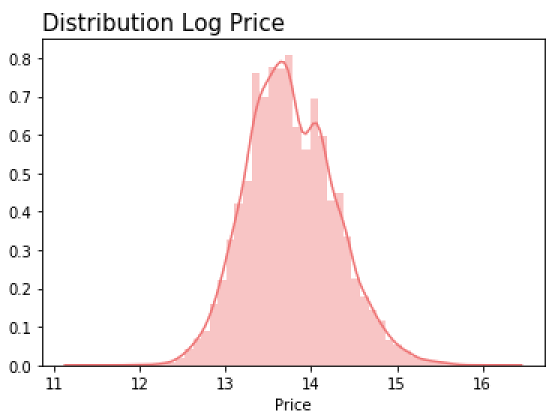
The goal of this project is to analyze this relationship on an example of the city of Melbourne – the second largest city in Australia with over four million habitants. With a thorough analysis, this paper estimates the effect of the number of rooms on housing prices which are additionally controlled by geographical features, namely longitude, latitude and the various council areas in Melbourne. For this, we first attain insight by analyzing the data in section 2. We then continue explaining the methodology of our estimation procedure. We conduct our estimations with two different approaches, the first one using the ordinary least squares method (OLS) and the second a modern estimation technique called post double selection (PDS) which introduces a formal sparsity procedure for eliminating expendable covariates. These are both explained in section 3. In section 4 the results of the OLS and PDS models are presented and discussed in section 5.

2. Data

To assess the effect of number of rooms to housing prices we use a data set provided on Kaggle, which contains scraped data from the website of the Domain Group - an Australian based real estate marketing business. The data contains various information which can be ordered in three main categories which are seller, building and geography related. For our study however, the necessary variables are solely the housing prices, number of rooms, and geographic related features such us longitude, latitude and all the council areas of Melbourne. We incorporate councils because we assume the geographical coordinates may not be able to capture clustered effects due to their nature of having the highest level of geographical granularity. After removing empty records for prices, the data set possesses a total of 20'993 observations. We apply Log transformation on the housing prices, create dummy variables for the 33 councils, and normalize latitude and longitude according to min max scaling, thereby creating values between 0 and 1. This leaves us with one target variable (Price), one variable of interest (Rooms), and 35 control variables.

The distribution of the transformed Prices can be examined in **Figure1**. The distribution closely resembles a gaussian form, with a skew of 0.27 and kurtosis of 0.1.

Figure 1: Distribution of log prices



The descriptive statistics of the number rooms are presented in **Table1.** According to the table the number of rooms range between 1 and 16. Median and mean are equivalent, hinting to a normal distribution. Nevertheless, since the third quantile equals to 4 this reveals that the observations containing 16 rooms may be outliers. However, there is no reason to exclude these observations as of now.

Table 1: Descriptive Statistics Rooms

Mean:	3.0591
Std. Error:	0.9499
Min:	1
Q1:	2
Median:	3
Q3:	4
Max.:	16

In **Figure 2** (Appendix I) the correlation plot of the features is presented. The correlation between the geographical variables and Price varies between -0.24 and 0.28, whereas with a value of 0.5 the plot suggests considerable positive correlation between Price and Room. This confirms the intuition that the number of rooms have a decisive impact on housing prices. This can be further assessed in **Figure3**, in which Price and Rooms are shown in a scatter plot. In **Figure3** it is possible to define a slight positive linear relationship between both variables. However, the relationship doesn't appear to be very strong. This is the reason we intend to use the councils as well as coordinates as controls, as they can impact both the number of rooms and the prices.

Figure 3: Scatter Plot Log Prices and Rooms

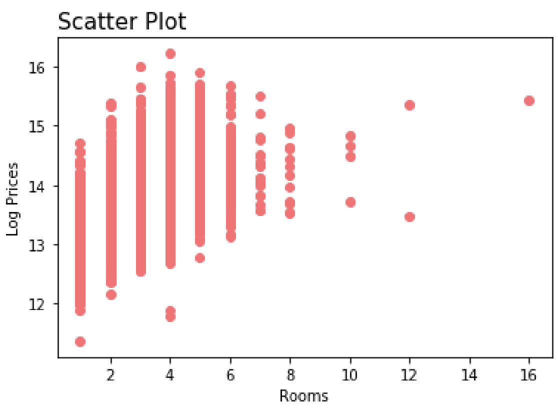
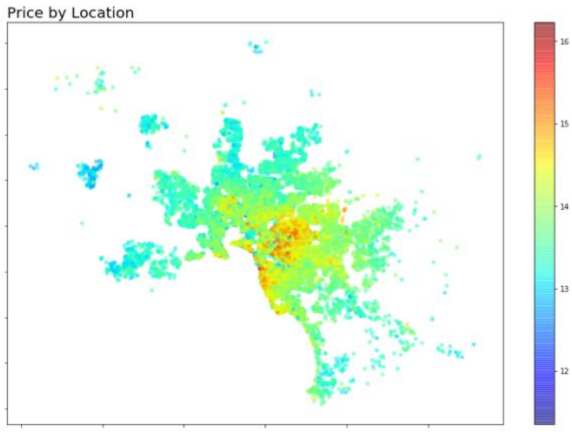


Figure 4 presents the potential influence of these geographical parameters. The map presents the different housing prices and their location in Melbourne. The map implies that prices are higher near the center and lower in the outskirts. This further justifies that the number of rooms should be controlled by geographical features.

Figure 4: Price by Location (in logs)



3. Methodology

To estimate the effect of number of rooms on housing prices we conduct two different approaches. We begin with an ordinary least squares (OLS) model specifying and testing three different variations with a varying number of control variables.

The general framework looks as follows (in vector notation):

$$y = \mu + \alpha_0 X + \beta_o Z + \varepsilon, \quad \mathrm{E}\,[\epsilon \mid X, Z] = 0, \qquad (1)$$

where y is the dependent variable (Price), X is the variable of interest (Nr. of Rooms), Z a group of covariates, μ the intercept, and ε a random error term.

We vary the number of covariates in a way such that the first model (R1) is a simple regression of Rooms on Prices, the second model (R2) includes the covariates Latitude and Longitude, and lastly the third model (R3) complements the prior models with CouncilAreas except for one to avoid the dummy trap. However, by following this approach we most probably encounter two main problems: On the one hand, in not considering potentially important covariates, we suffer from omitted variable bias (OVB). This occurs when we omit variables that correlate with the regressor and are determinants of the dependent variable. On the other hand, when using all the covariates at hand, the model may suffer from overspecification resulting in an overly complicated model. Additionally, more regressors bring more data and with it more noise, for example, deviating attributes and characteristics in individual observations which are not explained by the relationship we are interested in. (Stock and Watson, 2012)

When predetermining a selection on an informal basis, for example, based on intuition such as in model R2, said selection strategies often prove to be non-robust and lead to faulty inference. For this reason, researchers tend to employ ad hoc sensitivity analysis to test the outcome for robustness. However, according to Belloni et al. (2012) this may be useful, yet the process lacks rigorous justification.

This is the reason we determined to apply our second approach post double selection (PDS). Post double selection is an estimation method that involves two variable selection steps, followed by a final estimation step. The selection steps are designed in way that enables us to reduce the complexity of the model while tackling the OVB at the same time which has been successfully demonstrated on several examples by Urminsky et al. (2016).

The procedure builds upon the intuition of the Frisch-Waugh Theorem (1930) which offers us a successful strategy that prevents issues concerning bias when pretesting the covariates on their significance as this mostly results in a non-gaussian distribution of estimated parameters which again makes inference unfeasible. The theorem states that “Any predictor’s regression coefficient in a multivariate model is equivalent to the regression coefficient estimated from a bivariate model in which the residualised outcome is regressed on the residualised component of the predictor; where the residuals are taken from models regressing the outcome and the predictor on all other predictors in the multivariate regression (separately).” (Robinson, 2020, p.8).

In contrast to significance testing, however, post double selection employs double lasso variable selection best explained in Urminsky et al. (2016). Lasso regression is a non-parametric shrinkage technique based on the OLS model but with the addition of a penalty term to the minimization problem. In this way, it is possible for a Lasso regression to reduce parameters to zero. The formula for the Lasso estimate looks as follows, where $t \geq 0$ is a tuning parameter and alpha the penalty (James et al., 2013):

$$(\hat{\alpha}, \hat{\beta}) =$$

$$\arg\min\left\{\sum_{i=1}^{N}\left(y_i - \beta_0 - \sum_j \beta_j x_{ij}\right)^2\right\} \quad (2)$$

$$subject\ to \quad \sum_j |\beta_j| \leq t$$

With this in mind, we perform lasso regression on the partitioned regressions which are created according to the Frisch-Waugh Theorem. We then remove covariates for our final estimation if the corresponding parameters have been reduced to zero in both steps. This means that unlike in machine learning where Lasso is used to prevent overfitting by increasing bias, here the Lasso regression acts as a variable selector.

The complete PDS procedure is stated as follows (Belloni et al., 2012):

1. First Stage selection: Run Lasso on the regression of X on Z and select a set of control variables that are useful for predicting the treatment X.

2. Reduced Form selection: Run Lasso on the regression of y on Z and select additional variables by selecting control variables that predict y.

3.Final step: estimate the treatment effect α_0 of interest by the linear regression of y on the treatment X and the union of the selected control variables of the two variable selection steps.

The general framework of our PDS model can be described formally as follows (in vector notation):

$$y = \mu + \alpha_0 X + \beta_o Z + \epsilon, \quad \mathrm{E}\left[\epsilon \mid X, Z\right] = 0 \quad (3)$$

$$y = \gamma_o Z + \iota, \quad \mathrm{E}[\iota | Z] = 0, \quad (4)$$

$$X = \zeta_0 Z + \nu, \quad \mathrm{E}[\nu | Z] = 0, \quad (5)$$

where y is the dependent variable (log price), X is the variable of interest (Rooms), Z a group of covariates, μ the intercept, and ϵ, ι, and ν residual terms. By this means, we can perform a nonparametrically estimate in equations (4) and (5) with Lasso regression and use the sub selection for our final estimate in equation (3).

We repeat the PDS approach four times for four different alpha levels which controls the degree of sparsity of the estimated coefficients. This results in four different PDS models with different degrees of sparsity with model Z1 having the lowest and model Z4 the highest sparsity (Z1: alpha = 0.0005, Z2: alpha = 0.001, Z3: alpha = 0.005 Z4: alpha=0.01).

4. Results

Table 2: Results OLS (R-1 and PDS

	Effect	*Std. Error*	*Adj. R^2*	*Nr. of Controls*
Model R1	0.2711***	0.0033	0.2481	0
Model R2	0.2667***	0.0031	0.3271	2
Model R3	0.3205***	0.0026	0.5772	35
Model Z1	0.3206***	0.0026	0.5766	30
Model Z2	0.3197***	0.0026	0.5731	29
Model Z3	0.3147***	0.0027	0.5519	19
Model Z4	0.3041***	0.0028	0.5006	16

Table2 presents the results for the three OLS models and the four PDS models.

As we can see, all the estimators are significant on the $p < 0.01$ level. Model R2 possesses the lowest estimator ($a_{0,R2} = 0.2667$) which includes latitude and longitude as controls. This value is lower than the estimate for R1 ($a_{0,R1} = 0.2711$) which is the simple regression containing only the Rooms variable. However, R1 possesses a lower adjusted R^2 (= 0.2481) than R2 (= 0.3271). Additionally, the R1 model also carries the highest standard error (=0.0033). In fact, the standard error shrinks if more control variables are involved. However, this effect stagnates after a certain threshold, as all the models R3, Z1 and Z2 possess the same standard error (=0.0026) but have 35, 30, and 29 covariates.

The highest effect is achieved by model Z1 ($a_{0,Z1} = 0.3206$) followed closely by model R3 ($a_{0,R3} = 0.3205$) and model Z2 ($a_{0,Z2} = 0.3197$). Their adjusted R^2 is also almost equal with values of 0.5772 (Z1), 0.5766 (R3), and 0.5731(Z2) respectively. In other words, the removal of certain covariates didn't substantially impact the outcome. In comparison, model Z3 and Z4 reduced the complexity of the initial model immensely by removing 16 and 19 covariates, which for the latter equals to more than half of all the covariates. Nevertheless, the outcome on effect of Rooms doesn't change much compared to prior models ($a_{0,Z3} = 0.3147$, $a_{0,Z4} = 0.3041$) while having slightly lower adjusted R^2 ($adj.R^2{}_{Z3} = 0.5519$, $adj.R^2{}_{Z4} = 0.5006$). In other words, even though adjusted R^2 suffers slightly with the elimination of more than half of the covariates, in the end the effect of number of rooms on housing prices still seems to stabilize above a value of 0.3.

It the appendix, the reader finds the results of the lasso double selection in **Table3** and the detailed results of the final estimations for OLS & PDS in **Table4**.

5. Discussion

According to the results we were able to produce significant estimators in all models. Thus, we can properly assess the effect of number of rooms on housing prices. However, in contrast to the models using a high number of covariates, the estimates for R1 and R1 can be assumed to suffer from more latent OVB, as both their adjusted R^2 and estimated effect significantly differ. Moreover, we see that by adding the CouncilArea dummies in our OLS model R3 we are able to significantly increase the adj. R^2 whilst furthermore reducing the std. error. Yet, in comparison, our first PDS model Z1 is able to drop 5 control variables whilst retaining an almost identical adj. R^2 with the same estimate and Std. errors as R3. This means the model has been able to effectively eliminate unnecessary control variables with low to non-existent explanatory value. The other PDS models provide similar results with only minor differences. Thus, we interpret the PDS procedure as a success, since it enables to introduce sparsity into the model whilst keeping the outcome of the estimates nearly equal. In other words, we are able to retain the model's quality while reducing complexity and simplifying inference. But since the adjusted R^2 shrinks in accordance with higher penalization this hints that other explanatory variables, which aren't given in our specification should be added to the models.

Therefore, the limitations of the models are the persistent OVB and that we are merely analyzing a strongly simplified model of the reality of housing market prices. While geographical features should capture many other explanatory variables, such as public transport and shopping possibilities, and the number of rooms can be seen as main driver of prices some degree of OVB is still present in our analysis. The reason for this is the absence of other non-geographical related variables which have an effect on housing prices, for example, interest rates, the mortgage outstanding, accumulated maintenance and repair expenses. However, even if our dataset would encompass all the mentioned drivers of housing prices, it is unfortunately the case that OVB can never be ruled out entirely but merely minimized.

In conclusion, our findings suggest that the PDS procedure is a helpful tool when facing inference problems on models that contain an extensive number of controls. As a next step, other variables should be added to the data set which complement the information that the current variables possess so far. In addition, other methods of variable selection or preprocessing could be tested such as double machine learning which enables one to use different machine learning models than Lasso regression.

References

Alexandre B., Victor C., & Christian H. (2014). Inference on Treatment Effects after Selection among High-Dimensional Controls. *The Review of Economic Studies, 81*(2 (287)), 608-650.

Gaba, V. K., (2020), Decoupling of real economy from markets: Why stock prices don’t reflect economic reality, from website, https://www.moneycontrol.com/news/trends/expert-columns-2/decoupling-of-real-economy-from-markets-why-stock-prices-dont-reflect-economic-reality-5311021.html on 20.01.2021, 20.1.21.

James, G.; Witten, D.; Hastie, T. & Tibshirani, R. (2013**),** An Introduction to Statistical Learning, *Elsevier BV*.

Thomas S. R. (2020). *10 Fundamental Theorems for Econometrics*, from website, https://bookdown.org/ts_robinson1994/10_fundamental_theorems_for_econometrics/frisch.html, 20.1.21.

Oleg, U.; Christian, H. & Victor, C. (2016), Using Double-Lasso Regression for Principled Variable Selection, *SSRN Electronic Journal, Elsevier BV*.

Appendix

Figure 2: Correlation Matrix housing data

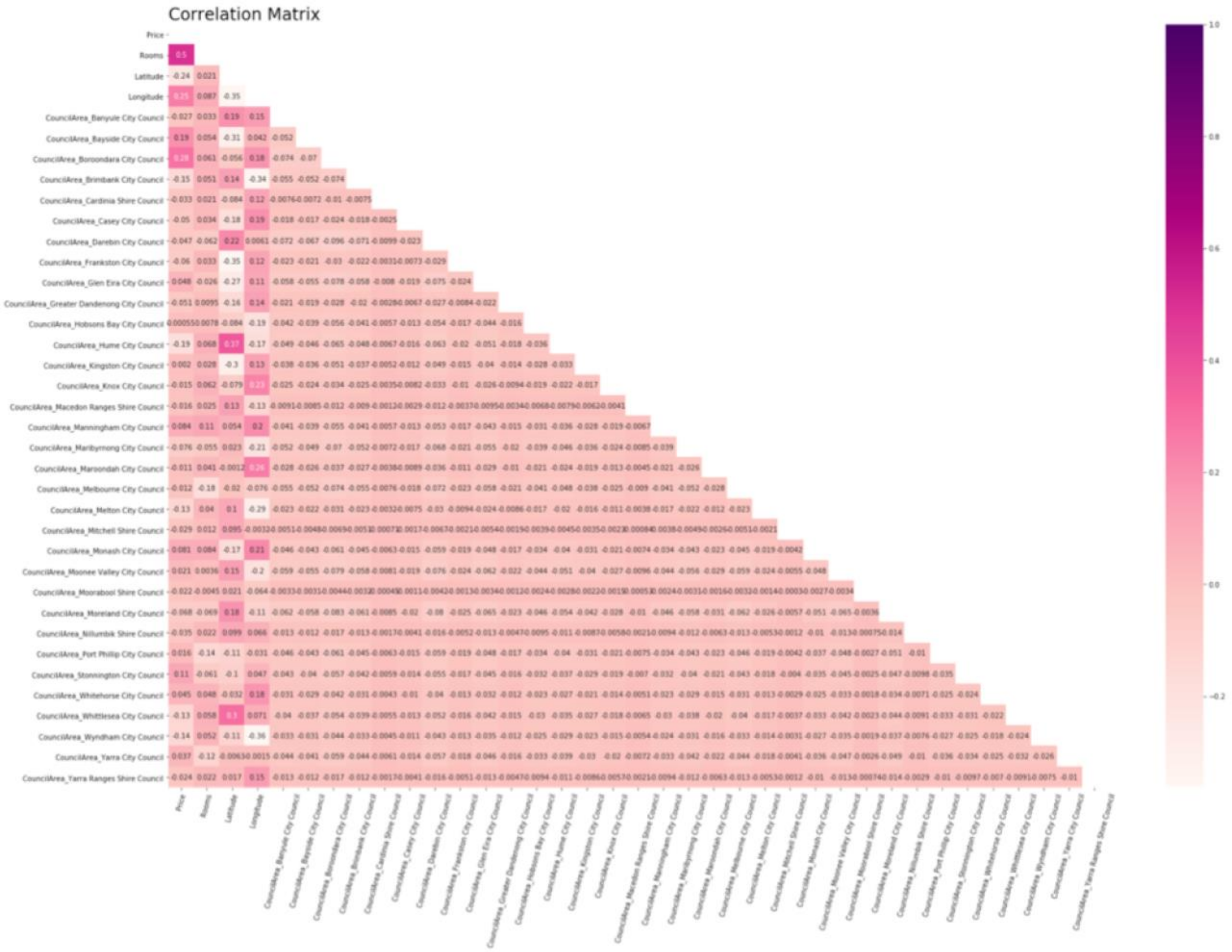


Table 3: Results Lasso Double Selection (Fs=First Selection, Rs=Reduced Selection, 1-4=Model Z1-Z4))

	Fs1	***Rs1***	***Fs2***	***Rs2***	***Fs3***	***Rs3***	***Fs4***	***Rs4***
Latitude	0.919127	-0.464143	0.443907	-0.278771	0.000000	-0.000000	0.000000	-0.000000
Longitude	0.625351	0.000000	0.311862	0.000000	0.000000	0.158130	0.000000	0.018549
CouncilArea_Banyule City Council	-0.167965	0.053504	-0.130128	0.002020	-0.000000	-0.000000	0.000000	-0.000000
CouncilArea_Bayside City Council	0.178686	0.448085	0.059492	0.442520	0.000000	0.388553	0.000000	0.275639
CouncilArea_Boroondara City Council	-0.010406	0.524050	-0.033873	0.497818	0.000000	0.440018	0.000000	0.385108

CouncilArea_Brimbank City Council	0.056455	-0.218742	0.000000	-0.245988	0.000000	-0.171186	0.000000	-0.104915
CouncilArea_Cardinia Shire Council	0.000000	-0.110192	0.000000	-0.000000	0.000000	-0.000000	0.000000	-0.000000
CouncilArea_Casey City Council	0.239768	-0.300727	0.067501	-0.187710	0.000000	-0.000000	0.000000	-0.000000
CouncilArea_Darebin City Council	-0.441719	0.029432	-0.435350	-0.000000	-0.290332	-0.000000	-0.170770	-0.000000
CouncilArea_Frankston City Council	0.385597	-0.384274	0.139296	-0.277634	0.000000	-0.000000	0.000000	-0.000000
CouncilArea_Glen Eira City Council	-0.184173	0.110941	-0.258684	0.100814	-0.168605	0.049032	-0.021091	0.000000
CouncilArea_Greater Dandenong City Council	0.000000	-0.251690	-0.000000	-0.173426	-0.000000	-0.000000	0.000000	-0.000000
CouncilArea_Hobsons Bay City Council	-0.000000	0.037257	-0.080805	0.006314	-0.000000	0.000000	-0.000000	0.000000
CouncilArea_Hume City Council	0.000000	-0.290853	0.010084	-0.340923	0.034902	-0.294545	0.000000	-0.192998
CouncilArea_Kingston City Council	0.107994	-0.000000	0.000000	-0.000000	0.000000	0.000000	0.000000	0.000000
CouncilArea_Knox City Council	0.262829	-0.000000	0.212994	-0.000000	0.000000	-0.000000	0.000000	-0.000000
CouncilArea_Macedon Ranges Shire Council	0.000000	0.000000	0.000000	-0.000000	0.000000	-0.000000	0.000000	-0.000000
CouncilArea_Manningham City Council	0.304602	0.324047	0.305297	0.276130	0.286738	0.101965	0.178953	0.000000
CouncilArea_Maribyrnong City Council	-0.347999	-0.083596	-0.403374	-0.099700	-0.284590	-0.011030	-0.117650	-0.000000
CouncilArea_Maroondah City Council	0.000000	0.000000	0.000000	0.000000	0.000000	-0.000000	0.000000	-0.000000
CouncilArea_Melbourne City Council	-0.860308	0.037366	-0.905539	0.005134	-0.783552	0.000000	-0.627341	0.000000

CouncilArea_Melton City Council	0.256423	-0.475908	0.118201	-0.469750	0.000000	-0.040381	0.000000	-0.000000
CouncilArea_Mitchell Shire Council	-0.000000	-0.000000	0.000000	-0.000000	0.000000	-0.000000	0.000000	-0.000000
CouncilArea_Monash City Council	0.236164	0.234934	0.174344	0.214781	0.118089	0.102436	0.039765	0.000000
CouncilArea_Moonee Valley City Council	-0.164484	0.143492	-0.189857	0.097999	-0.057612	0.015219	-0.000000	0.000000
CouncilArea_Moorabool Shire Council	-0.000000	-0.000000	-0.000000	-0.000000	-0.000000	-0.000000	-0.000000	-0.000000
CouncilArea_Moreland City Council	-0.466934	-0.011992	-0.473879	-0.043615	-0.328689	-0.006072	-0.190000	-0.000000
CouncilArea_Nillumbik Shire Council	-0.000000	-0.000000	0.000000	-0.000000	0.000000	-0.000000	0.000000	-0.000000
CouncilArea_Port Phillip City Council	-0.789209	0.075280	-0.851088	0.049129	-0.714471	0.000000	-0.516015	0.000000
CouncilArea_Stonnington City Council	-0.440199	0.341122	-0.485189	0.312779	-0.317893	0.190715	-0.099580	0.028448
CouncilArea_Whitehorse City Council	0.070324	0.211665	0.035001	0.162831	0.000000	0.000000	0.000000	0.000000
CouncilArea_Whittlesea City Council	-0.048507	-0.223233	0.000000	-0.267098	0.000000	-0.196291	0.000000	-0.019113
CouncilArea_Wyndham City Council	0.407905	-0.467002	0.209632	-0.448253	0.000000	-0.178988	0.000000	-0.000000
CouncilArea_Yarra City Council	-0.786509	0.162573	-0.812807	0.123959	-0.635507	0.000000	-0.426960	0.000000
CouncilArea_Yarra Ranges Shire Council	-0.000000	-0.000000	0.000000	-0.000000	0.000000	-0.000000	0.000000	-0.000000

Table 4: Results Final Estimation OLS & PDS (R1-R3 = OLS Models, Z1-Z4 = PDS Models)

	Model R1	***Model R2***	***Model R3***	***Model Z1***	***Model Z2***	***Model Z3***	***Model Z4***
const	12.9321***	13.0344***	13.2394***	13.5649***	13.3780***	12.8517***	12.2772***
	(0.0104)	(0.0242)	(0.1012)	(0.0789)	(0.0764)	(0.0259)	(0.0182)
Rooms	0.2711***	0.2667***	0.3205***	0.3206***	0.3197***	0.3147***	0.3041***
	(0.0033)	(0.0031)	(0.0026)	(0.0026)	(0.0026)	(0.0027)	(0.0028)
Latitude		-0.9029***	-1.6276***	-1.5724***	-1.3357***		
		(0.0273)	(0.0767)	(0.0712)	(0.0707)		
Longitude		0.6615***	0.0241	-0.2491***	-0.2352***	-0.3043***	0.5380***
		(0.0287)	(0.0866)	(0.0718)	(0.0695)	(0.0398)	(0.0289)
CouncilArea_Banyule City Council			0.3601***	0.1640***	0.2097***		
			(0.0489)	(0.0203)	(0.0210)		
CouncilArea_Bayside City Council			0.4287***	0.2328***	0.3395***	0.4933***	0.6476***
			(0.0555)	(0.0311)	(0.0299)	(0.0127)	(0.0126)
CouncilArea_Boroondara City Council			0.6772***	0.4850***	0.5580***	0.5280***	0.6464***
			(0.0498)	(0.0221)	(0.0218)	(0.0098)	(0.0097)
CouncilArea_Brimbank City Council			0.0189	-0.2389***	-0.1835***	-0.3458***	-0.0396***
			(0.0601)	(0.0313)	(0.0308)	(0.0155)	(0.0130)
CouncilArea_Cardinia Shire Council			-0.9119***	-0.9951***			
			(0.0873)	(0.0771)			
CouncilArea_Casey City Council			-0.6157***	-0.7382***	-0.6078***		
			(0.0588)	(0.0411)	(0.0401)		
CouncilArea_Darebin City Council			0.4235***	0.2092***	0.2585***	0.0862***	0.2517***
			(0.0513)	(0.0220)	(0.0222)	(0.0105)	(0.0099)
CouncilArea_Frankston City Council			-0.7953***	-0.9452***	-0.7803***		
			(0.0646)	(0.0464)	(0.0448)		
CouncilArea_Glen Eira City Council			0.2369***	0.0464*	0.1440***	0.2505***	0.3787***
			(0.0531)	(0.0277)	(0.0268)	(0.0116)	(0.0116)
CouncilArea_Greater Dandenong City Council			-0.3582***	-0.5103***	-0.3958***		
			(0.0565)	(0.0368)	(0.0359)		
CouncilArea_Hobsons Bay City Council			0.2113***	-0.0291	0.0545		0.3282***
			(0.0606)	(0.0344)	(0.0334)		(0.0157)
CouncilArea_Hume City Council			0.0674	-0.1793***	-0.1572***	-0.4972***	
			(0.0562)	(0.0263)	(0.0269)	(0.0149)	
CouncilArea_Kingston City Council			-0.0881	-0.2630***	-0.1444***		
			(0.0553)	(0.0327)	(0.0314)		
CouncilArea_Knox City Council			-0.0840*	-0.2261***	-0.1400***		
			(0.0503)	(0.0288)	(0.0285)		
CouncilArea_Macedon Ranges Shire Council			0.5514***				
			(0.0941)				
CouncilArea_Manningham City Council			0.4145***	0.2366***	0.2959***	0.1980***	0.2671***
			(0.0482)	(0.0217)	(0.0220)	(0.0151)	(0.0158)
CouncilArea_Maribyrnong City Council			0.2288***	-0.0108	0.0565*	-0.0303**	0.2244***

			(0.0578)	(0.0300)	(0.0294)	(0.0145)	(0.0130)
CouncilArea_Maroondah City Council			0.1463***				
			(0.0484)				
CouncilArea_Melbourne City Council			0.4993***	0.2798***	0.3509***	0.2956***	0.4871***
			(0.0545)	(0.0270)	(0.0265)	(0.0130)	(0.0123)
CouncilArea_Melton City Council			-0.2774***	-0.5838***	-0.5379***	-0.7782***	
			(0.0733)	(0.0456)	(0.0448)	(0.0301)	
CouncilArea_Mitchell Shire Council			0.2587**				
			(0.1211)				
CouncilArea_Monash City Council			0.2302***	0.0582**	0.1491***	0.2287***	0.3006***
			(0.0501)	(0.0250)	(0.0243)	(0.0139)	(0.0145)
CouncilArea_Moonee Valley City Council			0.4593***	0.2218***	0.2760***	0.1179***	
			(0.0559)	(0.0270)	(0.0268)	(0.0133)	
CouncilArea_Moorabool Shire Council			-0.1338	-0.4932***			
			(0.1868)	(0.1754)			
CouncilArea_Moreland City Council			0.3864***	0.1594***	0.2097***	0.0356***	0.2398***
			(0.0537)	(0.0245)	(0.0246)	(0.0122)	(0.0112)
CouncilArea_Nillumbik Shire Council			0.1924***		0.0250		
			(0.0635)		(0.0476)		
CouncilArea_Port Phillip City Council			0.4677***	0.2561***	0.3402***	0.3606***	0.5373***
			(0.0551)	(0.0291)	(0.0283)	(0.0145)	(0.0143)
CouncilArea_Stonnington City Council			0.6149***	0.4158***	0.4993***	0.5233***	0.6654***
			(0.0532)	(0.0273)	(0.0267)	(0.0148)	(0.0150)
CouncilArea_Whitehorse City Council			0.3259***	0.1574***	0.2312***		0.2663***
			(0.0497)	(0.0254)	(0.0253)		(0.0204)
CouncilArea_Whittlesea City Council			0.1189**	-0.0902***	-0.0693***	-0.3956***	
			(0.0511)	(0.0234)	(0.0246)	(0.0155)	
CouncilArea_Wyndham City Council			-0.4751***	-0.7580***	-0.6627***	-0.6185***	
			(0.0723)	(0.0467)	(0.0453)	(0.0234)	
CouncilArea_Yarra City Council			0.6053***	0.3949***	0.4642***	0.4036***	0.5671***
			(0.0533)	(0.0263)	(0.0259)	(0.0147)	(0.0147)
R-squared	0.2481	0.3271	0.5772	0.5766	0.5731	0.5519	0.5006
	0.2482	0.3272	0.5779	0.5773	0.5737	0.5523	0.5010
No. observations	20993	20993	20993	20993	20993	20993	20993