# A COMPARISON OF DEA AND STATISTICAL REGRESSION ANALYSIS IN EXPLORING PROPERTY VALUE EFFECTS OF FERRY TERMINALS IN BRISBANE

**BARBARA T.H. YEN**
Urban Research Program, Griffith University, 170 Kessels Road, Nathan, Brisbane, Queensland 4111, Australia, (corresponding Author)

**CHI-HONG (PATRICK) TSAI**
Institute of Transport and Logistics Studies, the University of Sydney Business School

**CORINNE MULLEY**
Institute of Transport and Logistics Studies, the University of Sydney Business School

**MATTHEW BURKE**
Urban Research Program, Griffith University

## ABSTRACT

*Property value effects of urban river ferries that run linear services to multiple stops are not well understood. Brisbane's CityCat, CityHopper and CityFerries together service 24 terminals in a connected system. This paper employs the Data Envelopment Analysis (DEA) approach to measure the property value effects of that system. The prediction performance of DEA is then compared with statistical regression methods. The research findings suggest that DEA has the capability of portfolio analysis to understand which types of properties benefit from the ferry service more. By contrast, the statistical regression models have the better flexibility in selecting the explanatory variables and greater power in identifying the magnitude of impact on property price from each of explanatory variable. This research sheds light on how DEA can be applied as a tool to evaluate the relationship between property value and price determinants, which has been conventionally undertaken by statistical regression methods.*

***Keywords****: data envelopment analysis (DEA); statistical regression, property value effects.*

## INTRODUCTION

Many cities have introduced linear passenger ferry systems providing frequent services along rivers or parallel to shorelines, servicing multiple stops (Thompson et al., 2006). Key examples are the East River Ferry in New York, the Chao Praya Express in Bangkok, and the CityCat ferries in Brisbane, Australia. In Brisbane and in New York, these systems are seen as helping stimulate urban land (re)development. This paper explores the relationships between property value and the ferries in Brisbane as part of broader research on public transport and land value effects.

In order to measure the property effects Data Envelopment Analysis (DEA) modeling is applied. DEA is a non-parametric technique that employs the linear programming method to determine relative efficiencies of a set of homogeneous and comparable observations (Chiou et al., 2012). DEA provides an objective measure of the observed units to two simultaneous perspectives: maximization of outputs and minimization of inputs. In this paper apply the value estimation made by DEA, Regression Analysis (RA), Geographically Weighted Regression (GWR) and real estate data over a geographic data basis of Brisbane ferry systems and analyze them comparatively. We use GIS to visualize the DEA analysis results.

The rest of this paper is organized as follows. Section 2 addresses the methodology adopted. The empirical study is conducted in Section 3. The paper concludes with a discussion as to what the results may mean for broader transport and land use policy and practice, and avenues for further research.

## METHODOLOGY: DEA VS. RA

*Data envelopment analysis*

DEA was initially developed as a method for assessing the comparative efficiencies of organizational units. The key feature which makes the units comparable is that they perform the same function in terms of the kinds of inputs they use and the types of outputs they produce.

DEA was first developed by Charnes, Cooper and Rhodes (1978), who established a linear combination of outputs and inputs to measure the efficiency for observations by integrating the outputs/inputs ration efficiency measure (henceforth, CCR model). CCR model is under the assumption of constant returns to scale (CRS) production technology. Banker, Charnes and Cooper (1984) relaxed this CRS constrain to variable returns to scale (VRS) technology by adding convexity constraint (henceforth, BCC model). Under VRS assumption, BCC model could identify each observation is performing in region of increasing, constant or decreasing returns to scale with multiple inputs and multiple outputs. These two model forms are illustrated as following.

[BCC] $\underset{z,\lambda_i}{Min} \quad z$ (1)

s.t. $zx_{kj} - \sum_{i=1}^{n} x_{ij}\lambda_i \geq 0$, $j = 1,2,\cdots,m$ (2)

$-y_{kr} + \sum_{i=1}^{n} y_{ir}\lambda_i \geq 0$, $r = 1,2,\cdots,s$ (3)

$\lambda_i \geq 0$, $i = 1,2,\cdots,n$ (4)

$\sum_{i=1}^{n} \lambda_i = 1$ (5)

The combination of equation from Eq. (1)-(4) and (1)-(5) form the CCR and BCC models, respectively.

*Statistical regression analysis*

In recent years, hedonic method is used for investigating the property value effect. More recently, hedonic method is developed in from using an RA, more specific ordinary least squares (OLS) methodology. However, OLS method fails to reveal spatial variability due to assumption of relationships between variables are consistent geographically.

In order to investigate property value effect with geographical effects, GWR is adopted in this study. GWR (Fotheringham et al., 2002) recognises that relationships between variables are likely to vary across space. For public transport infrastructure, Du and Mulley (2006) first applied GWR to investigate the relationship between transport accessibility and land value in England for light rail. GWR is most conveniently explained by reference to a traditional cross-sectional regression model (Fotheringham et al., 2002), with this model written as

$$y_i = \beta_0 + \sum_j x_{ij}\beta_j\beta_{ij} + \varepsilon_i \tag{7}$$

The GWR methodology expands this model to a form which allows for local variations in the parameter values which take account of the coordinates of individual regression points. If the dependent variable has the coordinates $(u_i, v_i)$ the model expressed in Eq.(7) above can be rewritten as the following GWR local model:

$$y_i(u_i, v_i) = \beta_0(u_i, v_i) + \sum_j x_{ij}\beta_j(u_i, v_i)\beta_{ij} + \varepsilon_i \tag{8}$$

The parameters are estimated at the location $(u_i, v_i)$ using a weighted least squares method and a predicted value of $y_i$ effectively giving a regression at each point of the dataset.

## EFFICIENCY ANALYSIS

### *Data*

Property prices are determined by three major characteristics: 1) the internal characteristics of the property; 2) the neighbourhood in which it is located and; 3) the accessibility of the property location. The interest of this paper is measured the property value effect from linear ferry system. Burke and Brown (2007) identified that the 85th percentile access to Brisbane ferry terminals was known to be 1.54km. Therefore, the study area was defined by a 2 km buffer either side of the centre line of the river. This was chosen to ensure that all properties that could be affected by accessibility to terminals were captured.

The output/dependent variable are 2011 transaction price data from RPdata which combines data from different sources to provide details of the transaction price, property type (house or unit), area size, number of bedrooms and bathrooms and the number of parking places. The study area included 2,832 observations on house prices for analysis. The sample of case study is identical as Tsai et al. (2014).

The internal characteristics of the property include the number of bedrooms, and whether the property is a house or unit. The neighborhood features were controlled by a number of demographic variables which is collected at the smallest geography available in Australia from the 2011 Census. Each sample in study area was linked to the associated demographic variables.

Accessibility variables were calculated using GIS and are the network distances for each variable including to the nearest bus rapid transit (BRT) stop, to the central business district (CBD), to the nearest ferry wharf and to the nearest train station. The distance to the nearest BRT stop was removed from the analysis since this is highly correlated with distance to the CBD (r=0.901, significant at the 1% level). Table 1 shows the descriptive statistics for the data.

**Table 1: Descriptive statistics of variables**

| Variable | Unit of measurement | Minimum | Maximum | Mean | Std. Deviation |
|---|---|---|---|---|---|
| Price | $ | 96,500.00 | 4,800,000.00 | 645,582.79 | 413,621.38 |
| Bedrooms | number | 0.00 | 9.00 | 2.65 | 1.09 |
| Type | House/Unit | 0.00 | 1.00 | 0.48 | 0.50 |
| Unemployment | % unemployed persons | 0.00 | 0.21 | 0.05 | 0.03 |
| BRT | km | 0.21 | 8.86 | 2.94 | 1.75 |
| CBD | km | 0.41 | 10.98 | 4.53 | 2.22 |
| ferry | km | 0.01 | 5.54 | 1.59 | 0.94 |
| train | km | 0.02 | 4.96 | 1.48 | 0.83 |

*DEA empirical results*

In line with formulae (1)-(5), Table 2 report the performance of each property types and compares the performance under different property types. It is clear from Table 2, small bedroom properties, 1 bedroom units and 1~2 bedroom houses, will have higher performance index. The research findings indicate that the DEA has the capability of portfolio analysis to understand which types of properties benefit from the ferry service.

**Table 2: Performance index and Distribution under each property type by DEA model**

| **Property type** | | **all** | **Unit** | | | **House** | | | |
|---|---|---|---|---|---|---|---|---|---|
| Number of bedrooms | | all | 1 | 2 | 3+ | 1-2 | 3 | 4 | 5+ |
| Number of properties | | 2837 | 264 | 841 | 270 | 309 | 657 | 334 | 162 |
| Mean | | 0.412 | 0.698 | 0.388 | 0.301 | 0.600 | 0.344 | 0.303 | 0.397 |
| Max | | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| Min | | 0.073 | 0.257 | 0.155 | 0.106 | 0.274 | 0.126 | 0.105 | 0.073 |
| Performance index distribution | <0.4 | 61.72% | 3.03% | 68.25% | 81.11% | 20.06% | 76.86% | 82.63% | 66.05% |
| | 0.4-0.6 | 20.66% | 29.92% | 23.31% | 8.89% | 38.83% | 16.74% | 10.48% | 13.58% |
| | 0.6-0.8 | 9.73% | 40.91% | 5.23% | 3.70% | 19.09% | 4.26% | 3.59% | 9.26% |
| | 0.8-0.99 | 3.88% | 18.94% | 1.78% | 2.59% | 8.41% | 0.61% | 0.90% | 3.09% |
| | 1 | 4.02% | 7.20% | 1.43% | 3.70% | 13.59% | 1.52% | 2.40% | 8.02% |

*GWR empirical results*

The house price is the dependent variable which is related to explanatory variables as shown in Eq.(9) where *C*, *N* and *T* are respectively vectors of the internal property characteristics, neighbourhood features and transport accessibility measures.

$$P_i = f(C,N,T) \cdot TE_i \tag{9}$$

where $P_i$ represents the transaction price for home $i$, $f(C,N,T)$ is the frontier and $TE_i$ defines the technical efficiency (the ratio of the observed selling price that the home sold for to the maximum feasible price) (Samaha and Kamakura, 2008). Thus, the observed selling price of a home, $P_i$, reaches its maximum feasible value of $f(C,N,T)$ only when $TE_i = 1$. Otherwise, $TE_i < 1$ provides a measure of the shortfall of observed price from the maximum feasible price. Table 3 reports the GWR results. From

Table 3, GWR has better flexibility in selecting the explanatory variables and greater power in identifying the magnitude of impact on property price from each of the explanatory variables.

**Table 3: GWR model results**

| Variables | Coefficient | Standard Error | t-value |
|---|---|---|---|
| Constant | 622645.50 | 44567.54 | -2.06 |
| Bedrooms | 110177.70 | 7521.89 | 14.65 |
| Type | -123879.00 | 13901.71 | -8.91 |
| Unemployment | 565172.60 | 261648.90 | 2.16 |
| CBD | -14858.00 | 2964.72 | -5.01 |
| ferry | -44268.30 | 6609.18 | -6.70 |
| train | 24128.91 | 7312.67 | 3.30 |
| Observations | 2,832 | Adj. R-squared | 0.53 |

Table 4 report the performance of each property types and compares the performance under different property types. It is clear from Table 4, small bedroom properties, 1 bedroom units and 1~2 bedroom houses, will have higher performance index.

**Table 4: Performance index and Distribution under each property type by GWR model**

| Property type | | all | Unit | | | House | | | |
|---|---|---|---|---|---|---|---|---|---|
| Number of bedrooms | | all | 1 | 2 | 3+ | 1-2 | 3 | 4 | 5+ |
| Number of properties | | 2837 | 264 | 841 | 270 | 309 | 657 | 334 | 162 |
| Mean | | 0.191 | 0.422 | 0.248 | 0.367 | 0.234 | 0.191 | 0.183 | 0.252 |
| Max | | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| Min | | 0.001 | 0.001 | 0.073 | 0.143 | 0.001 | 0.001 | 0.001 | 0.001 |
| Performance index distribution | <0.4 | 88.30% | 46.20% | 94.00% | 69.40% | 93.40% | 97.40% | 96.70% | 84.70% |
| | 0.4-0.6 | 10.00% | 46.80% | 5.40% | 26.50% | 5.30% | 1.80% | 2.40% | 11.60% |
| | 0.6-0.8 | 1.20% | 5.20% | 0.50% | 2.60% | 1.00% | 0.60% | 0.30% | 2.50% |
| | 0.8-0.99 | 0.10% | 0.00% | 0.00% | 1.10% | 0.00% | 0.00% | 0.30% | 0.50% |
| | 1 | 0.40% | 1.70% | 0.10% | 0.40% | 0.30% | 0.20% | 0.30% | 0.60% |

## DISCUSSIONS

Table 2 and Table 4 show that DEA has higher performance index than GWR. In order to compare the performance index, Figure 1 compares the performance index distribution of DEA and GWR. Obviously, the benchmarking power of DEA is much superior to that of GWR. It should be mentioned that the performance index cannot be compared directly because the problem structures are different—DEA is a boundary method while GWR is a non-boundary method. This different between DEA and GWR has very important repercussions for the results given by the two methods.

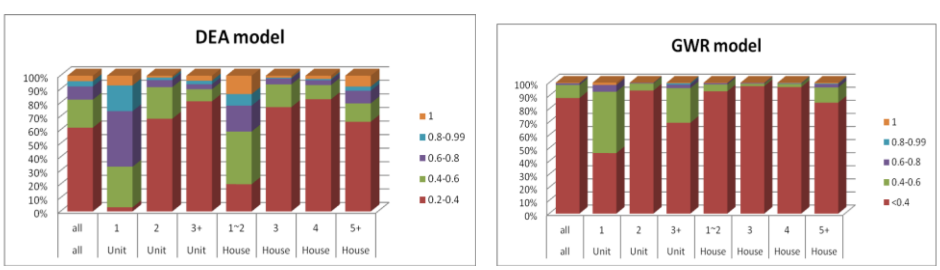


**Figure1 Distribution of performance index under different groups**

Figure 2 shows the geographical distribution of all properties within the study area and each property is coloured according to 75% quantile of performance index. The properties coloured in red is 75% quantile of performance index, that is, these properties have higher performance.

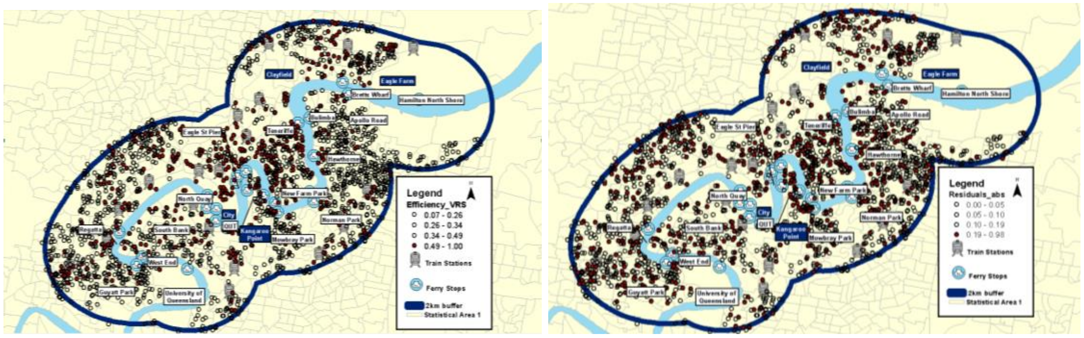


**Figure 2: 75% quantile of performance index by DEA and GWR**

Under both the DEA and GWR model, and focusing on the benefit of the ferry services, the land value gains are identified to be higher around the Bulimba, Hawthorne and Regatta terminals. This seems to be a result of the greater redevelopment potential and the maturity of new development surrounding the ferry terminals. More positive effects are also identified at the Mowbray Park terminal, which is on the CityCat route, as compared to the Norman Park terminal, which has lower frequency CityFerry services only. By contrast, very little land value uplift is identified around the Tenerife and North Shore Hamilton terminals as these have only entered service in the last two years and urban redevelopment remains quite limited at these locations.

This study introduces the application of DEA to identify the impact of the accessibility to the ferry services on residential property prices in the central area of Brisbane. The power of property value effect assessment using DEA is then compared with the GWR. The research findings in general confirm that property values in the study area do benefit from accessibility to ferry, especially in the areas where residential redevelopment has taken place around the ferry terminals. Being close to a ferry terminal in the inner Brisbane CBD adds to property values in contrast to being close to train which is designed more to serve outer suburbs. This suggests that like other rapid transit services, ferry can potentially generate substantial economic benefit as long as it is designed as a frequent and rapid service with a sufficiently wide geographical coverage, connecting major business centres and attractions.

The implications are that property developers are justified in seeking to secure ferry terminals to service their developments. Governments may also be justified in bringing in land value capture mechanisms to help pay for terminals, vessels or operating costs in appropriate locations.

Future research will focus on the investigation of other potential price determinants that may influence some particular areas as discussed, including the role of large university campuses, as well as developing the Geographically Weighted DEA model to consider geographical effects. A further segmentation of the study area such as creating catchment buffers around ferry terminals will also be considered as ways of further improving model fit and interpretation.

## ACKNOWLEDGMENTS

The acquisition of data has been supported by a University of Sydney Business School grant. Transport research at Griffith University's Urban Research Program is supported by the Academic Strategic Transport Research Alliance, involving the Queensland Government Department of Transport and Main Roads, the Motor Accident and Insurance Commission, and Queensland Motorways Limited. The views expressed are solely those of the authors, who are responsible for all errors and omissions.

## REFERENCE

[1] Banker, R.D., Charnes, A., Cooper, A.A., 1984. Some models for estimating technical and scale inefficiencies in data envelopment analysis. Management Science 30, 1078-1092.

[2] Burke, M., A.L. Brown, 2007. Distances People Walk for Transport. Road & Transport Research 16, 16-29.

[3] Charnes, A., Cooper, W.W., Rhodes, E., 1978. Measuring the efficiency of decision making units. European Journal of Operational Research 2, 429-444.

[4] Chiou, Y.C., Lan, L.W., Yen B.T.H., 2012. Route-based data envelopment analysis models. Transportation Research Part E 48, 415-425.

[5] Du, H., Mulley, C., 2006. Relationship between transport accessibility and land value: a local model approach with Geographically Weighted Regression, Transportation Research Record: Journal of the Transportation Research Board 301.

[6] Fotheringham, A.S., C. Brunsdon, Charlton, M.E., 2002. Geographically Weighted Regression: The Analysis of Spatially Varying Relationships. Chichester, UK: Wiley.

[7] Samaha, S.A., Kamakura, W.A., 2008. Assessing the market value of real estate property with a geographically weighted stochastic frontier model. Real Estate Economics 36, 717-751.

[8] Thompson, R., Burroughs, R., Smythe, T., 2006. Exploring the Connections Between Ferries and Urban Form: Some considerations before jumping on board. Journal of Urban Technology 13, 25-52.

[9] Tsai, C-H, Mulley, C., Burke, M., Yen, B.T.H., 2014. Exploring property value effects of ferry terminals: Evidence from Brisbane, Australia. World Symposium on Transport and Land Use Research, Delft, June.