APPENDIX C
Robust regressions

In linear regressions basic fundamental assumptions include:

1. The relationship is linear so that the errors all have an expected value of zero
2. Errors all have the same variance
3. Errors are independent of each other
4. Errors are all normally distributed

However, if the data has outliers present it can cause change in the slope line of the regression. In the Ordinary Least Squares (OLS) approach coefficient values are related to the mean, which is sensitive to outliers. A mean may not be the best measure. One option is to remove the outliers but simply removing the outliers can underestimate variances and the dependent variable values in this case are viewed as real values that reflect the spatial differences in prices. An alternative estimation technique is to down-weight the extreme observations rather than rejecting them.

Robust regressions provide an alternative to OLS when the data violates the fundamental assumptions. This procedure reduces the influence of outliers to estimate a better fit for the majority of the observations. It uses iterative methods to assign different weights to residuals until the regression converges.

In the analysis used in this paper a Huber-type M estimator was used. The basic algorithm for calculating M estimates is iteratively reweighted least squares. 'For each iteration, a set of weights for the observations is used in the least squares fit' (SAS 2009).

Akaike's Information Criterion (AIC) is a measure of the goodness of fit of the estimated model. It is not used for hypothesis testing but model selection. Competing models are compared through AIC with the lowest AIC representing the best fit. AIC also places a penalty on over-specification and attempts to represent the model that best explains the data. An alternative measure, but closely related to AIC, is the Bayesian Information Criterion (BIC). This again is a measure of goodness of fit but places a higher penalty for over-specification than AIC. In regards to the model presented the AICR and BICR are for robust regressions but the principle remains the same.

For a complete discussion on SAS robust regression refer to user manual SAS/STAT 9.2 User guide.