

Modeling Home Price Appreciation in Collaboration with US Bank

Ismail Abouamal
California Institute of Technology
abouamal@caltech.edu

Daniela Beckelhymer
University of Minnesota
beck1538@umn.edu

Jarrad Botchway
Missouri University of Science and Technology
jb24r@umsystem.edu

Jordan Pellett
University of Tennessee, Knoxville
jpellett@vols.utk.edu

Marshall Smith
University of Minnesota
smit5939@umn.edu

Leo Digiosia
Quantitative Model Analyst, US Bank
leo.digiosia@usbank.com

Chris Jones
Mortgage Model Manager, US Bank
christopher.jones6@usbank.com

July 2023

Contents

1 Executive Summary	3
2 Introduction	3
2.1 Data Used	3
2.2 Model Structure	4
2.2.1 Hyperbolic Sine Model	4
2.2.2 Piecewise Linear Model	6
3 Project Investigations	6
3.1 Adapting the Context	7
3.1.1 Modifying the Appreciation Window	7
3.1.2 Region and State-Level Analysis	8
3.1.3 Limiting the Training Window	13
3.2 Model Comparison	16
3.2.1 National Models Comparison	16
3.2.2 State models Comparison	19
3.3 Improving the Current Model	23
3.3.1 Modifying Demand: Grid Searching for Velocity Variables in the Linear Model	24
3.3.2 Modifying Demand: Incorporation of an Intensity Variable in the Hyperbolic Sine Model	25
4 Downstream Applications	26
5 Future Work	28
5.1 Further Directions for Regional and State Models	28
5.2 Further Modification of the National Model	28
5.3 Investigating Regional and State-Wide Discrepancies Contributing to National Model Error	28
A Appendix: Miscellaneous Plots	33
B Appendix: Regional Plots	34
C Appendix: State Model Plots	38

1 Executive Summary

US Bank's Home Price Appreciation model examines the relationship between price dynamics in the housing market relative to supply and demand forces, motivated by classical economic theory. This paper investigates the applicability and universality of this existing price model by examining its generalizations. Additionally, this paper establishes a better understanding of the model, its limitations, and potential avenues for improvement.

Reliable and sound models for home prices are critical inputs to downstream mortgage prepayment models. Prepayment models inform investment portfolio decision-making and interest rate risk hedging action, highlighting the importance of this model for US Bank's Corporate Treasury.

2 Introduction

Mortgages are fundamental to financial institutions as they represent valuable long-term assets in their investment portfolios. Mortgage loans offer diversification opportunities, allowing banks to balance risk across various lending instruments. Mortgages act as secured loans backed by collateral, reducing the potential risk for banks and ensuring they can recover the loan amount in case of borrower default. Collectively, these reasons underline the crucial role mortgages play in supporting banks' financial growth and stability.

When home prices rise, homeowners may see an opportunity to tap into their increased equity by refinancing their mortgages or selling their homes altogether. Refinancing allows borrowers to obtain larger loans, taking advantage of the improved value of their properties. Similarly, home sales triggered by price appreciation result in the early repayment of existing mortgages. These prepayments can be problematic for banks because they cut short the expected duration of the loans, reducing the overall interest income. Moreover, prepayments may force them to reinvest the returned principal at potentially lower interest rates, further affecting profitability. As a result, banks must continuously monitor market conditions and adjust their lending strategies to mitigate the impact of home price appreciation and manage their interest income effectively.

Conventional models for home prices often incorporate many correlated macroeconomic variables, leading to challenges in interpretation. In order to get a better understanding of the dynamics behind the fluctuations of home prices over time, our team from the Institute for Mathematics and its Applications (IMA) has expanded, simplified, and refined the current model developed by US Bank. This model employs a methodology for modeling home price appreciation from a differential point of view. Within academic economics research, the use of differential equations is often attributed to Samuelson [1947][3] who used simple differential equations to investigate the stability of equilibrium for supply-demand scenarios and built on the work of Walras and Marshall. Under the Walras assumption, price increases if excess demand is positive and price decreases when it is negative.

Regarding equilibrium price, efficient market and efficient price discovery theories have been active research areas from a theoretical point of view, see Gjerstad [2007][2]. Irrational price theory has been studied from an analytical view with early work given by Smith et al. [1988][4], and Caginalp et al. [2001][1]. However, the study of dynamics that describes the path to equilibrium is still in a developing stage.

We now discuss key definitions, our model investigations, results, and potential future directions.

2.1 Data Used

Various sources of data were used in modeling home price dynamics. At the national level, the Federal Housing Finance Agency's (FHFA) monthly seasonally adjusted purchase-only **Home Price Index** (HPI), directly sourced from FHFA, was employed. For state-level data, CoreLogic's seasonally adjusted monthly state indices were utilized. Each of these indices uses a repeat-sales methodology, which has the benefit of controlling for differences between properties. From these home price indices, we derived **Home Price Appreciation** (HPA) quantities, which represent the annual, quarterly, or monthly percentage change in

the reference index. April 2020 state population data reported by the US Census Bureau was also utilized in conjunction with this data.

Additionally, this model incorporates data on **Existing Home Sales** (EHS), which is a monthly measure of the seasonally adjusted annual rate of home sales for single-family homes, co-ops, and condos in the US. This data is collected and processed by the National Association of Realtors (NAR). The NAR also provides a regional breakdown of the national EHS variable into four geographic regions, which are used as inputs in generalizations of the main model. Sales are recorded after a seller accepts a sales contract on a previously owned property, hence the classification “existing”.

Furthermore, we utilize data on **Existing Home Inventory** (EHI), which represents the total number of existing single-family homes, co-ops, and condos listed or pending sale. This data, also provided by the NAR, is not adjusted for seasonality, but is seasonally adjusted in Python before use in the model. Unlike EHS, there is no regional breakdown available for EHI data provided by the NAR.

For a comprehensive view of the housing market, we consider data on the overall housing **stock**, encompassing occupied, vacant, seasonally vacant, and for-sale properties in the US. This information is provided quarterly by the US Census Bureau, and the model developers extend it to a monthly frequency through linear interpolation between quarters.

Additionally, we incorporate data on **New Home Sales** (NHS), which is reported on a monthly basis by the US Census Bureau as a seasonally adjusted annualized rate. Finally, the **New Home Inventory** (NHI) data, also provided monthly by the US Census Bureau, represents the seasonally adjusted quantity of new properties in various stages of completion (completed, under construction, or not started). All these data sources together, more concisely organized in Table 1, contribute to the comprehensive analysis and modeling of the housing market.

Table 1: Key Definitions

Data Abbreviation	Description
HPI	<i>Home Price Index</i>
HPA	<i>Home Price Appreciation</i>
EHS	<i>Existing Home Sales</i>
EHI	<i>Existing Home Inventory</i>
Stock	<i>Total number of properties in the US</i>
NHS	<i>New Home Sales</i>
NHI	<i>New Home Inventory</i>

Figure 1 shows the FHFA HPI over time along with the corresponding derived annual HPA as a year-over-year percent change:

$$\text{HPA}_t = 100 \cdot (\text{HPI}_t - \text{HPI}_{t-12}) / \text{HPI}_{t-12} \quad (1)$$

2.2 Model Structure

US Bank currently has two models for HPA. The first makes use of the structure of hyperbolic sine, and the second is a linear approximation of the first. In this section, we describe the structure of each model, including its parameters and key inputs.

2.2.1 Hyperbolic Sine Model

The hyperbolic sine model takes in a single input which developers refer to as *demand*. Demand measures purchase activity in the form of EHS relative to inventory, EHI. The reported units of the variables EHS and EHI are technically different: The former is a seasonally adjusted annualized rate, and the latter is a

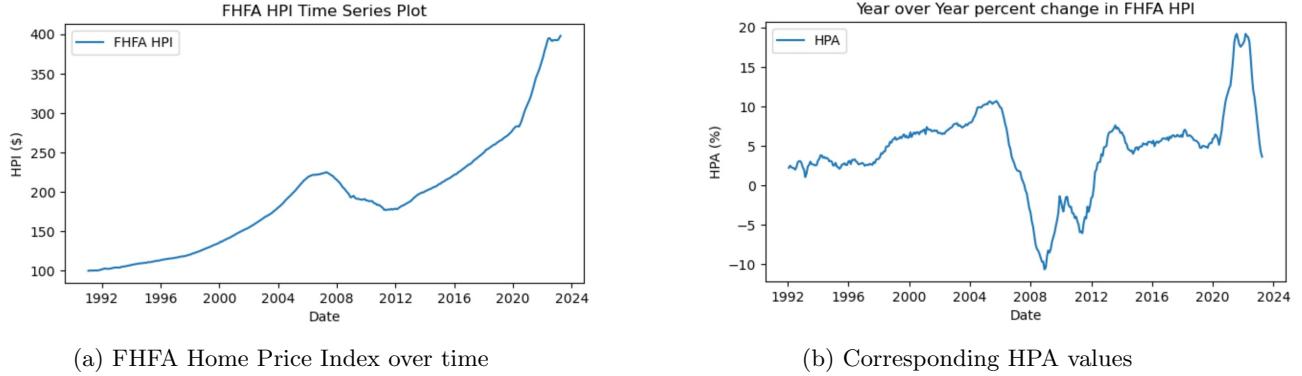


Figure 1

measure of absolute levels. More specifically, we define the value of demand at time t as

$$\text{Demand}_t = \frac{(\text{EHI}_t - b \cdot \text{EHS}_t)}{\text{Stock}_t} \quad (2)$$

Furthermore, we utilize **weighted moving averages** as a statistical method to control for noise in the demand variable. The formula for calculating the n -month weighted moving average for the variable $D = \text{Demand}$ is

$$\text{WMA}(D)_t = \frac{w_n \cdot D_t + w_{n-1} \cdot D_{t-1} + \dots + w_1 \cdot D_{t-n+1}}{w_n + w_{n-1} + \dots + w_1} \quad (3)$$

Here, w_i represents the weight assigned to a given lookback period. In our case, we use weights of $w_i = i$, with n initially set to 12. For notational convenience, Demand_t denotes $\text{WMA}(D)_t$ for a choice of n in this paper.

Taking Demand_t as a single input, HPA at time t can be modeled as

$$\text{HPA}_t \approx \alpha + \beta \sinh(\gamma(\text{Demand}_t - \delta)) \quad (4)$$

Figure 2 shows the output of the hyperbolic sine model with Figure 2a showing a time series of HPA and Figure 2b showing year-over-year HPA vs demand. Note that in Figure 2b, the black curve depicts the approximation given by Equation 4 while the points represent the observed $(\text{Demand}_t, \text{HPA}_t)$ pairs, color coded by date. Figure 2b illustrates one of the main reasons for using the hyperbolic sine model: its ability to effectively capture the extreme response of HPA to both high and low values of demand.

Table 2: Parameters

Parameter	Definition
b	A unit adjustment constant
$\alpha, \beta, \gamma, \delta$	The coefficients for hyperbolic sine model

Interpretations for the model parameters α, β, γ , and δ of Equation 4 are described in detail below.

α - represents the baseline level of home price appreciation at typical demand δ .

β - controls the magnitude of the price growth as demand moves away from the inflection point. Higher values of β correspond with larger increases in price for higher demand, and larger decreases for lower demand.

γ - controls the steepness of the price response to changes in demand. Higher γ means prices change more rapidly as demand moves away from the inflection point.

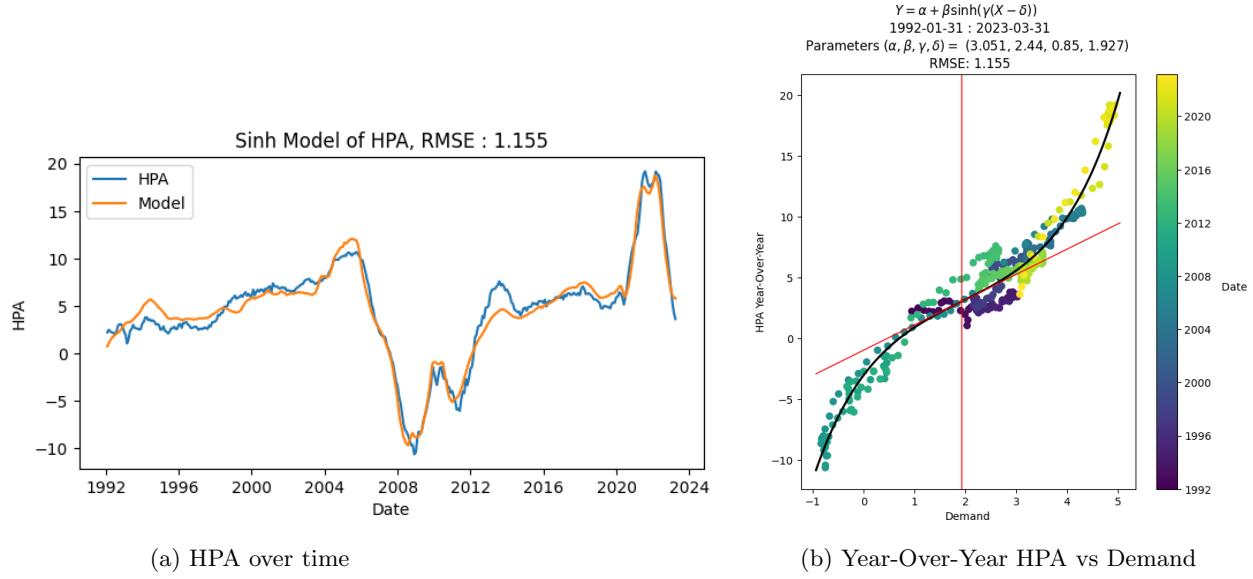


Figure 2: Hyperbolic Sine Model Output

δ - corresponds to the value of demand at the inflection point on the curve. At this point, prices grow at the baseline rate α . A demand that is higher or lower than δ causes prices to grow faster or slower, respectively.

(δ, α) - is the inflection point for the hyperbolic sine curve which can be seen in Figure 2b.

$\beta \cdot \gamma$ - is the slope of the tangent line that corresponds to the inflection point on the curve at (δ, α) . The tangent line can be seen in red in Figure 2b.

See Figure 32, in Appendix A for additional plots that aid in understanding and interpreting the model parameters.

2.2.2 Piecewise Linear Model

A linear regression model can be used to approximate the hyperbolic sine model. This formulation is based on the property captured in the hyperbolic sine model that HPA exhibits distinct responses to high, medium, and low levels of demand. This linear model produces four coefficients b , b_L , b_M , and b_H so that

$$\text{HPA}_t \approx b + b_L \cdot \text{Demand}_t^L + b_M \cdot \text{Demand}_t^M + b_H \cdot \text{Demand}_t^H \quad (5)$$

where HPA_t represents the Home Price Appreciation at time t , and Demand_t^L , Demand_t^M , and Demand_t^H represent the measures of demand at time t for low, medium, and high levels, respectively. The specific definitions for what constitutes low, medium, and high demand are provided in Section 3.2.

3 Project Investigations

Developers at the IMA focused on improving the model's accuracy in capturing housing market trends. We explored various levels of analysis, such as national, regional, and state-level HPA. Our efforts included limiting the training window to test model reliability, comparing the hyperbolic sine model to a linear model, and modifying demand to incorporate an intensity variable.

Our commitment to continuous improvement drives us to enhance the model's predictive capabilities and provide valuable insights for the housing market domain. Details of each investigation are presented in the following sections.

3.1 Adapting the Context

Developers considered ways in which the main model for national demand generalizes to other contexts. This included investigating how the model responds to modifications to the appreciation window, as well as its applicability to regional and state-level analyses.

3.1.1 Modifying the Appreciation Window

One of the project goals was to determine if the model adapts to different appreciation windows, and if so, to better understand how changing the appreciation window impacts the model parameters. Note, HPA is computed as

$$\text{HPA}_t = 100 \cdot (\text{Price}_t - \text{Price}_{t-\Delta}) / \text{Price}_{t-\Delta} \quad (6)$$

In Equation 6, taking $\Delta = 12$ months corresponds to annual appreciation, taking $\Delta = 3$ months corresponds to quarterly appreciation, and taking $\Delta = 1$ month corresponds to monthly appreciation. Further, Price is the FHFA House Price Index. Figure 3 and Figure 4 depict the results obtained for applying the hyperbolic sine model with $\Delta = 3$ and $\Delta = 1$, respectively. In each case, Demand was computed as a weighted moving average over Δ months with linear weights. Note Figure 3 and Figure 4 were generated using a b value of 1.35.

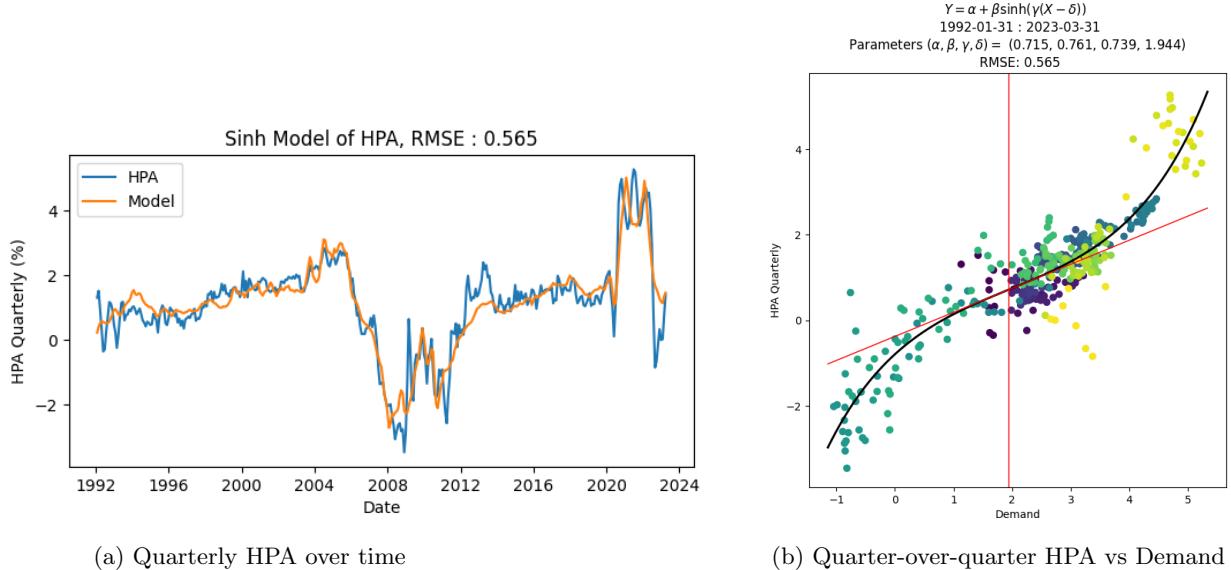


Figure 3: Quarterly Appreciation

To investigate the relationship between the parameters and the appreciation window, we ran the hyperbolic sine model for varying appreciation windows and recorded the model parameters α, β, γ , and δ . Figure 5 depicts each parameter vs. the appreciation window (in months). In general, we expect to see a linear response in the parameter value to time step and we have verified that expectation clearly for α and β . Although γ and δ exhibit some nonlinearity for longer appreciation windows, we note that δ varies over a small range of values and the slope $\beta\gamma$ still exhibits a linear relationship in response to time step. This demonstrates the generalizability of the existing US bank hyperbolic sine model.

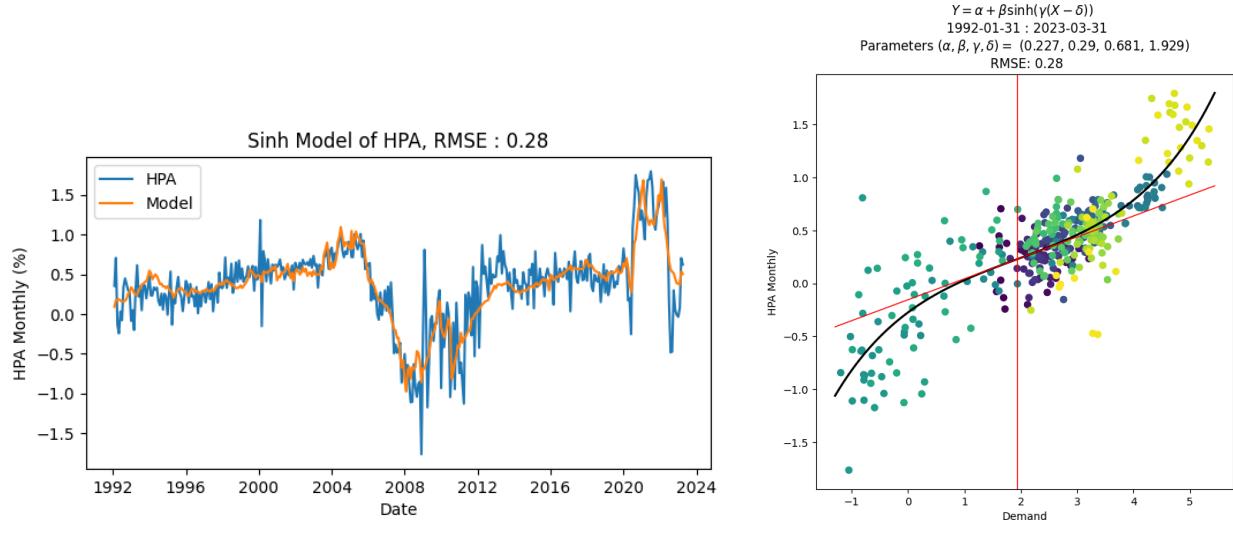


Figure 4: Monthly Appreciation

Along with looking at how α, β, γ , and δ change with appreciation window size, we also considered how b varies with appreciation window size. Figure 6 shows the optimal b value vs the appreciation window size in months. While the b value does depend on the appreciation window, we note that the range of b values is not very large.

3.1.2 Region and State-Level Analysis

Next, developers considered how the hyperbolic sine model adapts to regional and state-level data.

Regional Analysis

One challenge we incurred when conducting the regional analysis was that we did not have access to regional EHI or regional stock values. To address this challenge, we made the following assumption:

$$\frac{\text{EHI}_{\text{Regional}}}{\text{EHS}_{\text{Regional}}} = \lambda \cdot \frac{\text{EHI}_{\text{US}}}{\text{EHS}_{\text{US}}} \quad (7)$$

Here, $\frac{\text{EHI}}{\text{EHS}}$ represents months-of-supply, and we assume that regional months-of-supply trends with the national months-of-supply. This gives a regional EHI of

$$\text{EHI}_{\text{Regional}} = \lambda \cdot \frac{\text{EHI}_{\text{US}}}{\text{EHS}_{\text{US}}} \cdot \text{EHS}_{\text{Regional}} \quad (8)$$

Thus, regional demand is defined as

$$\text{Demand} = (\underbrace{\text{EHS}_{\text{Regional}} - b \cdot \text{EHI}_{\text{US}} / \text{EHS}_{\text{US}} \cdot \text{EHS}_{\text{Regional}}}_{\text{EHI}_{\text{Regional}}}) / \text{Stock}_{\text{US}} \quad (9)$$

Note that the value of the constant b was optimized for each region, and we therefore allow b to absorb the constant λ present in Equation 8.

Using April 2020 state population data obtained from the US Census Bureau, we computed the proportion of the regional population attributed to each state within a given US Census Bureau-defined region. Note,

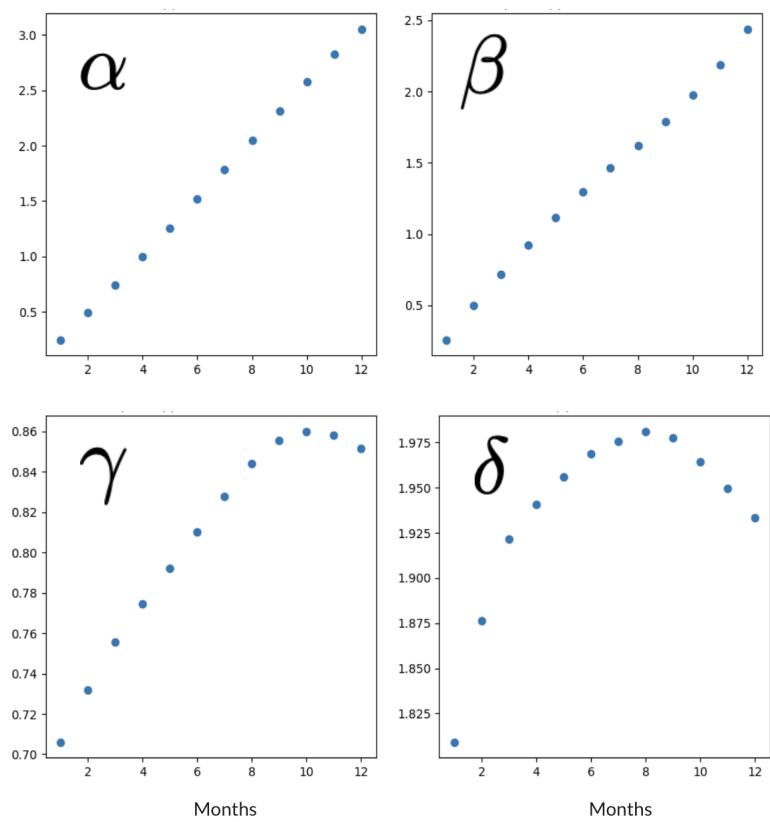


Figure 5: Model Parameters vs. Appreciation Window

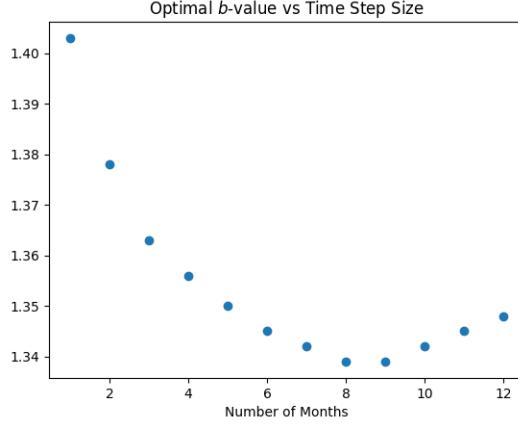


Figure 6: Optimal b-value vs. Appreciation Window

states that were not in the region under consideration had a proportion of 0. We then multiplied the regional proportion by the corresponding state HPI values obtained from CoreLogic. Summing up the resulting value for each state in the region at a given point in time, we were able to reconstruct a state-population-weighted HPI for each region (HPI_{Regional}). We then computed HPA_{Regional} using Equation 6 with HPI_{Regional} in place of Price. Figure 7 shows reconstructed HPI and HPA for each census region.

Analogously, we can reconstruct a national state-weighted HPI by computing the proportion of the US population attributed to each state and multiplying the state proportion by the corresponding state HPI values obtained from CoreLogic. Summing up the resulting value for each state at a given point in time, we can reconstruct a national state-population-weighted HPI. Figure 8 provides justification for the aforementioned reconstruction approach. Figure 8a and 8b compare our reconstructed HPI and HPA values with the FHFA HPI and HPA values, respectively. In Figure 8a, it is apparent that while on different scales, our reconstructed HPI follows the trend of the FHFA HPI. We similarly observe our HPA computed from the reconstructed HPI following the trend of the FHFA HPA, as depicted in 8b.

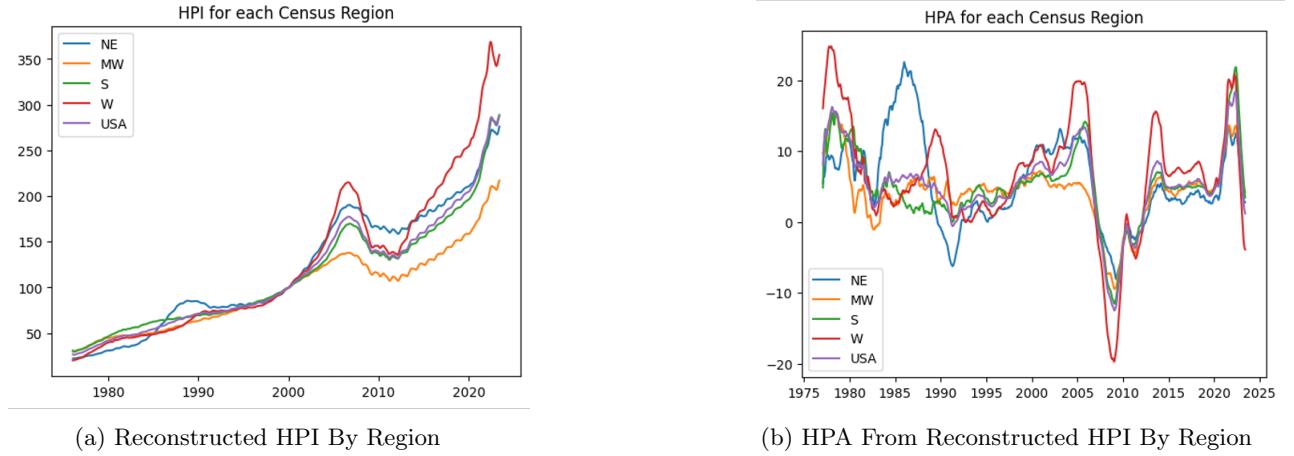


Figure 7: Reconstructed HPI and HPA for each census region

Results for the Midwest region are shown in Figure 9. Figure 9a shows the CoreLogic HPA model plot for the Midwest region and Figure 9b shows the plot of Midwest region CoreLogic HPA vs regional demand. We see that while there are some areas where the model doesn't fully capture the observed trends (notably Q1 2023), the model does remarkably well considering it does not utilize observed regional EHI or regional stock values. Plots for the remaining three regions can be found in Appendix B. Note, again, that the b-value was

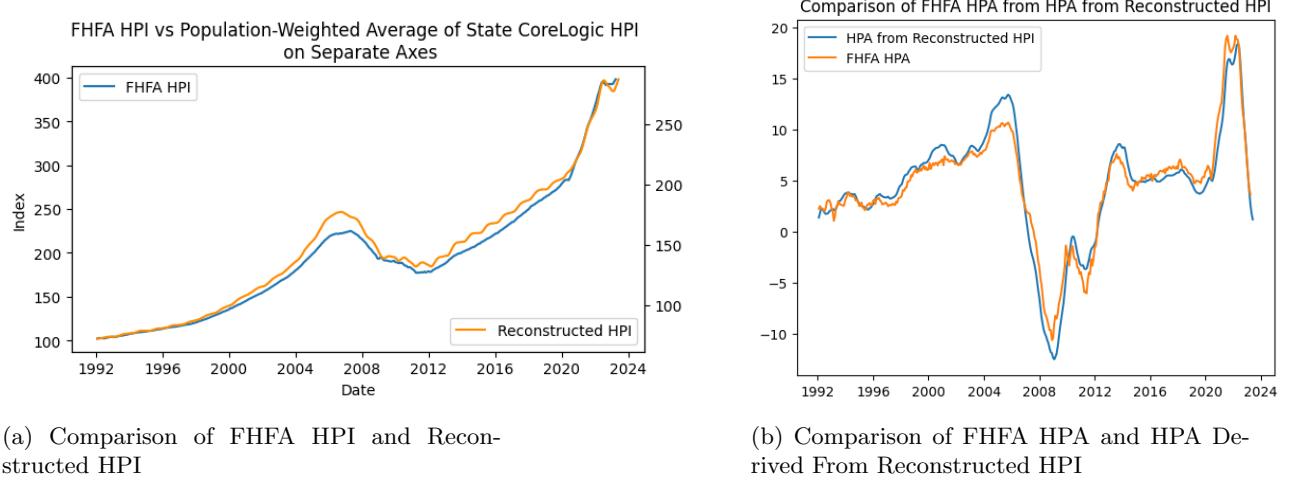


Figure 8: Reconstructed HPI and HPA for the US

optimized for each region.

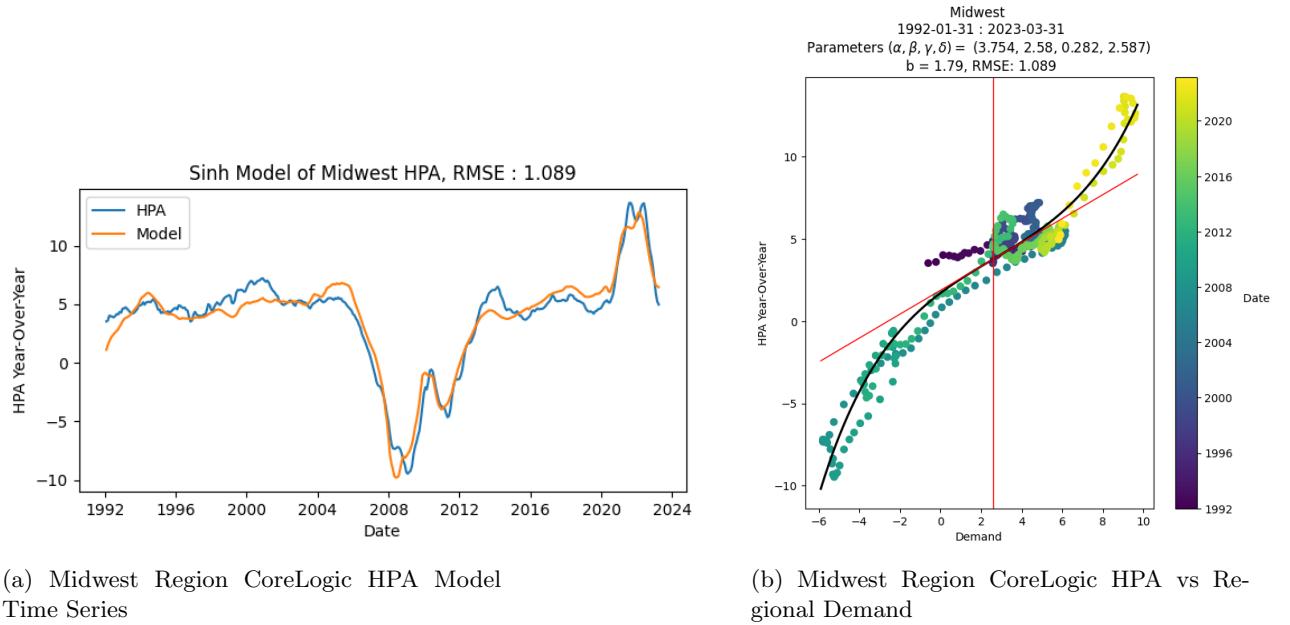


Figure 9: Midwest Region CoreLogic HPA Model

State-Level Analysis

For the state-level analysis we encountered an analogous issue to that encountered in the regional analysis: we did not have access to state EHI or state stock values. To address this issue, we utilized the regional demand (see Equation 9) for the state-level analysis, where the region used was the state's corresponding region. Additionally, we computed state HPA values from the reconstructed state CoreLogic HPI values. The process for obtaining the reconstructed state CoreLogic HPI values is detailed above in the discussion of the regional analysis. Like with the regional analysis, we optimized the b-value for each state.

Here we provide examples of the state model applied to two states. Figure 56 and Figure 83 show the results of the state model for Iowa and Tennessee, respectively. Looking more closely at Figure 56 we see that while

the state-specific model for Iowa does not fit the earliest data well, it matches the observed state HPA values quite well for the most recent years. Figure 83 shows a model fit that matches the observed HPA values fairly closely after approximately 2015, with the exception of Q1 2023 where the model is not dropping rapidly enough to match the observed HPA values. Plots for the remaining state and DC can be found in Appendix C. Overall, the state results illustrate the utility of the state-specific model and the adaptability of the original national hyperbolic sine model to state-level analyses.

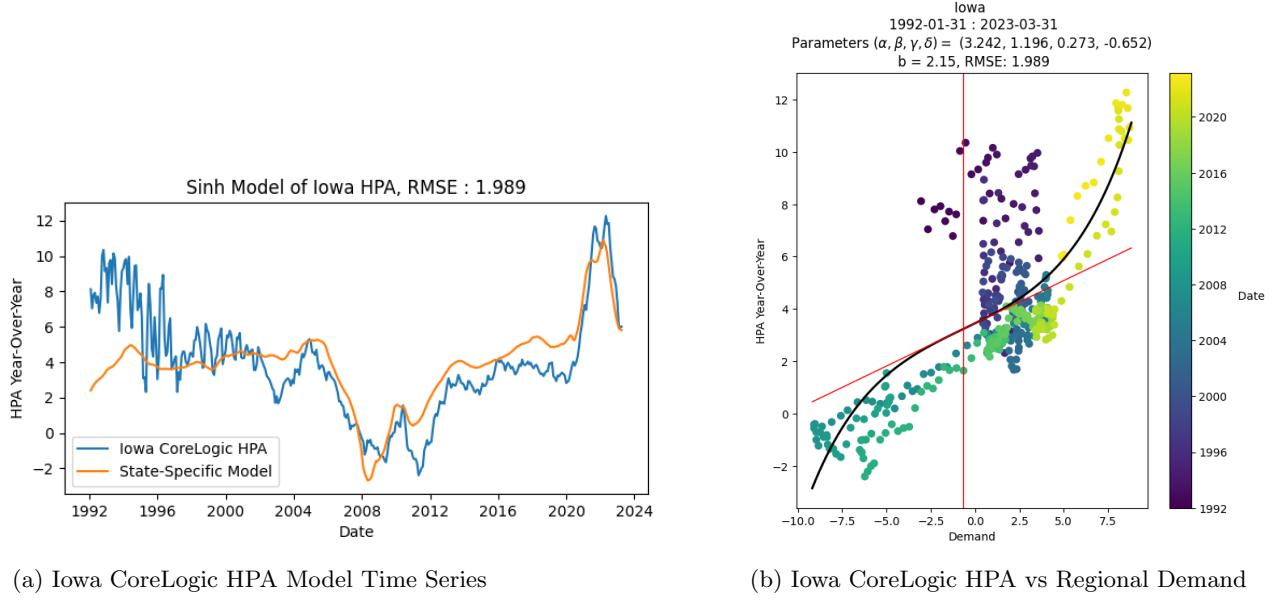


Figure 10: Iowa CoreLogic HPA Model

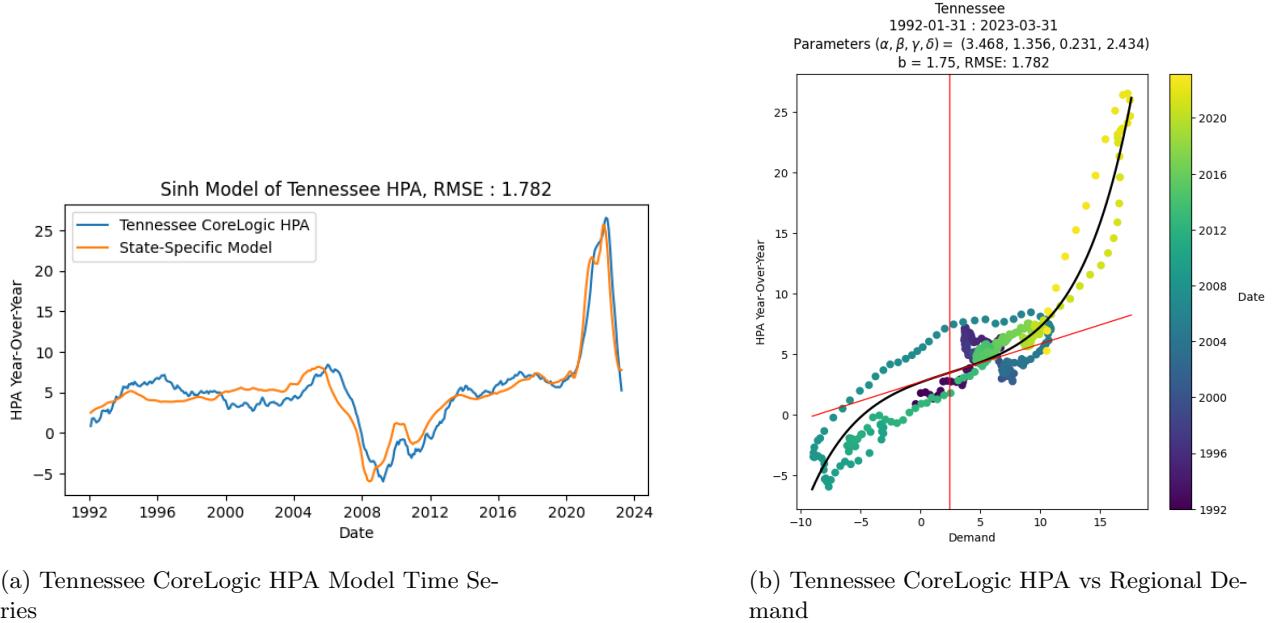


Figure 11: Tennessee CoreLogic HPA Model

3.1.3 Limiting the Training Window

In order to prevent over-fitting of the model and to test model reliability, it is useful to restrict the training data to a subset of that available and reserve the remaining for testing. In the context of the hyperbolic sine model for HPA, this amounts to limiting the data to a given time window prior to fitting a hyperbolic sine function, as is shown below in Figure 12, where restriction to the window of 1992-01-31 to 2013-12-31 results in a different model fit. We see that the value of α remains relatively stable when the training window is reduced, indicating that both each model predicts a similar appreciation rate at typical demand. Similarly, both γ and δ remain relatively unchanged by the reduction in the training window size, and the latter's stability means that the model trained on the reduced window yields nearly the same "typical" value of demand. We do, however, see a substantial change in the value of β , which increases when the window is reduced to ending in 2013. As a result, the model trained on the restricted training set predicts more dramatic changes in home appreciation at high and low demand than the model trained on a window of 1992 to 2023. One concern this could engender is that the model trained until 2013 might not perform well when tested on the out-of-sample data during the 2020-2023 time frame, where demand was exceptionally high.

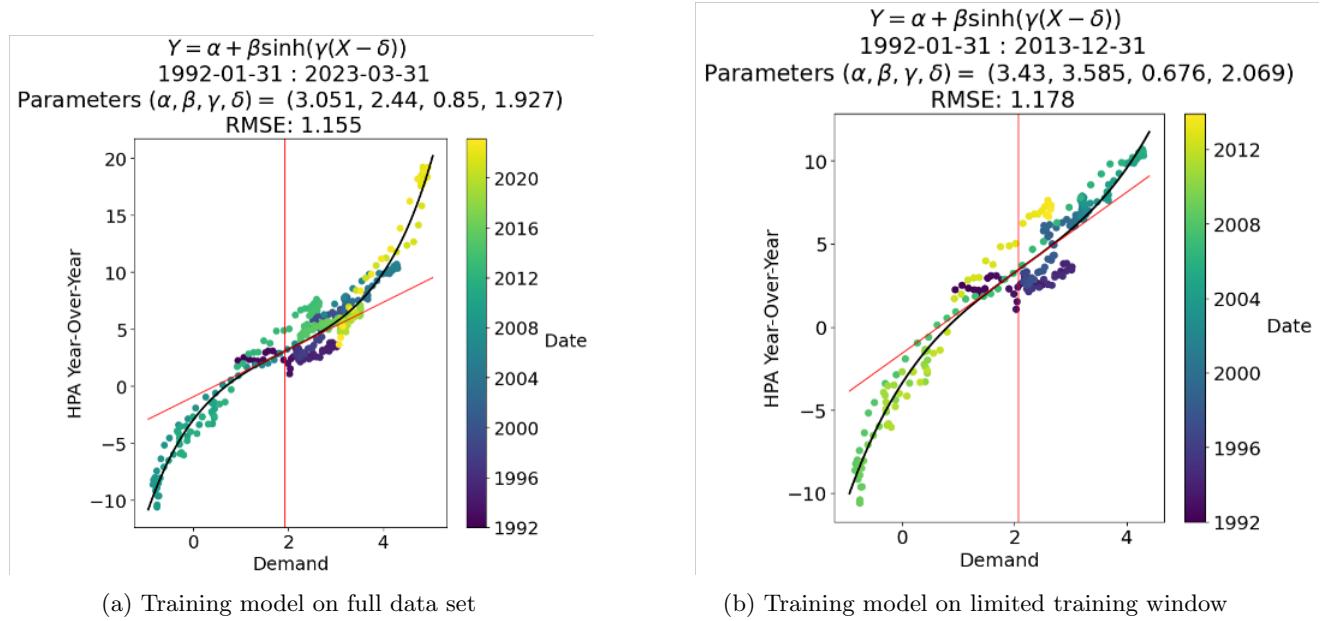


Figure 12: Restricting the Training Window

So, in order to test the resulting model's reliability beyond the year 2013, we then used the approximate formula

$$\text{HPI}_t = (\text{HPA}_{\text{Model}})_t \cdot \text{HPI}_{t-\Delta}/100 + \text{HPI}_{t-\Delta} \quad (10)$$

to iteratively model HPI from a given starting point. This is similar to using Euler's method to plot the solution to

$$\text{HPA} = \frac{d(\log \text{HPI})}{dt} \quad (11)$$

through a given point (t, HPI_t) , though it requires Δ initial HPI values, where Δ is the size of the appreciation window used for our model in months, in order to model HPI with monthly values.

As we see in Figure 13, both the monthly appreciation model and the annual appreciation model fit to this limited time window result in good models of HPI, indicating that the HPA models over-estimate and under-estimate the true values in a balanced manner.

Here, the 12-month seasonality present in Figure 13b is due to the necessity of using 12 seed values for HPI when using an appreciation window of $\Delta=12$ months. This shows that had this model been produced in

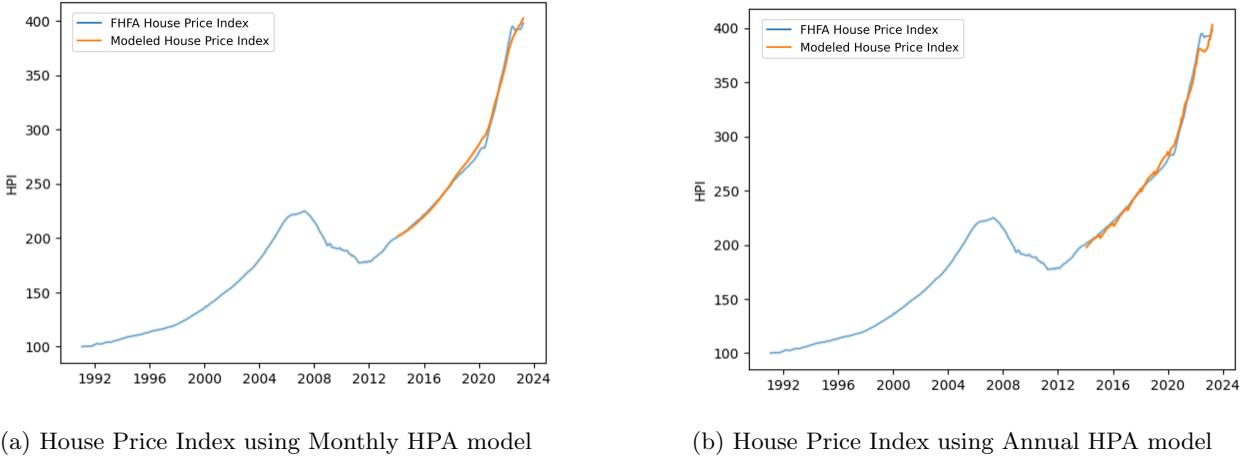


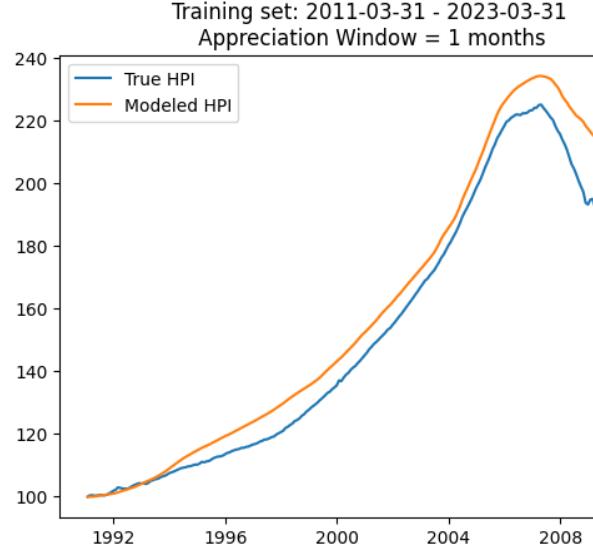
Figure 13: Reconstructing HPI from HPA Models trained on the 1992-01-31 to 2013-12-31 time window

2013, it would not have had to be updated for at least the following 10 years to maintain its usefulness.

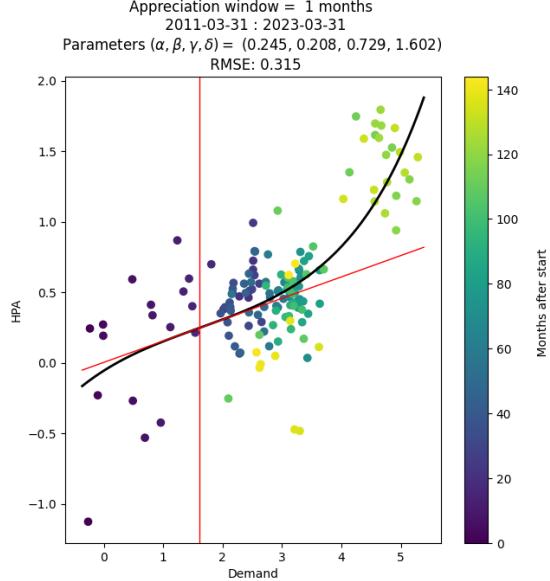
Further, Figure 13 shows that inclusion of the low demand period of the housing market crash in the training set for the model allows for an accurate model of HPI in the 2020-2023 time frame where demand is high. This is presumably due to the rotational symmetry present in the graph of the hyperbolic sine function about its inflection point, whereby fitting to periods of extremely low demand will induce the same fit (rotated 180°) where demand is high.

One can ask if this is reversible. That is, does training the model post-crash allow for a good fit through the crash? Figure 14 shows the result of training a monthly appreciation model post-crash (with a starting date of 2011-03-31) and then using the resulting models to iteratively model HPI starting at 1991-01-31 (when HPI is 100). We can see from Figure 14 that the inclusion of the period of high demand in the 2020-2023 window is insufficient to cause the resulting model to track the decline of HPI during the crash.

A distinguishing feature between Figure 12, where a training set including the crash reliably models the spike in HPI post-2020 and Figure 14, where we see a post-crash training insufficient to model the crash, is that in the case of Figure 12, the model is trained on data which contains an abundance of both positive and negative HPA values (with high and low demand, respectively), whereas in Figure 14, the training window did not include an adequate collection of low-demand, low HPA data points to accurately handle the housing market crash of 2008.



(a) HPI Model Resulting from Out-of-Sample
HPA Model



(b) Training the HPA Model Post-Crash

Figure 14: Modeling HPI using HPA Model Trained Post-Crash

Remark: Instead of using Euler's method to reconstruct HPI, we can instead solve the differential equation

$$\text{HPA} = 100 \cdot \frac{d \log(\text{HPI})}{dt} \quad (12)$$

as

$$\text{HPI}_t = \text{HPI}_{t_0} \cdot \exp \left(\int_{t_0}^t \frac{\text{HPA}}{100} dt \right). \quad (13)$$

As an example, in the sample Python code which follows (corresponding output shown in Figure 15a), we use the training window used in Figure 14b and construct HPI models using both Euler's method and via solving the O.D.E. and using Riemann sums (with an interval length of 1 month) to then approximate the integral present in the solution. For comparison, Figure 15b displays the analogous result where the training window is 1991-02-28 to 2014-01-31 and HPI is modeled from the end of that window to the present.

```

1 #Modeling HPI using Euler's Method
2 HPI_MODEL[0]=100
3 i=1
4 while i in range(1, len(HPA_MODEL)):
5     HPI_MODEL[i]=HPA_MODEL[i]*HPI_MODEL[i-1]/100+HPI_MODEL[i-1]
6     i+=1
7
8 start='1991-01-31'
9 stop='2011-03-31'
10 plt.plot(HPI[start]*np.exp(np.cumsum(HPA_MODEL[start:stop])/100), label='HPI Via
    solving the O.D.E.')
11 plt.plot(HPI_MODEL[:stop], label="HPI Via Euler's Method")
12 plt.plot(HPI[:stop], color='gray', alpha=.5, label='Observed HPI')

```

The resulting HPI models from these two disparate approaches are evidently extremely similar in both examples shown in Figure 15. This provides a glimpse into the vast mathematical toolbox made available to the developer through the novel approach of viewing home price appreciation through a differential framework

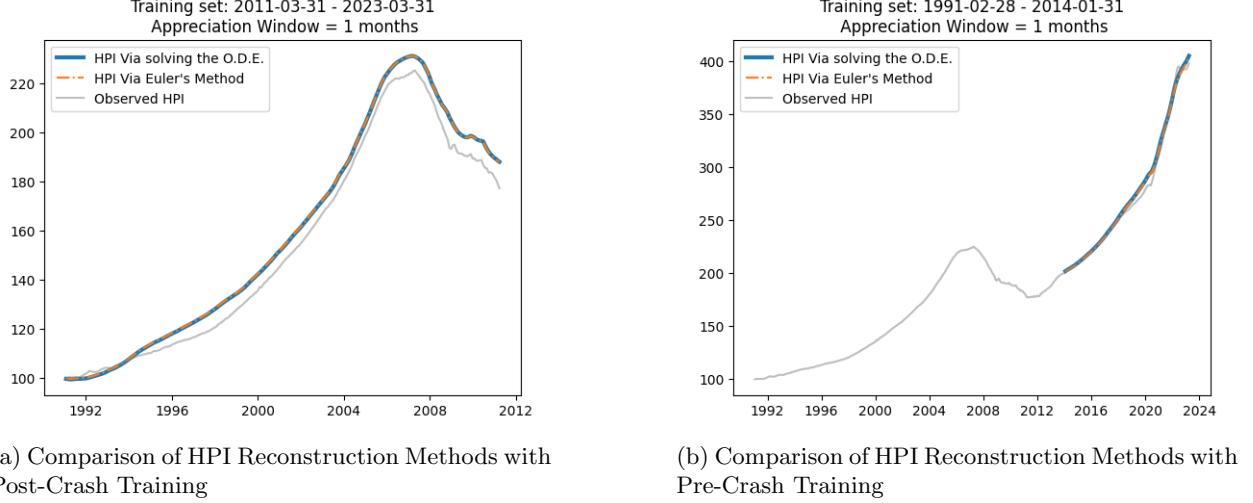


Figure 15: Comparison of Alternate Methods of HPI Reconstruction

as opposed to traditional linear constructs. Figure 15 further demonstrates the symmetry of the hyperbolic sine model with respect to low and high demand. In Figure 15a, we see that the high demand present in model’s the training window during the 2020-2023 time frame allows the model to fit well for the early parts of the crash, though it diverges from observed values when demand falls to exceedingly low values. Conversely, Figure 15b demonstrates the inclusion of the very low demand period of 2008 in the training window allows the model to track the 2020 spike in HPI (with associated high demand) closely.

When fitting a model, developers do not have access to data beyond the present, though would like to have an understanding of a model’s ability to remain functional into the future. As was shown in Figure 15b, a hyperbolic sine model fit to data up to and including 2014-01-31 functions well until at least 2023-03-31. We can expand upon this program; Figure 16 shows the result of progressively increasing the training window of the hyperbolic sine model of HPA on the corresponding HPI model, in each instance modeling HPI for 60 months beyond the end of the training window. The resulting models struggle to perform well for extended amounts of time near the beginning of the market crash, but adhere tightly to observed values when the training window includes at least the initial phase of the crash, indicating impressive model stability beyond that period. One can therefore reasonably expect a model trained on data up to and including the present to exhibit sustained future performance.

3.2 Model Comparison

3.2.1 National Models Comparison

Throughout our analysis, we explored two models for year-over-year HPA: the hyperbolic sine model as the primary approach and a linear regression model referenced as a reliable approximation (see Section 2.2). Although both models share theoretical similarities, the linear regression model’s advantage lies in its easily interpretable parameters. To gain deeper insights into the parameters of the hyperbolic sine model in Equation 4, we leverage the knowledge gained from the linear model referenced in Equation 5. This strategic approach enhances our understanding of the nonlinear model’s underlying characteristics.

To account for the varying responses of HPA in each of the three demand regions, we divide our model into three segments to facilitate the interpretation of year-over-year changes. We designate two breakpoints, A and B , which define the boundaries for these demand regions. In each segment, we utilize the new demand variable, labeled as Demand_t . The table provided below outlines the specific approach employed for each segment.

Based on the linear model and Table 3, we derive the three distinct models as follows:

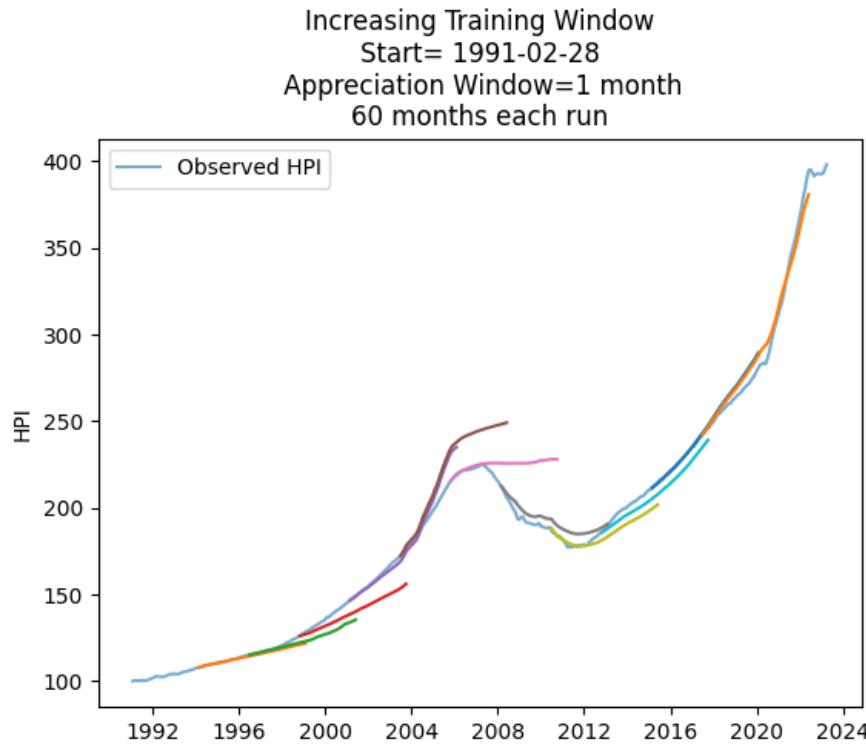


Figure 16: Modeling HPI using HPA Model Trained on Progressively Increasing Training Window

Region	Demand_t^L	Demand_t^M	Demand_t^H
Low: $\text{Demand}_t < A$	$(\text{Demand}_t - A)$	A	0
Medium: $A \leq \text{Demand}_t \leq B$	0	Demand_t	0
High: $\text{Demand}_t > B$	0	B	$(\text{Demand}_t - B)$

Table 3: Formulation for high, medium, and low demand

Low demand:

$$\text{HPA}_t \approx b + b_L(\text{Demand}_t - A) + b_M \cdot A = \{b + (b_M - b_L)A\} + b_L \cdot \text{Demand}_t \quad (14)$$

Medium demand:

$$\text{HPA}_t \approx b + b_M \cdot \text{Demand}_t \quad (15)$$

High demand:

$$\text{HPA}_t \approx b + b_M \cdot B + b_H(\text{Demand}_t - B) = \{b + (b_M - b_H)B\} + b_H \cdot \text{Demand}_t \quad (16)$$

Thus,

$$\text{HPA}_t \approx \begin{cases} \{b + (b_M - b_L)A\} + b_L \cdot \text{Demand}_t, & \text{Demand}_t < A \\ b + b_M \cdot \text{Demand}_t, & A \leq \text{Demand}_t \leq B \\ \{b + (b_M - b_H)B\} + b_H \cdot \text{Demand}_t, & \text{Demand}_t > B \end{cases} \quad (17)$$

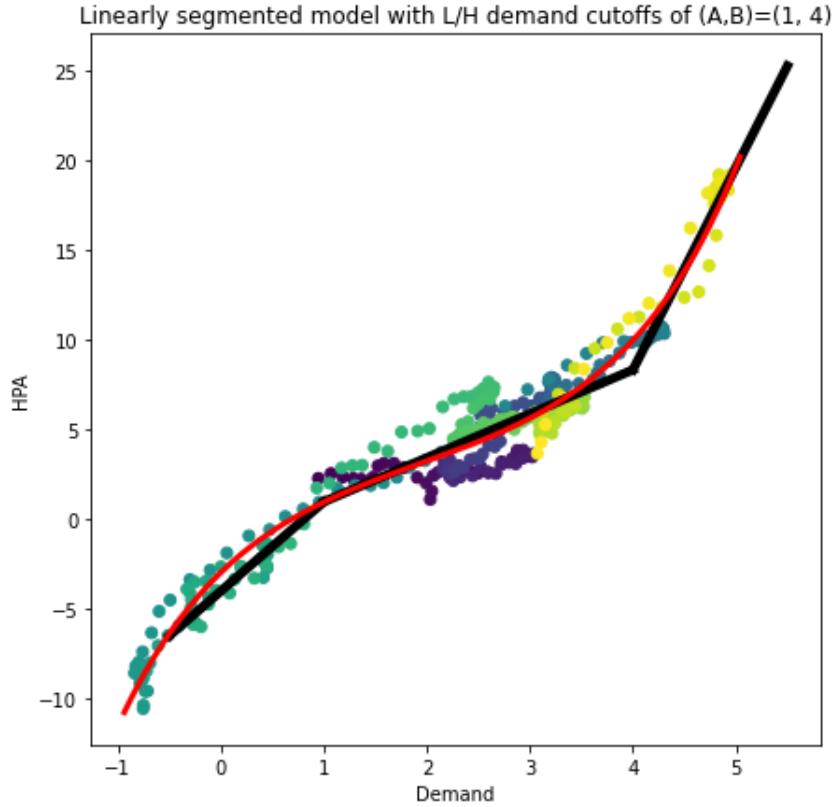


Figure 17

For illustrative purposes, we examine the plot shown in Figure 17. Equation 17 represents the regression models used for the linear approximation of the various demand regions.

$$\text{HPA}_t \approx \begin{cases} -4.04 + 4.98 \cdot \text{Demand}_t, & \text{Demand}_t < 1 \\ -1.51 + 2.45 \cdot \text{Demand}_t, & 1 \leq \text{Demand}_t \leq 4 \\ -36.93 + 11.31 \cdot \text{Demand}_t, & \text{Demand}_t > 4 \end{cases} \quad (18)$$

In general, the three segments provide valuable insights, indicating that HPA tends to increase when there is a rise in demand for homes, and conversely, it depreciates when there is little to no demand. Analyzing the results, we observe that, on average, home prices experience a substantial appreciation of approximately 11.31% for each unit increase in demand when demand is at the high end. Conversely, when demand is low, home prices typically depreciate by approximately 4.98% for each unit increase in demand. Lastly, HPA shows a steady rise, averaging around 2.45%, when the demand hovers around 1% to 4%.

By utilizing the breakpoints, intercepts, and slopes obtained from the three classical linear regression equations, we can estimate the coefficients in the nonlinear model using the following relationships:

- The medium range slope b_M and intercept b correspond to the linear part of the curve near the inflection point. Thus,
 - $\beta\gamma \approx b_M$ (slope at inflection point) and $\alpha \approx b$ (intercept at inflection point)
- $\delta = (A + B)/2$
- The low and high range slopes indicate how rapidly prices change as demand moves further from the inflection point. We can use the difference between b_M and b_L, b_H to estimate γ :
 - $\gamma = (b_M - b_L)/(A - C)$ (for low demand); where $C < A \in \text{Demand}_t$
 - $\gamma = (b_H - b_M)/(B - A)$ (for high demand)

In summary, the piecewise regression models offer estimates of the slope, intercept, and demand ranges, which can be used to substitute into the nonlinear model and determine its parameters. The presence of nonlinearity is captured by the varying slopes observed across different demand ranges.

3.2.2 State models Comparison

A linear state model

The state model analyzed in Section 3.1.2 takes one single variable as an input. Recall that the demand variable was obtained by confining the pair $(\text{EHS}_{\text{Regional}}, \text{EHI}_{\text{Regional}})$ into one single variable

$$\text{Demand} = \frac{\text{EHS}_{\text{Regional}} - b \cdot \text{EHI}_{\text{Regional}}}{\text{StockUS}}. \quad (19)$$

Recall also that the main advantage of this model is that it reduces the number of inputs needed to estimate $\text{HPA}_{\text{state}}$. By doing so, the model is easier to interpret and is less likely to have over-fitting issues. In order to test the robustness of this non-linear model, we choose to compare its performance with a simple linear model with two inputs.

More precisely, we assume the following linear model for $\text{HPA}_{\text{state}}$:

$$\text{HPA}_{\text{state}} = b_0 + b_1 \text{EHS}_{\text{Regional}} + b_2 \text{EHI}_{\text{US}} \quad (20)$$

for some real numbers b_0, b_1 and b_2 .

Recall that it was not possible to obtain reliable datasets for $\text{EHS}_{\text{State}}$ and $\text{EHI}_{\text{State}}$, so it is not possible to directly use these two variables to predict $\text{HPA}_{\text{state}}$. The linear model in equation (20) will overcome this issue by adjusting the parameters b_1 and b_2 accordingly. The results of this linear regression are summarized in Table 4.

One remarkable fact about this linear model is that an important relationship between $\text{HPA}_{\text{state}}$ and the two predictors was naturally captured in the parameters of the model. In fact, by inspecting the results of the linear regression state by state, we observe that the following inequalities hold for each state:

$$b_1 > 0 \quad \text{and} \quad b_2 < 0.$$

State Name	b_0	b_1	b_2	RMSE
AK	-10.32	17.47	-2.01	1.99
AL	-12.86	12.32	-2.97	1.67
AR	-6.13	8.88	-2.91	2.29
AZ	-25.46	53.59	-10.96	6.27
CA	-22.10	47.03	-9.43	4.58
CO	13.70	5.51	-6.18	3.17
CT	-7.17	35.65	-4.84	2.80
DE	-21.14	16.19	-2.31	2.94
FL	-21.45	23.33	-7.04	6.56
GA	-0.36	9.90	-6.11	2.80
HI	-28.17	43.89	-5.23	3.47
IA	2.15	6.32	-2.72	1.29
ID	-6.42	29.45	-8.06	6.47
IL	-16.00	23.31	-3.40	2.49
IN	7.17	7.05	-5.15	2.47
KS	-1.58	10.82	-3.16	1.77
KY	0.45	5.78	-3.42	1.28
LA	-12.82	9.25	-0.38	1.70
MA	6.28	20.23	-6.00	2.74
MD	-27.98	19.46	-1.87	5.03
ME	2.31	28.43	-6.12	2.52
MI	21.30	2.63	-9.12	3.81
MN	-2.08	15.90	-5.18	2.86
MO	-3.50	14.75	-4.24	1.31
MS	-7.52	8.96	-2.87	2.17

State Name	b_0	b_1	b_2	RMSE
MT	-3.51	19.62	-5.02	3.76
NC	-8.57	11.13	-3.51	2.81
ND	3.27	2.06	-0.59	2.52
NE	5.95	5.46	-3.67	1.39
NH	4.85	27.08	-7.03	2.30
NJ	-10.00	43.24	-5.16	2.38
NM	-7.90	20.74	-4.46	3.72
NV	-24.35	56.59	-13.01	6.43
NY	-10.52	34.93	-2.70	2.56
OH	7.09	7.51	-5.41	1.30
OK	-3.06	5.93	-1.97	1.99
OR	-11.89	28.88	-5.68	4.43
PA	-5.07	29.00	-3.59	1.52
RI	-2.90	42.93	-7.69	3.39
SC	-13.98	14.60	-3.88	2.57
SD	4.11	7.29	-3.37	2.49
TN	-5.22	10.75	-4.33	2.68
TX	2.71	5.66	-3.85	2.76
UT	0.29	17.72	-5.83	6.68
VA	-12.50	13.61	-3.57	6.24
VT	-1.02	28.06	-4.70	3.08
WA	-9.61	27.83	-5.99	4.63
WI	-5.07	15.64	-3.84	0.97
WV	-18.20	14.20	-1.94	4.32
WY	-7.68	16.75	-2.48	3.09

Table 4: Model parameters for the linear model in equation (20)

This is an indicator that linear regression, despite being a very simple model, was able to capture the following fundamental fact: for fixed inventory and sales, home appreciation increases with sales and decreases with inventory. This demonstrates naturally occurring signals in the data that inventory puts downward pressure on prices and sales put positive pressure on prices.

Note also that the relationship

$$|b_2| < |b_1|$$

holds for every state with the exception of Michigan. This difference in scale was to be expected since

$$|\text{EHS}_{\text{Regional}}| < |\text{EHI}_{\text{US}}|.$$

Two Examples

Figure 18 compares the predicted values of HPA_{WI} (orange plot) with the actual observed values over time (blue plot). The linear model performs fairly well from 2000 to 2020. The model starts underperforming around the end of 2020 but catches up with the observed values in recent years. One plausible explanation could be that the occurrence of the COVID-19 pandemic had a significant impact on the market dynamics during the period from 2021 to 2022 and the linear model could not fit one set of parameters to accommodate that region. On the other hand, it is important to observe that this model did catch the overall shape of the blue plot (with a minor shift to the left). This is remarkable and shows that linear regression is a strong model to use as a first approximation for other models with increased complexity.

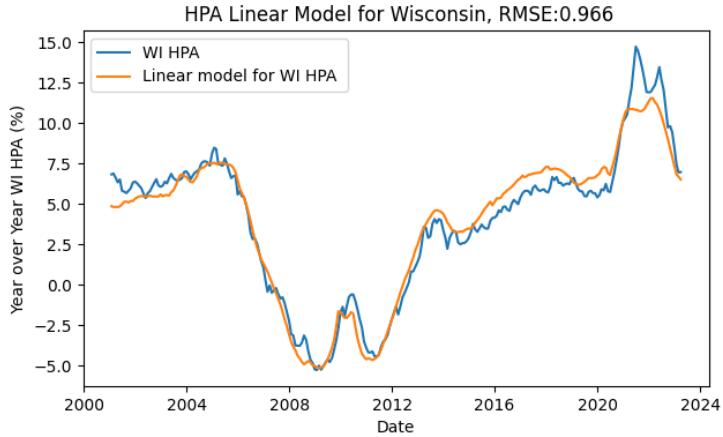


Figure 18: HPA Linear Model for Wisconsin, RMSE : 0.97

Similarly, Figure 19 compares the predicted values of HPA_{MO} with the observed values over time (blue plot). The RMSE of this model is larger than that of the model for Wisconsin depicted in Figure 18. Nevertheless, we see that again the linear model performs reasonably well outside the Great Recession time frame between 2008 and 2009 and the COVID-19 period between late 2020 and early 2022. This is an indicator that the linear model we used for each state is a good approximation of state home price appreciation in the absence of a major economic event. It is worth mentioning that as in the first example, the linear model did capture several small ‘bumps’ in the HPA_{MO} fluctuations.

Comparison with the hyperbolic sine state model

Figures 20 and 21 compare the results obtained by the linear model (20) with the hyperbolic sine model investigated in Section 3.1.2.

For Wisconsin, both the linear and hyperbolic sine models performed almost similarly from year 2000 to 2021. The short period of time following the crash in 2008 doesn’t seem to significantly diverge from the

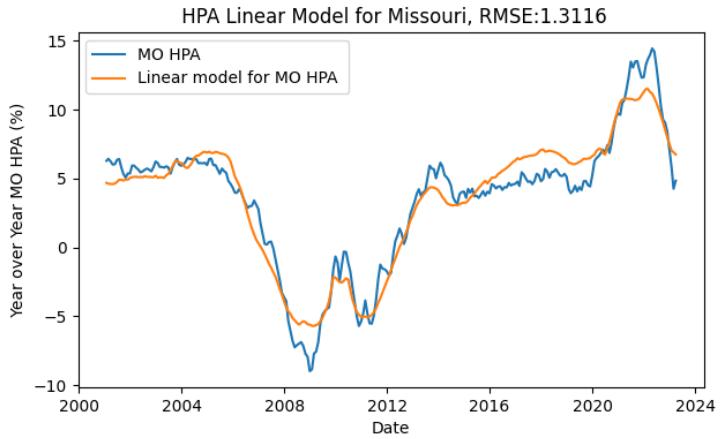


Figure 19: HPA Linear Model for Missouri, RMSE : 1.31

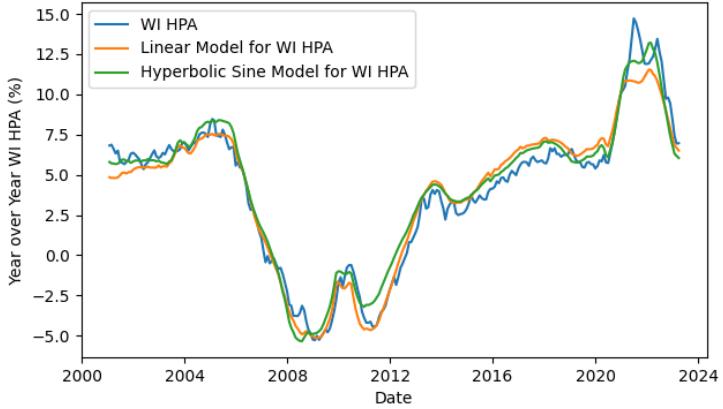


Figure 20: Linear vs Hyperbolic Sine Models for Wisconsin

linear fit. A plausible explanation could be the fact that Wisconsin experienced a relatively fast recovery from the financial crisis. We also observe a significant divergence in the predictive power of the two models between Q1 2021 and Q1 2023, with the highest difference around Q4 2020.

Similarly, Figure 21 compares the predictions of the two models for the state of Missouri. Unlike the hyperbolic sine model, the linear model misses again the full signal from Q1 2008 to Q1 2009 and from Q1 2020 to Q4 2022. Note also that both of these models were unsuccessful in capturing the observed values of home appreciation from Q2 2012 to Q3 2014 and around Q1 2023. Recall that this discrepancy between the predicted values and the observed values within those two time frames was also present in the main model used to predict HPA at the national level.

As a summary, the linear model is a simple yet powerful tool that can be used to better understand how state HPA responds to regional sales and national inventory. As for its predictive power, the results show that the linear model is overall successful in predicting the true observed data if the housing market is not disturbed. In the presence of major economic events the linear model fails to provide a reasonable estimate of the observed data and the hyperbolic sine model is an appropriate alternative in that case.

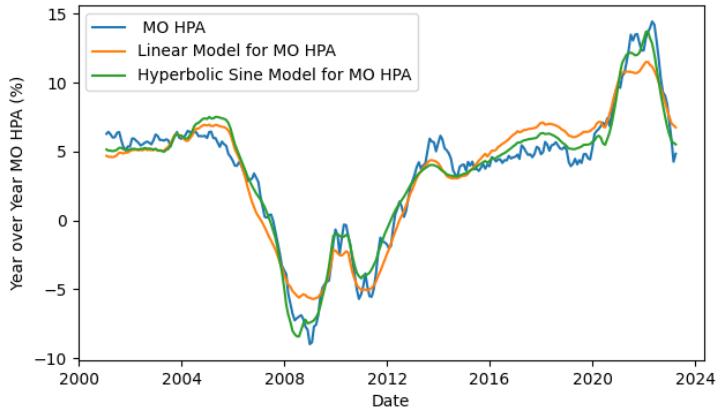


Figure 21: Linear vs Hyperbolic Sine Models for Missouri

3.3 Improving the Current Model

There are two time periods in which both existing models (hyperbolic sine and its linearly segmented approximation) do not accurately capture observed national home price appreciation.

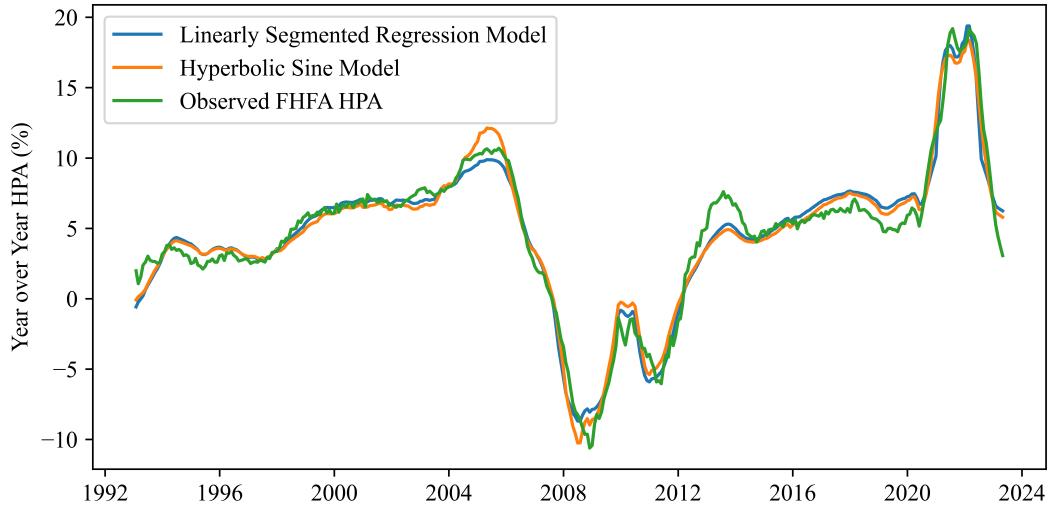


Figure 22: Both models underestimate HPA from 2012-2014 and overestimate HPA in 2023

- Both existing models underestimate observed home price appreciation between Q2 2012 and Q3 2014.
- Both existing models overestimate observed home price appreciation during Q1 2023.

These periods and the models' errors are shown in Figure 22.

The existing models for national home price appreciation are parsimonious in that they use few variables as inputs. It is expected that a simple model for a complex macroeconomic variable will not perform perfectly throughout the sample. Despite this, we notice a commonality between these periods that may suggest avenues for model improvement.

The second quarter of 2012 was witness to the onset of recovery in the housing market after the Great Financial Crisis (GFC). By March 2012, the bloated inventory of foreclosed properties had wound down from a peak of 3.5 million units to about 2 million units. Furthermore, consumer appetite for real estate by this point had returned, demonstrated by an increase in over 1 million units in existing home sales. These trends combined led the demand variable to increase rapidly during Q2 2012 - Q3 2014. Coincidentally, home prices began rising for the first time in four years during the same period.

By the end of 2022, the effects of the Federal Reserve's Quantitative Tightening policies had been realized in the housing market. Mortgage rates had almost doubled between mid-2021 and late 2022, severely damping the record-high existing home sales exhibited during the exciting Covid-19 housing frenzy. This massive drop in EHS (about a 25% decline) caused the demand variable to fall nearly as quickly as it had during the GFC. Coincidentally, the trend of unprecedented growth in home prices came to a halt, with the recent flattening of the FHFA Home Price Index.

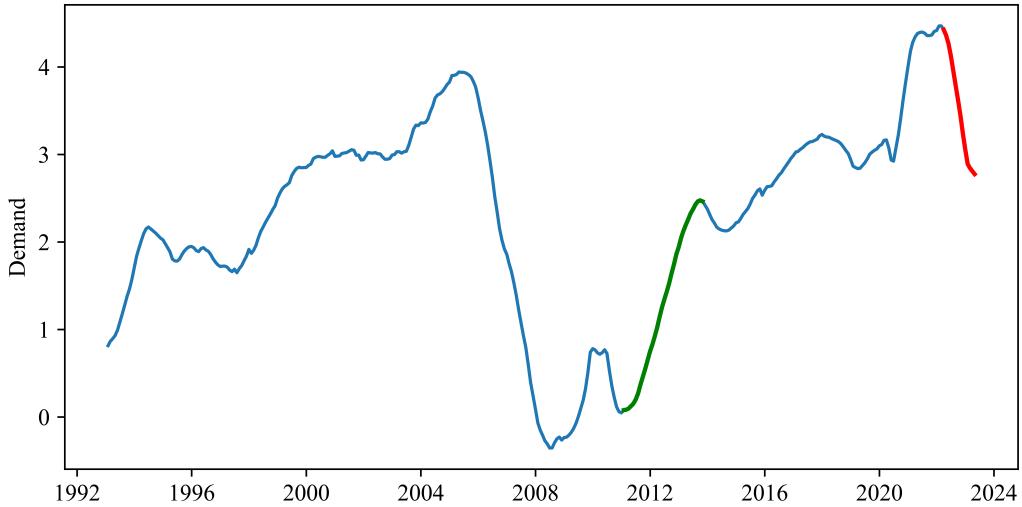


Figure 23: Demand rose rapidly before and during 2012-2014, while falling in the latter part of 2022

The shared feature of both periods, presented in Figure 23, is a rapidly moving demand variable, either upwards or downwards. This suggests that the *rate of change* of the demand variable, in addition to the absolute level of demand, may offer explanatory power in a model for home price dynamics.

3.3.1 Modifying Demand: Grid Searching for Velocity Variables in the Linear Model

In Section 3.2 of our study, we discuss a linear model version of the main model, dividing the demand variable into three segments. To define these segments, we used two breakpoints, A and B , set to 1 and 4, respectively, resulting in an initial RMSE value of 1.19.

To enhance the model's overall fit, we performed a grid search to find the best combination of A and B values. After a grid search operation over many combinations of (A, B) pairs, we discovered that the optimal values were $A = 0.74$ and $B = 4.01$, leading to a slight improvement with an RMSE value of 1.18.

In an effort to further improve the model's fit in the aforementioned periods, we introduced two predictor variables, designed to capture instances in which the velocity of demand is large in magnitude. For predetermined positive integers C and F , and real numbers $D > 0$ and $E < 0$, we construct the variables Velocity_t^{\pm} . For Velocity_t^+ :

1. Take a C -month difference of Demand _{t} ,

2. Create a variable which is 0 when the above does not exceed the positive constant D , and equals the positive difference between the above variable and D in the case that the above variable exceeds D ,
3. Take an F -month weighted moving average of the above variable.

Call this variable Velocity_t^+ . To create its opposite, Velocity_t^- :

1. Take a C -month difference of Demand_t ,
2. Create a variable which is 0 when the above exceeds the negative constant E , and equals the negative difference between the above variable and E in the case that the above variable is less than E ,
3. Take an F -month weighted moving average of the above variable.

By incorporating these two variables into the linear model with the existing three segmented demand variables, we hoped to resolve the periods of poor model fit.

However, despite conducting a second grid search to minimize the RMSE restricted to these periods over the quantities C , D , E , and F , the model's improvement in fit was not strong enough to warrant incorporation of these variables into a final model. Our attempts to fine-tune the model with these additional variables did not yield the desired outcome.

This indicates that the linear model may have limitations in capturing the complex dynamics of HPA, and further exploration or alternative modeling approaches may be necessary to enhance the accuracy of predictions during data gaps and periods with significant fluctuations.

3.3.2 Modifying Demand: Incorporation of an Intensity Variable in the Hyperbolic Sine Model

Model developers additionally considered ways to incorporate the velocity of demand into the hyperbolic sine model of home price appreciation. As before, define Demand_t to be a 12-month weighted moving average of the quantity

$$\frac{(\text{EHI}_t - b \cdot \text{EHS}_t)}{\text{Stock}_t}.$$

We introduce a variable, Intensity_t , informed by the velocity of demand. At time t , the value of Intensity_t equals

- U if $B \leq \text{Demand}_t - \text{Demand}_{t-24}$,
- 0, if $A < \text{Demand}_t - \text{Demand}_{t-24} < B$, and
- $-V$ if $\text{Demand}_t - \text{Demand}_{t-24} \leq A$

where $A < 0$ and $B, U, V > 0$. For appropriate choices of parameter values, Intensity_t identifies instances in which Demand_t has moved dramatically upwards or downwards over a two-year window. Figure 24 shows the relationship between this value and $\text{Demand}_t - \text{Demand}_{t-24}$.

We use Intensity_t to multiplicatively modify our definition of demand:

$$\text{Demand}_t^* := (1 + \text{Intensity}_t) \cdot \text{Demand}_t.$$

This new variable now captures a combination of current supply and demand forces in the market as well as their relative historical changes. Figure 25 illustrates this multiplicative effect on demand.

With a modified demand variable, we recalibrate the parameters of the hyperbolic sine model. Ultimately, we see an improvement in model fit near the two periods without compensating stability in parameter values. Side-by-side comparisons of the calibrated hyperbolic sine using these inputs are shown in Figures 26 and 27. Figure 28 shows a time series comparison of the model outputs for both.

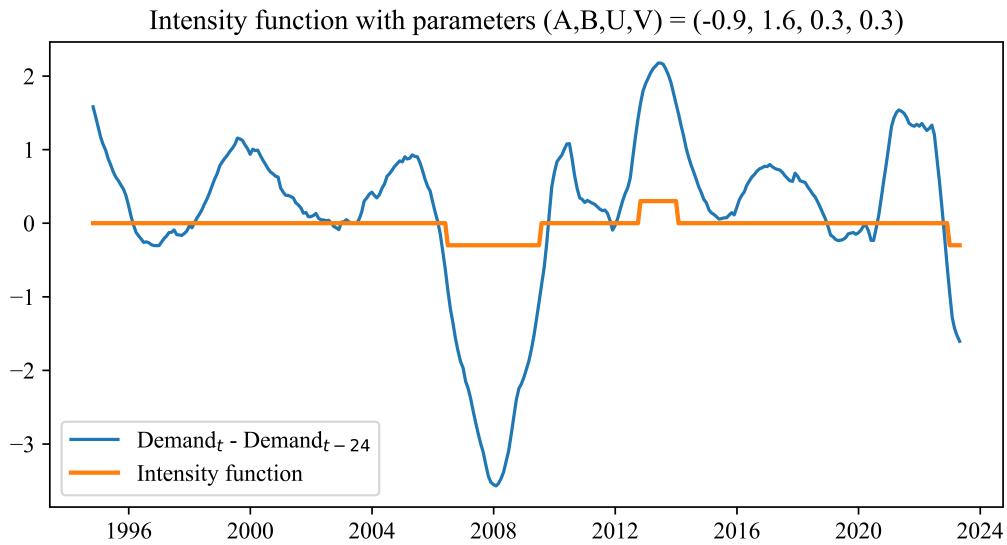


Figure 24: An example of an intensity plot derived from the value of $\text{Demand}_t - \text{Demand}_{t-24}$ for a choice of parameter values

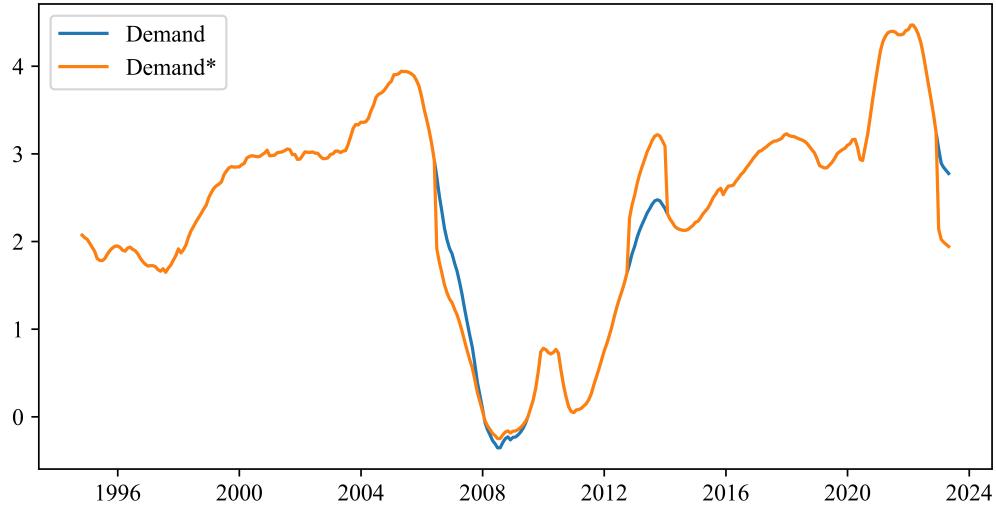


Figure 25: Comparing values of Demand_t and Demand_t^* for the chosen parameter combination

4 Downstream Applications

Embedded in most mortgage contracts between a lender and borrower is the optionality (the right, but not the obligation) on behalf of the mortgagor to pay the entirety of their remaining principal balance at any time. This *prepayment* can happen for a number of reasons, some of which include:

- The borrower is able to access accumulated equity in the property by means of a *cash out refinance*. In this case, the borrower secures a larger mortgage and uses the cash to pay off the principal balance

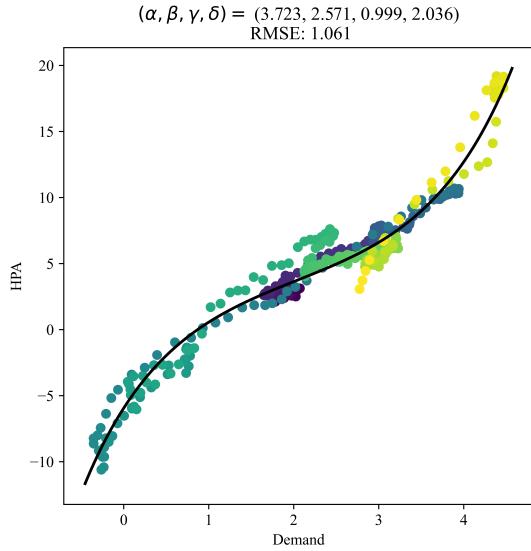


Figure 26: Model Calibrated using Demand_t

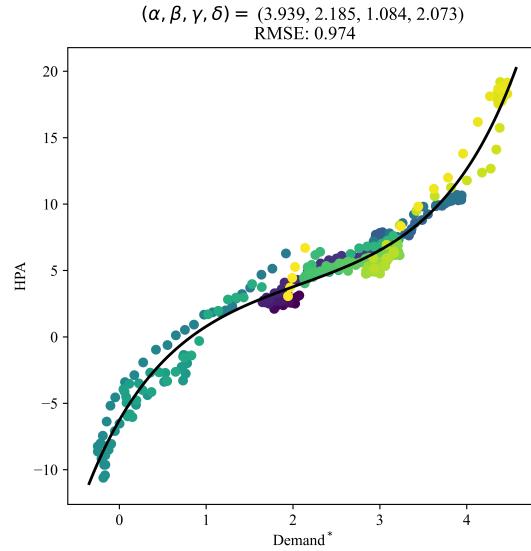


Figure 27: Model Calibrated using Demand_t^*

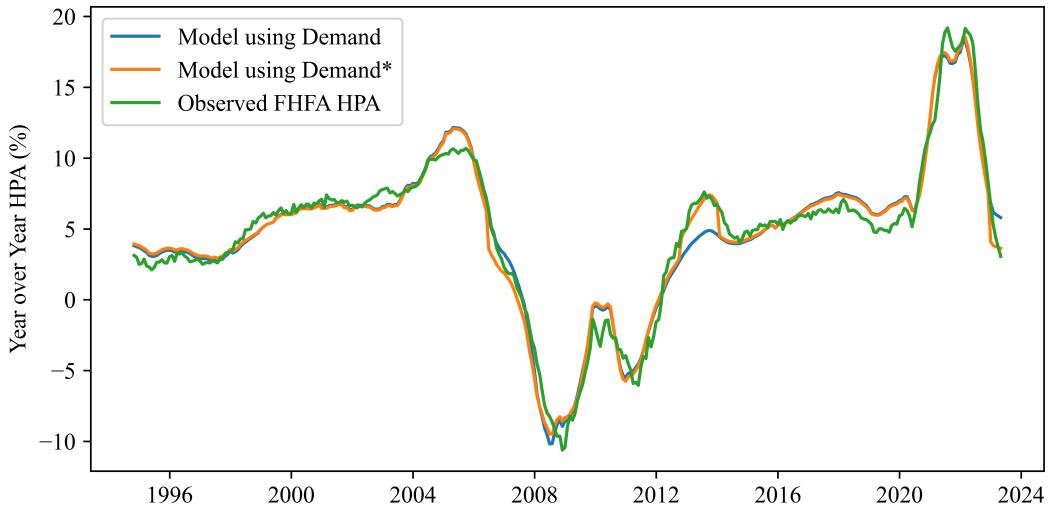


Figure 28: Comparing values of Demand_t and Demand_t^* for the chosen parameter combination. Note the relative improvement of model performance in 2012-2014 as well as 2023.

of the existing mortgage. The difference between the new mortgage and the unpaid principal of the existing mortgage is the equity extracted by the borrower.

- The borrower is able to secure cheaper financing (i.e. a lower mortgage rate) or restructure the length of the term of the loan with another lender. The execution of this option is known as a rate/term refinance.
- The borrower must relocate to a new property with a new mortgage, and uses the home sale proceeds to pay down the principal balance on the previous mortgage. This phenomenon is known as *turnover*

and is slightly less influenced by macroeconomic factors than the previous examples.

In each of the above cases, the lender receives the principal before the prescribed termination date of the original contract. Although prepayment events are common and expected¹ by banks and other lenders, the precise timing and rate of prepayments are carefully modeled by financial institutions.

The incentive for a borrower to execute a cash-out refinance is primarily driven by the Loan-to-Value ratio (LTV) of the loan. The LTV generally decreases over the term of the mortgage, and is influenced by two factors:

- The numerator: the principal balance of the mortgage decreases steadily as the borrower makes regularly scheduled monthly payments.
- The denominator: the value of the property fluctuates according to market forces.

As home prices rise, LTVs of borrowers decrease, providing more incentive for cash-out refinance action. Thus, all else equal, as home prices rise, banks and lenders can expect to see higher prepayment speeds. A key input to mortgage prepayment models is reliable forecasts for home price appreciation.

5 Future Work

Section 3.1 establishes the applicability of the main national model to regional and state housing markets. Further research and development of this model can elaborate on these discoveries. Section 3.3 outlines attempts made to improve the quality and performance of the main model for national home price appreciation. Further research and development of the model could expand in these directions.

5.1 Further Directions for Regional and State Models

Developers applied the linear and nonlinear versions of the HPA model to state-level HPA derived from the FHFA and Core-Logic's home price indices, detailed in Section 3.1. Developers may continue to employ these methods for more granular metrics of HPI; for example the FHFA's home price indices for Metropolitan Statistical Areas (MSAs).

5.2 Further Modification of the National Model

The incorporation of the velocity of demand as a multiplicative factor described in Section 3.3 can be refined. For example, a more robust analysis or optimization of the choice of non-estimated parameters A , B , U , and V could improve the strength and in-sample performance of the model. Furthermore, developers should consider an intensity factor that exhibits continuity, as opposed to a variable taking on three discrete values.

More generally, developers should investigate the conceptual soundness of the use of demand velocity as a way to improve model performance of the periods mentioned in Section 3.3, or should consider entirely alternative solutions. For example, in regards to the model's overestimation of HPA in 2023, the incorporation of NHI may explain recent negative pressure on home prices. Current Existing Home Inventory is at historic lows, providing an economic barrier to a collapse of home prices as observed during the GFC. However, New Home Inventory and construction are nearing recent historic highs to partially compensate for this relative scarcity. These trends are shown in Figure 29.

5.3 Investigating Regional and State-Wide Discrepancies Contributing to National Model Error

Developers fit hyperbolic sine models to the relationship between national demand

$$\text{Demand} = \frac{\text{EHS}_{\text{US}} - 1.35 \cdot \text{EHI}_{\text{US}}}{\text{Stock}_{\text{US}}} \quad (21)$$

¹The average length of a 30-year mortgage is between 7 and 10 years.

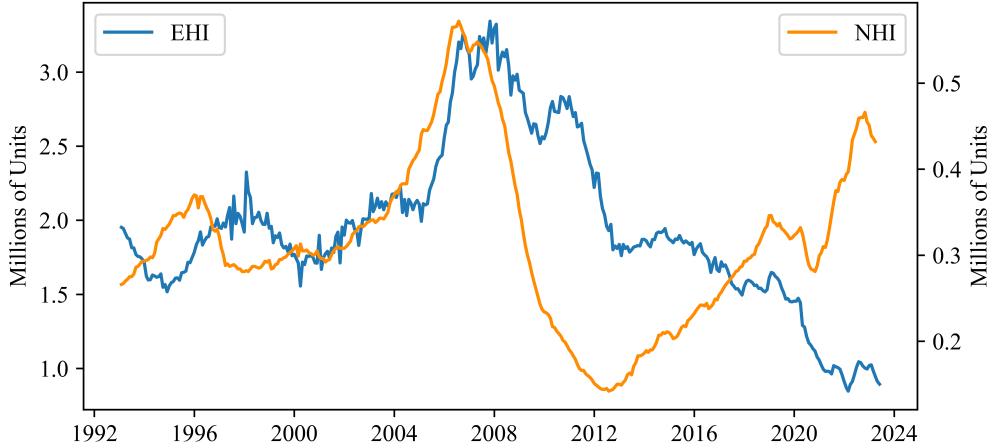


Figure 29: As EHI levels have reached historic lows, the number of new homes for sale has dramatically increased.

and state CoreLogic HPA values to see in which locations the national demand variable cannot sufficiently explain the 2012-2014 appreciation values. The results are shown in Figure 30. It is observed that the state models fit to national demand heavily underestimate HPA values in the Western part of the United States as well as in the states of Florida and Michigan. Further investigation may be prudent to discover whether some macroeconomic factor trended differently in those states during that time period, with the hope of incorporating that factor in an improved National Model.

Remark: This approach is different from computing the average signed difference between the National Model and observed state HPA values in that time window. We chose this alternate approach, as if the National Model produces HPA values consistently lower than observed values in State X, for instance, then if the National Model gives lower HPA values in the 2012-2014 window than observed in State X in that window, it does not necessarily indicate that any trend in State X at that time is contributing to the National Model's error.

Developers then applied this methodology to investigate the possibility of local trends contributing to the error observed in the National Model in 2023, computing the signed error present in the individual state models produced using national demand at the date of 2023-03-31. The results are shown in Figure 31. It is observed that the state models fit to national demand heavily overestimate HPA values in 2023 in the Western part of the United States as well as in the states of Massachusetts, Minnesota, and New York. Further investigation may show some macroeconomic factor is trending differently in those areas in 2023, with the hope of incorporating that factor in an improved National Model.

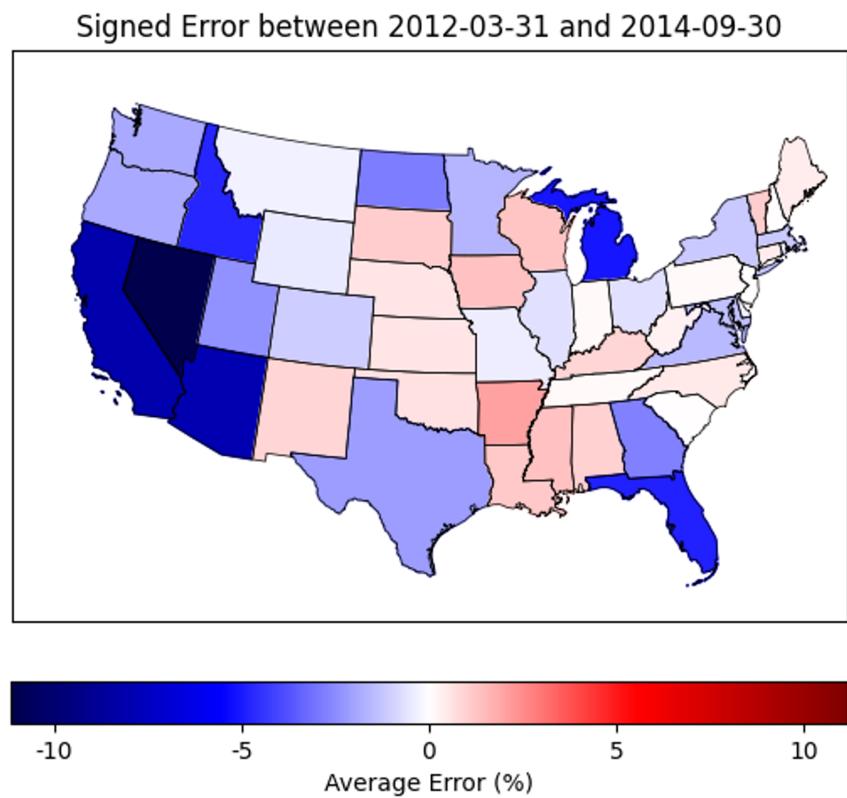


Figure 30: Signed error in Hyperbolic Sine Model Trained to CoreLogic State HPA and National Demand between 2012-03-31 and 2014-09-30

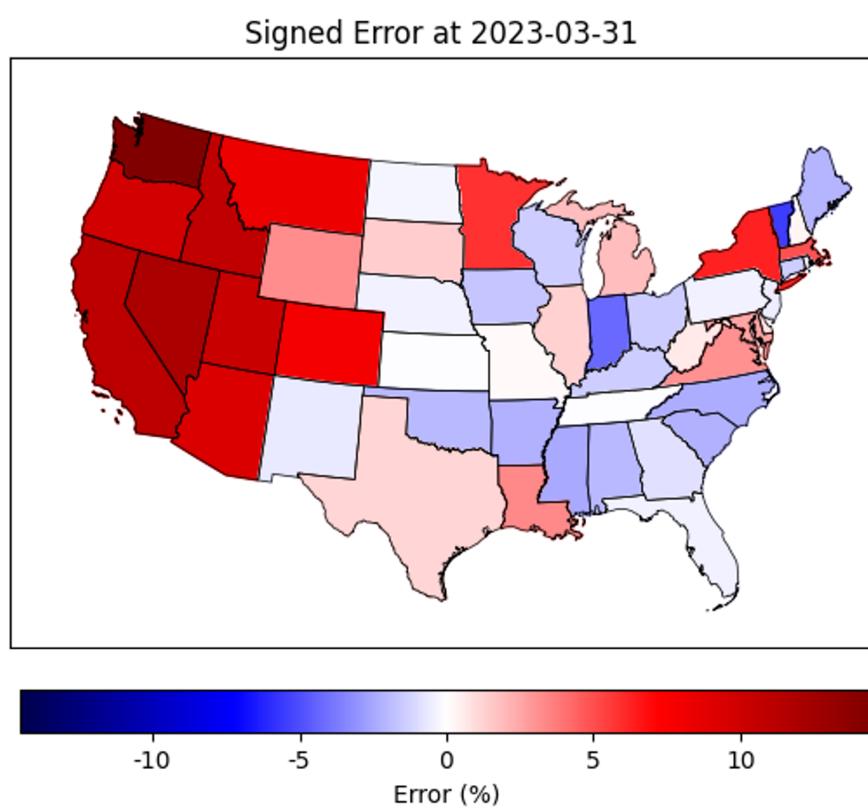


Figure 31: Signed error in Hyperbolic Sine Model Trained to CoreLogic State HPA and National Demand at 2023-03-31

References

- [1] G. Caginalp, D. Porter, and V. Smith, (2001). Financial bubbles: Excess cash, momentum, and incomplete information. *J. of Psychology and Financial Markets*, 2(2):80-99.
- [2] S. Gjerstad. (2007). The competitive market paradox, momentum, and incomplete information. *Econ. Theory*, (31)
- [3] P. Samuelson, (1947). *Foundations of Economic Analysis*, Harvard University Press, Cambridge.
- [4] V. Smith, G. Suchanek, and A. Williams. (1988). Bubbles, crashes, and endogenous expectations in experimental spot asset markets. *Econometrica*, 56(5):1119-1151

A Appendix: Miscellaneous Plots

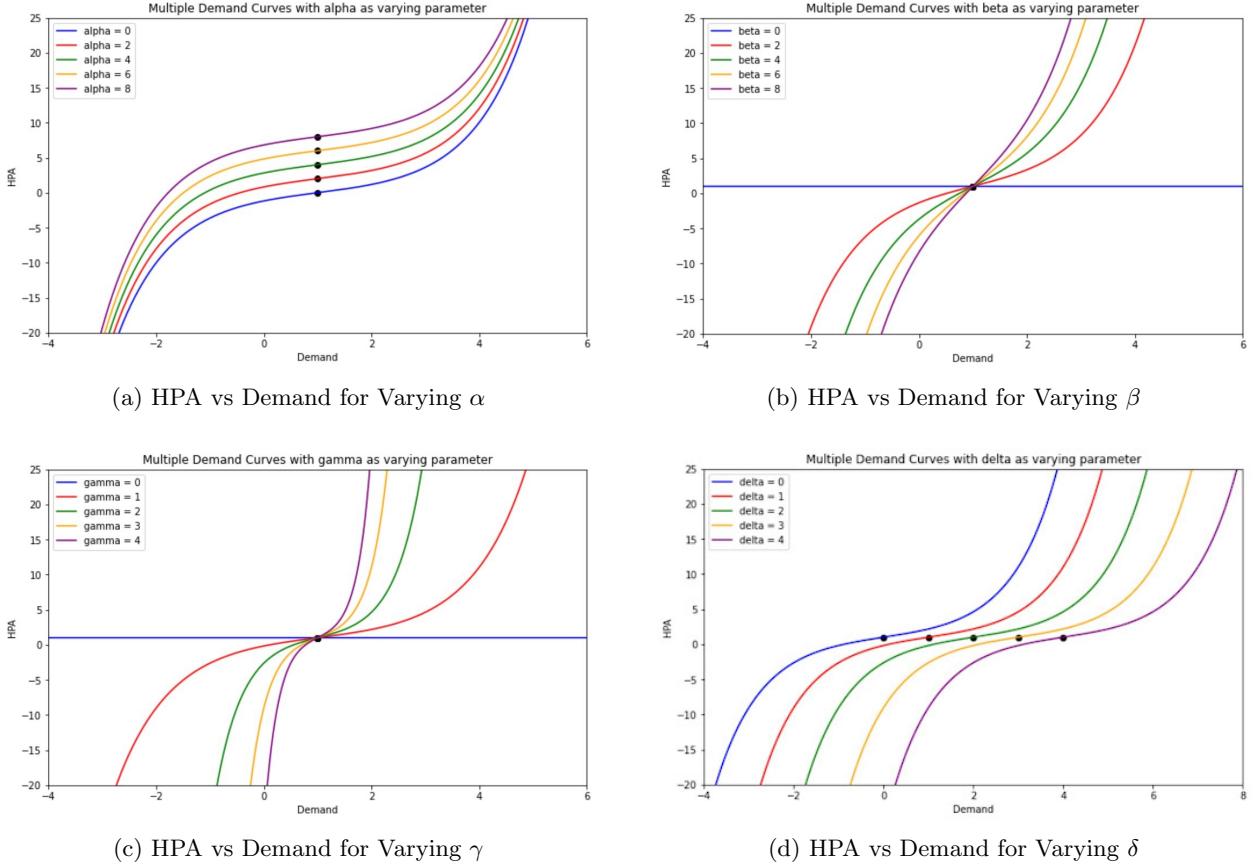


Figure 32: HPA vs Demand for Varying Hyperbolic Sine Model Parameters

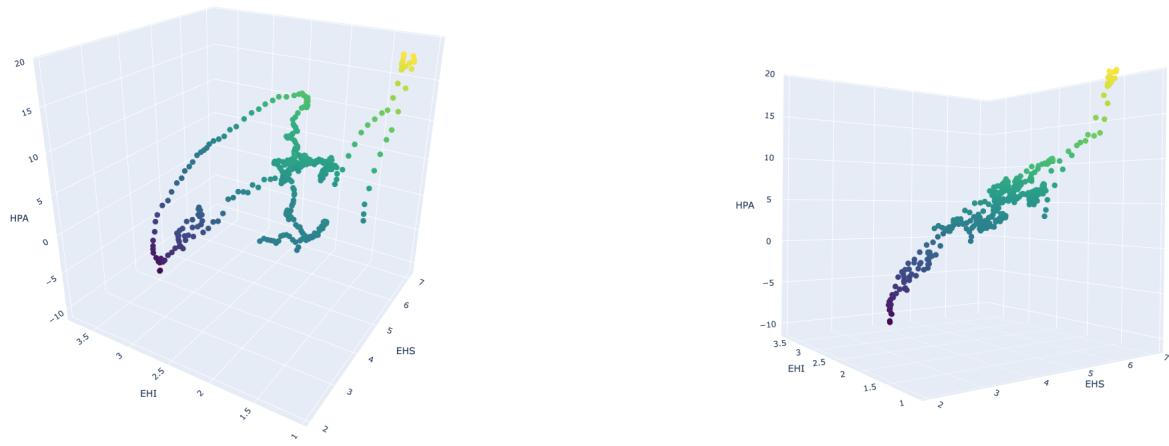


Figure 33: Evolution of $(\text{HPA}_t, \text{EHS}_t, \text{EHI}_t)$ from $t = 1992$ to $t = 2023$

B Appendix: Regional Plots

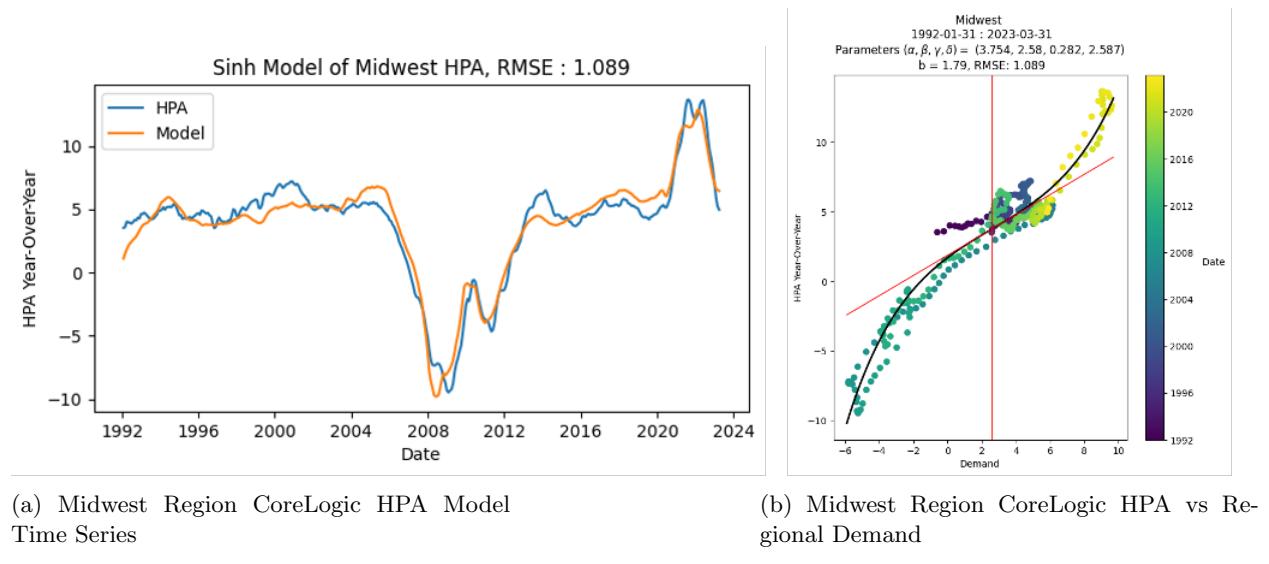


Figure 34: Midwest Region CoreLogic HPA Model

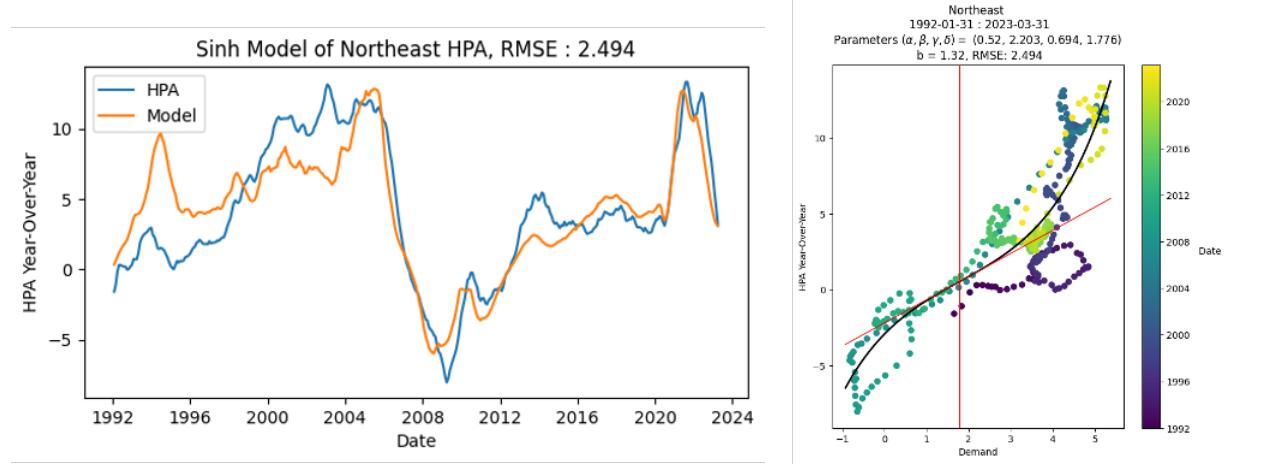
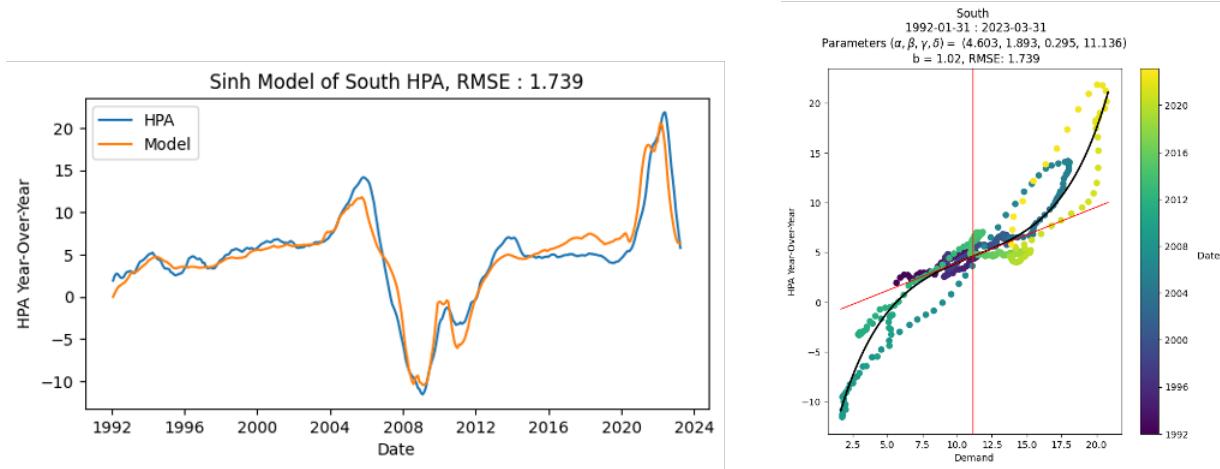


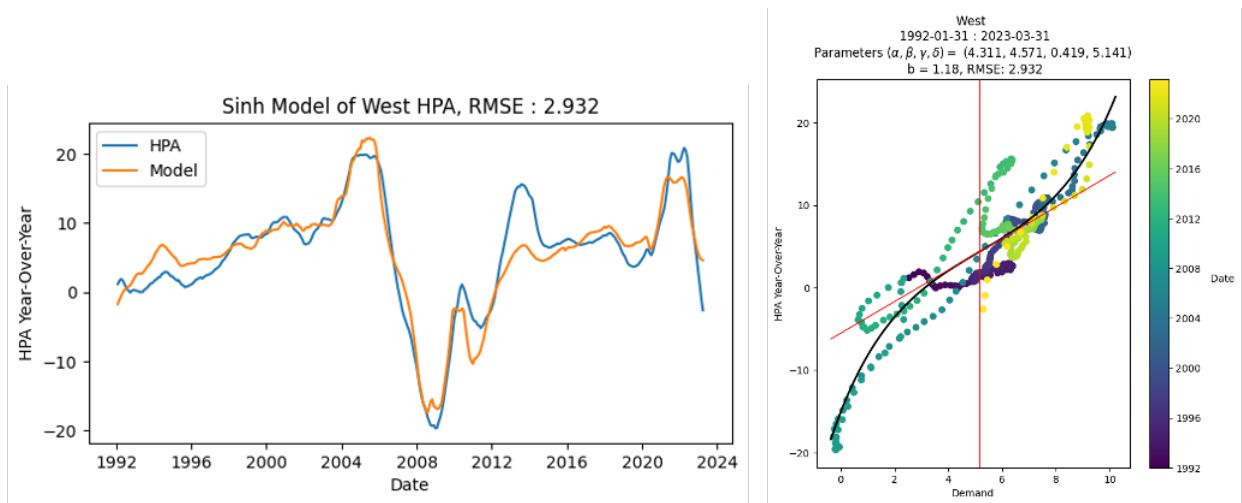
Figure 35: Northeast Region CoreLogic HPA Model



(a) South Region CoreLogic HPA Model Time Series

(b) South Region CoreLogic HPA vs Regional Demand

Figure 36: Northeast Region CoreLogic HPA Model



(a) West Region CoreLogic HPA Model Time Series

(b) West Region CoreLogic HPA vs Regional Demand

Figure 37: West Region CoreLogic HPA Model

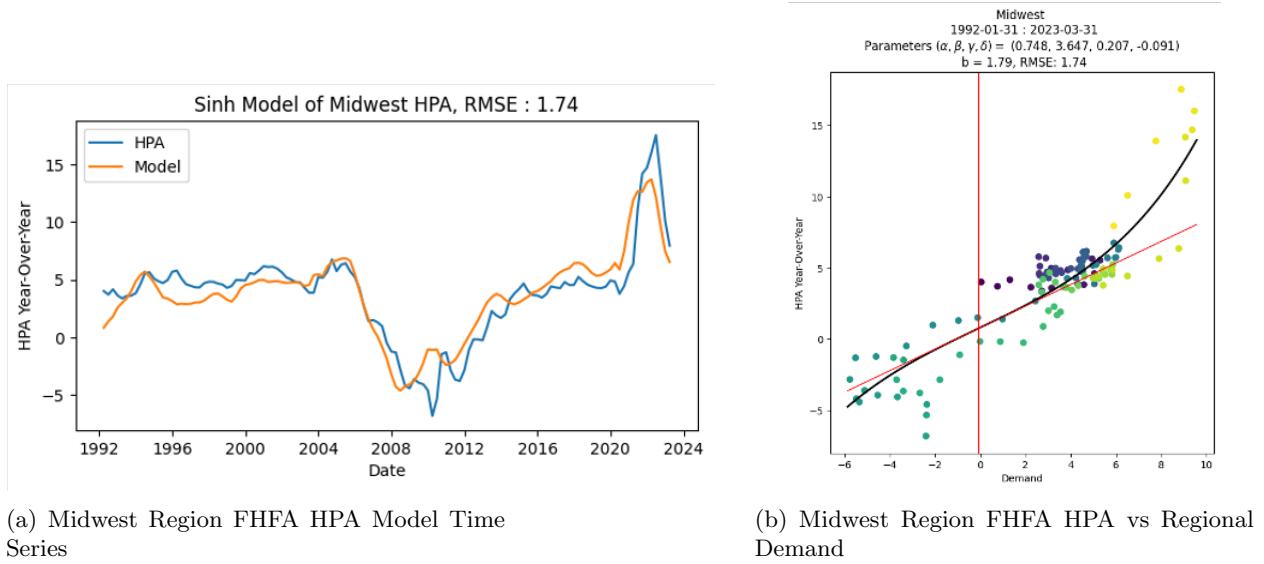


Figure 38: Midwest Region FHFA HPA Model

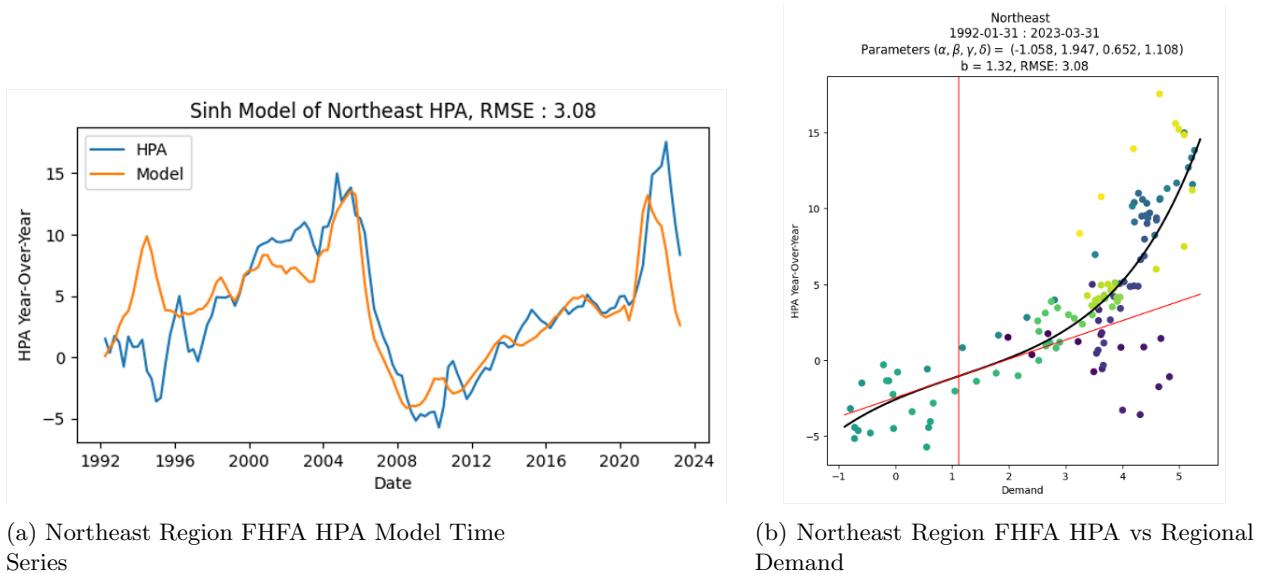


Figure 39: Northeast Region FHFA HPA Model

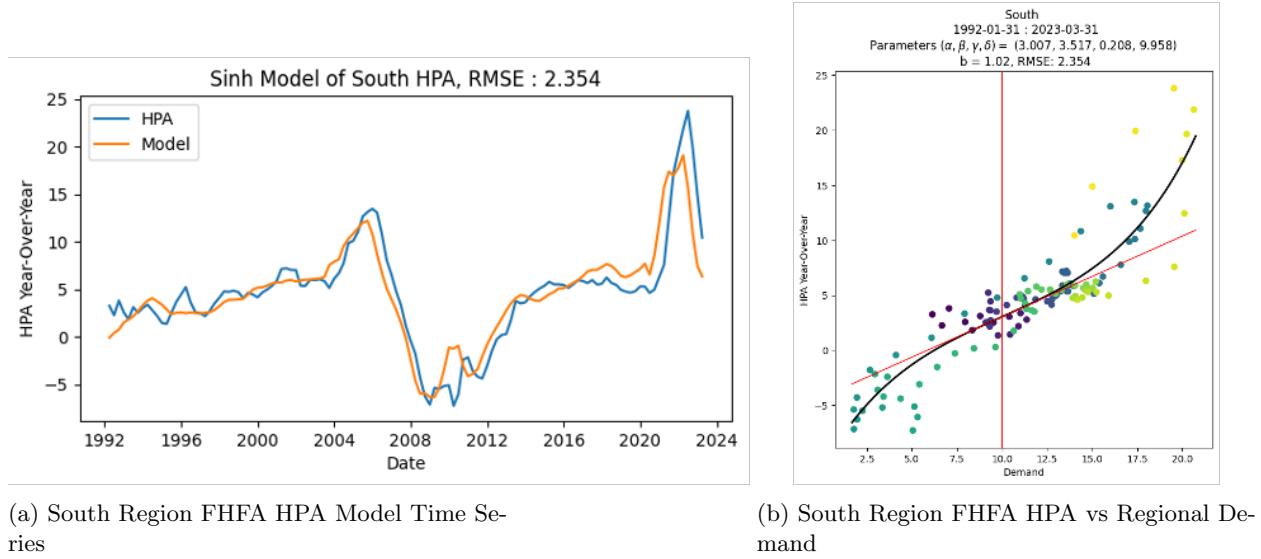


Figure 40: South Region FHFA HPA Model

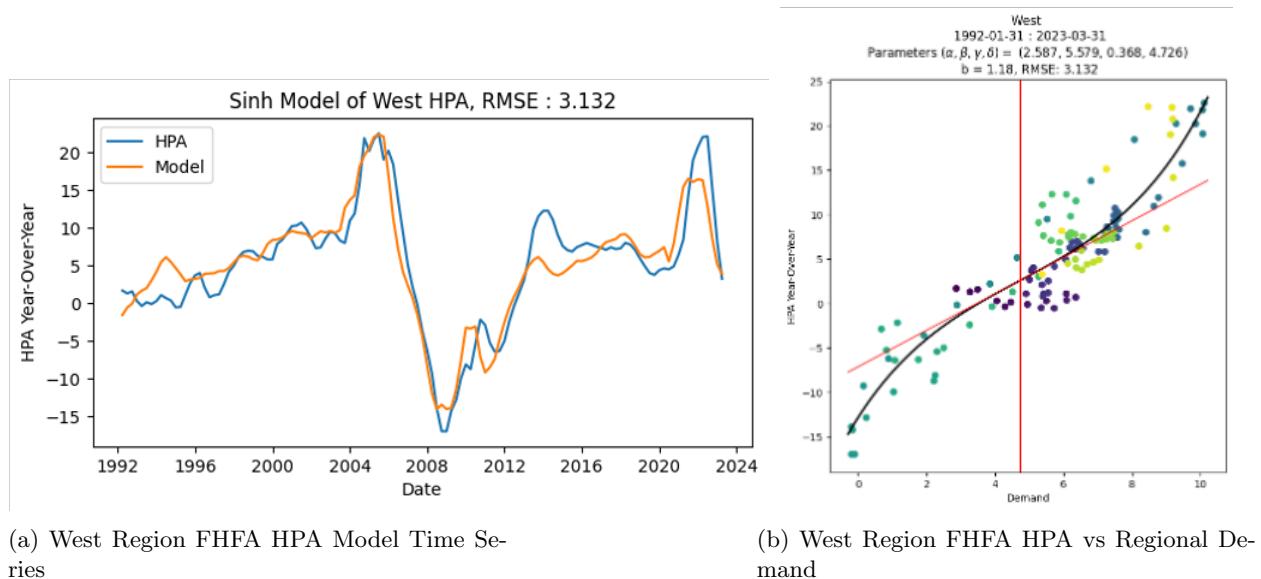


Figure 41: West Region FHFA HPA Model

C Appendix: State Model Plots

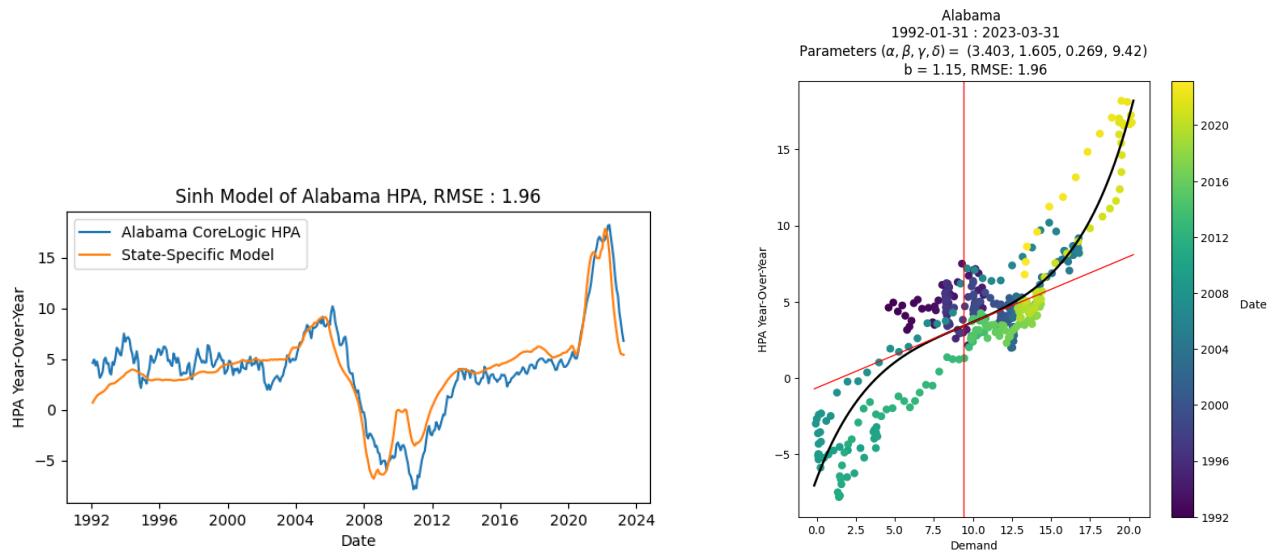


Figure 42: Alabama CoreLogic HPA Model

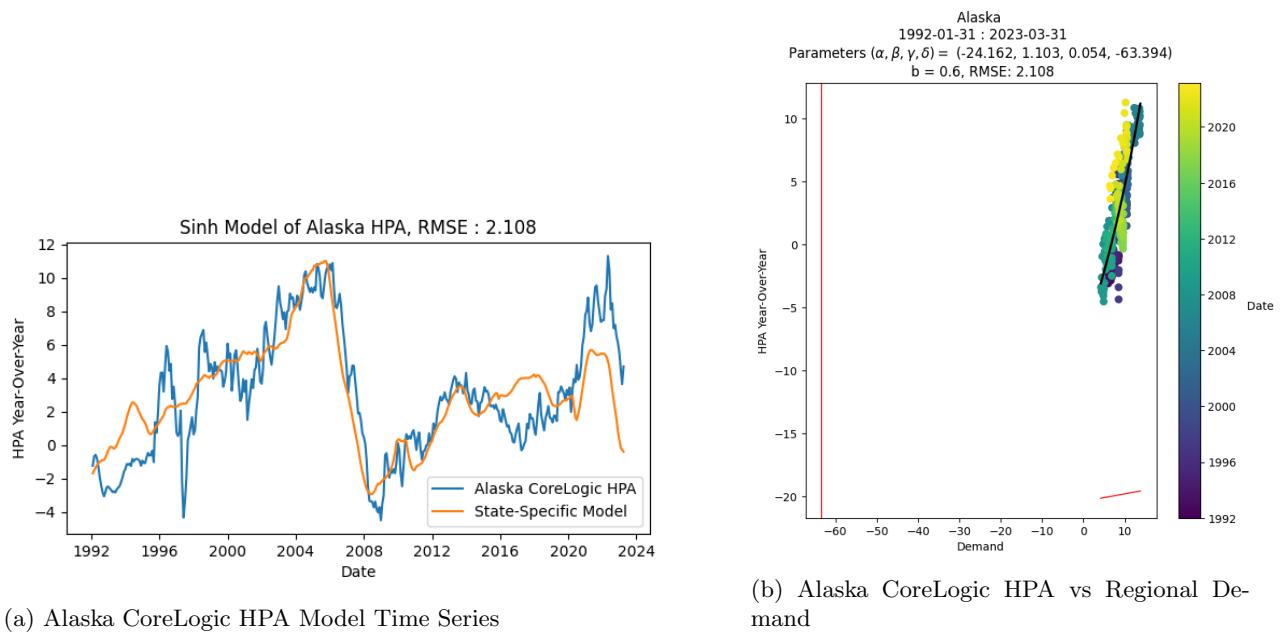


Figure 43: Alaska CoreLogic HPA Model

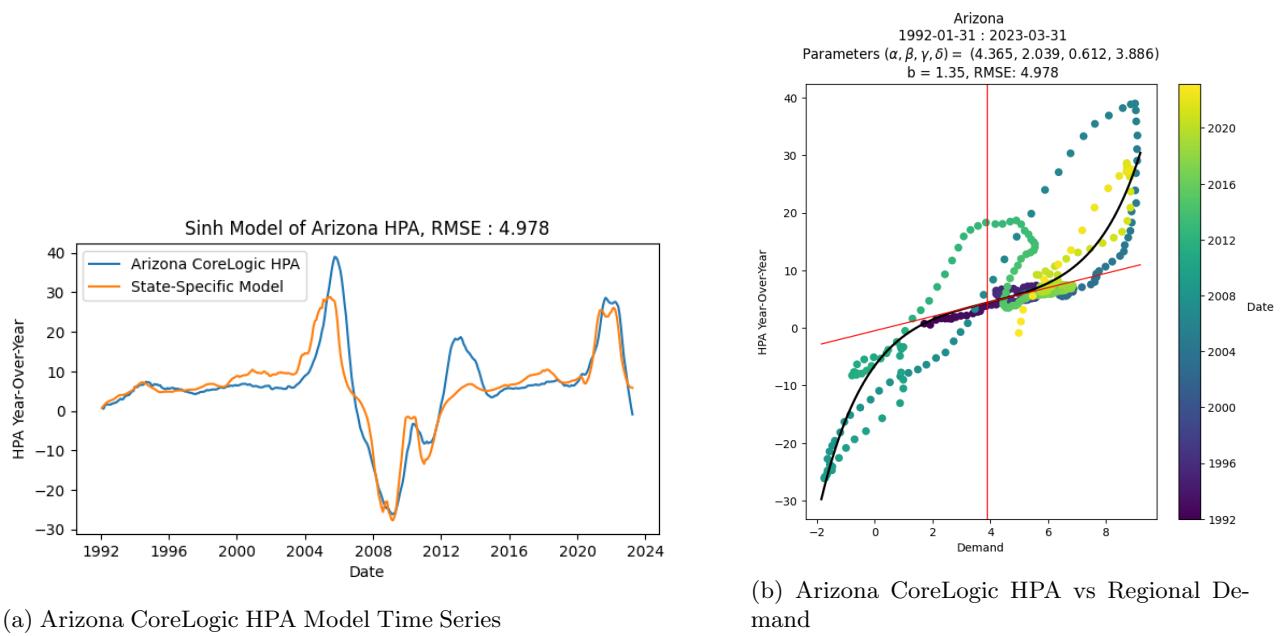


Figure 44: Arizona CoreLogic HPA Model

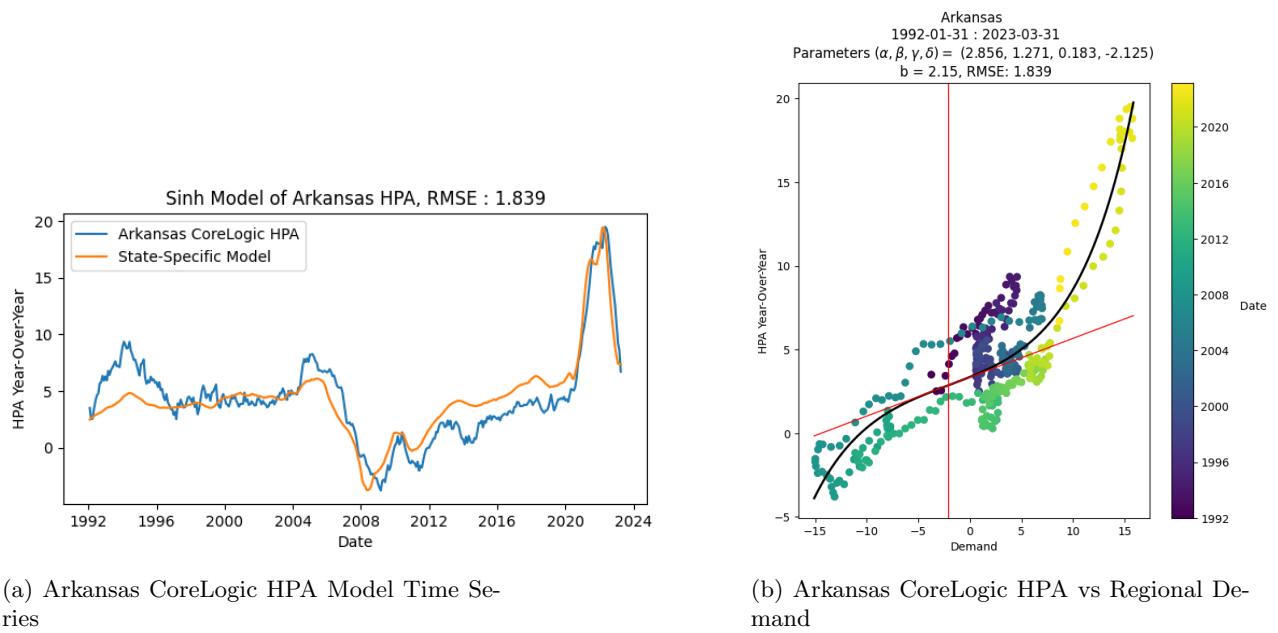


Figure 45: Arkansas CoreLogic HPA Model

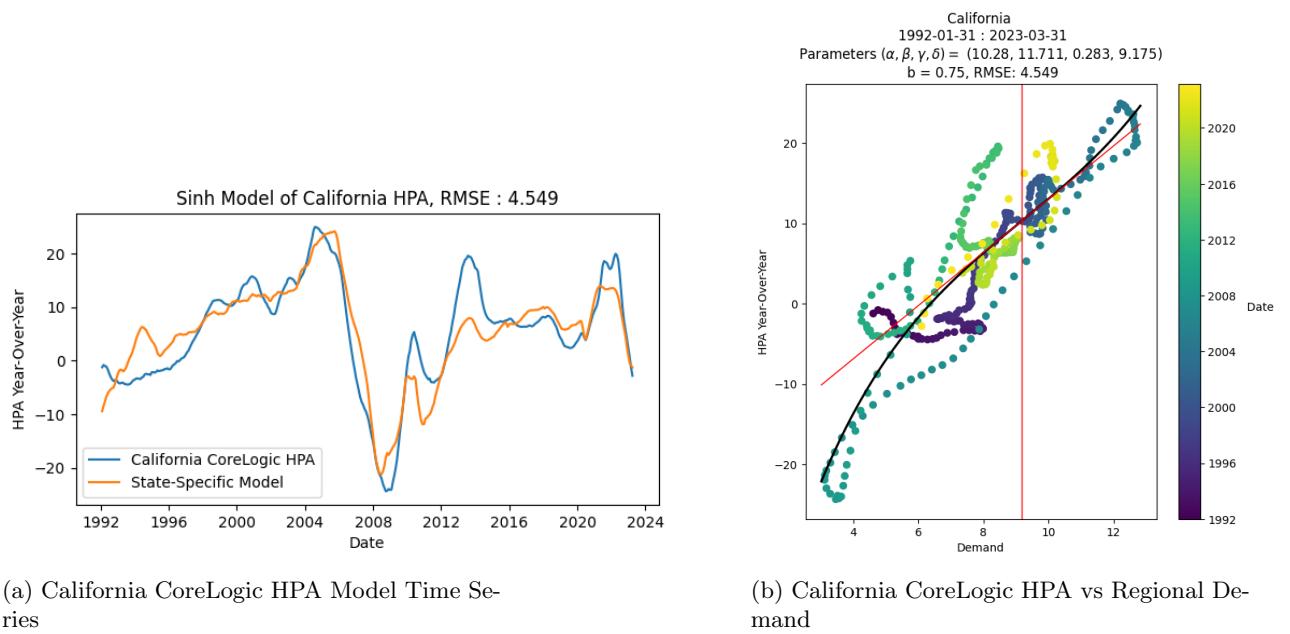


Figure 46: California CoreLogic HPA Model

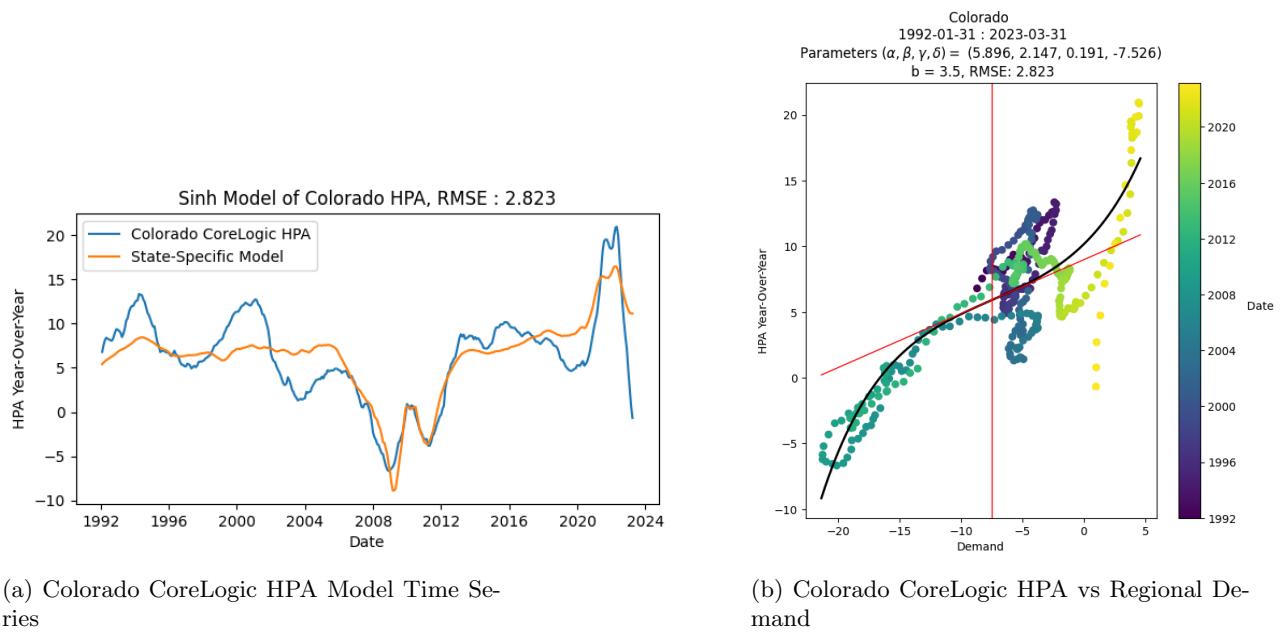


Figure 47: Colorado CoreLogic HPA Model

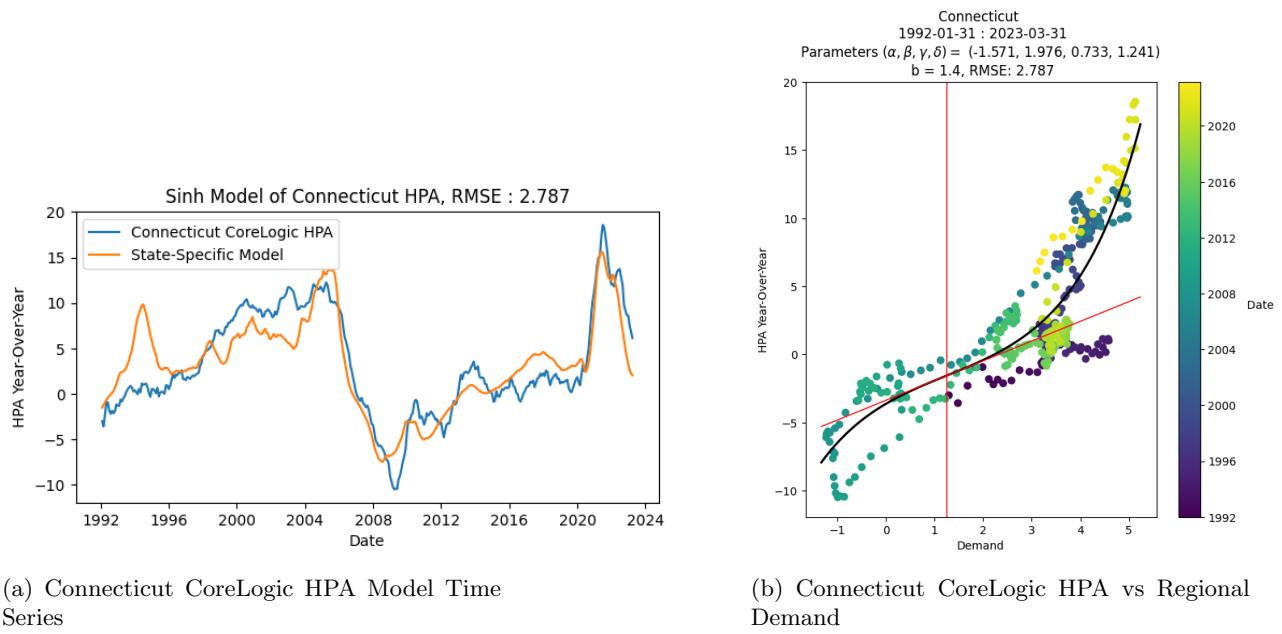


Figure 48: Connecticut CoreLogic HPA Model

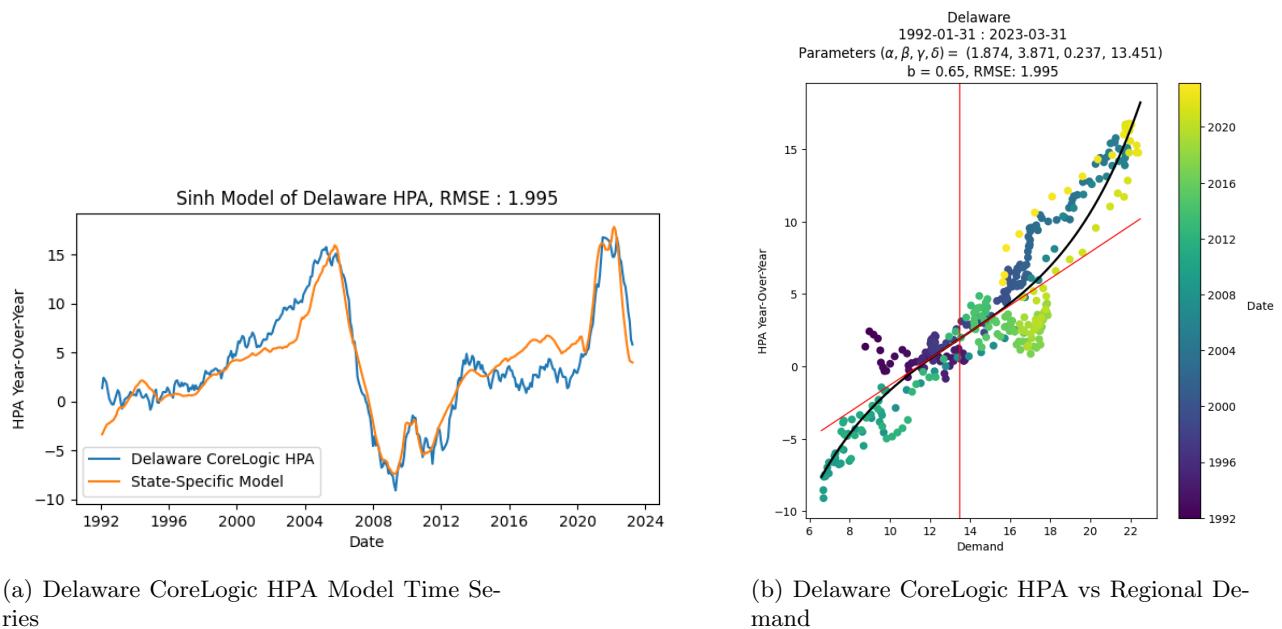


Figure 49: Delaware CoreLogic HPA Model

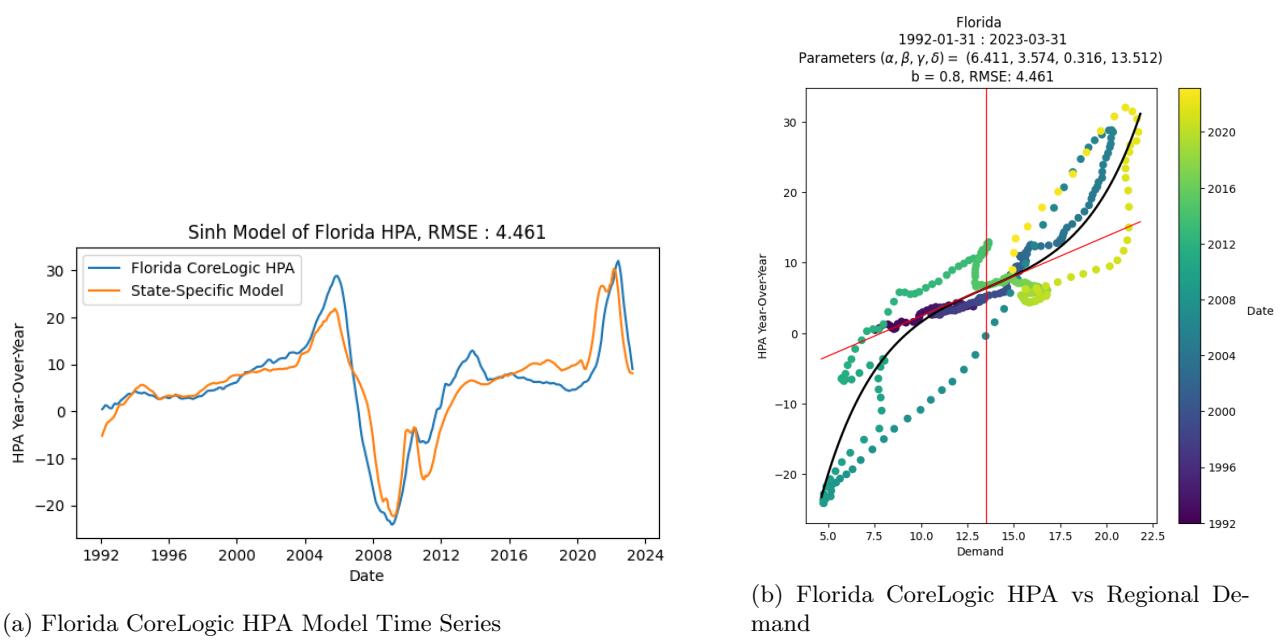


Figure 50: Florida CoreLogic HPA Model

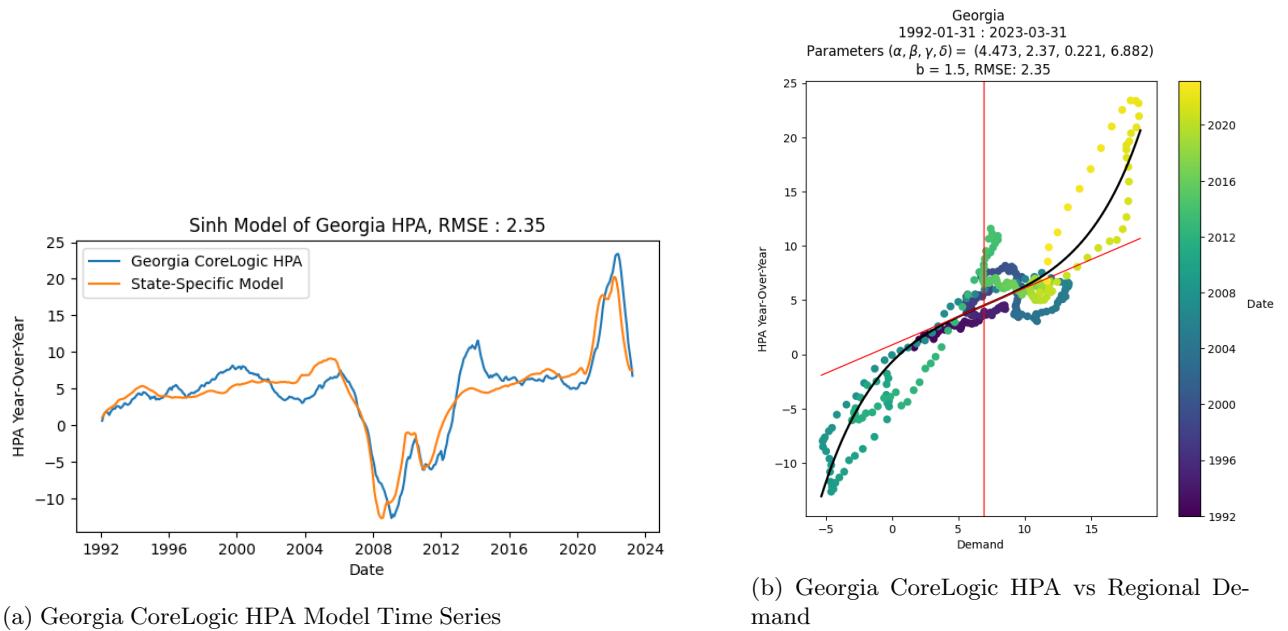


Figure 51: Georgia CoreLogic HPA Model

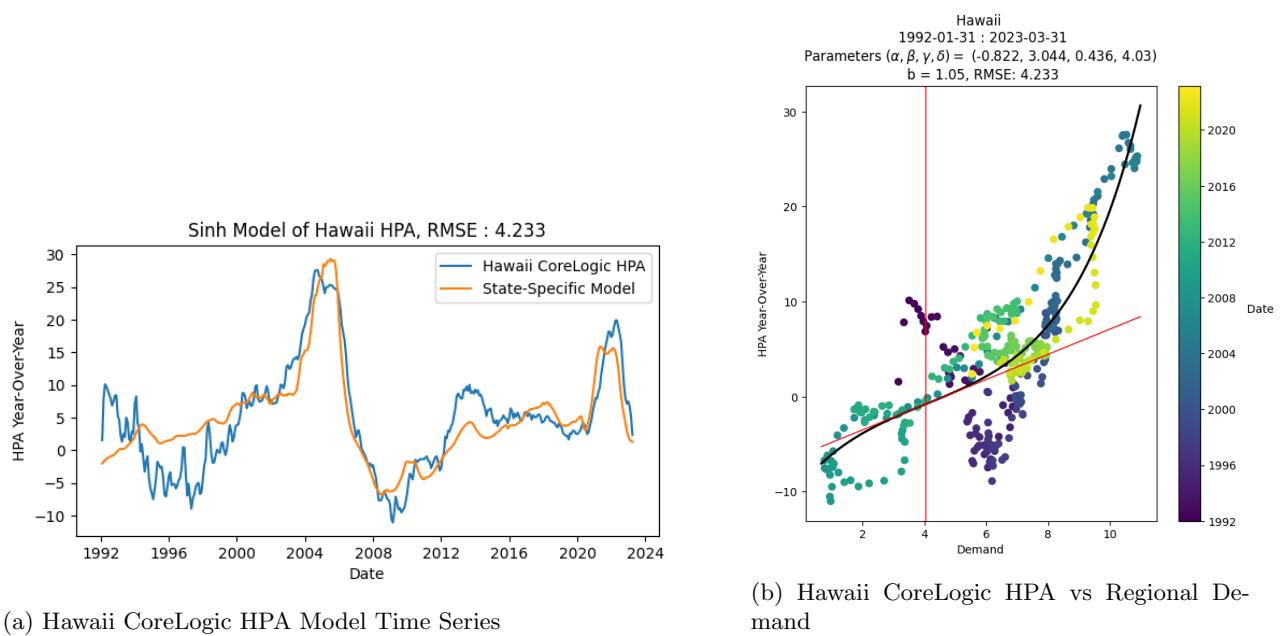


Figure 52: Hawaii CoreLogic HPA Model

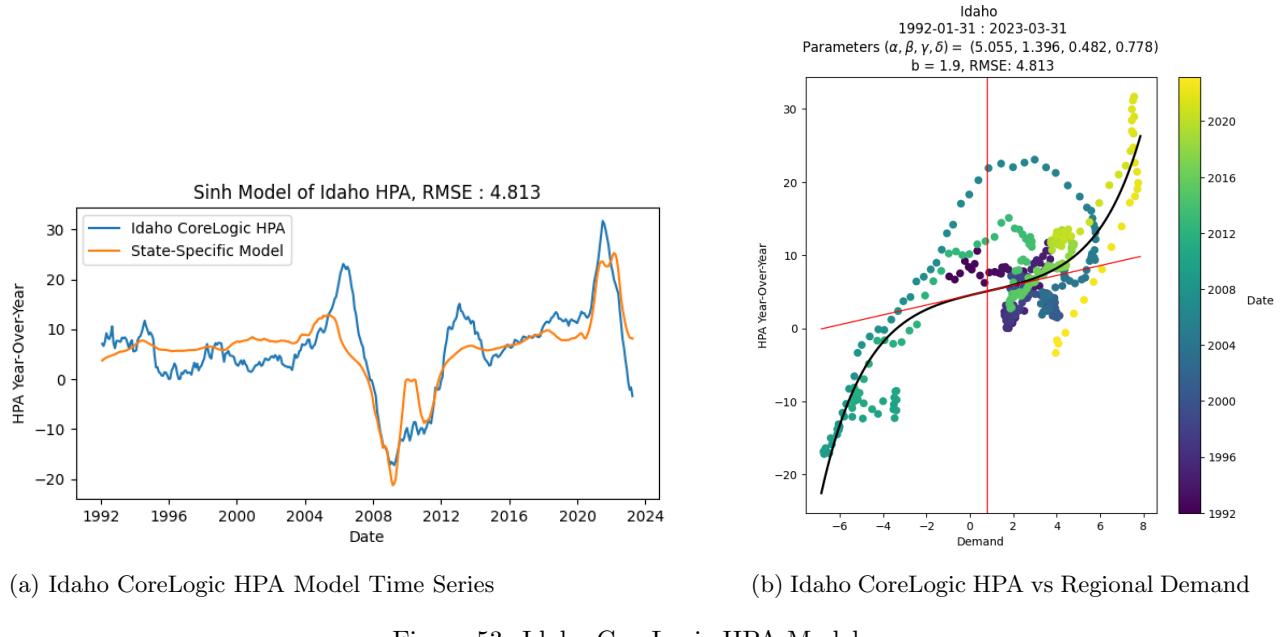


Figure 53: Idaho CoreLogic HPA Model

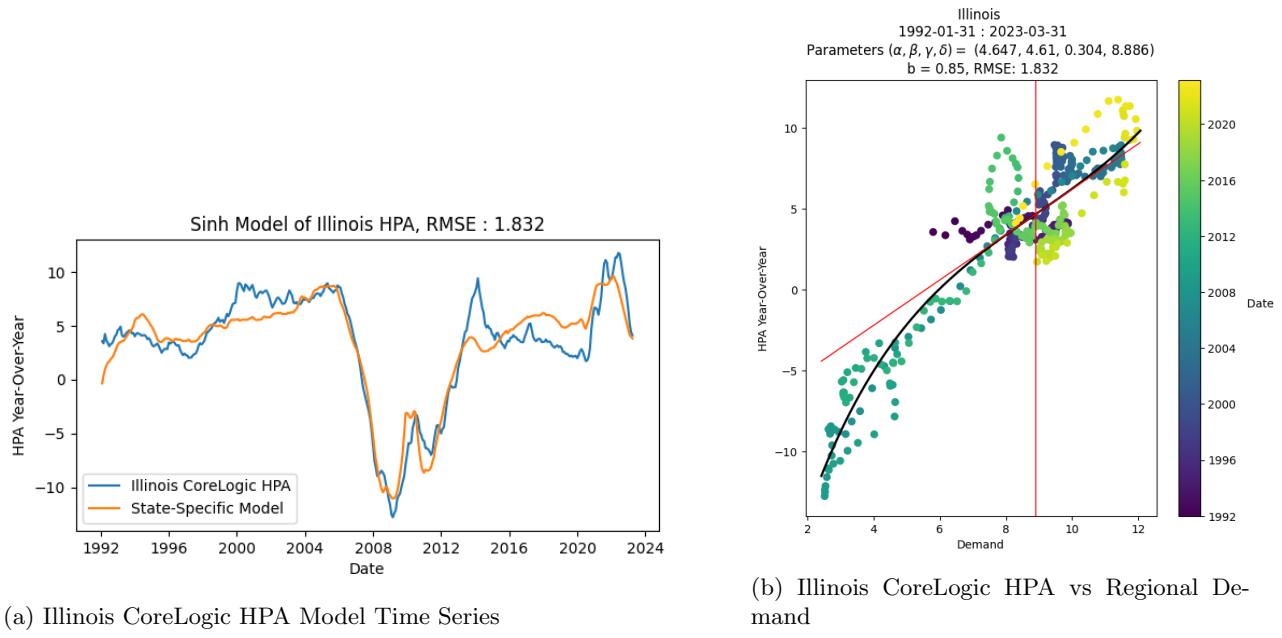


Figure 54: Illinois CoreLogic HPA Model

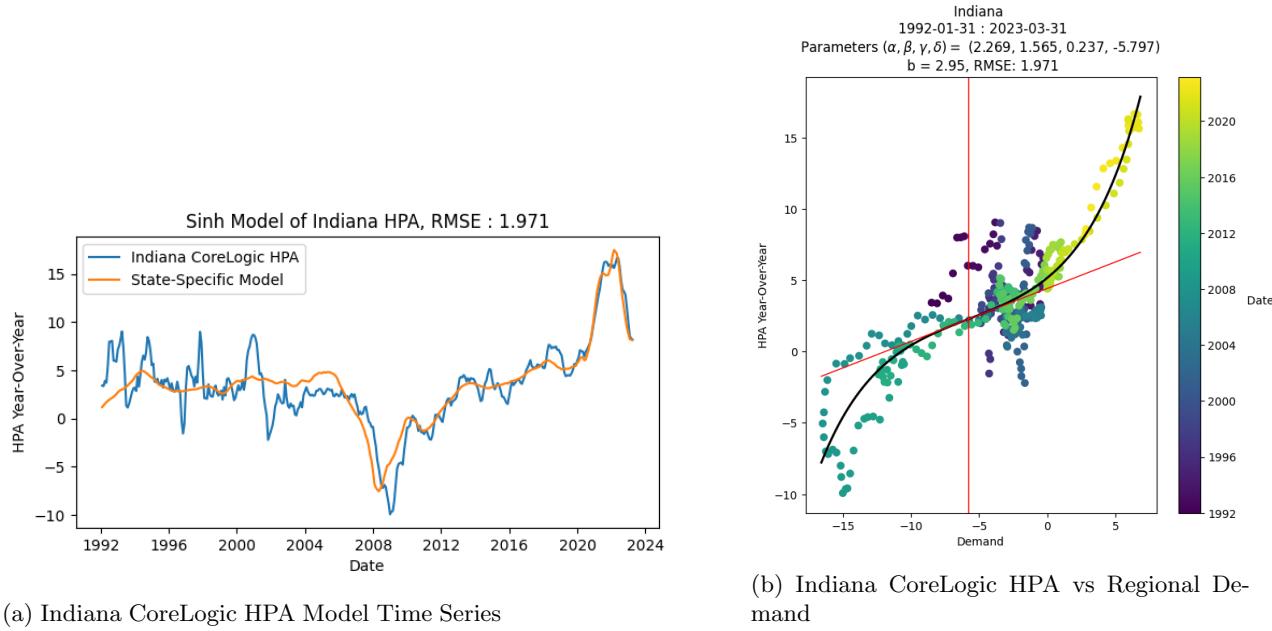


Figure 55: Indiana CoreLogic HPA Model

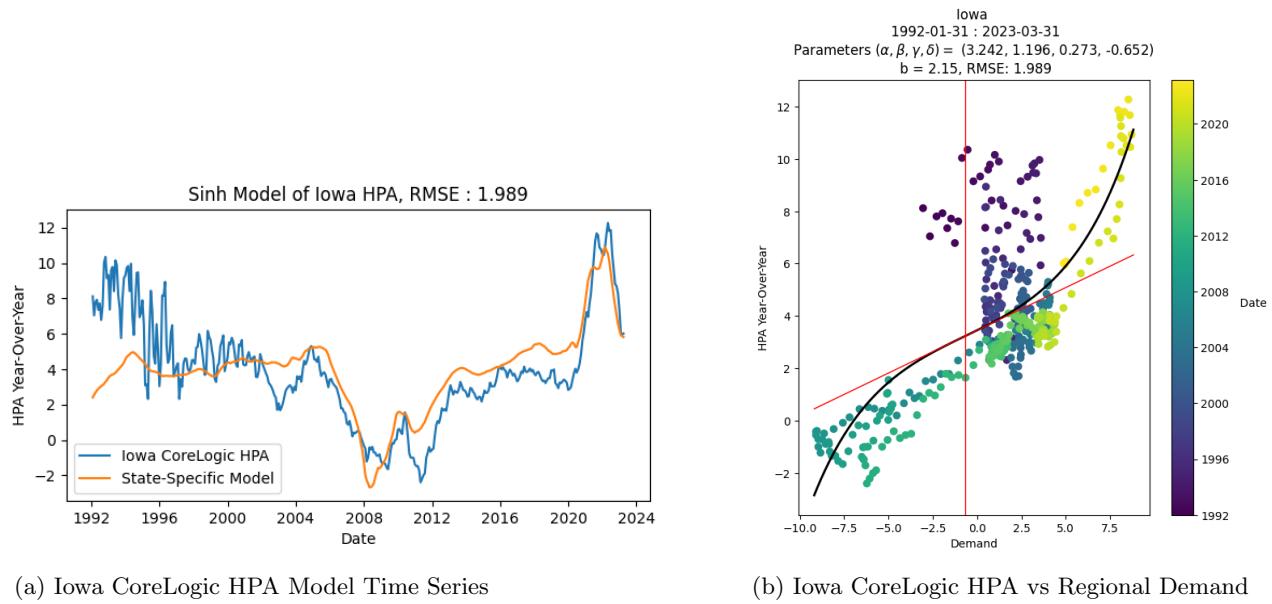


Figure 56: Iowa CoreLogic HPA Model

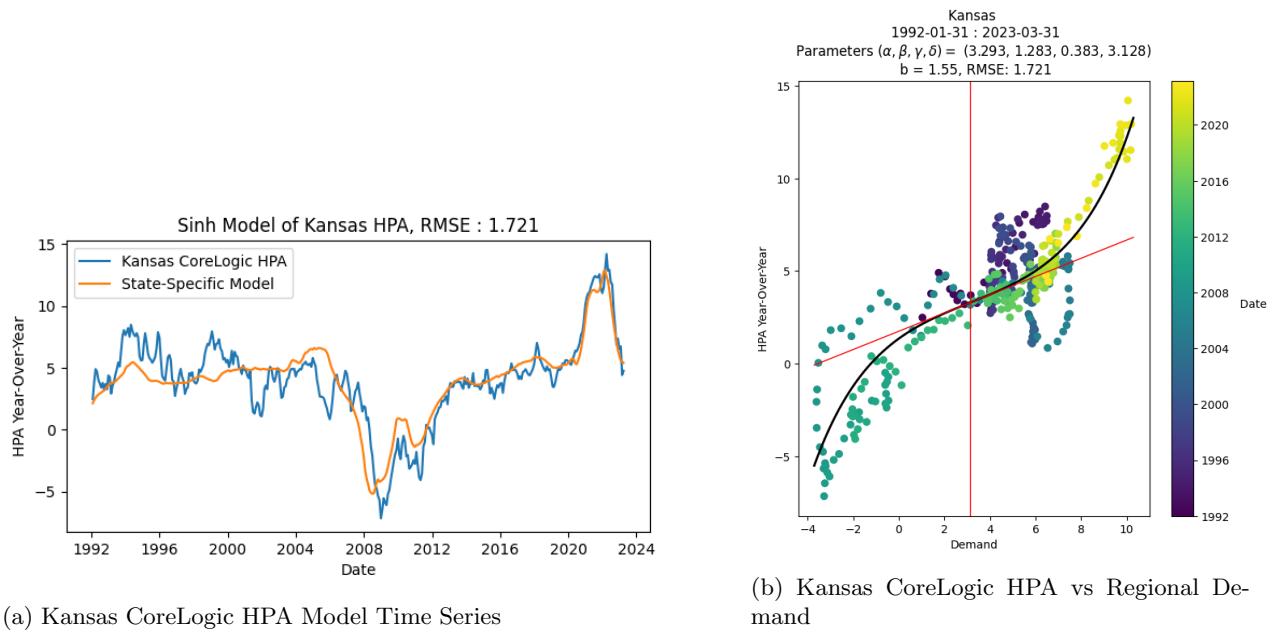


Figure 57: Kansas CoreLogic HPA Model

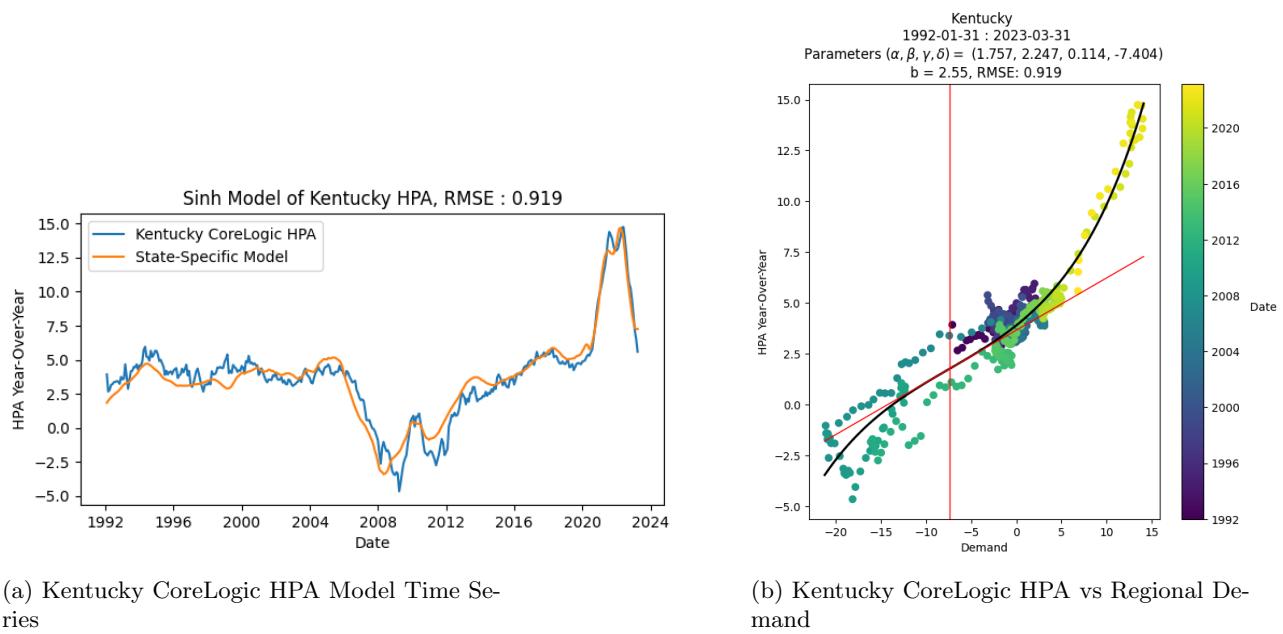


Figure 58: Kentucky CoreLogic HPA Model

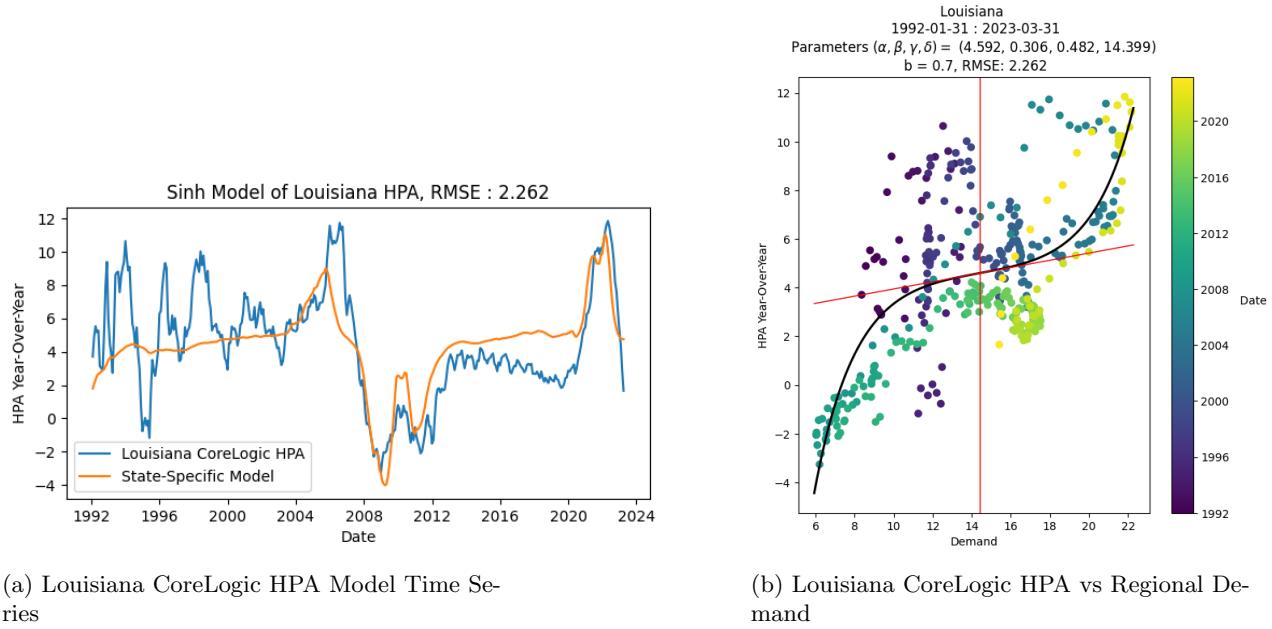


Figure 59: Louisiana CoreLogic HPA Model

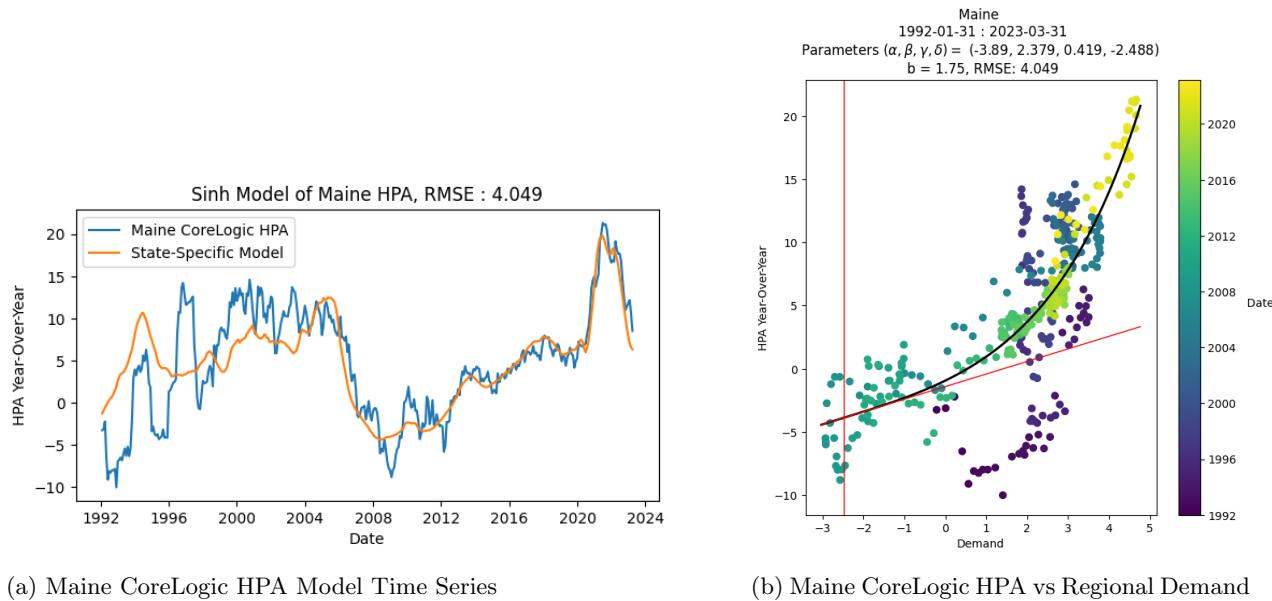


Figure 60: Maine CoreLogic HPA Model

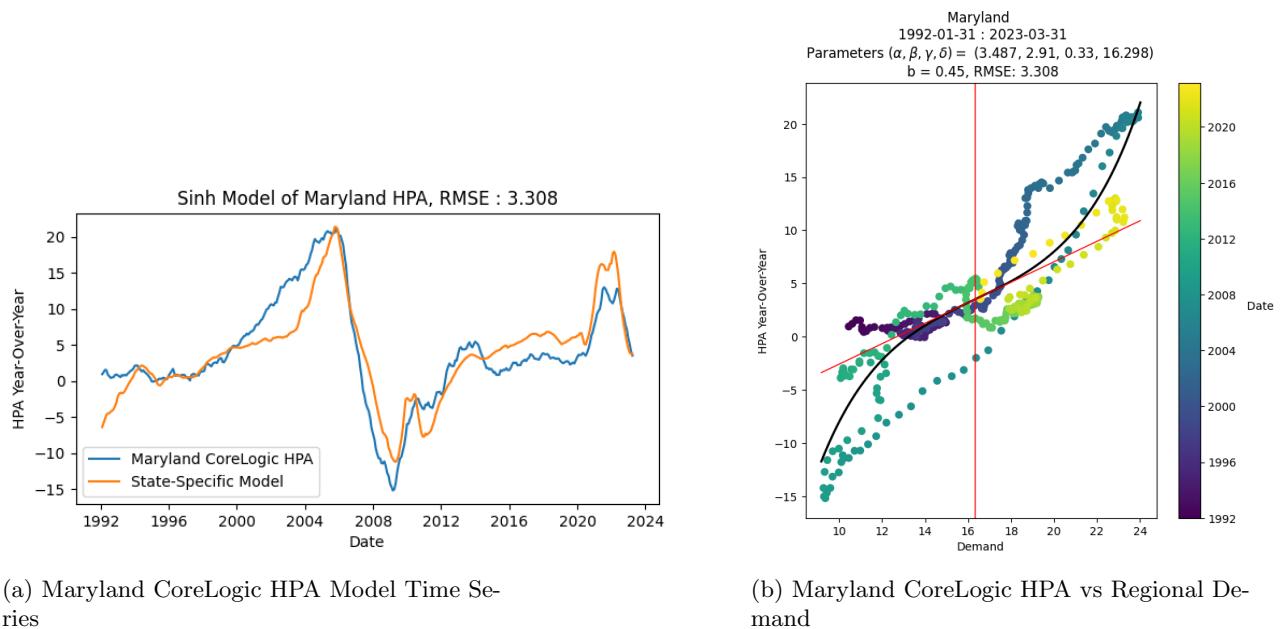


Figure 61: Maryland CoreLogic HPA Model

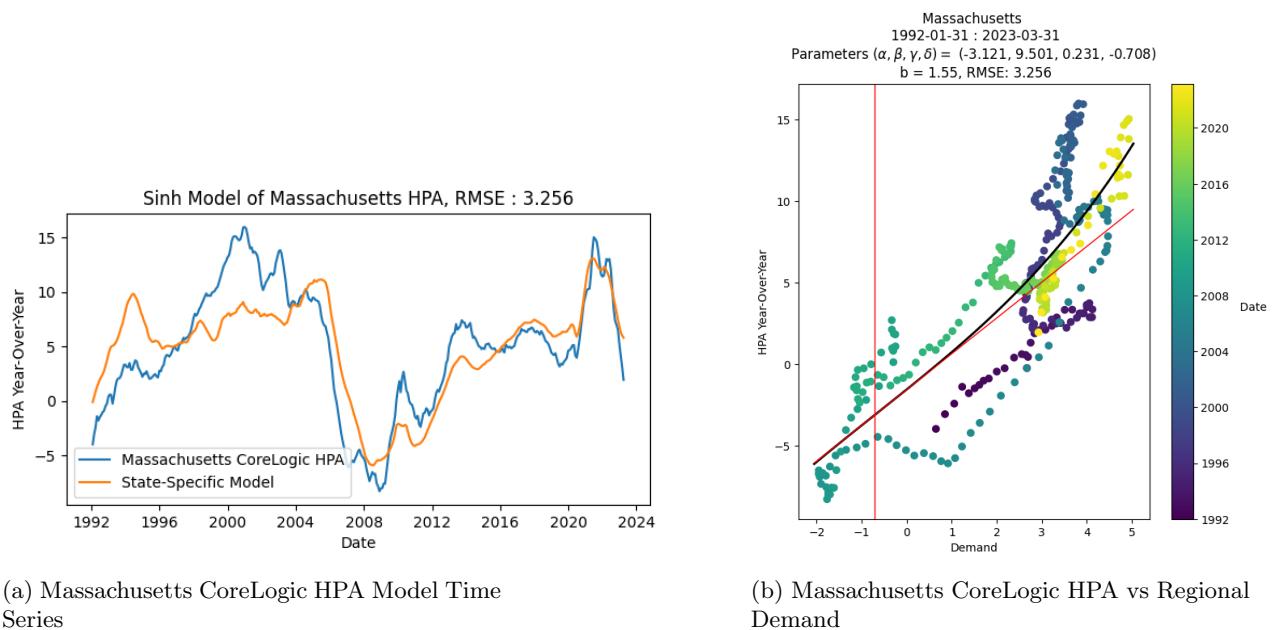


Figure 62: Massachusetts CoreLogic HPA Model

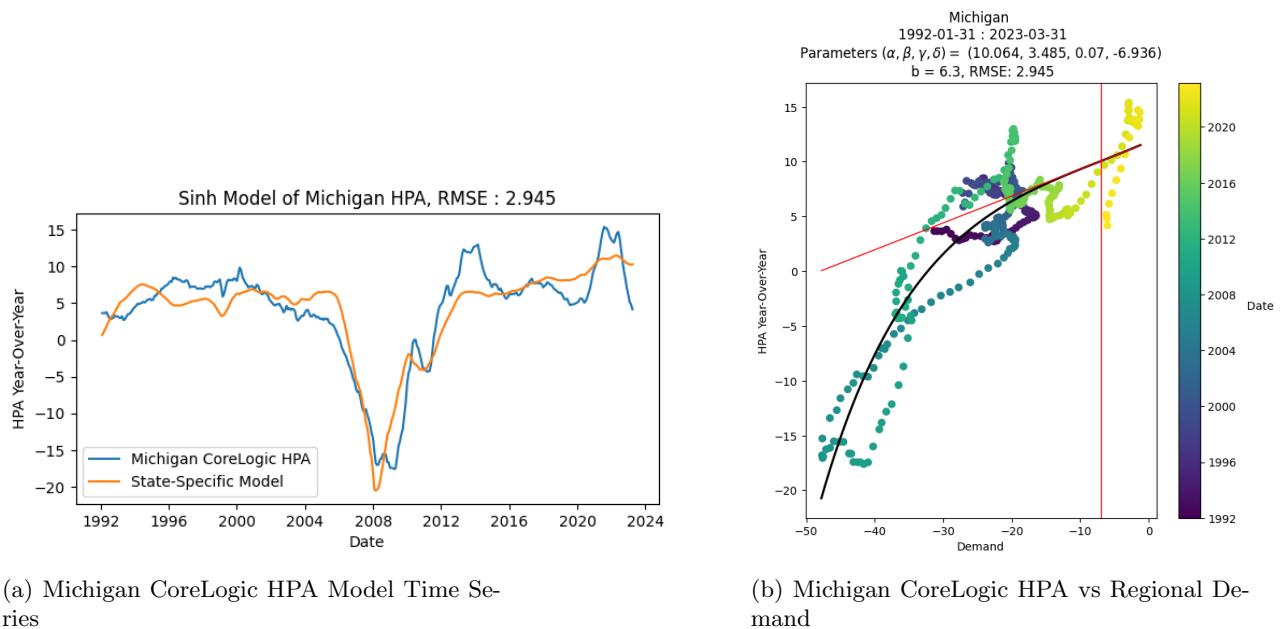


Figure 63: Michigan CoreLogic HPA Model

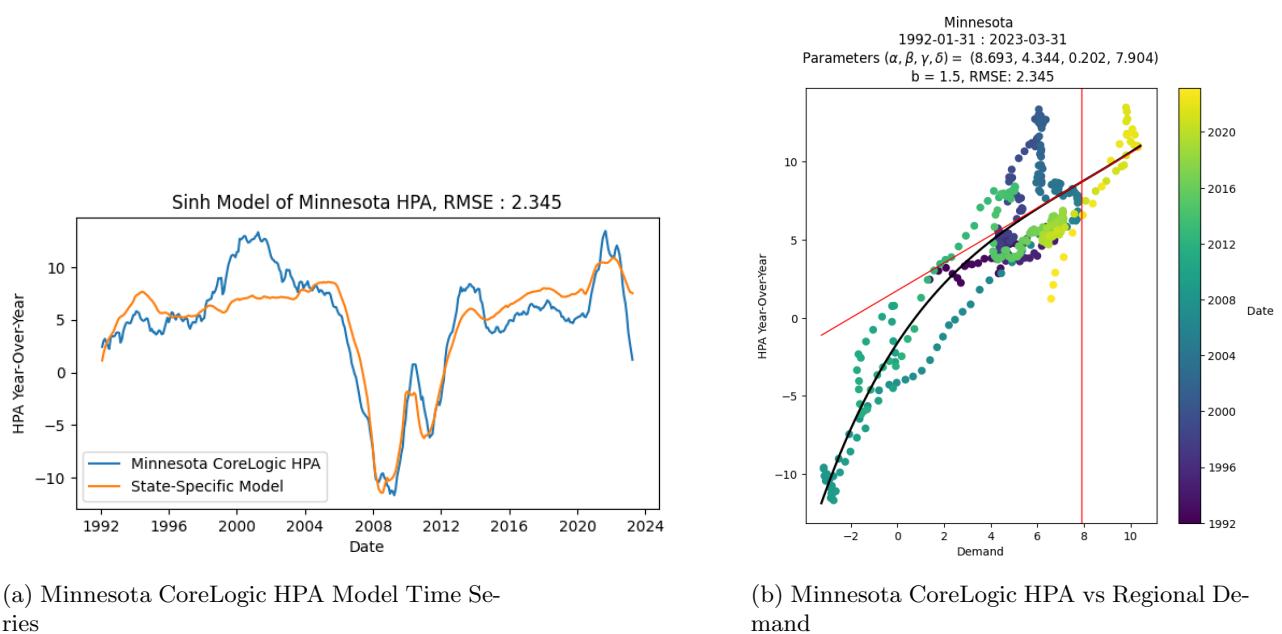


Figure 64: Minnesota CoreLogic HPA Model

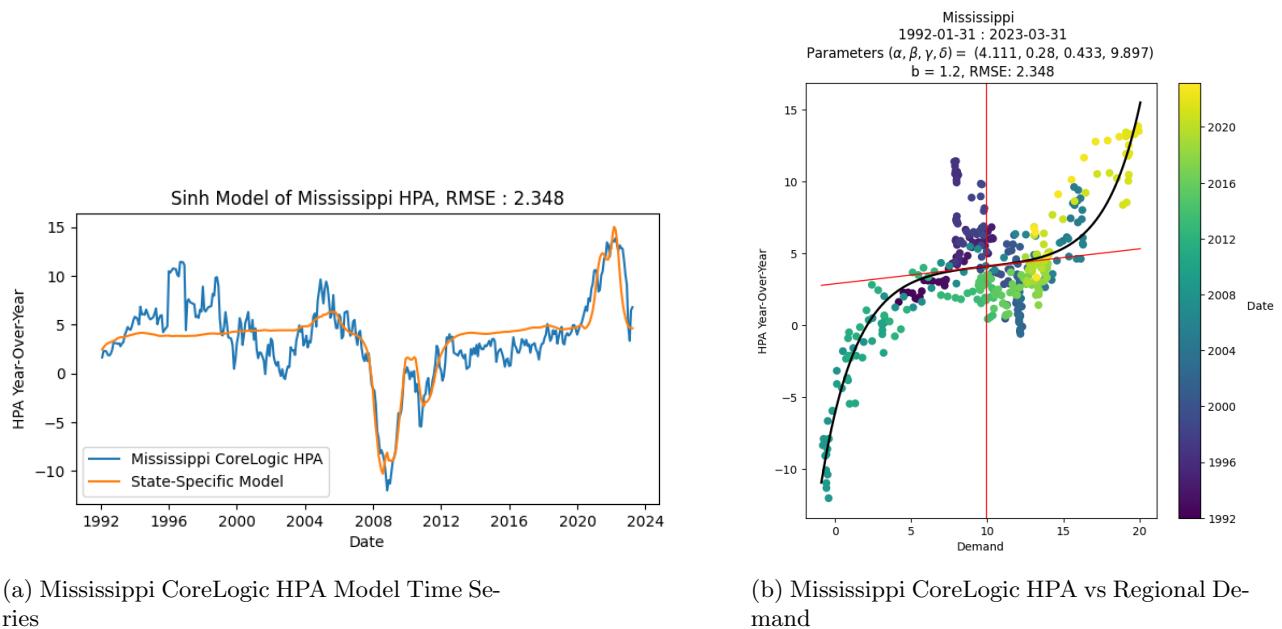


Figure 65: Mississippi CoreLogic HPA Model

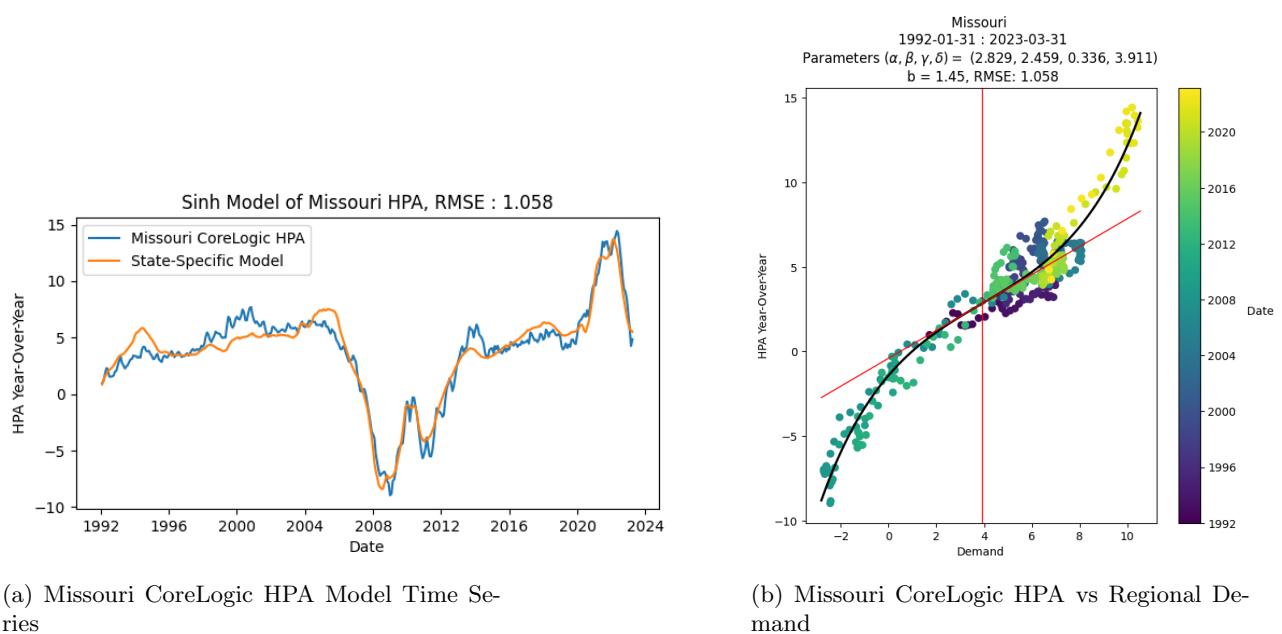


Figure 66: Missouri CoreLogic HPA Model

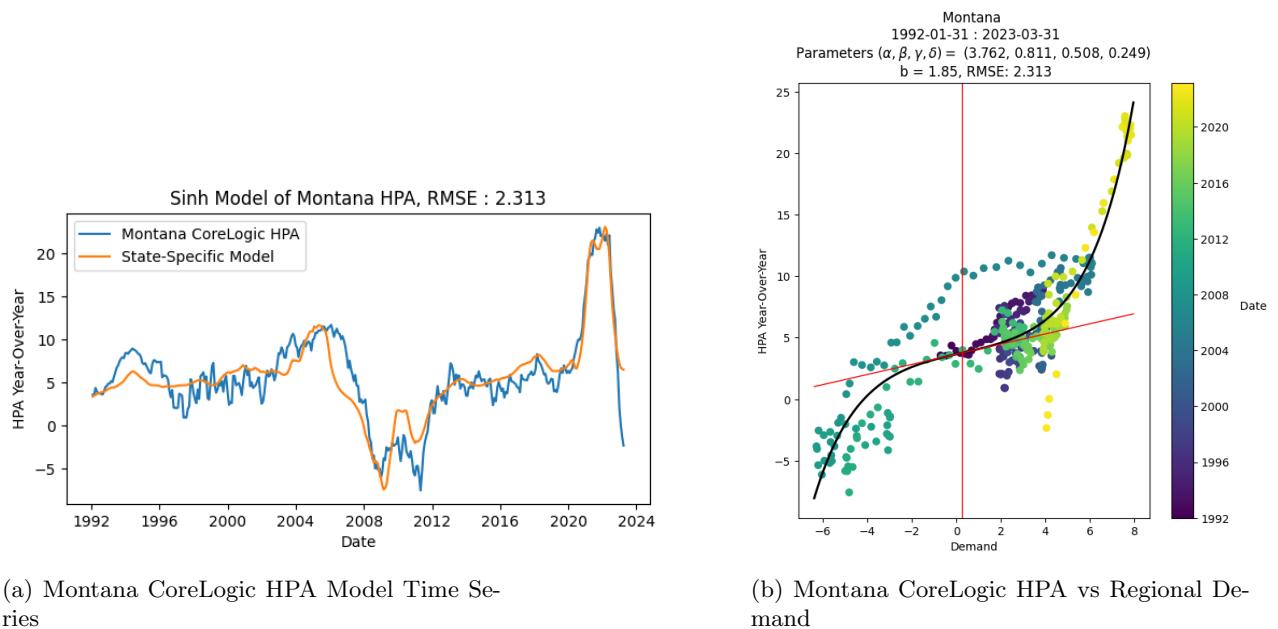


Figure 67: Montana CoreLogic HPA Model

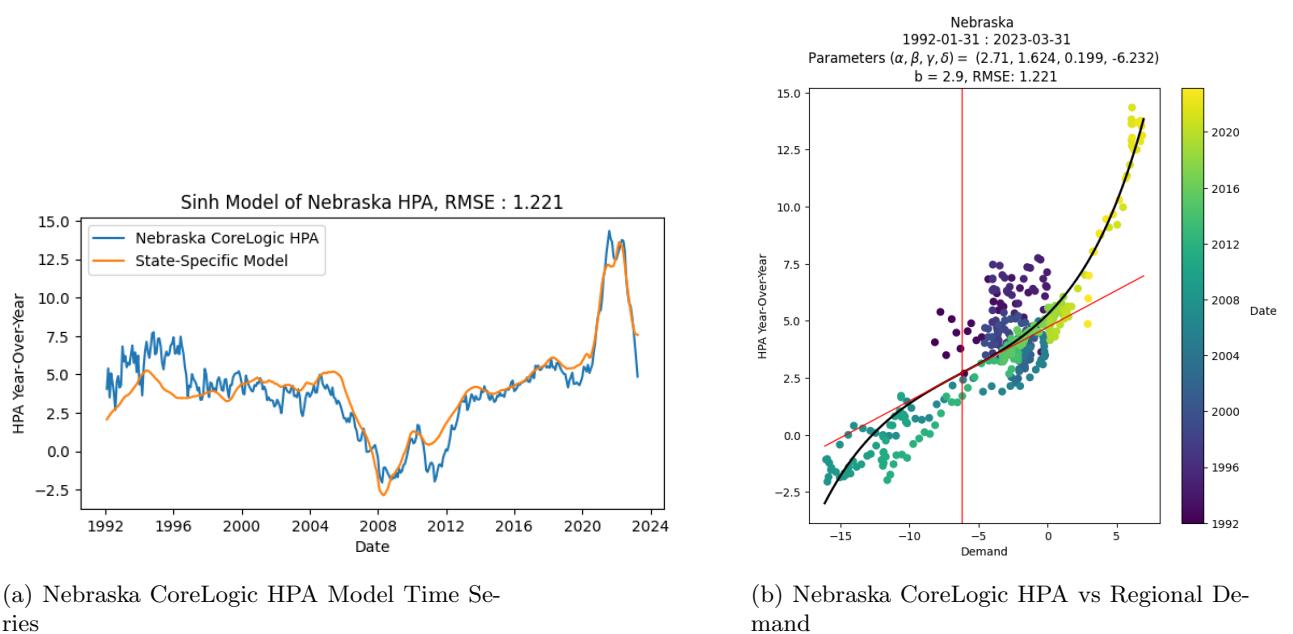
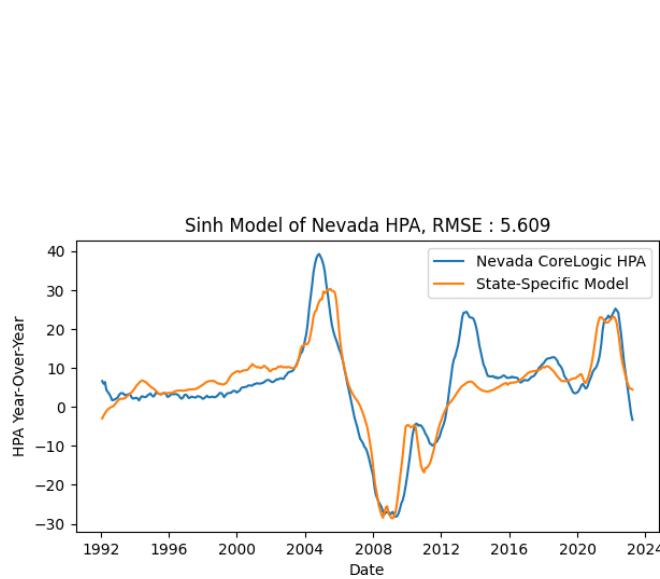
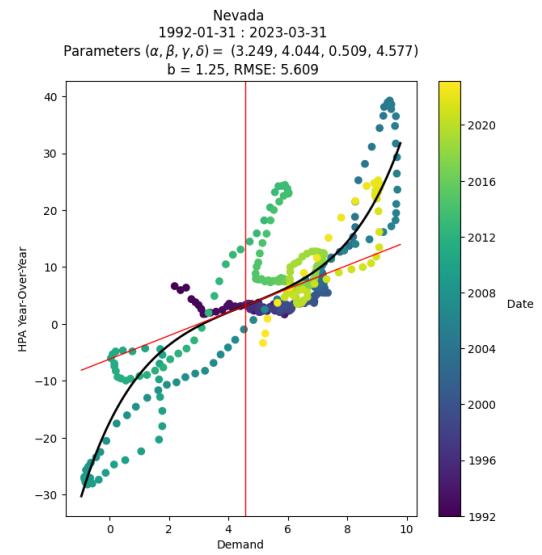


Figure 68: Nebraska CoreLogic HPA Model

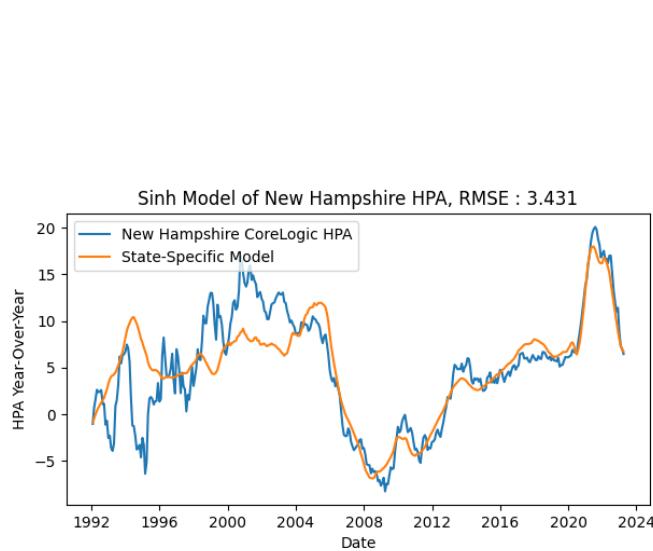


(a) Nevada CoreLogic HPA Model Time Series

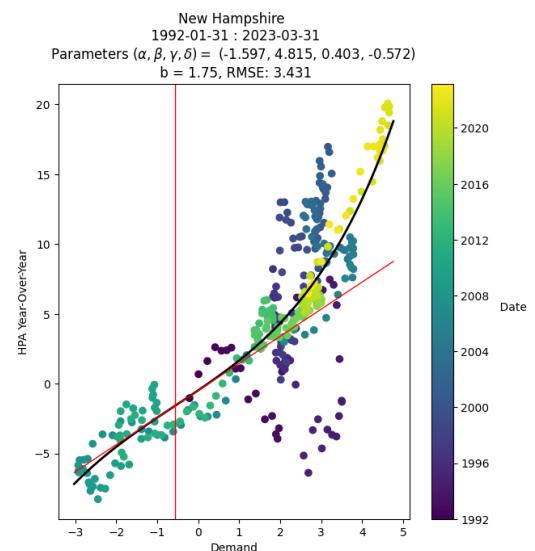


(b) Nevada CoreLogic HPA vs Regional Demand

Figure 69: Nevada CoreLogic HPA Model



(a) New Hampshire CoreLogic HPA Model Time Series



(b) New Hampshire CoreLogic HPA vs Regional Demand

Figure 70: New Hampshire CoreLogic HPA Model

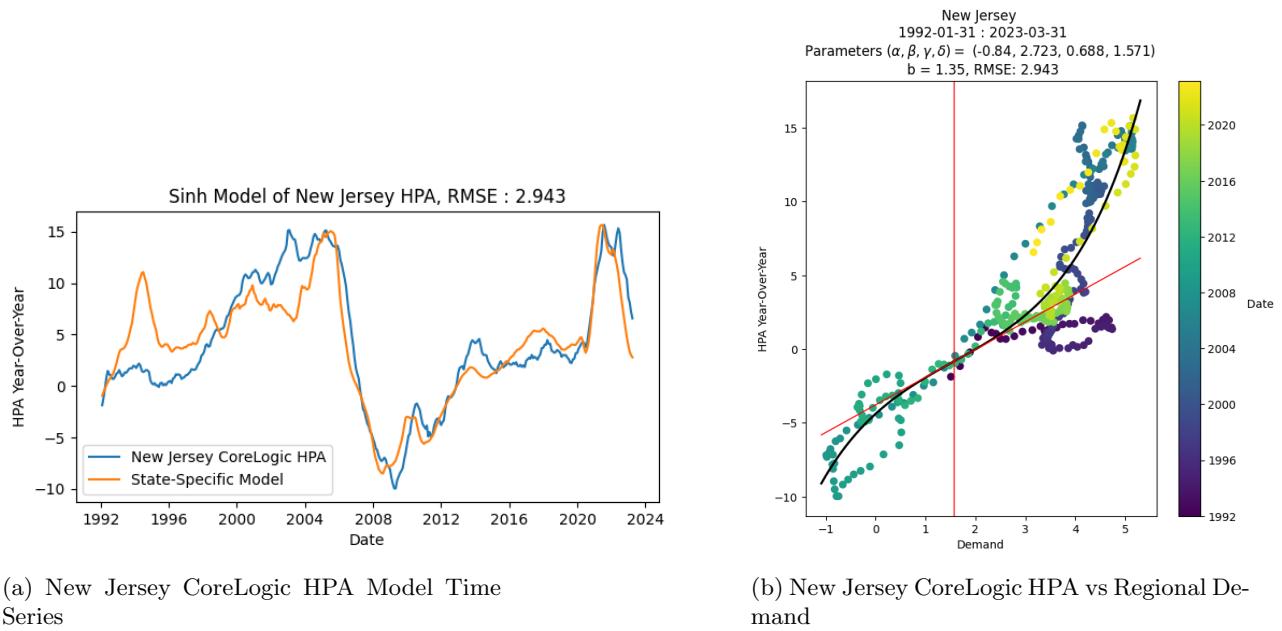


Figure 71: New Jersey CoreLogic HPA Model

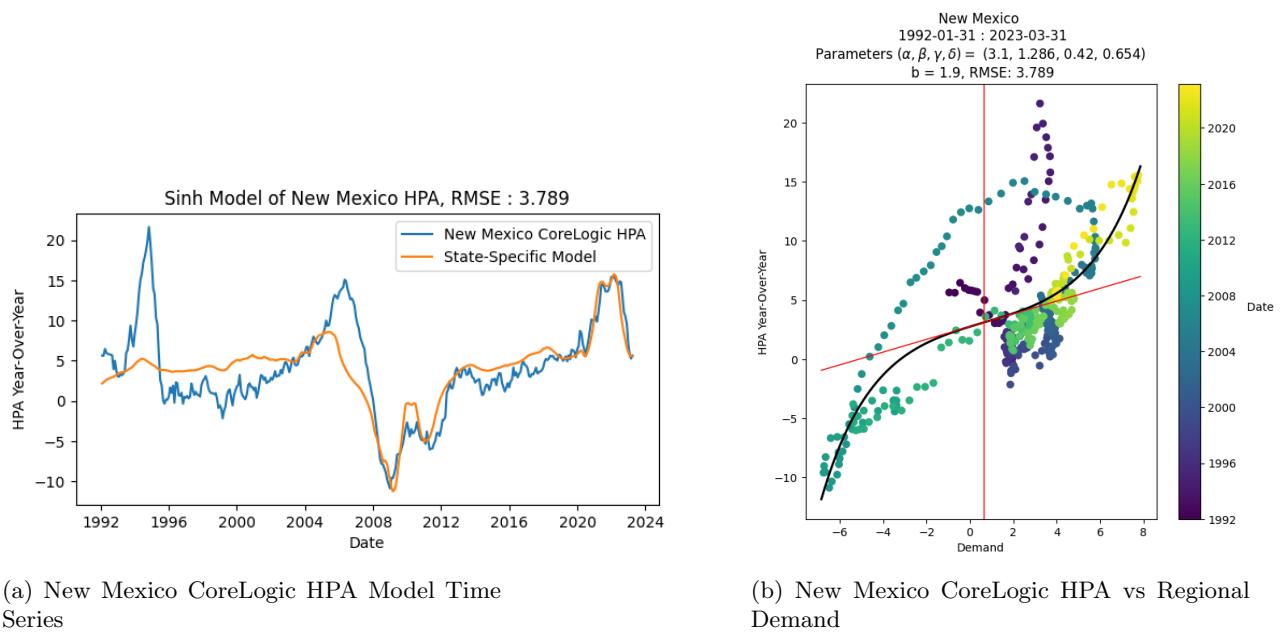


Figure 72: New Mexico CoreLogic HPA Model

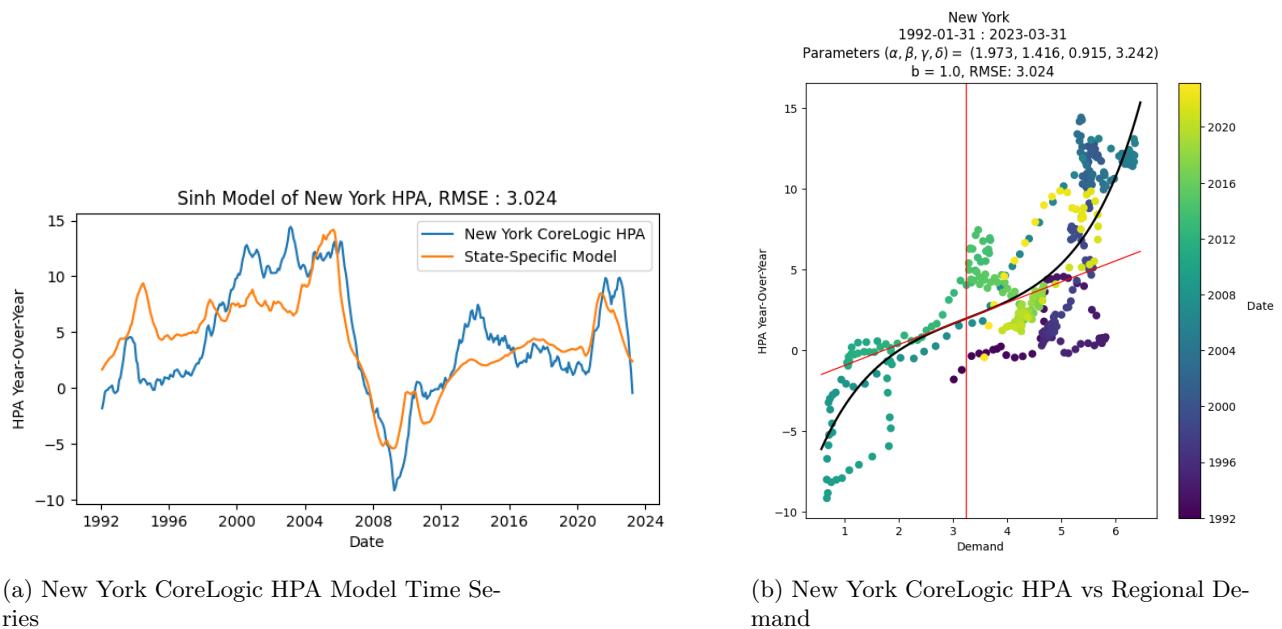


Figure 73: New York CoreLogic HPA Model

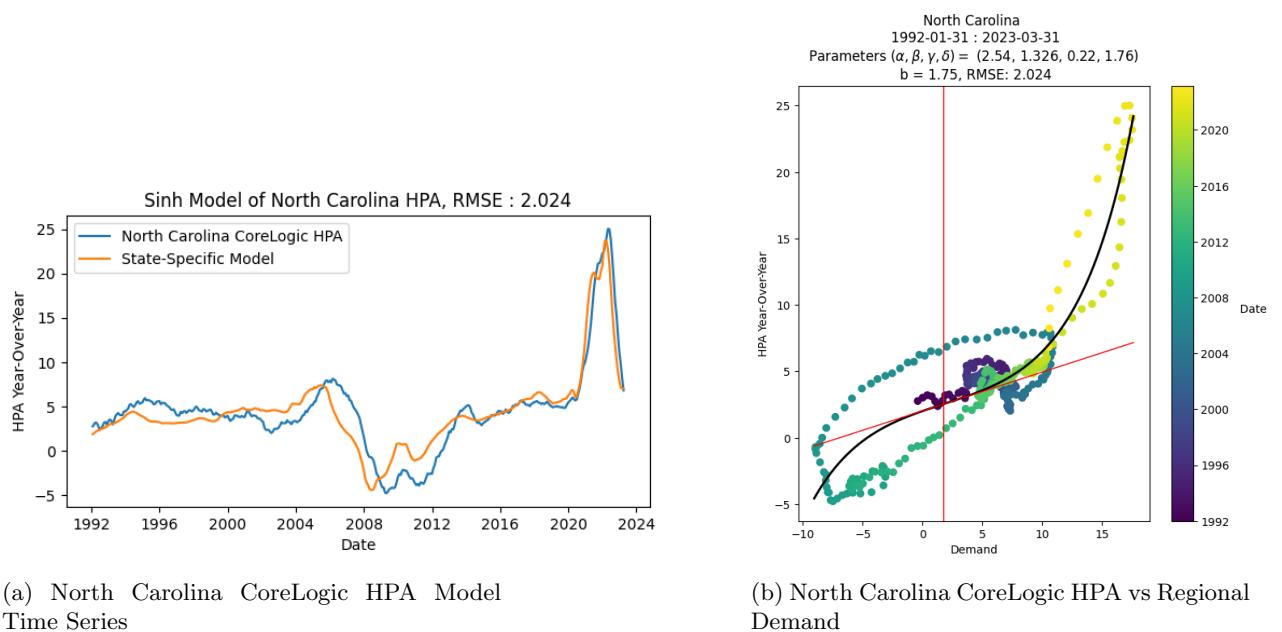


Figure 74: North Carolina CoreLogic HPA Model

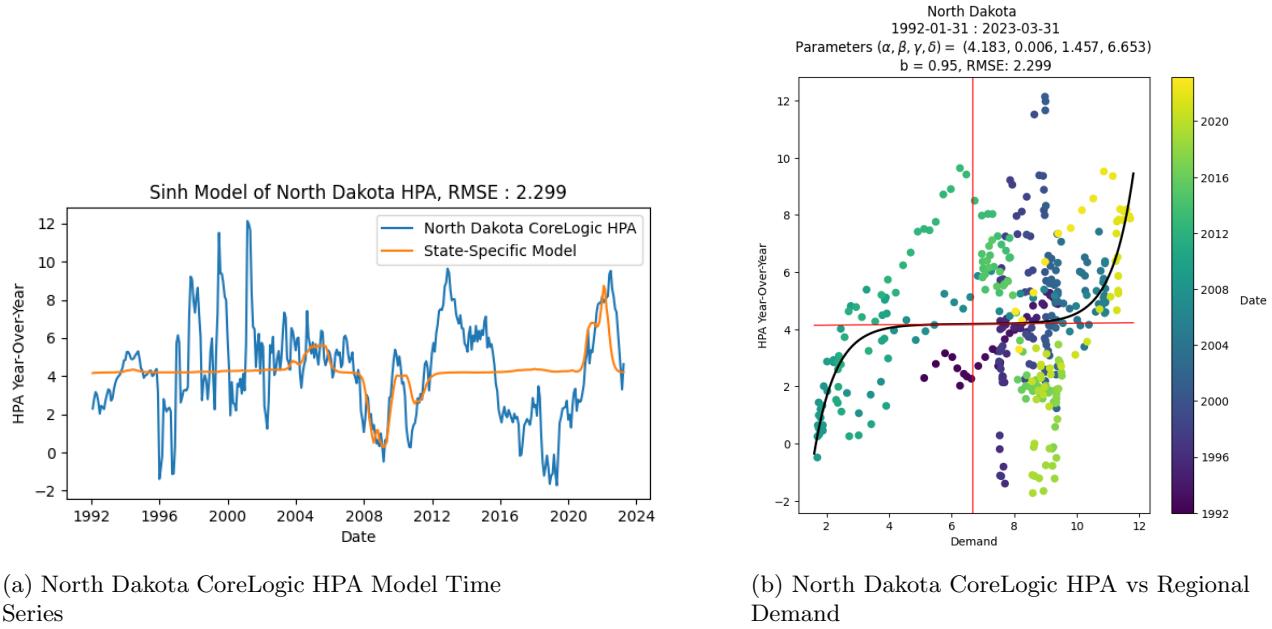


Figure 75: North Dakota CoreLogic HPA Model

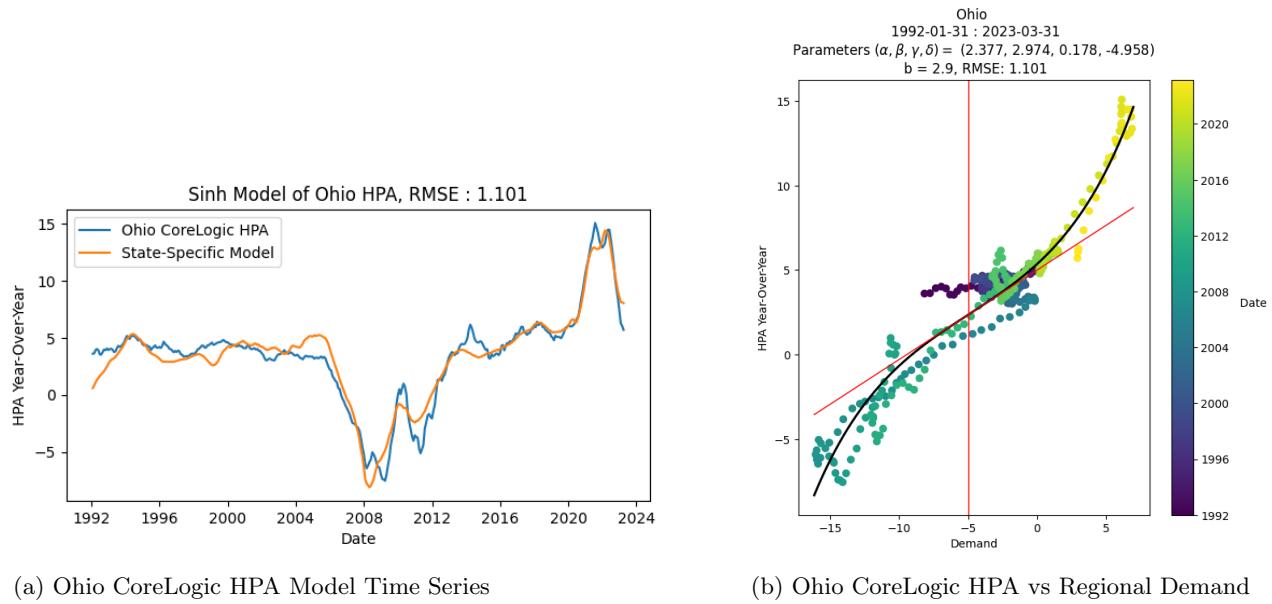


Figure 76: Ohio CoreLogic HPA Model

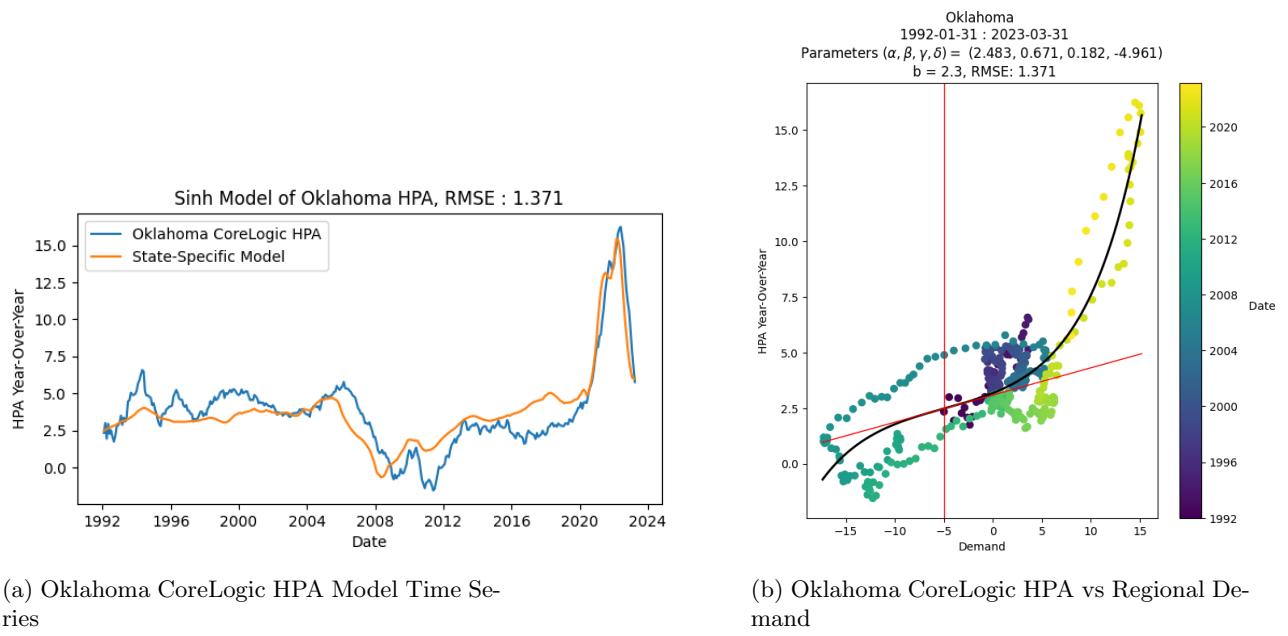


Figure 77: Oklahoma CoreLogic HPA Model

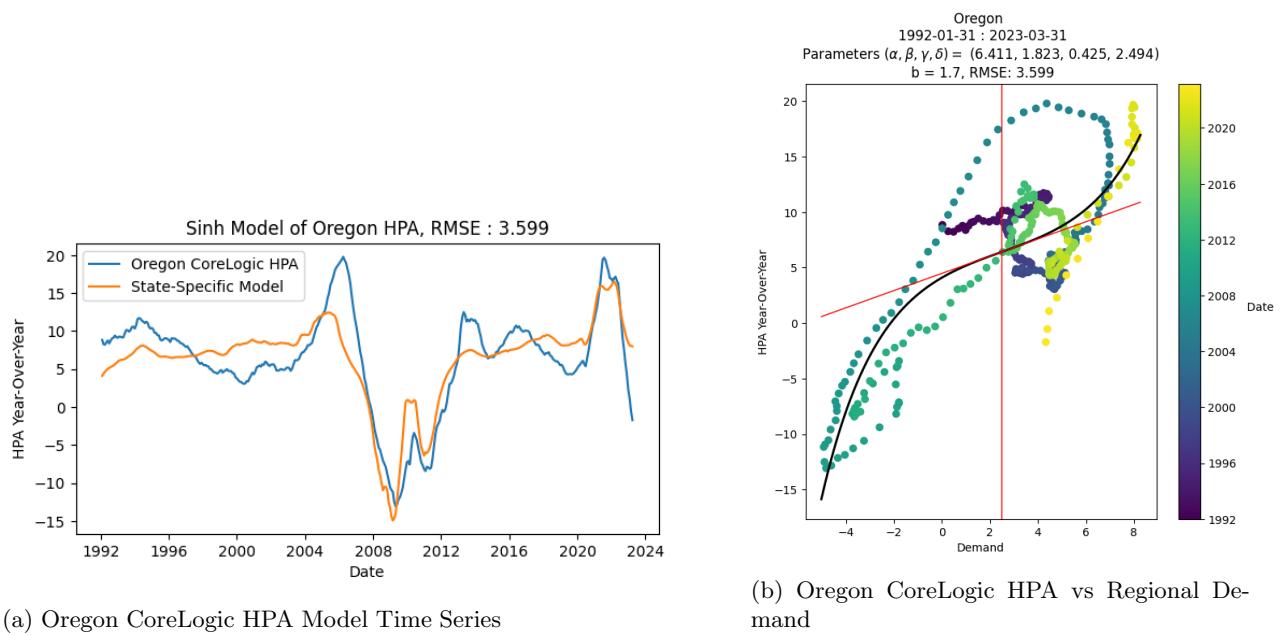


Figure 78: Oregon CoreLogic HPA Model

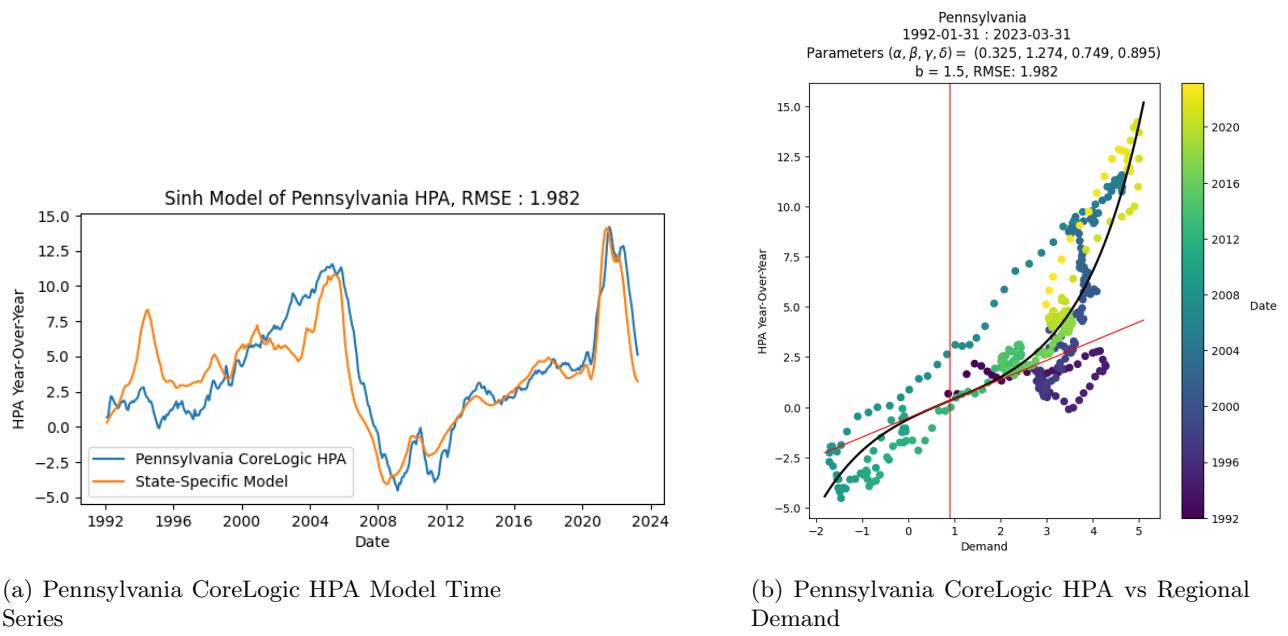


Figure 79: Pennsylvania CoreLogic HPA Model

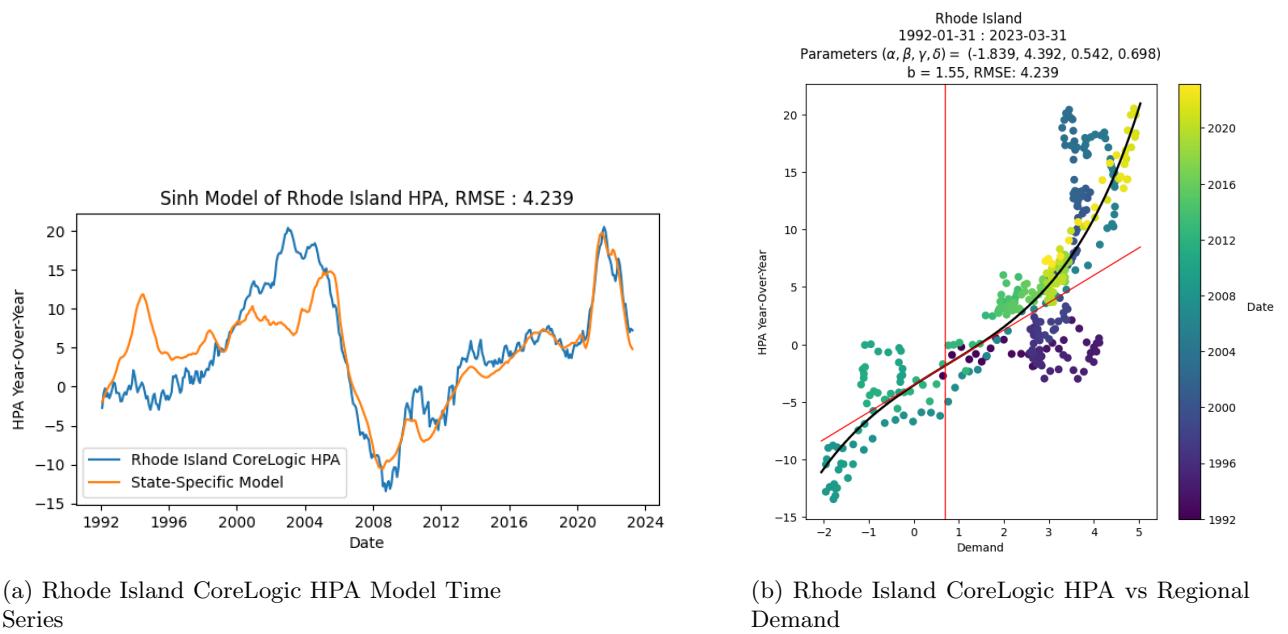


Figure 80: Rhode Island CoreLogic HPA Model

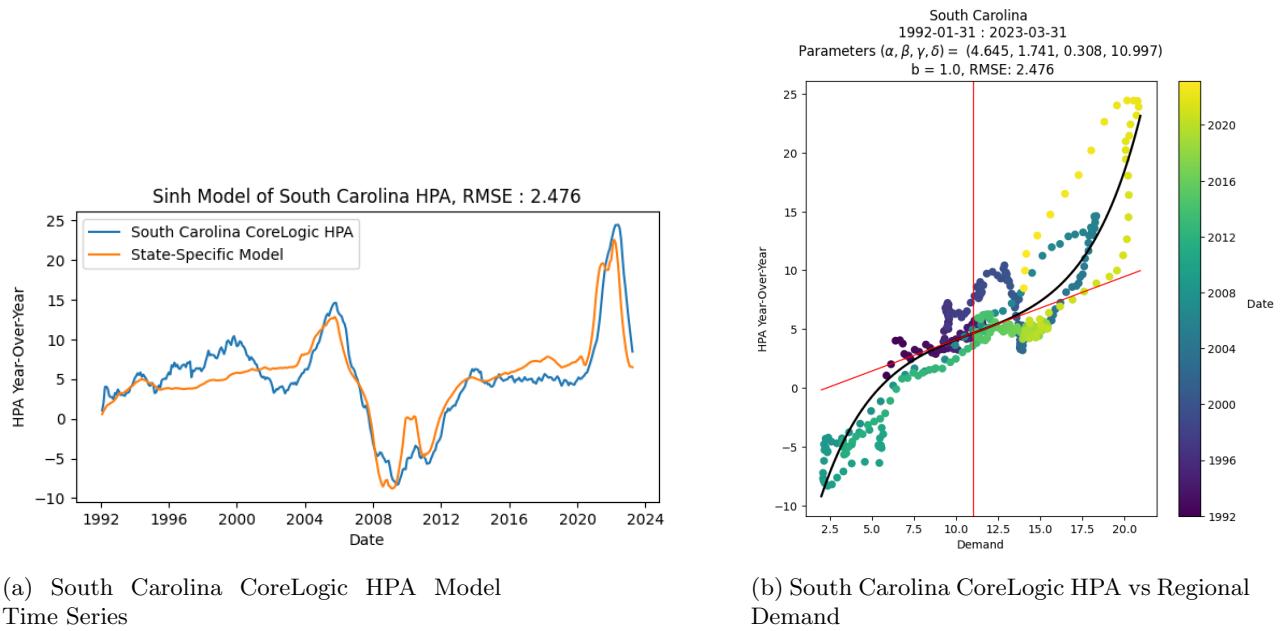


Figure 81: South Carolina CoreLogic HPA Model

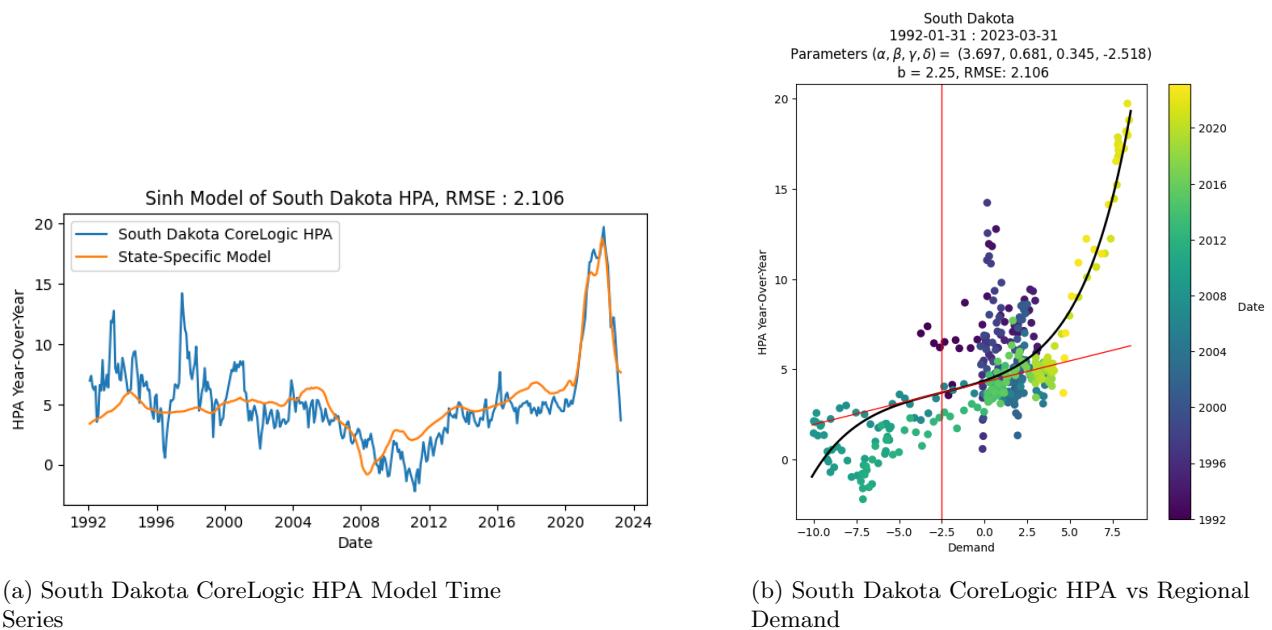


Figure 82: South Dakota CoreLogic HPA Model

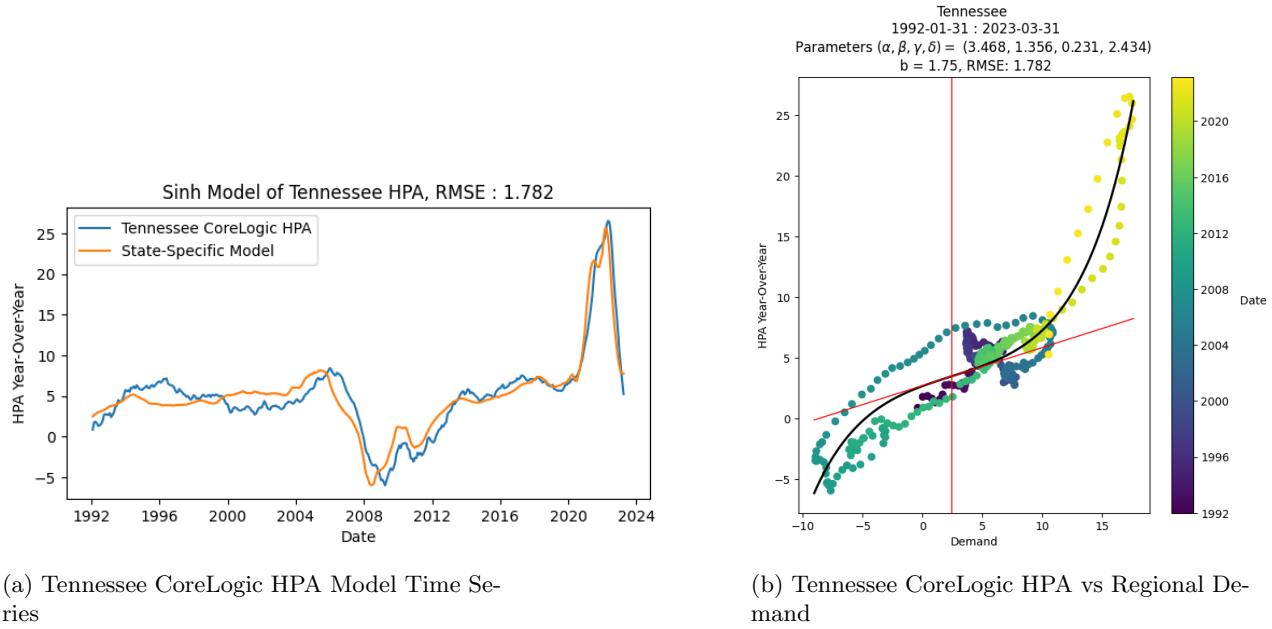


Figure 83: Tennessee CoreLogic HPA Model

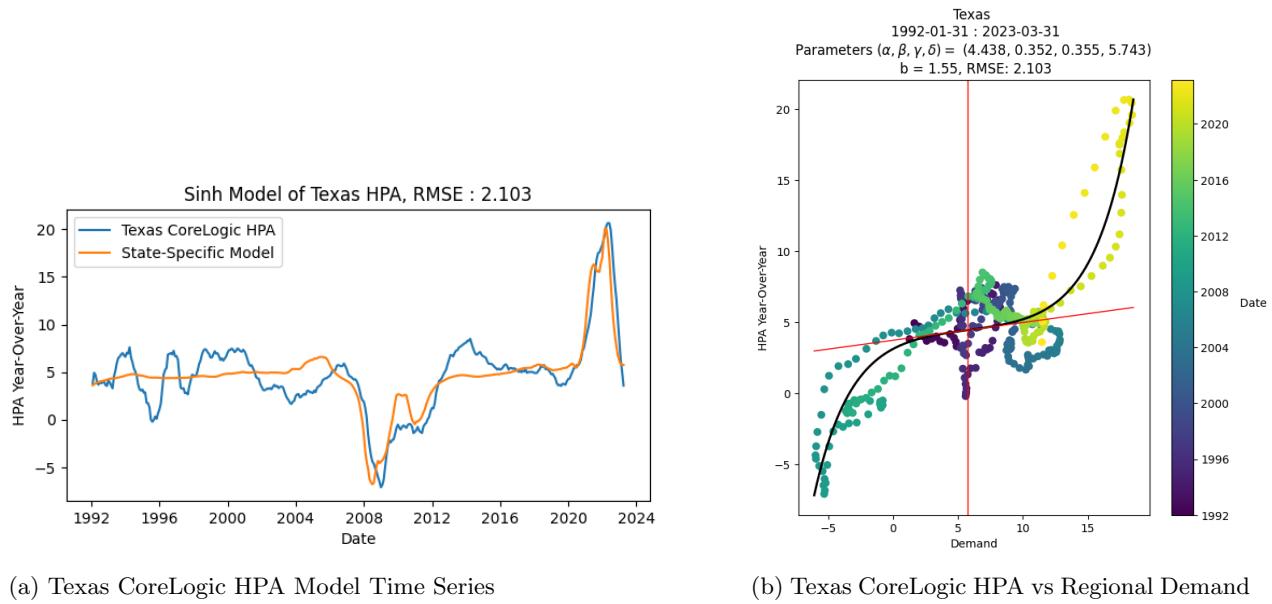
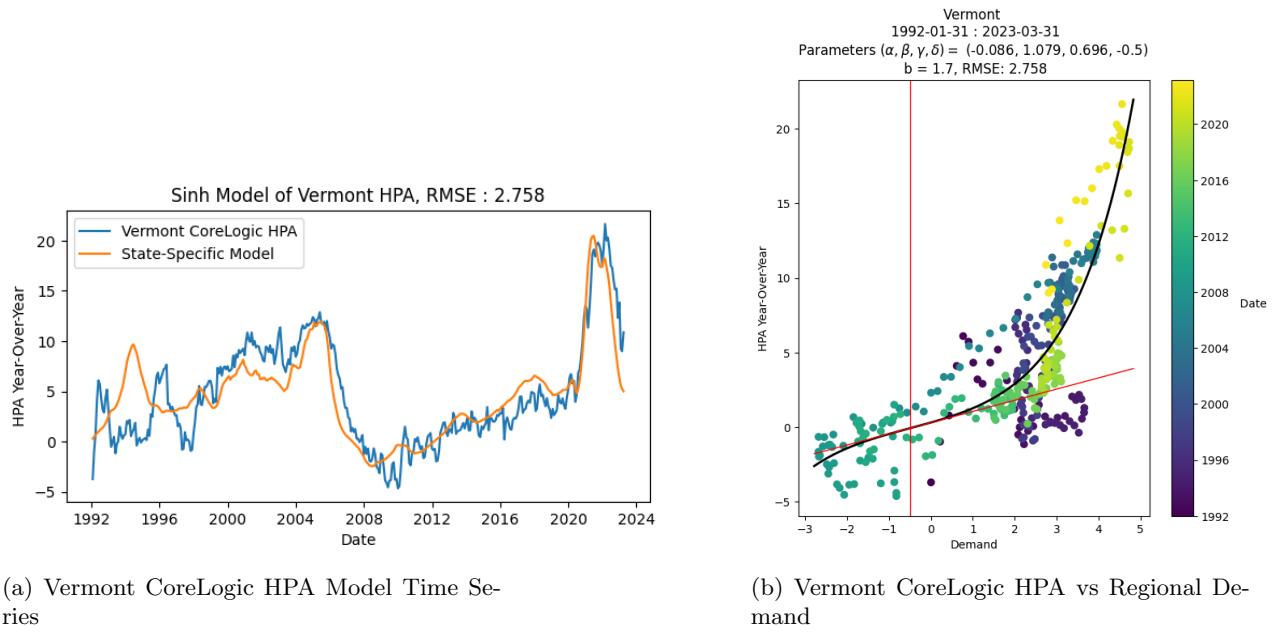
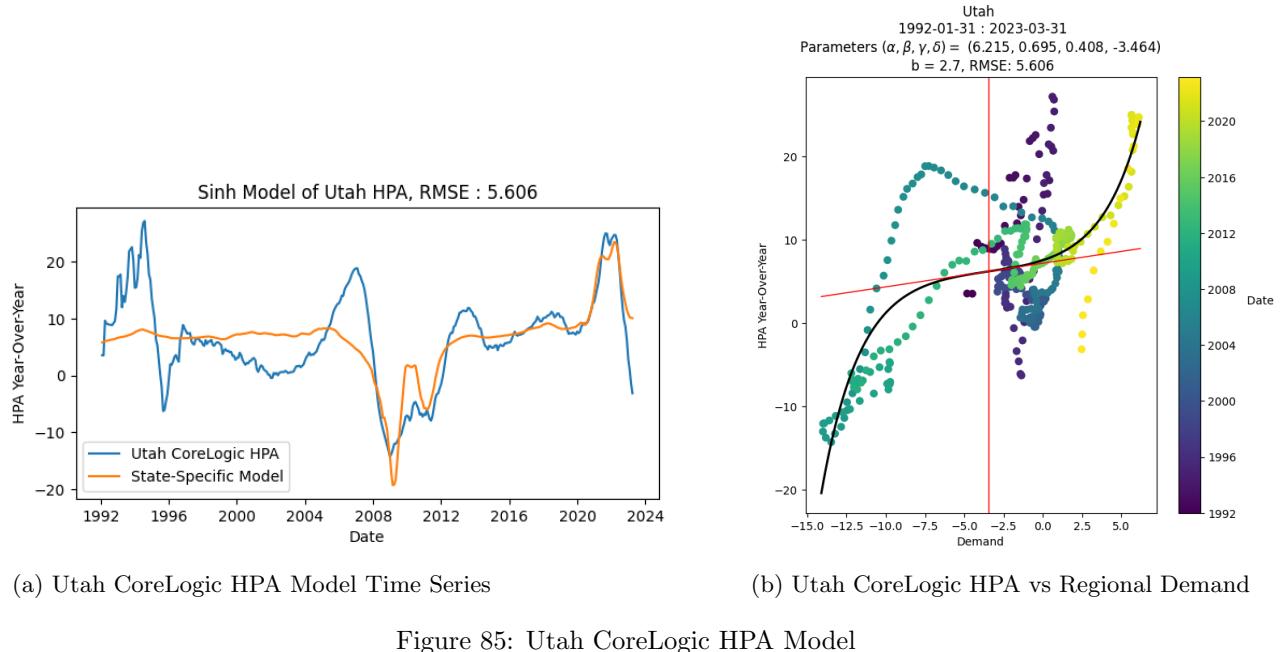


Figure 84: Texas CoreLogic HPA Model



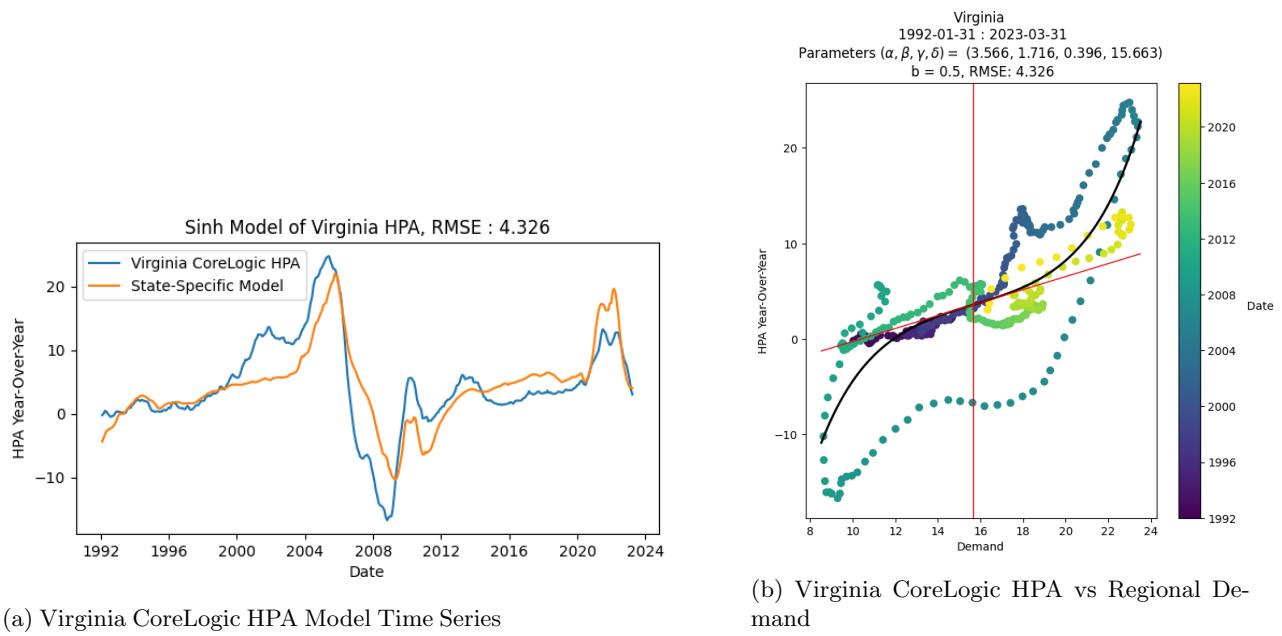


Figure 87: Virginia CoreLogic HPA Model

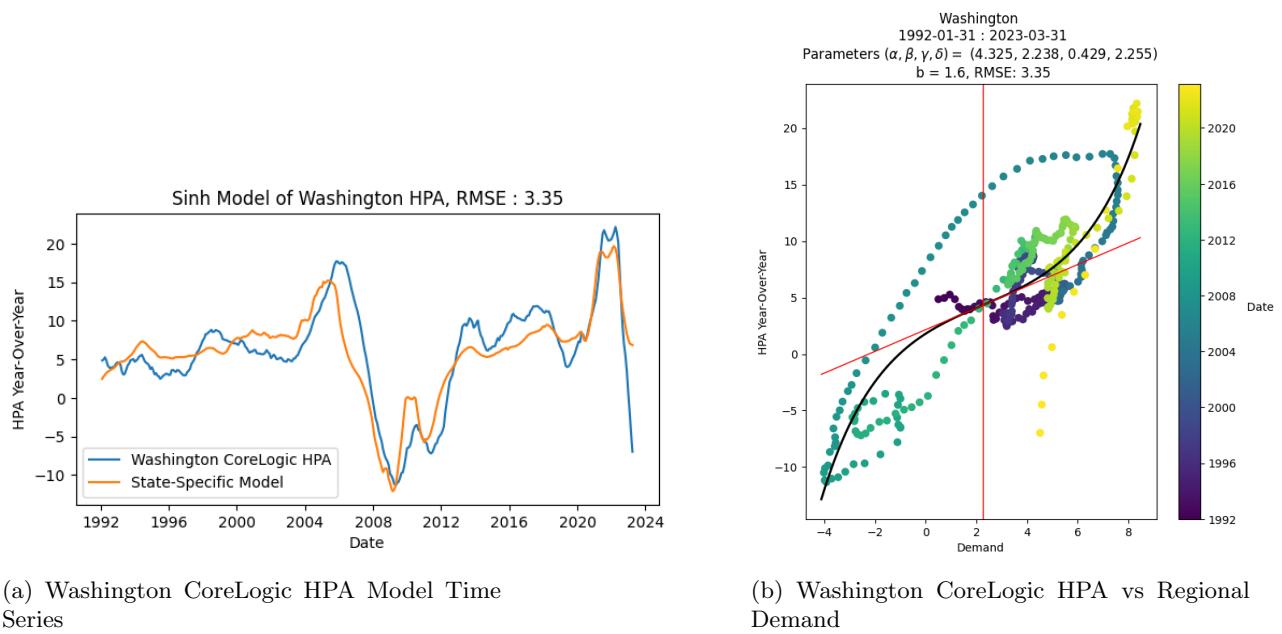


Figure 88: Washington CoreLogic HPA Model

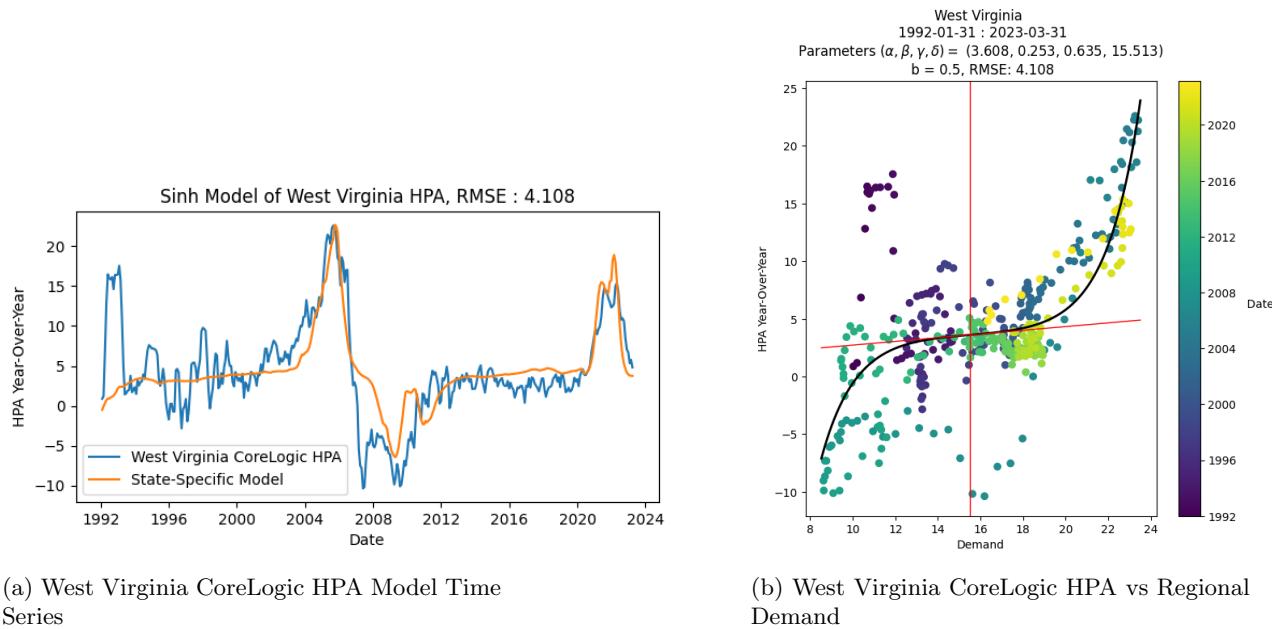


Figure 89: West Virginia CoreLogic HPA Model

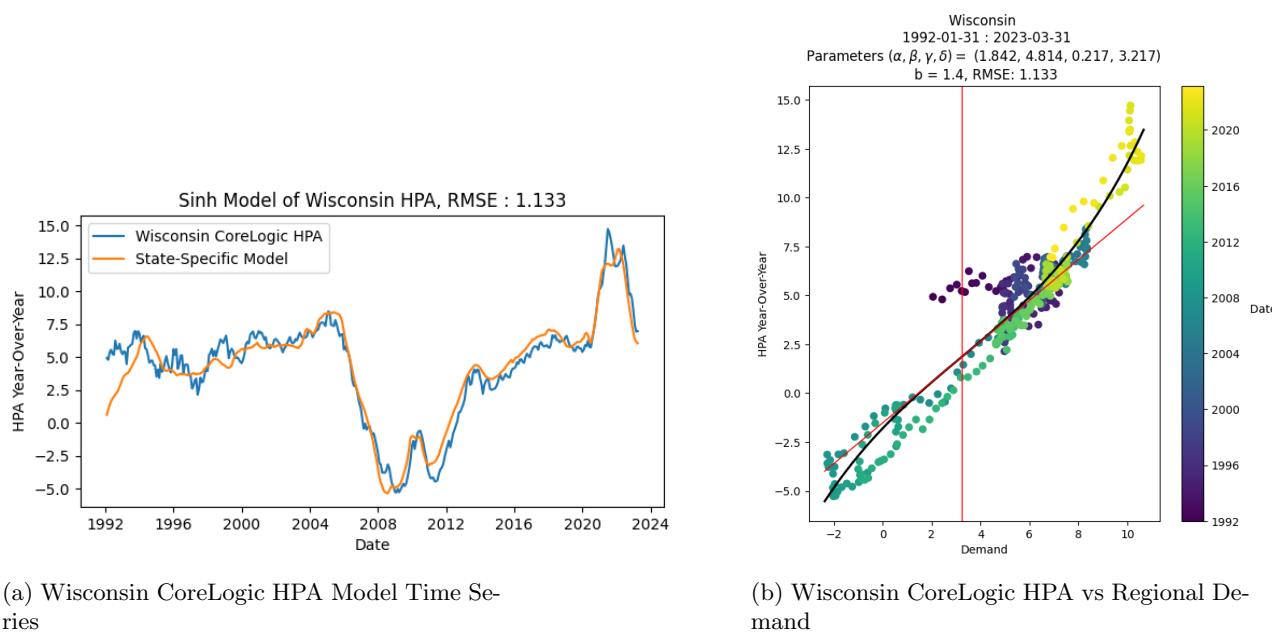


Figure 90: Wisconsin CoreLogic HPA Model

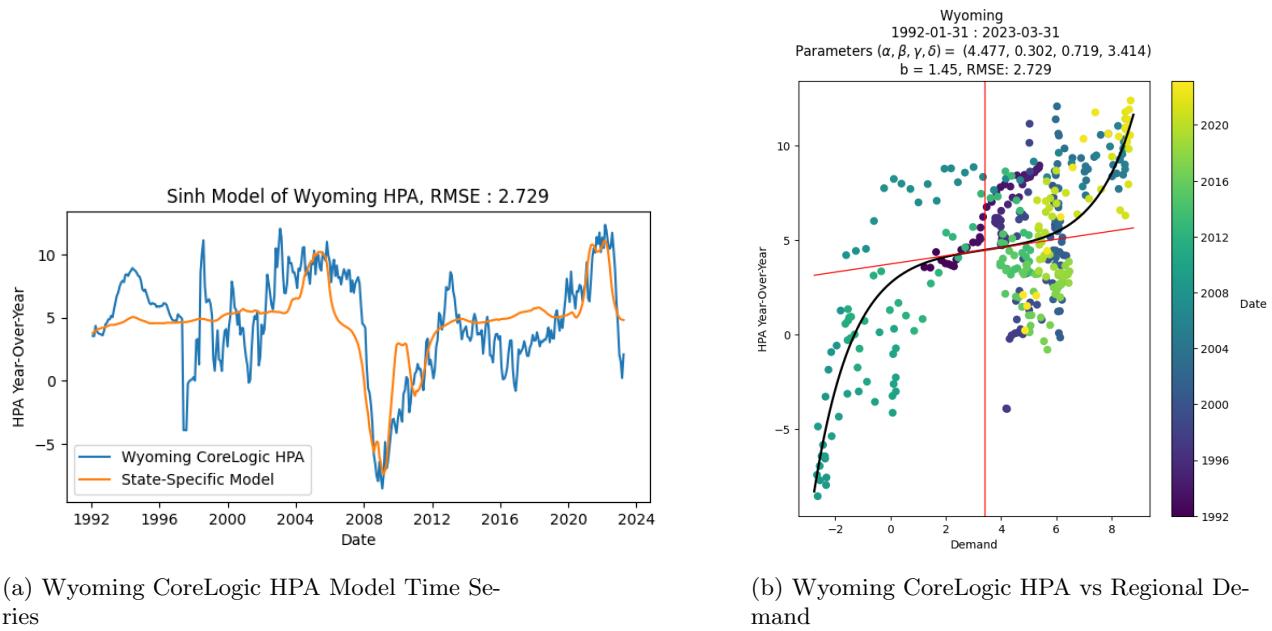


Figure 91: Wyoming CoreLogic HPA Model

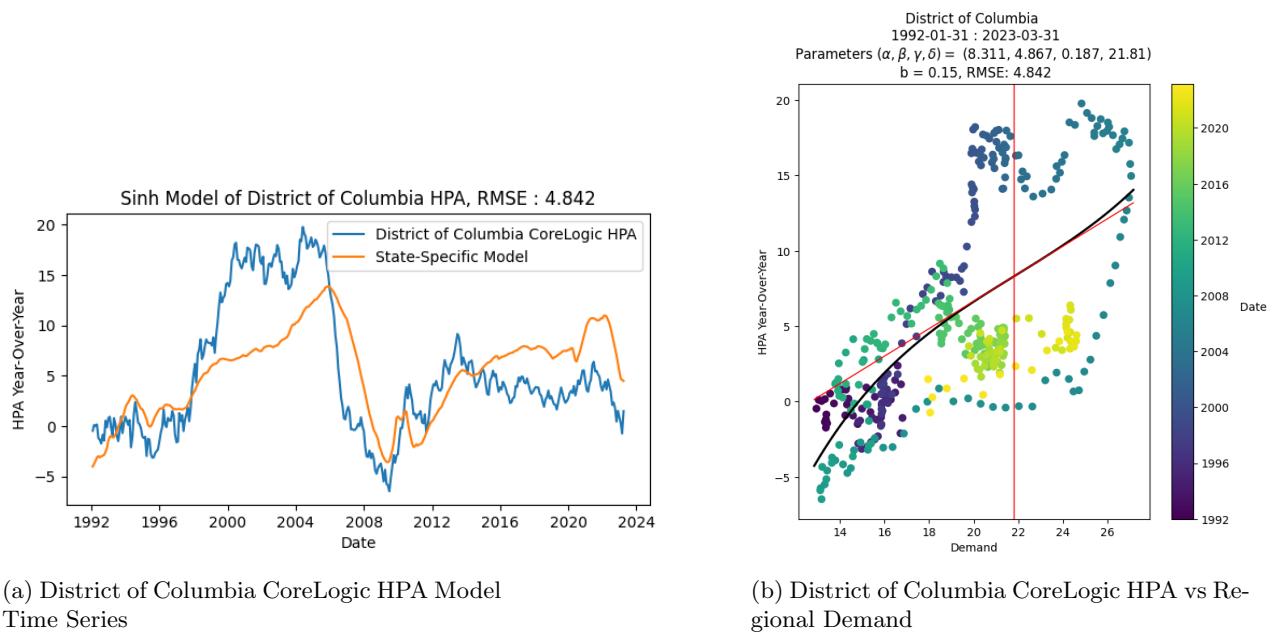


Figure 92: District of Columbia CoreLogic HPA Model