[Re] A New Approach to Estimating the Production Function for Housing

Aniruddha Dutta, Mansi Rathod, Mayur Gupta, Pradeep Prabhakar, Sreeja Vishaly Manohar

22 February, 2020

A reference implementation of:

Epple, Dennis, Brett Gordon, and Holger Sieg. 2010. *“A New Approach to Estimating the Production Function for Housing.”* American Economic Review, 100 (3): 905-24.DOI: 10.1257/aer.100.3.905

# Introduction

The purpose of the paper is to do a replication of the paper “A New Approach to Estimating the Production Function for Housing” (Epple, Gordon, and Sieg [2010](#ref-epple2010new)) that estimates the production function for housing. Based on two factors - the observed variation in land prices and housing values per unit of land, the original paper has provided an algorithm to identify the housing supply function per unit of land. This in turn is used to derive an estimate of the underlying production function. The approach yielded plausible estimates for the price elasticity of housing supply per unit of land, based on data from Allegheny County in Pennsylvania. In the paper,we propose a reference implementation by adding two additional models to replicate an estimation function that relates land price and home value per unit land and produce the corresponding tables and plots.Estimating housing production functions is challenging, as the quantity and price per unit of the housing services are not observed by the econometrician. Replicating the study helps us to understand how the underlying production function is estimated by treating prices and quantities of housing services as latent variables, without relying on strong functional form assumptions.

# Methods

The current methodology provided by the author for estimating house price supply function uses parametric models. The author in the paper aims to establish a function which relates land price and home value per unit land . The parametric functions used for the purpose is OLS with different variations in it. The paper has used multiple transformation on data to produce linear, log linear and polynomial linear models to best identify the relation.

Since, the function forms the root of further analysis and calculation, it is imperative that we replicate the study of these models to test the robustness and the assumptions behind the models. We decided two split the method of implementation in two direction. The first method is to use a different model for replication and the other is changing the loss function used in the model.

For the first case of different model, we decided use Generalized Linear Model with different distribution families. Using different families allows us to verify the condition of normality of error in the models used by authors. We noticed that the gaussian distribution family with log independent and dependent variable produced a line that fitted the best with the data and the corresponding coefficients were very similar to the log linear model used by the author.

In the case of different loss function, we wrote a function code to implement log linear regression with gradient descent loss function. The gradient descent loss function provides a more flexible approach because of the presence of hyperparameters *learning rate* and *number of iteration*. The author of the paper has not explored this methodology, therefore we had no information about the value of hyperparameters *learning rate* and *number of iteration*, therefore we ran a simulation study to determine the hyperparameters that delivered values of coefficients very similar to that of the models used by the author

# Results

The original paper had estimated the r(v) function which characterizes the relationship between the price of land and the value of housing per unit of land. Using OLS for log-linear, linear, quadratic and cubic models, it reports the estimates of the equilibrium locus. We have enclosed the same in Table 1. We have replicated the same results using gaussian log-linear model and gradient descent log linear model. Results are enclosed in Table 2. For both of these two models, they fit the main features of the residential data substantially well and their performance is similar to the earlier four models. All the p-values were calculated using heteroskedasticity-robust standard errors and they were significant at 1% level. For effective and comprehensive comparisons of all the six models used, we have also enclosed a plot of all our six models. As iterated, we see r(v) functions form all the six models yielding similarly shaped production functions

Table 1: Table

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| X | Log Linear | Linear | Quadratic | Cubic | Gaussian Generalized Linear Model | Gaussian Generalized Log Linear Model | Gradient Descent Log Linear Model |
| v |  | 0.1394 | 0.1685 | 0.1622 | 0.1526 |  |  |
| v2 |  |  | -2e-04 | -1e-04 |  |  |  |
| v3 |  |  |  | 3.9e-07 |  |  |  |
| Constant | -1.605 |  |  |  |  | -1.631603 | -1.631195 |
| log(v) | 0.909 |  |  |  |  | 0.918632 | 0.918491 |
| N | 6234 | 6234 | 6234 | 6234 | 6234 | 6234 | 6234 |

# Conclusion

# References

1.@epple2010new

Epple, Dennis, Brett Gordon, and Holger Sieg. 2010. “A New Approach to Estimating the Production Function for Housing.” *American Economic Review* 100 (3): 905–24.