House Price Indexes for Submarkets:
A Blending Approach

Andy Krause

23 October 2016

Abstract

Traditionally, literature on house price indexes has focused on the choice of method (repeat sales vs. hedonic vs. hybrid, etc.) or correcting for biases resulting from issues such as sample selection, aggregation, or temporal heterogeneity. While more recent research has explored the production of indexes at the sub-city or submarket level there remains, however, a lack of comparative analytics to determine the limits to market disaggregation in terms of generating reliable price indexes. The goals of this study are two-fold: 1) Compare index performance at different sub-market levels (county, ZIP, etc.); and 2) To test if blended indexes – between market levels and between methods – offer any improvement over more traditional non-blended models.

Introduction

Since the seminal Bailey et al. (1963) study, there has been considerable and sustained research effort put into comparing and improving competing methods for generating house price indexes. Published work in this subfield of housing economics is generally focused on one or more of four aims: 1) Comparison of model performance; 2) Identification and correction of an estimation issues or problems; 3) Development of a new model or estimator; and 4) Creation of local or submarket indexes.

Studies comparing various indexing method were popular in the 1980s and 1990s (Mark & Goldberg 1984; Case et al 1991; Clapp et al 1991; Crone & Voith 1992; Gatzlaff & Ling 1994; Meese & Wallace 1997). In most cases, work published during this period compared indexes generated with some permutation of repeat sale, hedonic, median/mean, assessed/appraised value and/or hybrid methods. The results suggest little agreement in terms of the preferred method, though the median/mean based methods used in industry are generally shown to be less preferred (Mark & Goldberg 1984; Crone & Voith 1992; Meese & Wallace 1997). Clapp et al (1991) suggest that hedonic method may outperform repeat sales models in short time periods, while a more recent study by Nagaraja et al (2014) show better results from a hybrid, autoregressive model (expanding on Case & Quigley (1991)) than from more traditional repeat sales models. Dorsey et al (2010) show substantively different results between hedonic and repeat sales models in identifying the peak of the market during the housing boom of the mid 2000s. Common across all studies, but most explicitly made by Case et al (1991), are discussions of the various shortcomings and issues that plague each model type.

Out of this arose a suite of research that set out to catalog, measure and, often, correct for the issues identified in the earliest comparative work. The most commonly addressed issue is that of sample selection bias (Abraham & Schauman 1991; Haurin & Henderschott 1991; Clapp et al 1992; Case et al 1997; Steele & Goy 1997; Gatzlaff & Haurin 1997, 1998; Munneke & Slade 2000), an issues affecting nearly all of the available models. Issues stemming from the necessity to revise indexes over time, especially in the case of repeat sales models, have received attention in the literature (Clapp & Giacotto 1999; Butler et al 2005; Clapham et al 2006; Deng & Quigley 2008). Property age and depreciation and the biases associated with it have also been raised and methods presented to control for it have been offered (Goodman & Thibodeau 1997; Cannaday et al 2005; Chau et al 2005). Adding to the list of biases are potential problems arising from seller reservation bias (Goetzmann & Peng 2006) and submarket (or product) aggregation bias (Guttery and Sirmans 1998).

A related set of work tackles issues stemming from the treatment of time in the various price indexing models. In short, the aggregation of time periods required to keep per period sample sizes high enough may create biased results (Englund et al 1999; Dreiman & Pennington-Cross 2004), as may the assumption in standard hedonic approach of time-consistent coefficients for the non-temporal independent variables (Knight et al 1995, Zabel 1999). To combat violations of the temporally stability in coefficient estimates a number of time-varying parameters approaches have been proposed (Knight et al 1995; Zabel 1999; Munneke & Slade 2001). Alternative time parametizations have, likewise, been presented to provide alternatives to excessive temporal aggregation (Schwann 1998; McMillen & Dombrow 2001).

Studies addressing spatial dependence in the models represent a final set of issues that have been raised (Can & Megbolugbe 1997; Gelfand et al 2004; Tu et al 2004; Nappi-Choulet & Maury 2009; Dorsey et al 2010). The primary goal here is to improve the model estimates, and therefore, the index values by better accounting for the inherent spatial structure of housing markets.

A third category of studies present novel improvements to existing data generation, models or estimators. The first and most common are those presenting hybrid hedonic and repeat sales models aimined at capturing the benefits and negating the shortcomings of both individual model types (Case & Quigley 1991; Quigley 1995; Hill et al 1997; Englund et al. 1998, Nagaraja et al 2014; Guntermann et al 2016). A number of more recent improvements have been offered to create pseudo-repeat sales (McMillen 2012; Guo et al 2014), estimate robust statistical models (Bourassa et al 2013; Bourassa et al 2016), produced chained indexes (Clapham et al 2006) and address small samples sizes with frequency a conversion method (Bourassa & Hoesli 2016).

The final broad category of research aims are those dealing with indexes for submarkets. In this context submarkets can be strictly geographic (Goodman 1978; Hill et al 1997; Gunterman et al 2016) or non-spatial in nature, based on price quantiles or housing product – 2br vs 3br, for example. (Pollakowski et al 1991; Guttery & Sirmans 1998; McMillen & Thorsnes 2006; Coulson & McMillen 2007; Jansen et al 2008; Prasad & Richards 2009). The most straightforward approach to dealing with submarkets involves simply segmenting the data by submarket definition (geographic or otherwise) and using an established method to generate an index. More complex methods have been proposed as well. These primarily include the use of weighting observations to better match the submarket for which the index is being generated (Meese & Wallace 1991; Goetzmann & Spiegel 1997; McMillen 2003; Clapp 2004; Jansen et al 2008). Spatially weighted approaches like the semi-parameteric method proposed by Clapp (2004) allow for a 'surface' of price indexes to be created that vary continuously over space.

Submarkets and Hierarchy

Though submarkets naturally exist in some hierarchical manner, at the very minimum in a two-level hierarchy with the overall market, few studies have acknowledged or leveraged this condition. A set of recent research has generated indexes at various market levels – county, city, neighborhood, ZIP code, etc. – (Dorsey et al 2010; Francke 2010; Bogin et al 2016; Bourassa & Hoesli 2016), but only Bourassa and Hoesli (2016) have provided comparative statistics on the various performance at each level[1].

A related set of work uses a blended (and weighted) combination of local level indexes to create a weighted overall market index series. Haurin et al (1991) blend three different sources of indexes (with various permutations of these) at the local level to create a local index and then blend these local indexes 'upward' to create weighted regional and national series. Hill et al (2008) do likewise and blend their 15 local indexes from Sydney to create a weighted metropolitan index. Even greater aggregation and upward blending is done by Dorsey et al (2010) who combine over 290 individual zip code level series to create metropolitan wide indexes for Los Angeles and San Diego.

In sum, sub-markets are well recognized to play an important role in real estate market analysis. Recent studies (Dorsey et al 2010; Francke 2010; Bogin et al 2016; Bourassa & Hoesli 2016) have decreased the level of spatial aggregation at which house price indexes are calculated. Estimates of price indexes for these small submarkets, while of considerable interest to analysts, policy-makers and market participants, are often troubled by low sample sizes resulting in high volatility and poor out-of-sample predictive performance (unless significant temporal aggregation is performed (Bogin et al 2016)). Leveraging the hierarchy of submarkets, the spatial dependence in housing markets and ensemble approaches popular in other 'predictive' endeavors, I propose two 'blending' approaches, described below, to help address these sample size/volatility issues.

[1] Note that Francke and Vos (2004) utilize a hierarchical trend modeling approach but do not specifically create indexes at different market levels

Methods

A variety of methods to estimate home price indexes have been proposed in the literature. De Haan and Diewert (2013) offer an in-depth review of these methods. In classifying price index methods, there are three key components or dimensions: 1) Data; 2) Model; and 3) Estimator. Not all permutations of the three create functional models, or rather, some combinations of data, model and estimators do not work together.

Data

In choosing a price index to use, the availability of data is the first consideration. Three major categories of data are used to develop house price indexes. A number of models and estimators use a compination of the three types.

1. Transactional observations: All recorded transactions of housing units in the (sub)market over the period of analysis.

2. Repeat transactions: Paired transactions of homes that have transacted more than once during the period of analysis. Recent work has also produced psuedo-repeat transaction dataset based on propensity or other matching techniques (McMillen 2012; Guo et al 2014).

3. Assessment or Appraisals values: Home values from assessements, valuations or appraisals of homes that may or may not have transacted.

Models

Three basic model types are used throughout the published literature, 1) Repeat Sales; 2) Hedonic Price; and 3) Appraised or Assessed Value. A fourth method, median or mean, is often used in industry, but due to changing characteristics of sold homes over time (Case and Shiller 1987) this method can be severly biased and is ignored in this study.

1. Repeat Sales

The repeat sales model was originally proposed by Bailey et al (1963). In this model, only sales that have sold more than once are considered. The difference in price between paired sales of a single property are regressed on the difference in time between the two sales. This model operates under the assumption that no significant changes have occurred to the property between the two sales. If this assumption holds, then the difference in price between the two sales can be solely attributed to the changing market. Case and Shiller (1989) offer a three-step weighting procedure to downweight paired sales with greater time differences as these properties are more likely to experience either significant depreciation or a renovation.

The biggest benefit to repeat sales models are that only the price and date of sale are needed. The model can be biased if the sample of sales that transaction more than once are not representative of the population of homes in the area or if the constant quality across time assumption is violated.

2. Hedonic Price

Hedonic price models are the second most common approach to creating house price indexes. A set of binary or indicator variables are specified for each time period and the coefficients of these are then used to derive a price index. The hedonic price index uses all sales transactions and, therefore, is more informationally efficient than a repeat sales approach. By controlling for property characteristics, renovations, depreciation and overall changes to the set of homes selling in a market are more adequately controlled for in this approach. Model misspecification, however, may impact coefficients, leading to a biased result. The hedonic price method requires a richer set of data that includes characterstics of the properties, data that may not always be available.

3. Appraised/Assessed Value Ratio

Appraisal or assessed value ratio models take a number of formats (Clapp and Giacotto 1992; Bourassa et al 2006; de Vries et al 2009). The most straightforward, and the one used in this paper, is the to compare sales prices over a time period to a constant (often current) assessed or appraised value. The change in this ratio – sales price divided by assessed value – over time produces an estimate of the general movement in the market. These models benefit by using all transactional information for which there is an associated sales price, but can be severly hampered by inconsistencies in the assessed or appraised values being used as a base.

Estimators

Finally, for each method there are different tatistical estimation techniques that can be used. While there are many to choose from, the three most common and those examined in this paper are: 1) Basic; 2) Weighted; and 3) Robust.

1. Basic

Basic estimator are the simplest form of each of the models listed above. For the repeat sales and hedonic price models this means that all data points are treated equally and the models are run in the simplest specification (usually a linear regression model). For the assessed value models, basic estimator uses all data points, regardless of outliers.

2. Weighted

Weighted estimators place weights on the individual observations based on a value of one of the independent variables. The most common example is the weighted approach used by Case and Shiller (1987, 1989) to downweight repeat sales with greater distance between sales.

3. Robust

Robust estimators use an iterative approach to down weighted observations that are having an overly large impact on the estimator. These methods have been applied relatively recently to house price indexes (e.g. Bourassa et al. 2013; 2016) and will be tested in this research.

Geographic Coverage

At least three different levels of submarkets will also be considered:

1. County level
2. Random spatial aggregation (1/10 of county)
3. ZIP code level

The county and ZIP code levels are obvious and need little explanation. In order to produce a mid-level of spatial aggregation that does not have any relation to the ZIP codes, I have used a constrained K-means clustering algorithm to divide the county up into 10 areas, each with a minimum of 5% of the total number of sales in the county. These zones are 'velocity' driven – meaning they have some relation to sales volumes – as opposed to ZIP codes which are merely artifacts of postal delivery.

Blending

After the the above models with associated estimators are created, two blending methods will be performed. For the highest level (county in this case) no hierarchal blending is available.

1. **Hierarchical Blending:** Theoretically, prices in small geographic areas (or for specific markets segments such as studio condominiums) are driven by both demand factors in that specific submarket as well as by trends in the wider market. A hierarchical blending approach blends indexes from a higher market level (ex. County) with those of smaller submarkets (ex. ZIP code) to allow local data to speak but also to control volatility.

2. **Model Blending:** Ensemble models are becoming popular in many fields such as weather forecasting and sports betting and have produced winning entries in multiple Kaggle contests. The same approach can be uses for creating house price indexes. Estimated indexes from a variety of models (repeat sale, hedonic, SPAR, hybrid, etc.) can be combined to produce an ensemble or blended model.

For all approaches, weights in the blending process can be determined based on a measure of volatility. More specifically, for each index, the average volatility will measured by the standard deviation of the index value for the three months prior. This volatility value is calculated for all index time periods 4 through n, where n is the total number of periods estimated. These n-3 estimates of volatility are then averaged to create an index-wide volatility measure. When blending, each index is weighted inversevely based on its volatility measure.

Judging Model Performance

Model performance is judged by out-of-sample prediction error, where the prediction is done on the second sale in a resale pair. The predicted value is created by applying the estimated index values to the first sale of the resale to generate a predicted value at the time of the resale. This predicted value is compared to the actual re-sale value to generate a prediction error. I do this error calculation two ways. First, I hold out 20% of all sale-resales and use the 80% training data to create indexes and make predictions on the 20% hold out sample. Second, I forecast indexes one period ahead, for each period i to n, where i starts at some training period distance (24 months). For instance, at month 60, I use data from months 1 to 59 to create an index, use a simple forecasting method[2] to create an estimated index value at month 60 and the predict the resale value for all homes reselling in time period 60. The holdout method has the benefit of not requiring a forward prediction, while the forecasting method has the benefit of being more informationally efficient both in index creation and in the number of prediction errors calculated.

Data

The data for this study come from Zillow.com. I have obtained sales and home characteristic data from 2006 to 2016 for the following eight counties:

- King, Pierce and Snohomish, WA
- Pinellas and Hillsborough, FL
- Chatham, GA
- San Diego, CA
- Denver, CO

The data have full information on transaction details, over fifteen property characteristics and lat/long coordinates.

[2] Using the AAA method in the R forecast package.

Initial Results

Disclaimer: *The analysis of submarket indexes and blending methods is one component of a larger project on house price indexes that I'm undertaking with data and some consultation from Zillow. The views and results expressed below are my own and are not indicative of the models being used at Zillow.*

Initial results have been derived for 10 different data/model/estimator combinations. For each geographic level (county, random-10 and ZIP) model blending has been done. For County to Random-10 and County-to-ZIP hierarchical blending has been calculated. Indexes are estimated at a monthly frequency. All data cleaning, analysis and plotting is done in R.I have included a few choice plots from the working set of results below. In the plot below the following abbreviations are used:

- AVR: Assessed/appraised value ratio model
- HED: Hedonic price model
- SRS: Sale-resale (repeat sale) model
- RME: Relative match estimator (pseudo repeat sales model)
- BASE: Basic estimator
- WGT: Weigthed estimator (Case Schiller 1987)
- ROB: Robust estimator
- HRB: Hierarchical blend
- BLD: Method Blend
- Rand.10: Submarket disaggregation done with a kmeans method creating 10 random submarket (spatially clustered)

An online dashboard https://andykrause.shinyapps.io/zpiAPP/ is also available customized visualization of the results.

Predictive Ability by Market Level by Blend Type

The plot below shows the predictive ability of each individual index at each level for all eight counties. The left panel shows indexes that have no hierarchical blending, while the right panel does. The far right boxplot in each chart shows the model blended results. The top panels are county level indexes (no hierarchical blending at the county level), the middle panels are a random 10 spatial unit breakdown of the county and the bottom panels are at the ZIP code level. The violin plots are boxplot but with the full density of results shown by the shape of the figure.

A number of key findings arise from this plot. For most methods (combinations of data, model and estimator) the smaller geographic units are more accurate (but also more variable in their accuracy) than the county level.

This trend, however, does not hold for the sale-resale (Blue) models without hierarchical blending (left panels). Next, overall, the hedonic models (Green) and the model blended (Black) offer the most accurate precitions in the base condition (left hand panels). The superiority of these two models holds with hierarchical blending (right panels) but their performance gain over the other models are greatly reduced. The hierarchial blending does improve the predictive ability of most data/model/estimator combinations, but the improvement is greatest in the repeat sale models (Blue). The repeat sale models with relative matcheds (or pseudo repeat sales) (Red) perform comparably to the other models and are relatively consistant regardless of estimator. Finally, the assessed value ratio models perform worse than the others.

The superiority of the hierarchical blended models is more evident in looking at results on a county by county level. The particular data/model/estimator this is most accurate in each county does vary by county and geographic level, but in general the hedonic and model blended approaches are most accurate and the repeat sales and assessed value the least. Robust estimators are generally better than weighted or base estimators. In terms of geographic levels, ZIP codes are usually the most accurate, but only when hierarchically blended. Non-blended ZIP code models are often worse than the non-blended Random-10 and County models. Rarely is the county the most accurate level at which to create indexes.

Below I will show and discuss the King County results, identical plots for the other seven counties are shown in the appendix.

King County Specific Error Results

Here, specific results from King County, WA are shown. Error from a holdout approach are shown on the left hand column, while those from a forecast are shown on the right. The general findings here mirror those of the eight county combined shown in Figure 1. Namely, the hedonic and model blended methods perform the best, with hierarchical blending ('x's and'∼'s) performing better than the basic models in nearly all instances. The repeat sales models (blue) at the county level ('o's) do quite well but their relative performance degrades at the zip code level, especially in the absence of hierarchical blending. Additionally, with the hedonic (green) and relative match (red) we see an interesting trend in that the hierarchical blending makes the random 10 models a bit worse, but makes the ZIP code models better. This suggests that at the random 10 level the hierarchical blending removes too much local variation making the model worse, but at the ZIP code level (a finer disaggregation than random 10) it removes enough volatility to make the models better. Future testing will determine if this particular relationships holds across space and time.

Finally, the last plot highlights the different indexes as estimated by the 10 models from 2006 to 2016. The largest deviations are in the assessed value ratio models (purple).

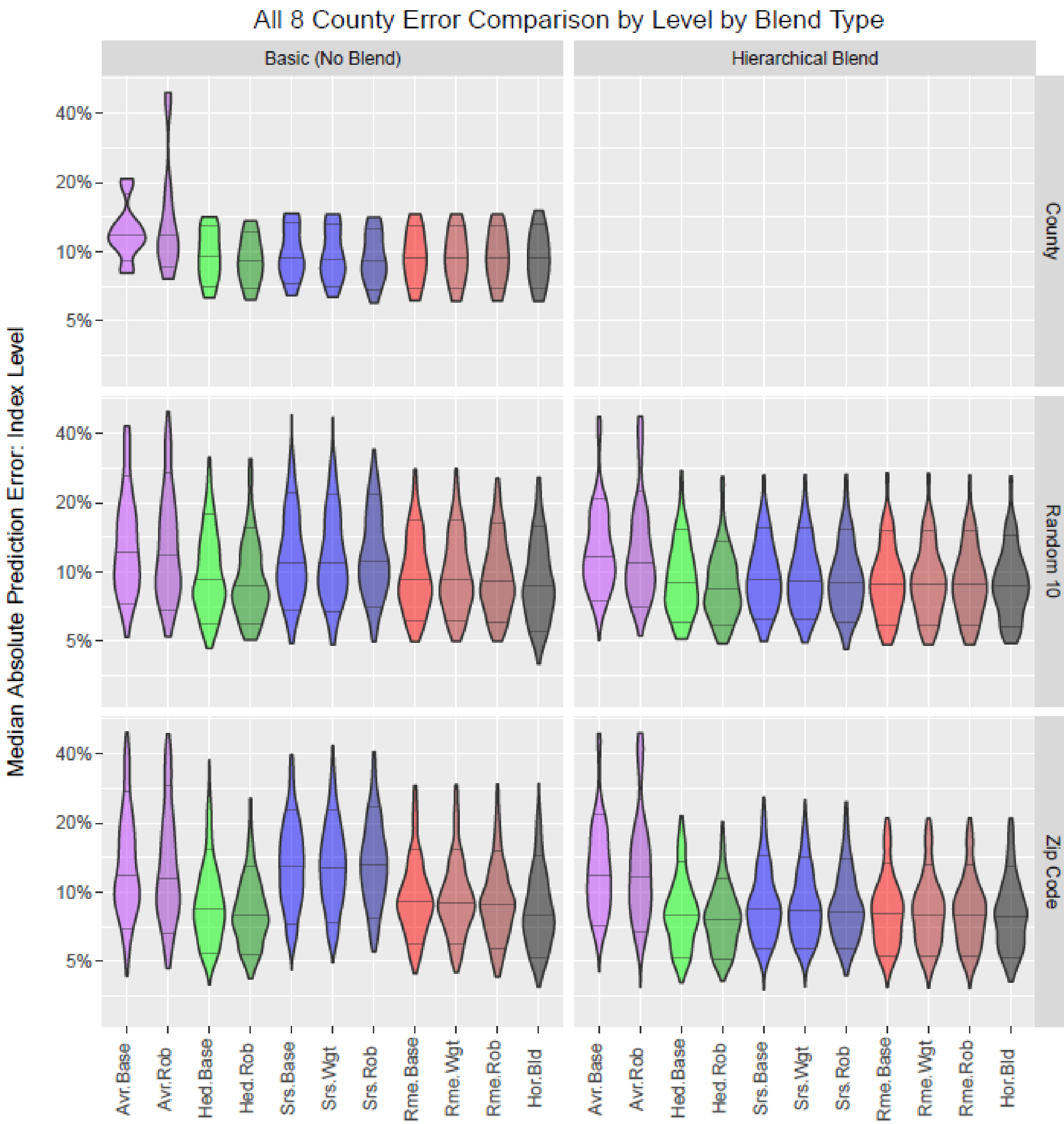

Figure 1: All 8 Counties - Errors (Index Level)

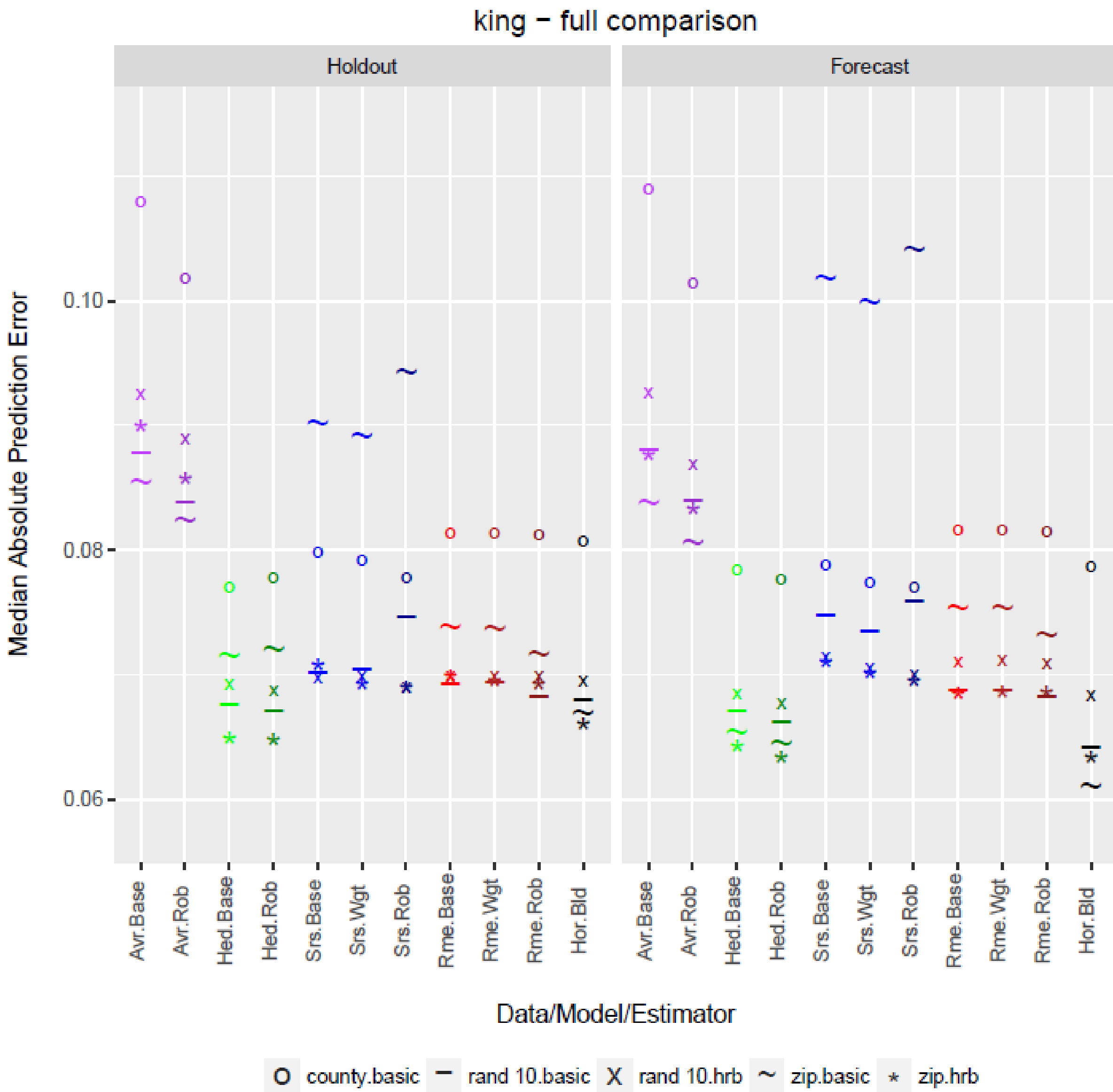

Figure 2: King County, WA - Error Analysis

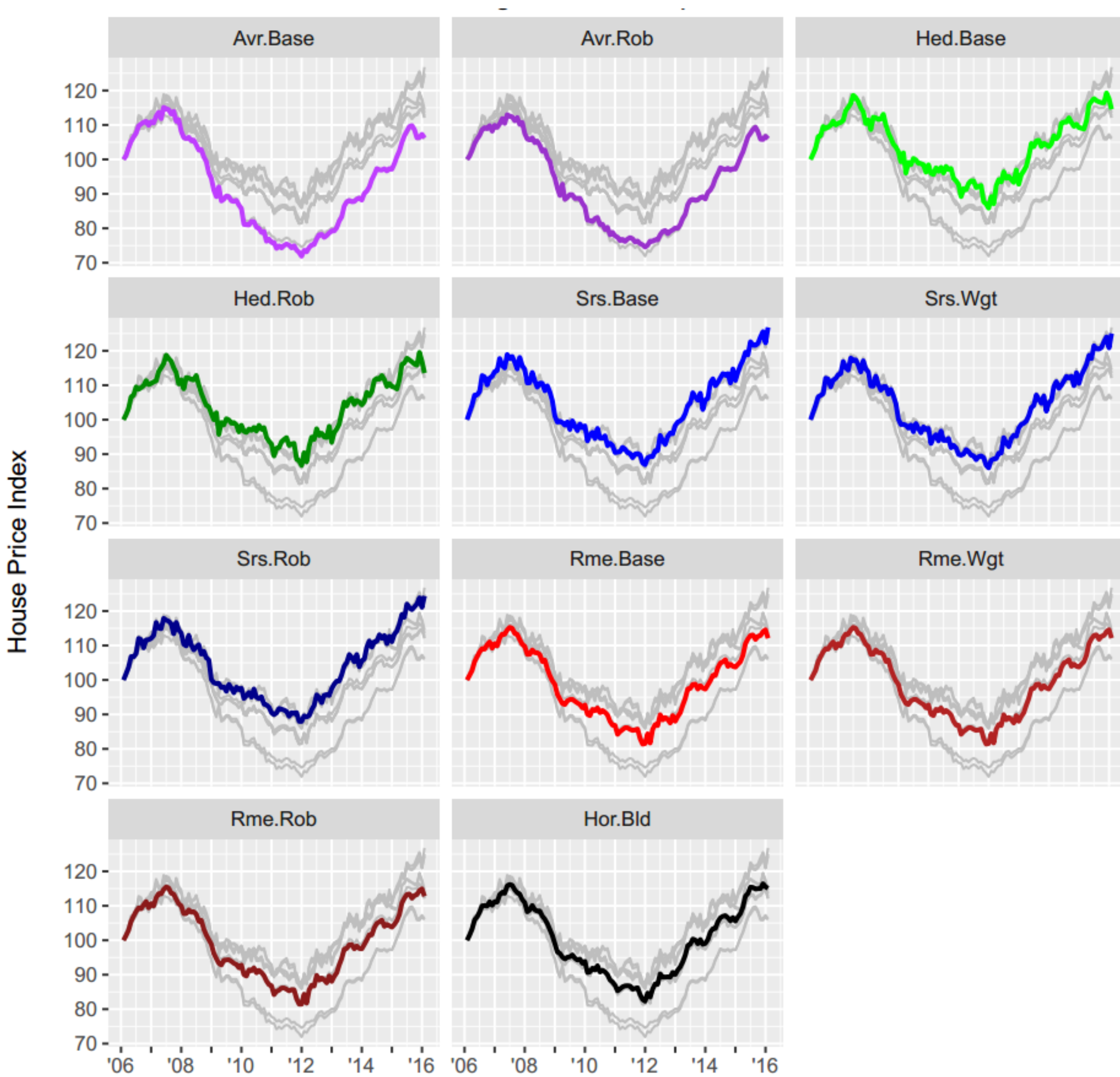

Figure 3: King County, WA - Index Differences

Conclusions

This paper set out to test if model blending both within a submarket and between market levels can improve house price index estimation. By testing a suite of models and estimators over eight U.S. counties the finding do indicate that both model and hierarchical blending do hold potential for house price indexes. Hierarchical blending shows the largest improvements on out-of-sample prediction when blending ZIP codes with county estimates. The often highly volatile sale-resale models (when used in small geographic areas) show the most gain from the blending process. Within submarket model blending tends to produce less volatile and equally or slighly more accurate house price indexes than single model approaches.

These finding suggest that single model/single geographic region approaches, long the mainstay in house price indexes analyses, could gain from blending, or ensemble, approaches that are becoming common in many data science applications.

Bibliography

Abraham, J. M., & Schauman, W. S. (1991). New evidence on home prices from Freddie Mac repeat sales. *Real Estate Economics*, 19(3), 333-352.

Bailey, M., Muth, R., & Nourse, H. (1963). A Regression Method for Real Estate Price Index Construction. *Journal of the American Statistical Association*, 58, 933-942.

Bogin, A. N., Doerner, W. M., Larson, W. D., & others. (2016). Local House Price Dynamics: New Indices and Stylized Facts *FHFA Working Paper*.

Bourassa, S., Cantoni, E., & Hoesli, M. (2013). Robust repeat sales indexes. *Real Estate Economics*, 41(3), 517-541.

Bourassa, S., Cantoni, E., & Hoesli, M. (2016). Robust hedonic price indexes. *International Journal of Housing Markets and Analysis*, 9(1), 47-65.

Bourassa, S. C., & Hoesli, M. (2016). High Frequency House Price Indexes with Scarce Data. *Swiss Finance Institute Research Paper*, (16-27).

Bourassa, S. C., Hoesli, M., & Sun, J. (2006). A simple alternative house price index method. *Journal of Housing Economics*, 15(1), 80-97.

Butler, J. S., Chang, Y., & Cutts, A. C. (2005). Revision bias in repeat-sales home price indices. *Freddie Mac Working Paper*

Can, A., & Megbolugbe, I. (1997). Spatial dependence and house price index construction. *The Journal of Real Estate Finance and Economics*, 14(1-2), 203-222.

Cannaday, R. E., Munneke, H. J., & Yang, T. T. (2005). A multivariate repeat-sales model for estimating house price indices. *Journal of Urban Economics*, 57(2), 320-342.

Case, B., Pollakowski, H. O., & Wachter, S. M. (1991). On choosing among house price index methodologies. *Real Estate Economics*, 19(3), 286-307.

Case, B., Pollakowski, H. O., & Wachter, S. (1997). Frequency of transaction and house price modeling. *The Journal of Real Estate Finance and Economics*, 14(1), 173-187.

Case, B., & Quigley, J. M. (1991). The dynamics of real estate prices. *The Review of Economics and Statistics*, 50-58.

Case, K., & Shiller, R. (1987). Prices of Single Family Homes Since 1970: New Indexes for Four Cities. *New England Economic Review*, Sept/Oct, 45-56.

Chau, K. W., Wong, S. K., & Yiu, C. Y. (2005). Adjusting for non-linear age effects in the repeat sales index. *The Journal of Real Estate Finance and Economics*, 31(2), 137-153.

Chinloy, P. T. (1977). Hedonic price and depreciation indexes for residential housing: A longitudinal approach. *Journal of Urban Economics*, 4(4), 469-482.

Clapham, E., Englund, P., Quigley, J. M., & Redfearn, C. L. (2006). Revisiting the past and settling the score: index revision for house price derivatives. *Real Estate Economics*, 34(2), 275-302.

Clapp, J. M. (2004). A semiparametric method for estimating local house price indices. *Real Estate Economics*, 32(1), 127-160.

Clapp, J., & Giaccotto, C. (1992). Estimating price indices for residential property: a comparison of repeat sales and assessed value methods. *Journal of the American Statistical Association*, 87(418), 300-306.

Clapp, J. M., & Giaccotto, C. (1999). Revisions in Repeat-Sales Price Indexes: Here Today, Gone Tomorrow? *Real Estate Economics*, 27(1), 79-104.

Clapp, J. M., Giaccotto, C., & Tirtiroglu, D. (1991). Housing price indices based on all transactions compared to repeat subsamples. *Real Estate Economics*, 19(3), 270-285.

Clapp, J. M., Giaccotto, C., & Tirtiroglu, D. (1992). Repeat sales methodology for price trend estimation: an evaluation of sample selectivity. *Journal of Real Estate Finance and Economics*, 5(4), 357-374.

Coulson, N. E., & McMillen, D. P. (2007). The dynamics of intraurban quantile house price indexes. *Urban Studies*, 44(8), 1517-1537.

Crone, T. M., & Voith, R. (1992). Estimating house price appreciation: a comparison of methods. *Journal of Housing Economics*, 2(4), 324-338.

de Haan, J., & Diewert, W. E. (2013). Handbook on Residential Property Price Indices. Eurostat, Belgium.

De Vries, P., de Haan, J., van der Wal, E., & Mariën, G. (2009). A house price index based on the SPAR method. *Journal of Housing Economics*, 18(3), 214-223.

Deng, Y., & Quigley, J. M. (2008). Index revision, house price risk, and the market for house price derivatives. *The Journal of Real Estate Finance and Economics*, 37(3), 191-209.

Dorsey, R. E., Hu, H., Mayer, W. J., & Wang, H. (2010). Hedonic versus repeat-sales housing price indexes for measuring the recent boom-bust cycle. *Journal of Housing Economics*, 19(2), 75-93.

Dreiman, M. H., & Pennington-Cross, A. (2004). Alternative methods of increasing the precision of weighted repeat sales house prices indices. *The Journal of Real Estate Finance and Economics*, 28(4), 299-317.

Englund, P., Quigley, J. M., & Redfearn, C. L. (1998). Improved price indexes for real estate: measuring the course of Swedish housing prices. *Journal of Urban Economics*, 44(2), 171-196.

Englund, P., Quigley, J. M., & Redfearn, C. L. (1999). The choice of methodology for computing housing price indexes: comparisons of temporal aggregation and sample definition. *The Journal of Real Estate Finance and Economics*, 19(2), 91-112.

Francke, M. K. (2010). Repeat sales index for thin markets. *The Journal of Real Estate Finance and Economics*, 41(1), 24-52.

Francke, M. K., & Vos, G. A. (2004). The hierarchical trend model for property valuation and local price indices. *The Journal of Real Estate Finance and Economics*, 28(2), 179-208.

Gatzlaff, D. H., & Haurin, D. R. (1997). Sample Selection Bias and Repeat-Sales Index Estimates. *The Journal of Real Estate Finance and Economics*, 14, 33-50.

Gatzlaff, D. H., & Haurin, D. R. (1998). Sample Selection and Biases in Local House Value Indices. *Journal of Urban Economics*, 43, 199-222.

Gatzlaff, D. H., & Ling, D. C. (1994). Measuring Changes in Local House Prices: An Empirical Investigation of Alternative Methodologies. *Journal of Urban Economics*, 35, 221-244.

Gelfand, A. E., Ecker, M. D., Knight, J. R., & Sirmans, C. F. (2004). The Dynamics of Location in Home Prices. *Journal of Real Estate Finance and Economics*, 29(2), 149-166.

Goetzmann, W., & Spiegel, M. (1997). A Spatial Model of Housing Returns and Neighborhood Substitutability. *Journal of Real Estate Finance and Economics*, 14(1), 11-31.

Goetzmann, W., & Peng, L. (2006). Estimating house price indexes in the presence of seller reservation prices. *Review of Economics and Statistics*, 88(1), 100-112.

Goodman, A. C. (1978). Hedonic prices, price indices and housing markets. *Journal of Urban Economics*, 5, 471-484.

Goodman, A. C., & Thibodeau, T. (1997). Dwelling-age-related heteroskedasticity in hedonic house price equations: an extension. *Journal of Housing Research*, 8, 299-317.

Guntermann, K. L., Liu, C., & Nowak, A. D. (2016). Price Indexes for Short Horizons, Thin Markets or Smaller Cities. *Journal of Real Estate Research*, 38(1), 93-127.

Guo, X., Zheng, S., Geltner, D., & Liu, H. (2014). A new approach for constructing home price indices: The pseudo repeat sales model and its application in China. Journal of Housing Economics, 25, 20-38.

Guttery, R. S., & Sirmans, C. F. (1998). Aggregation Bias in Price Indices for Multi-Family Rental Properties. *Journal of Real Estate Research*, 15(3), 309-325.

Haurin, D. R., & Hendershott, P. H. (1991). House price indexes: issues and results. *Real Estate Economics*, 19(3), 259-269.

Haurin, D. R., Hendershott, P. H., & Kim, D. (1991). Local house price indexes: 1982–1991. *Real Estate Economics*, 19(3), 451-472.

Hill, R. C., Knight, J. R., & Sirmans, C. F. (1997). Estimating capital asset price indexes. *Review of Economics and Statistics*, 79(2), 226-233.

Jansen, S. J. T., de Vries, P., Coolen, H., Lamain, C. J. M., & Boelhouwer, P. J. (2008). Developing a house price index for the Netherlands: A practical application of weighted repeat sales. *The Journal of Real Estate Finance and Economics*, 37(2), 163-186.

Knight, J. R., Dombrow, J., & Sirmans, C. F. (1995). A varying parameters approach to constructing house price indexes. *Real Estate Economics*, 23(2), 187-205.

Leishman, C., & Watkins, C. (2002). Estimating local repeat sales house price indices for British cities. *Journal of Property Investment & Finance*, 20(1), 36-58.

McMillen, D. P. (2003). Neighborhood house price indexes in Chicago: a Fourier repeat sales approach. *Journal of Economic Geography*, 3(1), 57-73.

McMillen, D. (2012). Repeat sales as a matching estimator. *Real Estate Economics*, 40(4), 745-773.

Mcmillen, D. P., & Dombrow, J. (2001). A Flexible Fourier Approach to Repeat Sales Price Indexes. *Real Estate Economics*, 29(2), 207-225.

Mcmillen, D. P., & Thorsnes, P. (2006). Housing Renovations and the Quantile Repeat-Sales Price Index. *Real Estate Economic*s, 34(4), 567-584.

Mark, J. H., & Goldberg, M. (1984). Alternative housing price indices: an evaluation. *Real Estate Economics*, 12(1), 30-49.

Meese, R. A., & Wallace, N. (1991). Nonparametric estimation of dynamic hedonic price models and the construction of residential housing price indexes. *Journal of the American Real Estate & Urban Economics Association*, 19(3), 308-332.

Meese, R. A., & Wallace, N. (1997). The construction of residential housing price indices: a comparison of repeat-sales, hedonic-regression, and hybrid approaches. *The Journal of Real Estate Finance and Economics*, 14(1), 51-73.

Munneke, H. J., & Slade, B. A. (2000). An empirical study of sample-selection bias in indices of commercial real estate. *The Journal of Real Estate Finance and Economics*, 21(1), 45-64.

Munneke, H. J., & Slade, B. A. (2001). Metropolitan Transaction-Based Commercial Price Index : A Time-Varying Parameter Approach. *Real Estate Economics*, 29(1), 55-84.

Nagaraja, C., Brown, L., & Wachter, S. (2014). Repeat sales house price index methodology. *Journal of Real Estate Literature*, 22(1), 23-46.

Nappi-Choulet, I., & Maury, T.-P. (2009). A spatiotemporal autoregressive price index for the Paris office property market. *Real Estate Economics*, 37(2), 305-340.

Pollakowski, H. O., Stegman, M. A., & Rohe, W. (1991). Rates of return on housing of low-and moderate-income owners. *Journal of the American Real Estate & Urban Economics Association*, 19(3), 417-426.

Prasad, N., & Richards, A. (2009). Improving median housing price indexes through stratification. *Journal of Real Estate Research*, 30, 45-71.

Quigley, J. M. (1995). A simple hybrid model for estimating real estate price indexes. *Journal of Housing Economics*, 4(1), 1-12.

Schwann, G. M. (1998). A real estate price index for thin markets. *The Journal of Real Estate Finance and Economics*, 16(3), 269-287.

Steele, M., & Goy, R. (1997). Short holds, the distributions of first and second sales, and bias in the repeat-sales price index. *The Journal of Real Estate Finance and Economics*, 14(1), 133-154.

Tu, Y., Shi-Ming, Y., & Sun, H. (2004). Transaction-based office price indexes: A spatiotemporal modeling approach. *Real Estate Economics*, 32(2), 297.

Zabel, J. E. (1999). Controlling for quality in house price indices. *The Journal of Real Estate Finance and Economics*, 19(3), 223-241.

Appendix

Pierce County

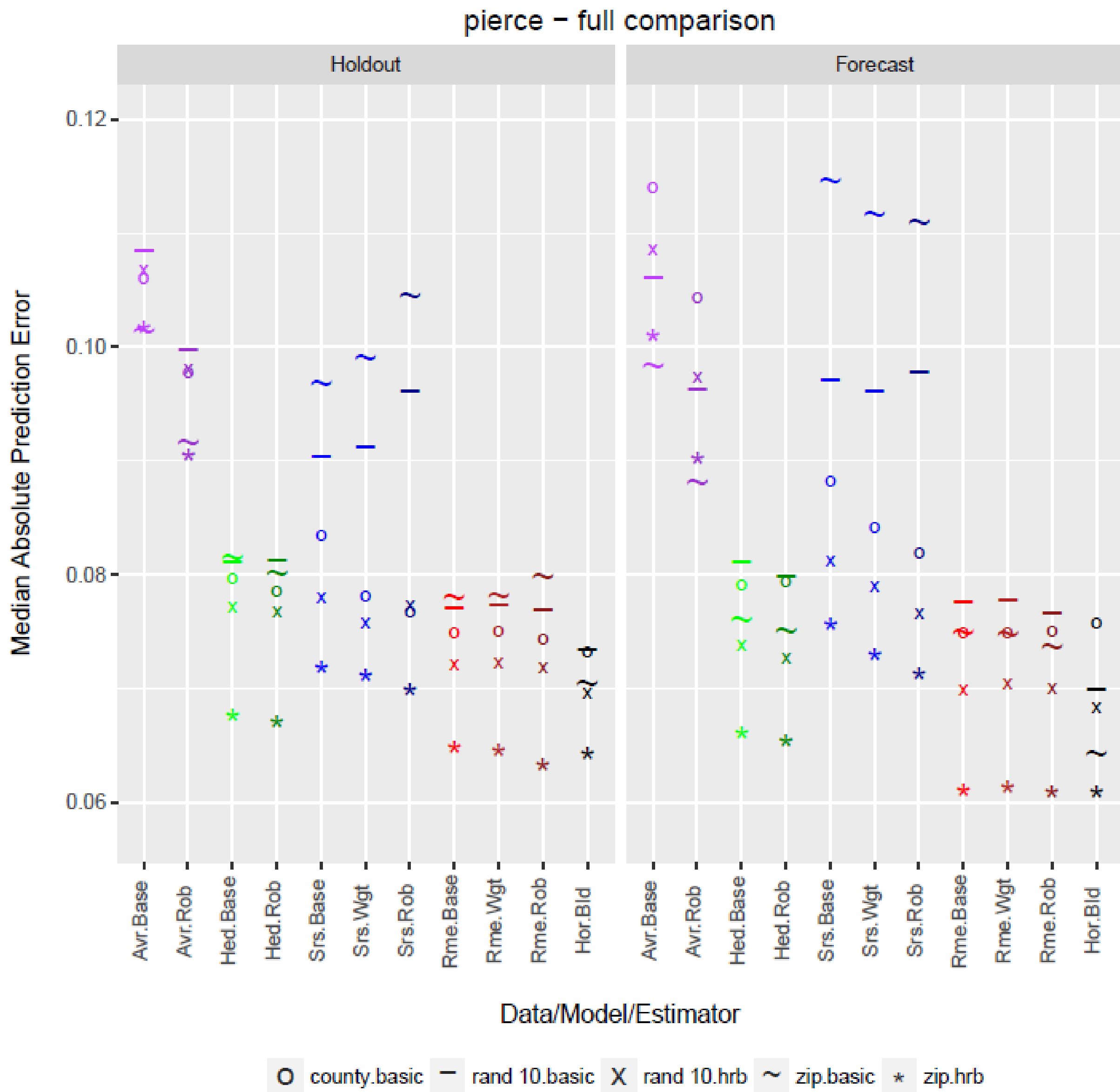

Figure 4: Pierce County, WA - Error Analysis

Pinellas County

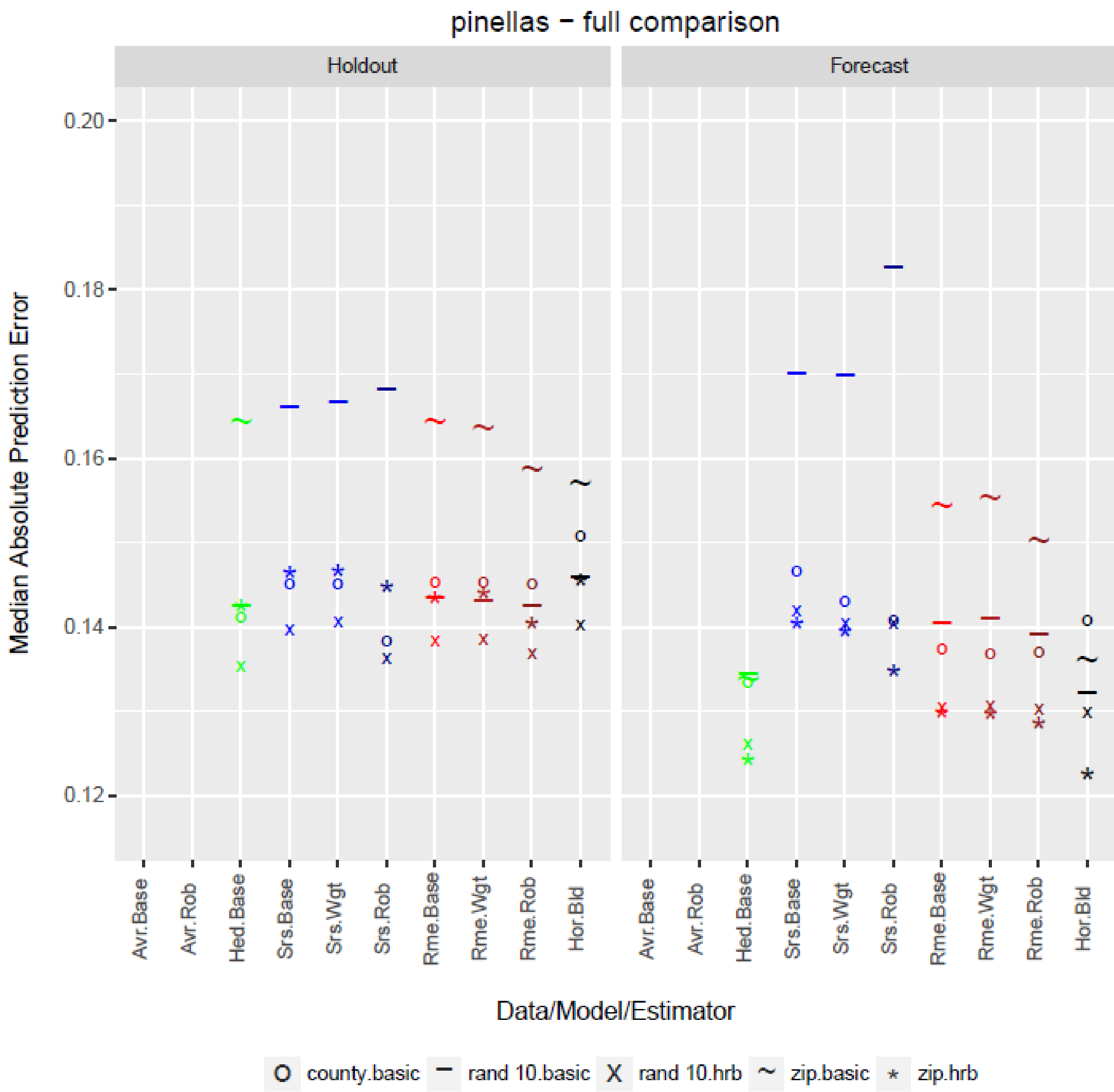

Figure 5: Pinellas County, FL - Error Analysis

Hillsborough County

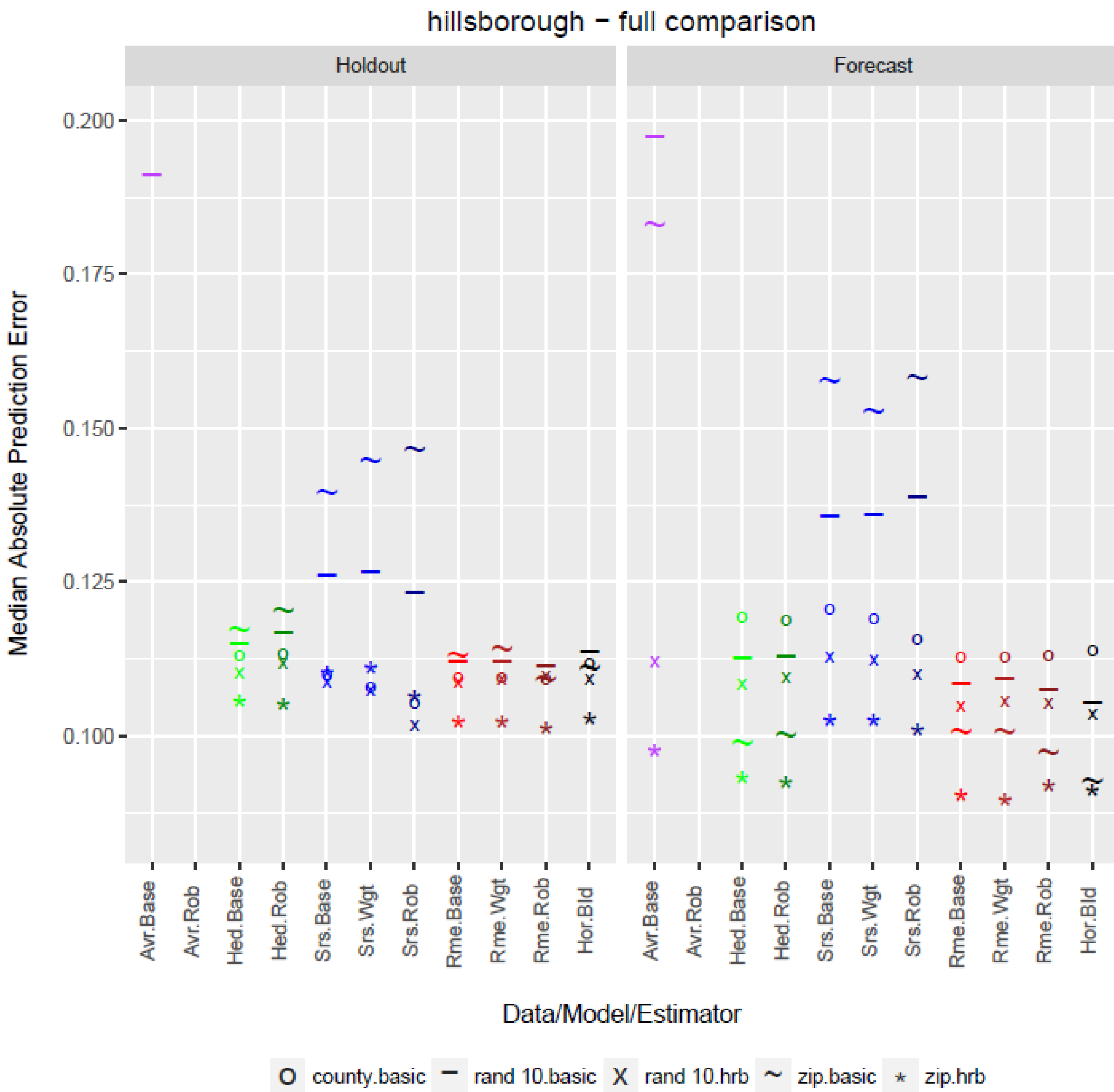

Figure 6: Hillsborough County, FL - Error Analysis

Chatham County

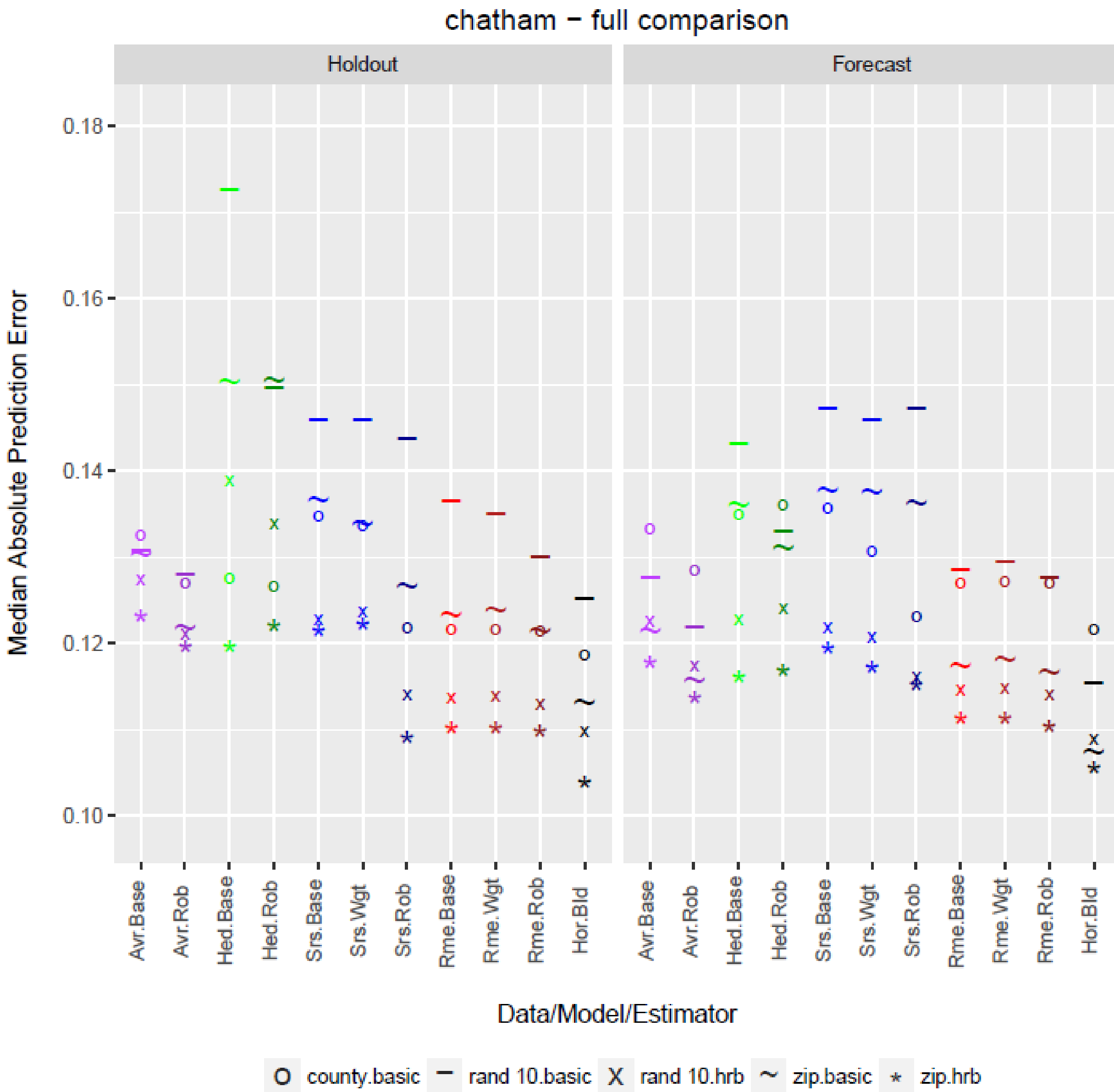

Figure 7: Chatham County, GA - Error Analysis

Denver County

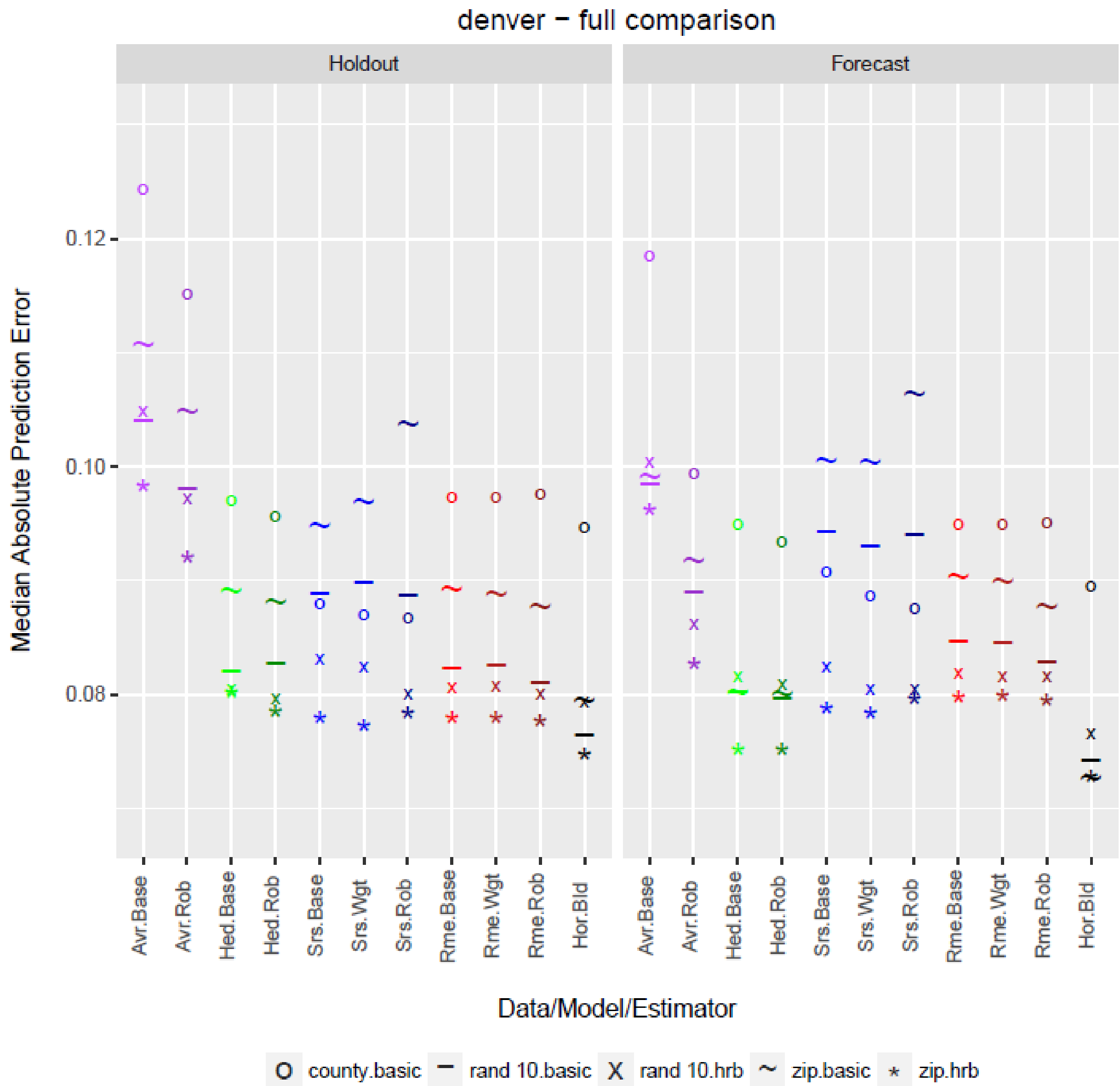

Figure 8: Denver County, CO - Error Analysis

San Diego County

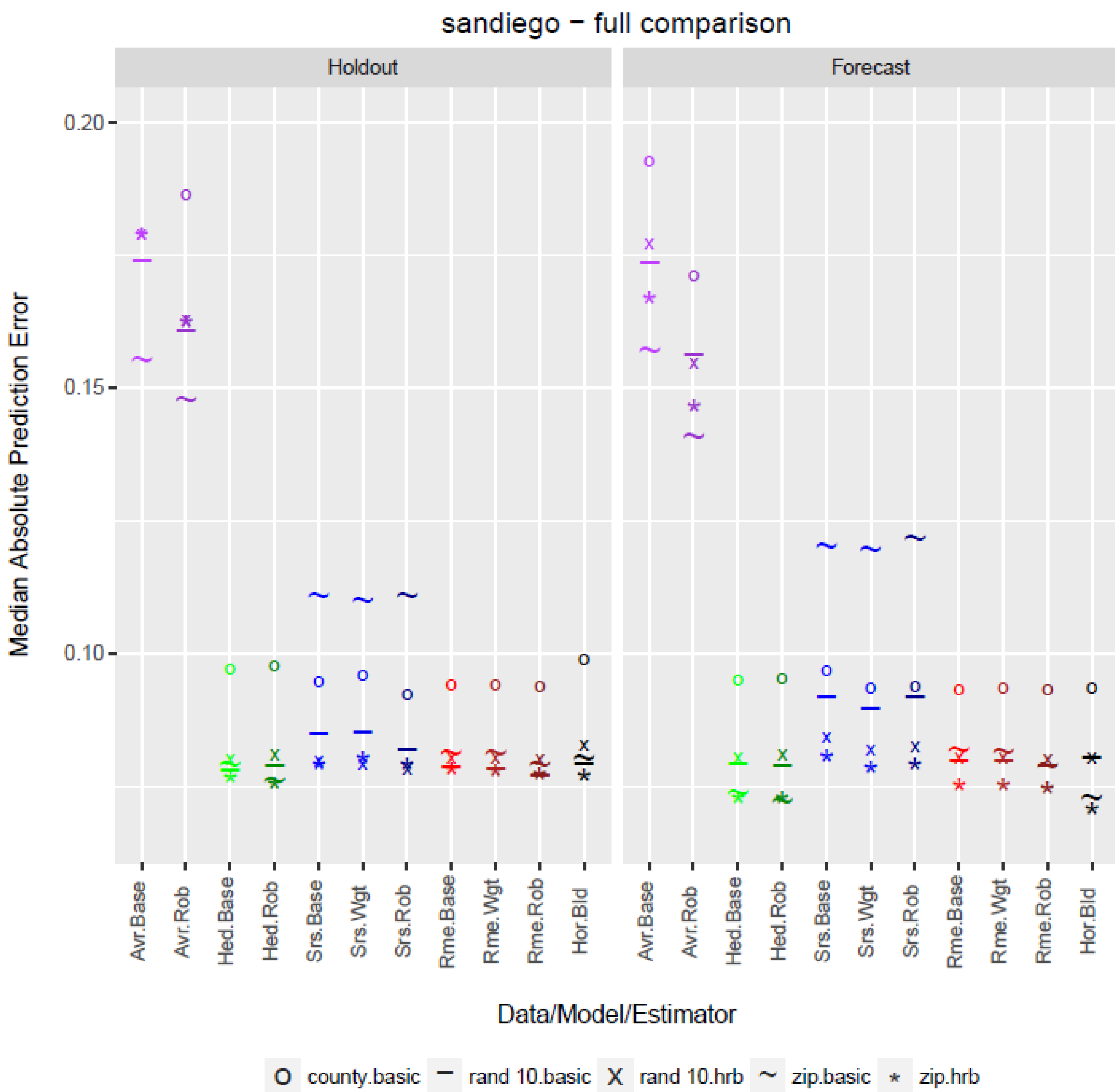

Figure 9: San Diego County, CA - Error Analysis