Check for
updates

A Bayesian piecewise linear model for the detection of breakpoints in housing prices

Jabed H. Tomal[1] (ID) · **Hafizur Rahman**[2]

Abstract

Statistical thresholds occur when the changes in the relationships between a response and predictor variables are not linear but abrupt at some points of the predictor variable values. In this paper, we defined a piecewise-linear regression model which can detect two thresholds in the relationships via changes in slopes. We developed the corresponding Bayesian methodology for model estimation and inference by proposing prior distributions, deriving posterior distributions, and generating posterior values using Metropolis and Gibbs sampling algorithm. The parameters in our model are easy to understand, highly interpretable, and flexible to make inferences. The methodology has been applied to estimate threshold effects in housing market pricing data in two cities - Kamloops and Chilliwack - in British Columbia, Canada. Our findings revealed that the implementation of changes in the government property tax policies had threshold effects in the market price trend. The proposed model will be useful to detect threshold effects in other correlated time series data as well.

Keywords Breakpoint detection · Bayesian method · Time series data · Home prices · Canada

1 Introduction

Purchase of residential housing constitutes the largest expenditure for most households [10]. For that reason, unlike other commodities purchased by households, home purchase is considered an investment item in national income accounting. The housing market provides many important linkages for various sectors of the economy generating significant impact on the level and rate of change in aggregate income of an economy, most often measured by gross domestic product (GDP) [27].

✉ Jabed H. Tomal
   jtomal@tru.ca

   Hafizur Rahman
   hrahman@tru.ca

1   Department of Mathematics and Statistics, Thompson Rivers University, Kamloops, BC V2C 0C8, Canada

2   School of Business and Economics, Thompson Rivers University, Kamloops, BC V2C 0C8, Canada

As with all commodities in the market, housing prices are determined by the forces of demand and supply [17,32]. However, specific factors on demand and supply differ depending on the nature of the commodity in question. While at the general equilibrium level all markets and all factors influencing individual market are interrelated, some of the most relevant determinants of housing prices on the demand side are: interest rates, the state of the economy, demographics including immigration, availability of mortgage financing, affordability of housing, expectations, and government policies and legislation including tax incentives, tax deductibility and subsidies [16]. On the supply side, the important determinants are: number of active listings reflecting availability of homes for sale, new home construction reflecting increase in the stock of housing, and land use and zoning reflecting the availability of land for different types of buildings permitted [18].

A wide array of studies has attempted to capture the impacts of various demand and supply side factors on housing prices. Nistor and Reianu [33] examined the effect of immigration on housing prices in the ten largest metropolitan areas in Ontario, Canada. Using quarterly data for Canada, Hossain and Latif [24] found that housing price volatility is affected significantly by GDP growth rate, housing price appreciation rate and the rate of inflation. Another study of the factors influencing home prices in the greater Toronto area in Canada [6] identify a strong economy, low interest rates and favourable mortgage insurance rules on the demand side, and limited supply of housing stock on the supply side as contributing factors for price increases.

The housing markets of large Canadian cities have been attractive places of investment for foreign buyers [11]. It has not been more pronounced anywhere than in Vancouver which has attracted a large inflow of foreign buyers due to its proximity to the Asia-Pacific region, mild weather and natural beauty [12,23]. Limited supply of land and strong demand resulted in a rapid increase in real estate prices which prompted for calls from the public to curtail the demand from foreign buyers [22,25]. In response to that, the provincial government enabled the cities and municipalities to impose an additional property transfer tax on foreign buyers for purchases of residential properties. The tax was first introduced in Vancouver and some other cities soon followed suit[1]. This additional tax currently stands at 20% of the fair market value of the property. Our study attempts to fill a void in literature by examining the impact of change in government policy on the housing prices of the Province of British Columbia, Canada. If all else remains unchanged, we expect that the imposition of the additional foreign buyer's tax in one city would impact the housing prices of other cities in the province by shifting demand away from the city with the tax, to cities without. [1]

We propose a statistical model to measure association between housing price and time demarcated by thresholds. Two variables are considered to be related to each other when there is a change in the distribution of one variable for a change in the other variable [15]. In a regression set up, the variables are considered to be related if the conditional mean of the response variable changes for a change in the predictor [14]. The variables are considered to be related with threshold effect when the slope of the regression relationship changes abruptly at a certain level of the predictor variable, which is known as the breakpoint [44]. Estimation of breakpoint is very popular in many applications. In environmental ecology, for example, the piecewise-linear regression model (PLRM) is often used as a tool to calculate environmental thresholds resulting from human induced disturbances to nature [4,19,45]. In Business and Economics, estimation of breakpoints in the housing market is known as the estimation of structural breaks [3,9]. Our goal in this paper is to use the PLRM to estimate the

---

[1] https://www2.gov.bc.ca/gov/content/taxes/property-taxes/property-transfer-tax/additional-property-transfer-tax (Accessed on March 7, 2021).

breakpoints in housing prices trends resulting from the adoption and changes in government policies.

A piecewise-linear regression model (PLRM) with one breakpoint represents two linear lines with differing slopes demarcating the data into two segments [4]. Muggeo [31] proposed a PLRM (also know as segmented regression model) to estimate breakpoints using classical statistical method. Toms and Lesperance [44] developed several PLRMs where two linear lines are connected at the breakpoint using various mechanisms. Examples of estimating a single breakpoint can be found at [19,38,44,45]. A PLRM with two breakpoints has three slopes, defined at three segments of the data, where two slopes are connected at a single breakpoint. Tomal and Ciborowski [42] developed a PLRM with two breakpoints and applied the model to estimate environmental thresholds from human induced disturbances to nature. All of the above studies are based on either classical statistical or quantile regression methods.

There is considerable interest in developing Bayesian models for the detection of statistical thresholds via the changes in slopes. Qian et al. [35] proposed a Bayesian hierarchical model (BHM) for the detection of environmental threshold effects of total phosphorus on macroinvertebrate composition in wetlands. Liang et al. [28] developed a Bayesian change point model to estimate threshold effects of nutrients-phytoplankton relationship in lakes. Ouyang et al. [34] extended BHM to improve the estimates of build-up areas from night time light across globally distributed cities. Bucci et al. [5] proposed a Bayesian segmented regression model and applied their method to forensic age estimation using the assumption that the data points were independent to each other. All of the above Bayesian models were developed to detect a single threshold via the changes in slopes using a PLRM. In this paper, we develop a Bayesian PLRM to detect two thresholds via changes in slopes.

Chien [9] used a Lagrange Multiplier unit root test [26] to examine the issue of whether regime changes had broken down the stability of the ripple effect in Taiwan's housing market. Ravazzolo et al. [36] utilized Bayesian model averaging, an ensemble method, to estimate structural breaks in the regression parameters, uncertainty about the inclusion of forecasting variables, and uncertainty about parameter values with application in the stock market. Vizek and Posedel [47] used threshold autoregressive (TAR) and momentum TAR (M-TAR) models, defined thresholds in terms of the changes in the error term, to test if housing prices in the United States, United Kingdom, Spain and Ireland were characterized by threshold effects. On the other hand, Begiazi and Katsiampa [3] used return $[(ln(p_t) - ln(p_{t-1})) \times 100]$ and the GARCH model to detect structural breaks in housing prices in the United Kingdom. Their transformation and the GARCH model are useful for examining volatility in price, breaks in variance-covariance, and conditional heteroscedasticity. Unlike Begiazi and Katsiampa [3], the objectives of our paper are proposing a Bayesian model for time series data to estimate breaks in linear trend and interpreting the linear relationships estimated from the slopes of the piecewise linear model. Therefore, our methodology and objectives are different than the methodology and objectives of the studies mentioned above. Moreover, we have transformed our data using "return" as in Begiazi and Katsiampa [3], and found that there is no noticeable breaks in volatility, variance-covariance, and conditional heteroscedasticity. Additionally, the transformation using return causes to lose one data point, especially for the first month, which might appear crucial in statistical estimation and model building when the amount of data is limited.

Time series data is collected over time and, as a result, the successive data points are correlated as opposed to independent. Djafari and Féron [30] used a Bayesian method for change point detection in variables that are collected over time. Their method can be used to detect change points via the distribution of the variables. Ruggieri [37] used a Bayesian approach for the detection of change points in climatic records through the probability density function.

Instead of using a regression model that is defined for the entire data series, the author used chain rule of probability theory by multiplying the conditional and marginal probabilities to obtain the joint probability. On the other hand, we have defined a piecewise-linear regression model with two breakpoints via the changes of slopes which is defined for the entire set of data. Unlike Bucci et al. [5], we have extended the Bayesian piecewise linear regression model to time series data where the errors are correlated. Instead of assuming independence of data points, we proposed an auto-regressive correlation structure of the residuals which allowed us to calculate correlations between successive data points. We further showed how these correlations can be estimated from MCMC samples using Metropolis algorithm. Note that similar classical statistical models developed for time series data to capture breakpoints are available in the literature [1,39,40,46]. Our proposed model is a Bayesian and extended version of the standard classical models. The beauty of our model lies in its simplicity, less involvement of model parameters, and evidence-based intuitive assumptions. The main strengths of our paper are the interpretability of the model parameters (slopes and breakpoints), and simple and straightforward statistical inference procedures.

The proposed Bayesian piecewise-linear regression model has been applied to data of two housing markets—(1) Chilliwack, BC, and (2) Kamloops, BC. We hypothesize that the changes in government policies in large cities, such as imposition of the new non-resident property transfer tax, cause threshold effects on home prices in nearby cities. We also hypothesize that when the government changes policies such as imposition of new non-resident property transfer tax in the smaller cities, it also causes similar ripple threshold effects in nearby cities of similar size.

Therefore, the first objective of our paper is geared by the application for the detection of threshold effects of the introduction of government's tax policy changes in Vancouver, British Columbia, Canada on housing prices in two similar sized cities in the province (similar in terms of population, business capacity and size of universities). The second objective is to extend the Bayesian piecewise linear regression model (segmented regression model) for the detection of thresholds in time series data. Our findings showed that the changes in government tax policies caused significant threshold effects in housing prices.

The model building procedure is summarized as follows. Firstly, we visualize the data to observe the type of relationships between variables and define the model. The model is finalized by adaptive imposition of the appropriate correlation structure evident in the data points. We then define appropriate prior distributions for the model parameters, and derive their posterior distributions. We propose the Metropolis and Gibbs sampling algorithm to generate Markov Chain Monte Carlo (MCMC) samples from the posterior distributions. Lastly, we propose a two-step procedure to determine appropriate proposal distribution for the parameters of interest.

2 Housing price/time series data

We collect time series data of monthly average home prices for two small cities in British Columbia—Chilliwack, which is about 100 km east to metropolitan Vancouver, and Kamloops, which is about 350 km north-east of Vancouver. These two cities are of similar size in terms of landmass, business capacity and population, and each is the home to one similar-sized university. The data are collected from January 2011 to July 2020 for a total of 115 months from the monthly statistical report of the Canadian Real Estate Association (CREA) (https://
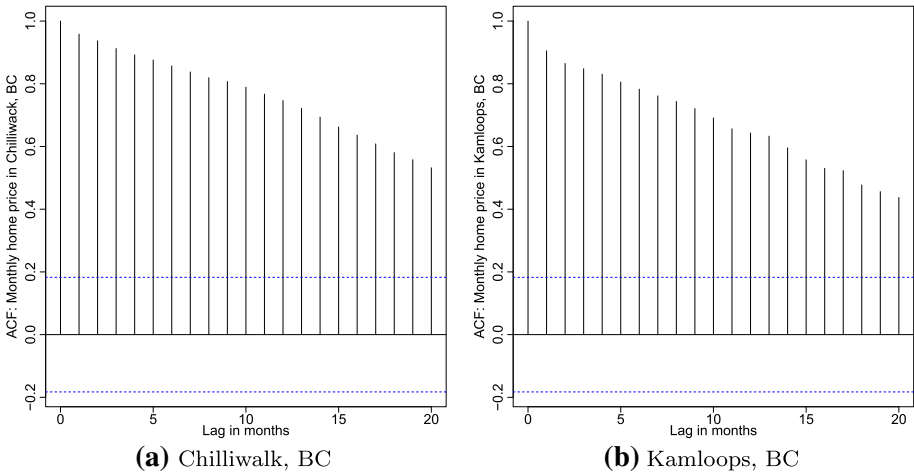
Fig. 1 Auto-correlation functions for monthly home prices against time

creastats.crea.ca/en-CA/). In our data, the monthly average home prices are the average of all types of houses such as detached and semidetached houses, townhouses, and condos.

3 Methods

Before we propose the model, we introduce the autocorrelation function (ACF) of the monthly average home prices in Chilliwack (left panel) and Kamloops (right panel), British Columbia (BC) against time lag (months) in Fig. 1. We can see that the monthly home prices are highly correlated. Even at the time lag of month 21, the autocorrlation is above 0.40 and highly statistically significant. This observation leads us to propose a model with positively correlated errors as defined below.

3.1 The model

Let Y and x be the average monthly home price and the linear trend variable, in month, respectively. We define our model as:

$$Y_t = \theta_0 + \theta_1 x_t + \theta_2(x_t - \theta_4)I(x_t - \theta_4) + \theta_3(x_t - \theta_5)I(x_t - \theta_5) + \epsilon_t, \tag{1}$$

where θ_4 and θ_5 are the first and second breakpoints, respectively, and

$$I(x_t - \theta_j) = \begin{cases} 1 & \text{for } x_t \geq \theta_j \\ 0 & \text{for } x_t < \theta_j, \end{cases}$$

for $j \in \{4, 5\}$ and $\epsilon = (\epsilon_1, \epsilon_2, \ldots, \epsilon_n)^T \sim$ Multivariate Normal$(\mathbf{0}, \Sigma)$. The linear trend part of our proposed model deals with *non-stationarity* in time series data modeling. In this formulation, θ_1, $\theta_1 + \theta_2$ and $\theta_1 + \theta_2 + \theta_3$ are the slopes in the first, second and third segments of the model demarcated by the thresholds θ_4 and θ_5, respectively. When θ_2 and θ_3 are 0, the thresholds θ_4 and θ_5 are statistically insignificant and, thus, there is no threshold via the changes in slope. The beauty of this model lies in the easy interpretation of the linear trends via the slopes.

Given the conditional mean of the response variable against linear trend of time denoted as $E(\mathbf{Y}|\mathbf{x}, \boldsymbol{\theta}, \Sigma) = \theta_0 + \theta_1\mathbf{x} + \theta_2(\mathbf{x} - \theta_4)I(\mathbf{x} - \theta_4) + \theta_3(\mathbf{x} - \theta_5)I(\mathbf{x} - \theta_5)$, the conditional distribution of the response variable can be written as

$$\mathbf{Y}|\mathbf{x}, \boldsymbol{\theta}, \Sigma \sim \text{Multivariate Normal}\left(E(\mathbf{Y}|\mathbf{x}, \boldsymbol{\theta}, \Sigma), \Sigma\right), \tag{2}$$

where Σ is the variance-covariance matrix of the vector of error term. From the auto-correlation function plot of the average home price, we consider that the error terms are not independent, but temporally autocorrelated. As a result, the structure of the variance-covariance matrix Σ needs to reflect positive autocorrelation between sequential observations. As the autocorrelation gradually decreases over time, we consider the following first-order autoregressive structure for the variance-covariance matrix of the error term:

$$\Sigma = \sigma^2 C_p = \sigma^2 \begin{pmatrix} 1 & \rho & \rho^2 & \cdots & \rho^{n-1} \\ \rho & 1 & \rho & \cdots & \rho^{n-2} \\ \rho^2 & \rho & 1 & \cdots & \rho^{n-3} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \rho^{n-1} & \rho^{n-2} & \rho^{n-3} & \cdots & 1 \end{pmatrix},$$

where ρ is the correlation between successive error terms ϵ_t and ϵ_{t+1}. Under this covariance structure, the variance of ϵ_t is σ^2, and the autocorrelation between ϵ_t and ϵ_{t+d} is ρ^d, which decreases as the time-lag d grows larger.

3.2 The sampling distribution of data

Let $\boldsymbol{\theta} = (\theta_0, \theta_1, \theta_2, \theta_3, \theta_4, \theta_5)^T$ be the vector of regression coefficients. Then the sampling distribution of data is defined as

$$P(\mathbf{Y} = \mathbf{y}|\mathbf{x}, \boldsymbol{\theta}, \Sigma) \equiv P(\mathbf{Y} = \mathbf{y}|\mathbf{x}, \boldsymbol{\theta}, \sigma^2, \rho) \sim \text{MVN}\left(E(\mathbf{Y}|\mathbf{x}, \boldsymbol{\theta}, \Sigma), \Sigma\right), \tag{3}$$

where MVN stands for the multivariate normal distribution with mean vector $E(\mathbf{Y}|\mathbf{x}, \boldsymbol{\theta}, \Sigma)$ and variance-covariance matrix Σ.

3.3 Prior distributions

The prior distribution for the error variance is considered as Inverse-Gamma as following

$$\sigma^2 \sim \text{Inverse-Gamma}\left(\frac{v_0}{2}, \frac{v_0\sigma_0^2}{2}\right), \tag{4}$$

where the prior mean is $E(1/\sigma^2) = 1/\sigma_0^2$ with strength of prior belief v_0. Here, larger values of $v_0 > 0$ indicate stronger prior belief and vice versa.

The prior distribution for $\boldsymbol{\theta}$ is considered as Multivariate-Normal

$$\boldsymbol{\theta} \sim \text{Multivariate Normal}\left(\boldsymbol{\mu}_0, \Sigma_0\right), \tag{5}$$

with prior mean $\boldsymbol{\mu}_0$ and variance-covariance matrix Σ_0.

As noted before, we consider that the successive observations are positively correlated, and use a flat prior for correlation coefficient ρ on the range from 0 to 1:

$$\rho \sim \text{Unif}[0, 1], \tag{6}$$

where Unif stands for the density of Uniform distribution defined over $[0, 1]$.

3.4 The posterior distributions

The posterior distribution for the error variance σ^2 is obtained by combining the sampling distribution (Eq. 2) and the prior distribution of σ^2 (Eq. 4) as following:

$$\sigma^2|\mathbf{y}, \mathbf{x}, \boldsymbol{\theta}, \rho \sim \text{Inverse-Gamma} \left(\frac{v_0 + n}{2}, \frac{v_0\sigma_0^2 + \text{SSE}_\rho(\boldsymbol{\theta})}{2} \right), \tag{7}$$

where the sum of squares of errors is expressed as

$$\text{SSE}_\rho(\boldsymbol{\theta}) = (\mathbf{y} - E(\mathbf{Y}|\mathbf{x}, \boldsymbol{\theta}, \Sigma))^T C_\rho^{-1} (\mathbf{y} - E(\mathbf{Y}|\mathbf{x}, \boldsymbol{\theta}, \Sigma)).$$

Unfortunately, the posterior distribution for $\boldsymbol{\theta}$ has no closed form expression. We write

$$\begin{aligned} \boldsymbol{\theta}|\mathbf{y}, \mathbf{x}, \Sigma &\sim P(\mathbf{y}|\boldsymbol{\theta}, \mathbf{x}, \Sigma) \times P(\boldsymbol{\theta}) \\ &\sim \text{MVN}(E(\mathbf{Y}|\mathbf{x}, \boldsymbol{\theta}, \Sigma)) \times \text{MVN}(\boldsymbol{\mu}_0, \Sigma_0), \end{aligned} \tag{8}$$

where MVN stands for the Multivariate Normal distribution.

As like the posterior distribution for $\boldsymbol{\theta}$, the posterior distribution for ρ has no closed form expression. Thus, we write

$$\begin{aligned} \rho|\mathbf{y}, \mathbf{x}, \boldsymbol{\theta}, \Sigma &\sim P(\mathbf{y}|\boldsymbol{\theta}, \mathbf{x}, \Sigma) \times P(\rho) \\ &\sim \text{MVN}(E(\mathbf{Y}|\mathbf{x}, \boldsymbol{\theta}, \Sigma)) \times \text{Unif}(0, 1), \end{aligned} \tag{9}$$

where Unif stands for the Uniform distribution defined over [0, 1].

3.5 Metropolis and Gibbs sampling algorithm

As we do not have a closed form expression for some of the posterior distributions, we cannot use Monte Carlo method or Gibbs sampling to generate posterior values. Instead, we propose the following Metropolis and Gibbs sampling algorithm to generate samples from the full posterior distribution of the parameters. Here, we have fused the Metropolis and Gibbs sampling algorithms together in a single algorithm to estimate the set of parameters of our model which has been designed for the estimation of breakpoints (i.e., structural breaks). For a given set of parameter values $\{\boldsymbol{\theta}^{(s)}, \sigma^{2(s)}, \rho^{(s)}\}$ for sth MCMC iteration, the next iteration $s + 1$ is generated as following:

1. Update σ^2 using Gibbs sampling:

   (a) Calculate the sum of squares of error given $\boldsymbol{\theta}^{(s)}$ and $\rho^{(s)}$

   $$\text{SSE}_{\rho^{(s)}}(\boldsymbol{\theta}^{(s)}) = \left(\mathbf{y} - E(\mathbf{Y}|\mathbf{x}, \boldsymbol{\theta}^{(s)})\right)^T C_{\rho^{(s)}}^{-1} \left(\mathbf{y} - E(\mathbf{Y}|\mathbf{x}, \boldsymbol{\theta}^{(s)})\right).$$

   (b) Generate a new value

   $$\sigma^{2(s+1)} \sim \text{Inverse-Gamma} \left(\frac{v_0 + n}{2}, \frac{v_0\sigma_0^2 + \text{SSE}_{\rho^{(s)}}(\boldsymbol{\theta}^{(s)})}{2} \right),$$

   where the posterior distribution is given by Eq. 7.

2. Update $\boldsymbol{\theta}$ using Metropolis algorithm:

   (a) Given the current state of $\boldsymbol{\theta}^{(s)}$, propose $\boldsymbol{\theta}^*$ from a symmetric proposal distribution $J_1(\boldsymbol{\theta}|\boldsymbol{\theta}^{(s)})$

(b) Given the parameter values $\theta^{(s)}$, $\sigma^{2(s+1)}$, $\rho^{(s)}$, and the proposal θ^*, calculate the acceptance ratio for the proposal

$$r_{\theta *} = \frac{P\left(\mathbf{y}|\mathbf{x}, \theta^*, \sigma^{2(s+1)}, \rho^{(s)}\right) P\left(\theta^*\right)}{P\left(\mathbf{y}|\mathbf{x}, \theta^{(s)}, \sigma^{2(s+1)}, \rho^{(s)}\right) P\left(\theta^{(s)}\right)},$$

where the expression for the numerator and denominator is given in Eq. 8.

(c) Accept or reject the proposal as following

$$\theta^{(s+1)} = \begin{cases} \theta^* & \text{with probability } \min(r_{\theta *}, 1) \\ \theta^{(s)} & \text{with probability } 1 - \min(r_{\theta *}, 1). \end{cases}$$

3. Update ρ using Metropolis algorithm:

(a) Given $\rho^{(s)}$, propose ρ^* from a symmetric proposal distribution $J_2(\rho^*|\rho^{(s)})$

(b) Given the parameter values $\theta^{(s+1)}$, $\sigma^{2(s+1)}$, $\rho^{(s)}$, and the proposal ρ^*, calculate the acceptance ratio for the proposal

$$r_{\rho *} = \frac{P\left(\mathbf{y}|\mathbf{x}, \theta^{(s+1)}, \sigma^{2(s+1)}, \rho^*\right) P\left(\rho^*\right)}{P\left(\mathbf{y}|\mathbf{x}, \theta^{(s+1)}, \sigma^{2(s+1)}, \rho^{(s)}\right) P\left(\rho^{(s)}\right)},$$

where the expression for the numerator (and denominator) is given in Eq. 9.

(c) Accept or reject the proposal as following

$$\rho^{(s+1)} = \begin{cases} \rho^* & \text{with probability } \min(r_{\rho *}, 1) \\ \rho^{(s)} & \text{with probability } 1 - \min(r_{\rho *}, 1). \end{cases}$$

3.6 Proposal distributions

1. The proposal distribution for θ is considered as

$$J_1\left(\theta|\theta^{(s)}\right) = \text{MVN}\left(\theta^{(s)}, \Sigma_p\right),$$

where Σ_p is the proposal variance-covariance matrix of θ.

2. We set the proposal distribution for ρ as

$$J_2\left(\rho|\rho^{(s)}\right) = \text{Unif}\left(\rho^{(s)} - \delta, \rho^{(s)} + \delta\right).$$

If $\rho^* < 0$, reassign it to be $|\rho^*|$. If $\rho^* > 1$, reassign it to be $2 - \rho^*$.

3.7 Prior specification

In order to specify the hyper-parameters for the prior distribution of σ^2, we depend on the *loess* fits. We calculated the residual variance from the loess fit, and assign that as σ_0^2 to be $17{,}037.93^2$ and $12{,}080.95^2$ for Chilliwack and Kamloops data, respectively. We further use a non-informative prior and assign $\nu_0 = 1$ for both sets of data. This ensures minimal effects of our prior knowledge on the posterior model.

To specify the hyper-parameters for the prior distribution of θ, we depended on the prior knowledge, the hypotheses, and the raw data. The hyper-parameter vector $\mu_0 = (\theta_{00}, \theta_{10}, \theta_{20}, \theta_{30}, \theta_{40}, \theta_{50})$ is specified as follows. On July 25, 2016, the provincial government of British Columbia introduced Bill 28, Miscellaneous Statutes (Housing Priority

Initiatives) Amendment Act, 2016. The bill amended the Property Transfer Tax Act to include an additional 15% transfer tax on foreign entities buying property in Metro Vancouver. We hypothesize that this bill introduces threshold effects in the surrounding cities of the Lower Mainland of British Columbia and reaching as far as Kamloops. This hypothesis helps us specifying θ_{40} to be 67 (as July 2016 is the 67th month counting from January 2011). On February 20, 2018, the government of British Columbia extended the foreign buyer property transfer tax to the following regions (i) Capital Regional District, (ii) Fraser Valley Regional District, (iii) Metro Vancouver Regional District, (iv) Regional District of Central Okanagan, and (v) Regional District of Nanaimo. We further hypothesize that the extension of the foreign buyer tax to these regions introduces threshold effects in the cities within these regions. This hypothesis helps us to specify θ_{50} to be 86 (as February 2018 is the 86th month counting from January 2011). The fitted non-parametric *loess* model helped us to determine the other hyperparameters for Chilliwack, BC data. We used the fitted loess values at some index time points (e.g., the first, 67th, 86th and last months) to calculate the slopes reasonable in the three segments of the data. We then used the relationships between the parameters and slopes to produce the following vector of hyperparameters for Chilliwack, BC as $\boldsymbol{\mu}_0 = (275573.80, 1860.78, 3859.67, -3589.30, 67, 86)$. Following the same steps, we specify the hyperparameters for Kamloops, BC as $\boldsymbol{\mu}_0 = (308941.00, 462.53, 1456.33, 866.66, 67, 86)$.

To specify the hyper-parameter Σ_0, we use statistical knowledge and common sense. In every specification, we make the prior distribution flat and diffuse leading to large prior variance. We define the prior covariance matrix for Chilliwack first. We consider that the covariances between θ_i and θ_j are zero. The standard deviation of θ_0 is considered as 10,000. For the parameters θ_1, θ_2 and θ_3 (parameters specific to the slopes), we consider the standard deviation such as 1900, 3900 and 3600, respectively. Such specification ensures that each of the prior distributions of θ_1, θ_2 and θ_3 includes 0 well within the high probability region. For the threshold parameters θ_4 and θ_5, we consider prior distributions by assigning enough masses within 24-month time spans. This gives us a prior covariance matrix for Chilliwack as

$$\text{diag}(\Sigma_0) = \left(10000^2, 1900^2, 3900^2, 3600^2, 8^2, 8^2\right).$$

Similarly, the prior covariance matrix for Kamloops is considered as

$$\text{diag}(\Sigma_0) = \left(10000^2, 500^2, 1500^2, 900^2, 8^2, 8^2\right).$$

Such specification of hyperparameters make the prior distributions for $(\theta_0, \theta_1, \theta_2, \theta_3)$ diffuse and flat leading to negligible effect on the posterior distributions. On the other hand, the priors variances for θ_4 and θ_5 (the breakpoints) make their priors informative. The reason for using informative prior is that [2] suggested to use informative prior instead of vague, flat or diffuse prior for the parameters whenever prior information is available.

3.8 Initial values of the parameters

In order to start the Metropolis and Gibbs sampling algorithm, we use the following initial values $\{\boldsymbol{\theta}^{(s)}, \sigma^{2(s)}, \rho^{(s)}\} = \{\boldsymbol{\mu}_0, \sigma_0^2, \hat{\rho}_1\}$, where $\hat{\rho}_1$ is the lag-1 autocorrelation of the residuals obtained from the fitted loess. The lag-1 autocorrelations of the loess residuals are calculated as $\hat{\rho}_1 = 0.30$ for Chilliwalk and $\hat{\rho}_1 = 0.12$ for Kamloops.

To determine the proposal covariance matrix Σ_p, we ran the algorithm in two phases. In the first phase, we considered $\Sigma_p = \Sigma_0/k$, and determined a reasonable value of k using trial-and-error method. In the first phase, a total of 22,000 MCMC iterations were generated

and the first 2000 iterations were used as burn-in. Among the 20,000 MCMC iterations after burn-in, we saved every 20th iteration for the $\boldsymbol{\theta}$ vector as part of thinning. The empirical covariance matrix of the thinned values for $\boldsymbol{\theta}$ was used as the proposal covariance matrix Σ_p for the second phase of the algorithm. In the second phase, the algorithm was run for a total of 120,000 iterations and the first 20,000 iterations were used as burn-in. After throwing out the burn-in iterations, in the second phase, every 100th iteration was saved as part of thinning. This provided us exactly 1000 values for each parameter after burn-in and thinning.

3.9 MCMC diagnostics

The second phase of the algorithm gave us an acceptance rate of 21.58% and 77.05% for $\boldsymbol{\theta}$ and ρ, respectively, for the data from Chilliwack, BC. Fig. 2 shows the autocorrelation functions for each of the θ's. As usual, the generated MCMC iterations are highly autocorrelated. This justifies our approach of thinning the generated MCMC iterations. After thinning, the effective sample sizes for θ_0, θ_1, θ_2, θ_3, θ_4, θ_5, σ, and ρ are 1000, 1000, 1000, 1000, 1000, 1000, 1000, and 995, respectively. Fig. 3 shows the autocorrelation functions for each of the θ's for the data obtained from Kamloops, BC. As usual, the generated MCMC iterations are highly autocorrelated. After thinning, the effective sample sizes for θ_0, θ_1, θ_2, θ_3, θ_4, θ_5, σ, and ρ are 1000, 1000, 1000, 479, 1000, 444, 1000, and 1000, respectively.

Figure 4 shows the trace plots of the generated regression coefficients after burn-in and thinning for Chilliwack data. The plots show that the Markov Chains achieved stationarity. Similarly, the stationarity for the MCMC iterations of the coefficients were also checked for the Kamloops data (Fig. 5).

Note that the four parameters (θ_0, θ_1, θ_2 and θ_4) in Fig. 3 decayed quickly, but not θ_3 and θ_5. This is because, for Kamloops data, only the first threshold is statistically significant. Precisely, there is no significant difference between the slopes in the second and third segments, demarcated by the threshold θ_5, represented by $\theta_1 + \theta_2$ and $\theta_1 + \theta_2 + \theta_3$ (details are reported in Fig. 6 and Table 2). From computational perspective, the estimates of θ_3 were scattered around 0 and showed greater autocorrelation. This indicates strong evidence towards insignificant second threshold effect in Kamloops, BC. The corresponding, traceplots in Fig. 5 show that the MCMC iterations mixed nicely and their distributions achieved stationarity. To our support, the ACFs of the six parameters for Kamloops data (figure not shown) after burn-in and thinning showed negligible autocorrelation between the saved MCMC iterations. Finally, after burn-in and thinning, the effective sample sizes for θ_3 and θ_5 became 479 and 444, respectively, which are reasonably large to effectively measure various statistical properties.

4 Results

For all the parameters of our model, the point estimates represent the median of the posterior distribution of the parameters calculated from the thinned MCMC samples. Moreover, the lower and upper limits of the 95% credible intervals are calculated using the 2.5th and 97.5th percentiles, respectively, of the posterior distribution.

Figure 6a shows the 95% credible intervals (equivalent to confidence intervals in classical statistics) for θ_2 and θ_3 for Chilliwack data. None of the credible interval of θ_2 and θ_3 includes 0 value inside, making the thresholds of θ_4 and θ_5 statistically significant. Table 1 shows the 95% credible intervals for the breakpoints θ_4 and θ_5. The breakpoints occurred in months of 54 (April 2015) and 87 (March 2018) with non-overlapping intervals. Hence, the two
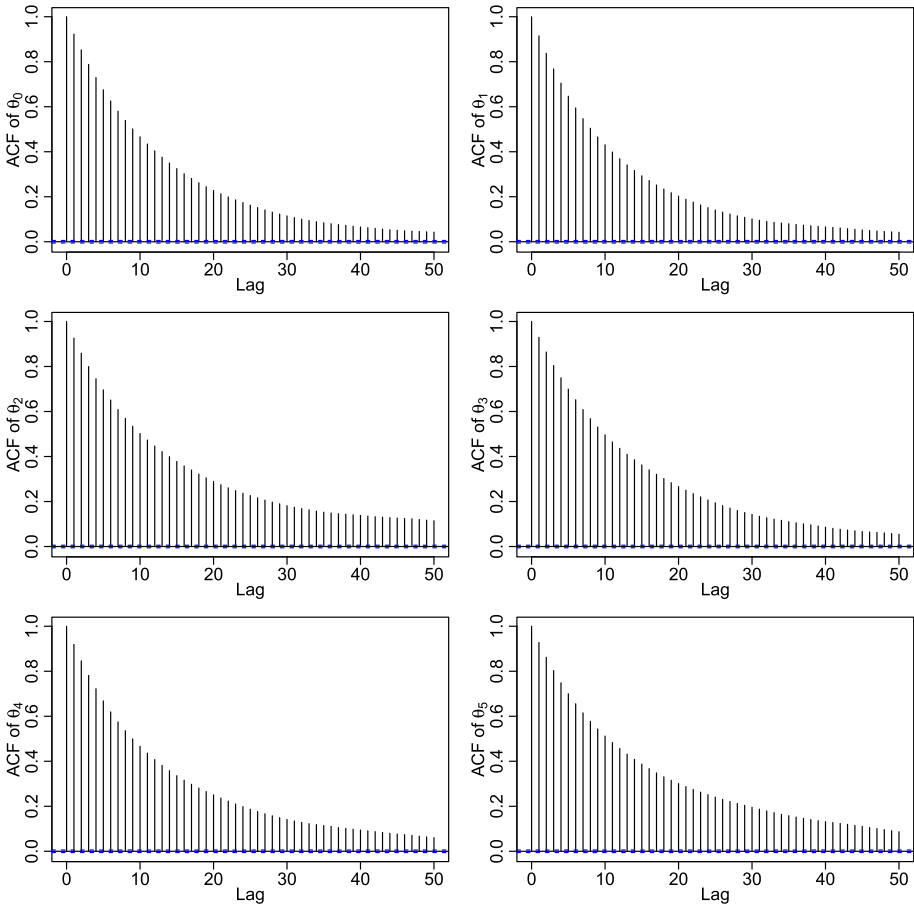
Fig. 2 Auto-correlation functions for $\boldsymbol{\theta}$ against MCMC iterations for the data from Chilliwalk, British Columbia, Canada

breakpoints are significantly different from each other. Fig. 6b and bottom part of Table 1 show the 95% credible intervals of the slopes. In the first segment of the model (from January 2011 to April 2015), the slope θ_1 is 686.75 with 95% credible interval ranging from 288.17 to 1112.54: The average price increase per month is $686.75. In the second segment of the model (from April 2015 to March 2018), the slope $\theta_1 + \theta_2$ is 5681.86 with 95% credible interval ranging from 4965.46 to 6857.23: The average price increase per month is $5681.86. In the third segment of the model (from March 2018 to July 2020), the slope $\theta_1 + \theta_2 + \theta_3$ is 1281.52 with 95% credible interval ranging from 381.15 to 2232.62: The average price increase per month is $1281.52. The price increases in the first and third segments are not significantly different as the two credible intervals overlap each other.

Figure 6c shows the 95% credible intervals for θ_2 and θ_3 for the Kamloops data. The credible interval for θ_2 does not include zero inside, and, thus, the threshold effect θ_4 is statistically significant. On the other hand, the credible interval of θ_3 includes zero inside, and the threshold effect of θ_5 is statistically insignificant. Table 2 shows the 95% credible intervals for breakpoint θ_4 which occurred in month of 58 (August 2015) with 95% credible
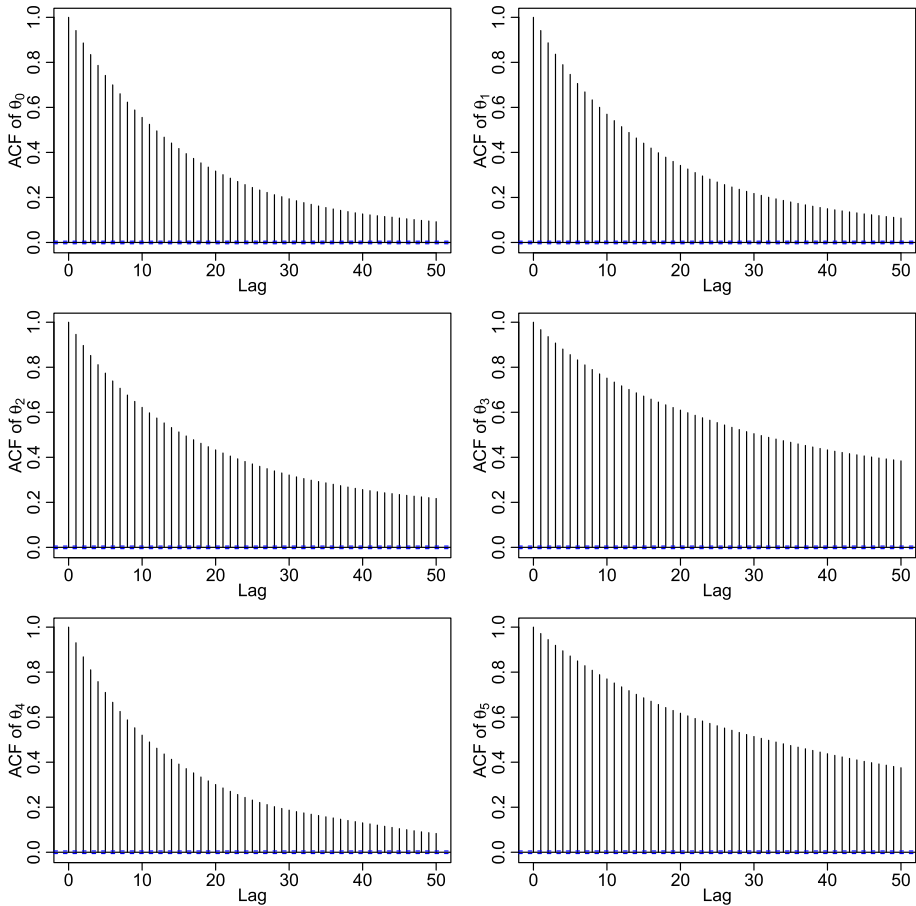
Fig. 3 Auto-correlation functions for $\boldsymbol{\theta}$ against MCMC iterations for the data from Kamloops, British Columbia, Canada

interval ranging from 50.17 to 65.82. Fig. 6d and bottom part of Table 2 show the 95% credible intervals for the slopes. In the first segment of the model (from January 2011 to August 2015), the slope θ_1 is 407.87 with 95% credible interval ranging from 161.37 to 623.11: The average price increase per month is \$407.87. In the second segment of the model (from August 2015 to February 2018), the slope $\theta_1 + \theta_2$ is 1991.66 with 95% credible interval ranging from 1449.35 to 2608.90: The average price increase per month is \$1991.66. Although the effect is insignificant in the third segment of the model defined from February 2018 to July 2020, the slope $\theta_1 + \theta_2 + \theta_3$ is 2093.71 with 95% credible interval ranging from 1500.72 to 2889.25: The average price increase per month is \$2093.71. The price increases in the second and third segments are not significantly different from each other as the two credible intervals are highly overlapping.

Having seen the numbers and figures relating the regression coefficients, we produce the expected regression model (expected home price over linear trend in time) with 95% credible intervals. Fig. 7a shows the expected home price plotted against linear trend in time (in month) for Chilliwack data highlighting the breakpoints along the horizontal axis
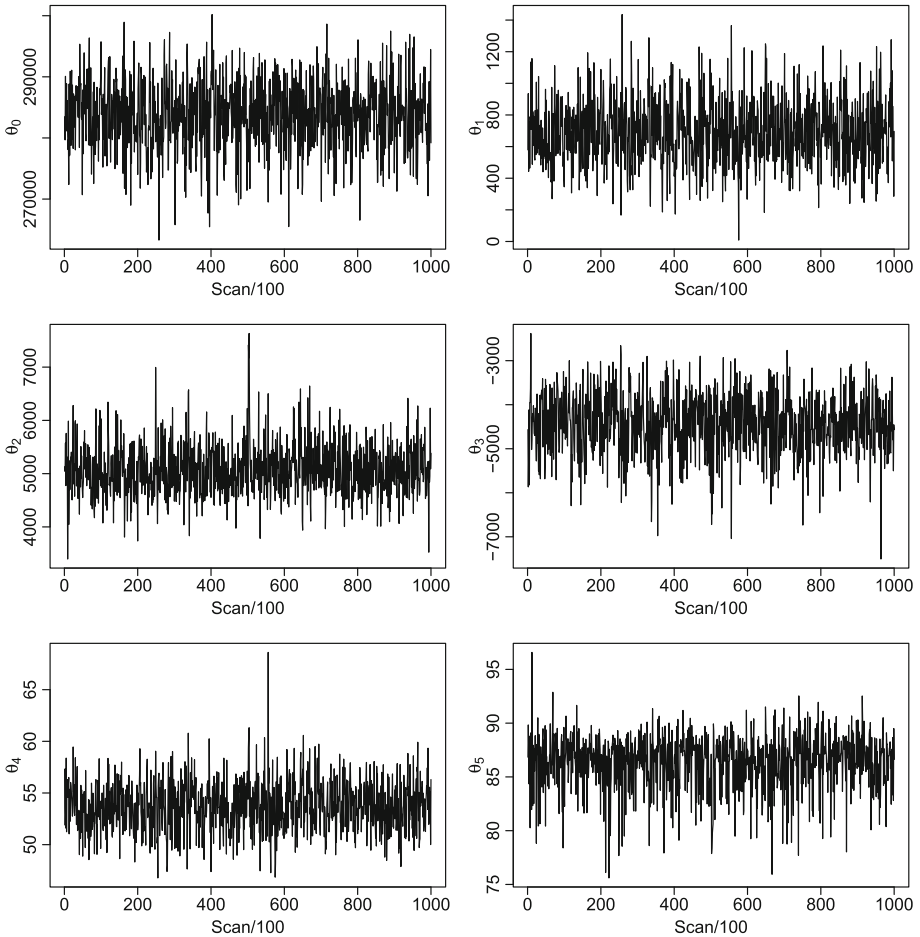
Fig. 4 Traceplots for θ against thinned MCMC iterations for the data from Chilliwack, BC

with 95% credible intervals. As noted before, the first and second breakpoints occurred in June 2015 and March 2018, respectively. After the second breakpoint, the slope in the third segment becomes similar to the slope in the first segment. Fig. 7b shows the expected home price plotted against linear trend in time (in month) for Kamloops data highlighting one breakpoint along the horizontal axis with 95% credible interval. The breakpoint occurred in August 2015 and the slope remained steep ever since.

It is interesting to note that the first breakpoint in Chilliwack occurred in June 2015 while the foreign buyer's tax was introduced in Vancouver on August 2, 2016. It appears that the debate and discussions surrounding this policy change and the resulting expectations of the impact on the cost of buying homes in Vancouver led to the shift in demand for housing in other cities in BC even before the tax was finally introduced. The breakpoint in Kamloops came with a 2-month time lag in August 2015 possibly for it being further away from Vancouver than Chilliwack.

To better understand the reasons for shifting the breakpoint in Chilliwack earlier than the date of the introduction of higher property transfer tax, we may need to consider the context
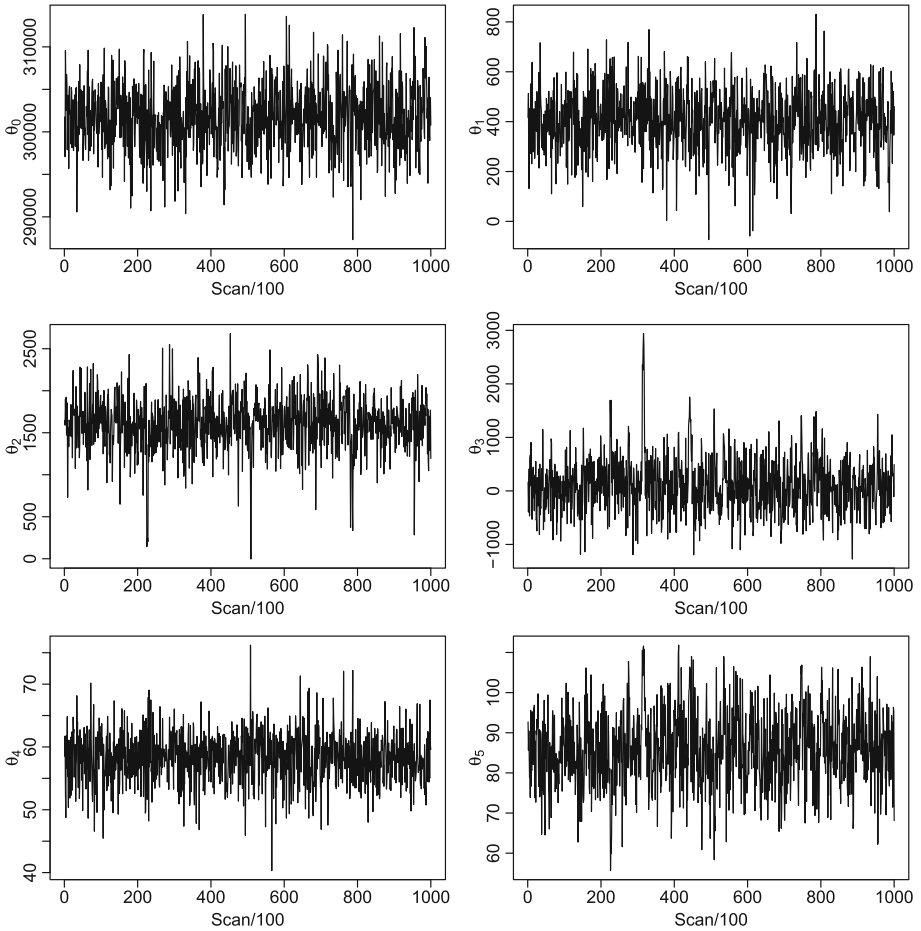
Fig. 5 Traceplots for $\boldsymbol{\theta}$ against thinned MCMC iterations for the data from Kamloops, BC

Table 1 Estimates and 95% credible intervals for breakpoints and slopes for Chilliwack data

Names	Parameter	Median	95% credible interval	
			Lower limit	Upper limit
Breakpoints	θ_4	53.66	48.99	58.75
	θ_5	86.84	80.24	90.50
Slopes	θ_1	686.75	288.17	1112.54
	$\theta_1 + \theta_2$	5681.86	4965.46	6857.23
	$\theta_1 + \theta_2 + \theta_3$	1281.52	381.15	2232.62

(a) Theta, Chilliwalk

(b) Slopes, Chilliwalk

(c) Theta, Kamloops

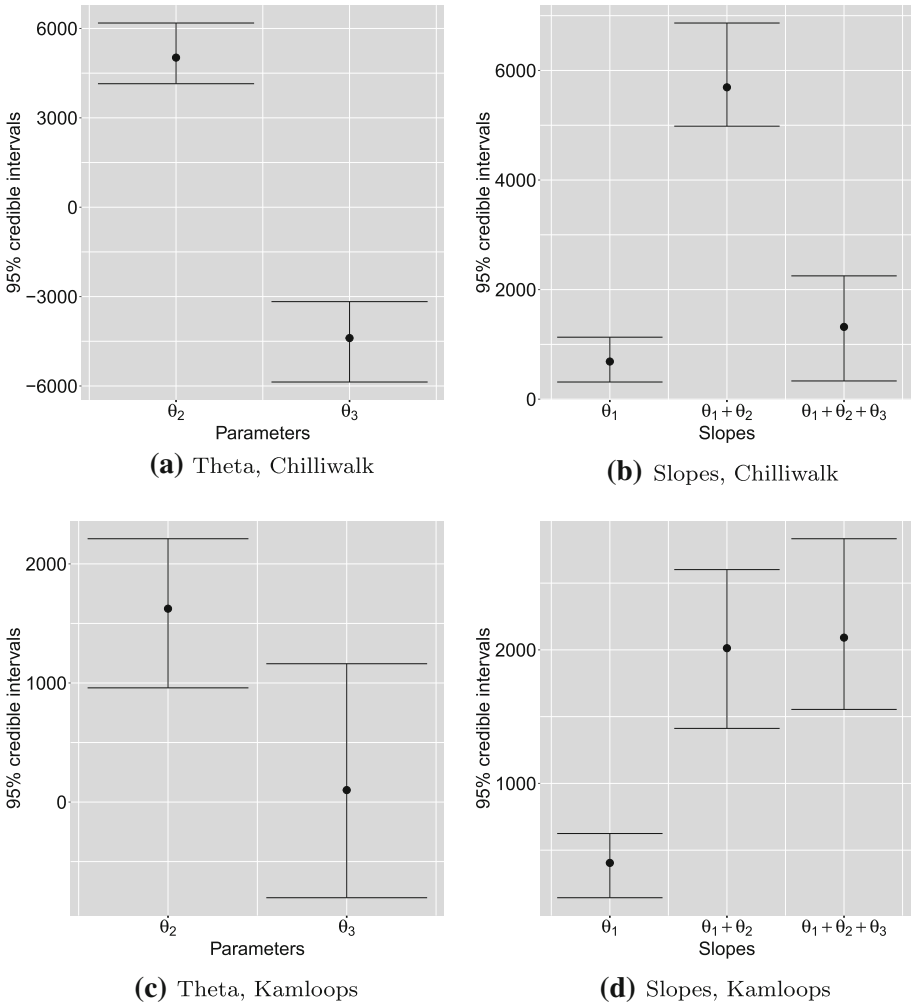**(d)** Slopes, Kamloops

Fig. 6 The 95% credible intervals for θ's and slopes

for which the Government of British Columbia introduced the higher property transfer taxes to foreign buyers. Note that the home prices in Metro Vancouver had been increasing sharply in the years preceding 2016. When the home prices already became unbearable, some buyers might have shifted their interests of buying homes from Metro Vancouver to the nearby cities like Chilliwack. This shift of interest might have caused the rapid increase of home prices in Chilliwack earlier in 2016. Later, when the Government of BC introduced the higher property transfer taxes for foreign buyers it would have caused increased influx of buyers to the neighbouring city of Chilliwack, resulting in further rapid changes in home prices there. In that light, the early structural break in home prices in Chilliwack may therefore have been caused by these additional contributing factors along with the strong expectations of the tax to be introduced by the government.

Finally, we finish the results section by reporting the numbers regarding the lag-1 autocorrelation ρ among the residuals. The median of the posterior realization of ρ is 0.274

Table 2 Estimates and 95% credible intervals for breakpoints and slopes for Kamloops data

Names	Parameter	Median	95% credible interval	
			Lower Limit	Upper Limit
Breakpoints	θ_4	58.18	50.17	65.82
Slopes	θ_1	407.87	161.37	623.11
	$\theta_1 + \theta_2$	1991.66	1449.35	2608.90
	$\theta_1 + \theta_2 + \theta_3$	2093.71	1500.72	2889.25

with 95% credible interval ranging from 0.064 to 0.485 for Chilliwack data. For Kamloops data, the median of the posterior realization of ρ is 0.168 with 95% credible interval ranging from 0.024 to 0.361. The numbers show non-negligible autocorrelation among successive residuals of the Bayesian piecewise-linear regression model.

Also, the median of the posterior standard deviation of residuals is 16,659.21 with 95% credible interval ranging from 14,449.39 to 19,674.93 for Chilliwack, BC. The median of the posterior standard deviation of residuals is 12,514.00 with 95% credible interval ranging from 10,933.89 to 14,422.36 for Kamloops, BC.

Note that the robustness of the results has been assessed by making the prior distributions of $\boldsymbol{\theta}$ further diffuse. Specifically, we multiplied the prior variances for θ_0, θ_1, θ_2 and θ_3 by 2^2, and for θ_4 and θ_5 by 4^2. These changes caused a few small changes in the results, here and there, but did not alter the inference. Therefore, we consider that the results obtained from our model are robust against changing priors from informative to non-informative.

5 Discussion

The Provincial Government of British Columbia introduced Bill 28, Miscellaneous Statutes (Housing Priority Initiatives) Amendment Act, 2016 on July 25, 2016. The introduction of an additional Property Transfer Tax on foreign entities buying property in Metro Vancouver caused threshold effects in housing prices in nearby cities such as Chilliwack, and as far as Kamloops. The threshold effect in Chilliwack happened relatively earlier than the threshold effect in Kamloops. This happened most likely because Chilliwack is situated within a closer proximity of the Metropolitan Vancouver than Kamloops which is 350 km farther. We iterate that the threshold effects in both cities occurred more or less at the same time as the two credible intervals (one to each city) highly overlapped each other.

Most importantly, on February 20, 2018, the Government of British Columbia extended the foreign buyer additional property transfer tax to the following regions: Capital Regional District, Fraser Valley Regional District, Metro Vancouver Regional District, Regional District of Central Okanagan, and Regional District of Nanaimo. These extended regions for additional property transfer tax to foreign entities include Chilliwack, but not Kamloops. As a result of the implementation of this policy Chilliwack experienced a second threshold in March 2018 with 95% credible interval extending from September 2017 to July 2018. This policy implementation from the local government of British Columbia did not cause any threshold effects via the changes in slopes in the Kamloops housing market.

The housing prices in Chilliwack show *three* linear trend segments demarcated by the two thresholds. The first threshold caused the housing prices to shoot up very quickly, whereas the second threshold caused a price correction by stabilizing the growth in housing prices back
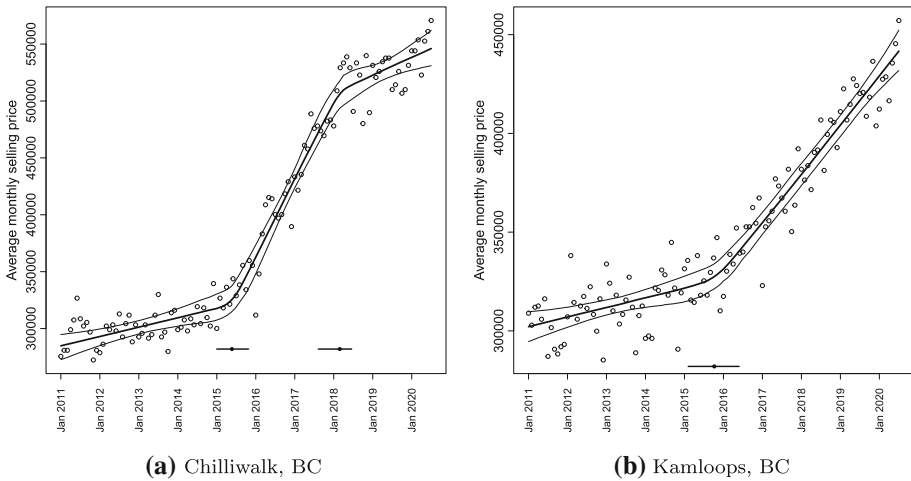
(a) Chilliwalk, BC   **(b)** Kamloops, BC

Fig. 7 Expected price with 95% credible intervals

to normal as in the first segment. In contrast, there is only one threshold in Kamloops housing market due to the change in government policy back in July 2016. The housing market in Kamloops is sky rocketing since then with rapid growth in home prices quantifying 513.33% higher after threshold effect than the price growth before October 2015.

In a recent paper, Tomal and Ciborowski [42] developed a piecewise linear regression model with two breakpoints, where the model is defined through the conditional mean of the response variable given the predictor, and applied the model to two ecological datasets to detect environmental thresholds caused by human induced disturbances to aquatic habitat for fishes in the Great Lakes Coastal Wetland. They used classical method such as non-linear least squares for parameter estimation and inference. Moreover, they developed a piecewise linear quantile regression model which is defined through the conditional quantile of the distribution of the response variable given the predictor. In this paper, we proposed the Bayesian counterpart of the two-thresholds piecewise linear regression model, defined the priors, derived the posteriors, and presented the Metropolis and Gibbs sampling algorithm to generate data from the posterior. The proposed model is further extended to incorporate non-independent correlated error terms suitable for time series data. The methods have been applied to the housing market data of two nearby cities in British Columbia, Canada to estimate the threshold effects resulting from the implementation and changes of government tax policies.

There are methods in Business and Economics, Lagrange Multiplier unit root test [9] and Bayesian model averaging [36] and GARCH model [3], which can detect thresholds in the relationships between variables. These models detect thresholds in the distribution of the time series variables, and are not very useful in terms of interpretation of the relationship. On the other hand, there are non linear autoregressive models [39,40] and partial structural change model [1] for time series data that can estimate the slopes for the relationships between variables while testing for the presence of breakpoints. However, these latter models are developed using classical statistical method. In contrast, our proposed model is fully Bayesian in nature. Besides, our proposed model comes with the strength of estimating the thresholds while performing statistical tests to determine if the estimated breakpoints are significant or not. While estimating the breakpoints, our model estimates the relationships via slopes which

are easy to interpret. Moreover, our model comes with capabilities to test if the estimated slopes in the different segments of the data are statistically significant or not. We have proved the ability of our model via two applications in housing price trend analysis. Last but not the least, the proposed model is equipped with abilities to estimate the appropriate amount of autocorrelation that is present in the data points defined through the autocorrelation of error terms.

In a recent article, Tomal et al. [43] developed a missing value imputation method for a Bayesian hierarchical model to measure the effect of COVID-19 on students' marks. In that model, the authors used conjugate prior distributions and, eventually, proposed Gibbs sampling to generate data and missing values from the posterior distributions. In this paper, we have used conjugate prior for the variance parameter σ^2, and non-conjugate priors for the regression coefficients $\boldsymbol{\theta}$ and correlation coefficient ρ for the error term. As a result, we have used Gibbs sampling for the variance parameter, and Metropolis algorithm for the regression coefficients and correlation coefficient for the error terms.

In order to initiate the algorithm, one needs to come up with a good proposal distribution to yield a reasonable acceptance rate. It is difficult to come up with a reasonable proposal variance-covariance matrix for the regression coefficients vector especially when the dimension of the matrix is large. Even if one comes with a reasonable positive definite variance-covariance matrix, a compact matrix may produce high acceptance rate and a sparse matrix may produce low acceptance rate. Note that both low and high acceptance rates generate highly autocorrelated MCMC iterations, which contain less amount of information in the generated samples. To overcome this problem, we have proposed a two-step method. In the first step, we run the algorithm for a small number of iterations and obtain a reasonable approximation for the proposal covariance matrix. The variance-covariance matrix for the regression coefficients is then used as the proposal distribution in the second step of the algorithm. The results show that this proposed method produces a nice balance in the acceptance rate and reasonable autocorrelation in the generated MCMC iterations. This method of dealing with the proposal distribution in higher dimension may become useful in many computational algorithms used in Bayesian statistical methods. Our two-step updating method for the proposal distribution is similar in nature to the adaptive Metropolis and Hastings algorithm proposed by [8,21].

Note that the other two cities within the proximity of Vancouver, British Columbia are Victoria and Kelowna. These cities have experienced their own threshold effects from the introduction of property transfer taxes by the provincial government of British Columbia. Importantly, the patterns of threshold effects in these two cities are different than those observed in Chilliwack and Kamloops. The reasons for different patterns could be that the cities of Victoria and Kelowna are much different in size and business capacity than Chilliwack and Kamloops. Also, each of them houses a different sized university than the universities in Chilliwack and Kamloops. In summary, the differences of threshold patterns in these two cities make the proposed model not directly applicable to the data from Victoria and Kelowna. However, at this moment, the authors of this paper are working on developing models to detect threshold effects in Victoria and Kelowna as well. All of these results, when available and published, might help the readers to gain a better understanding of spatial effects of the introduction of property transfer taxes on housing prices in cities nearby Vancouver, BC.

In this paper, we considered an Inverse-Gamma prior for σ^2 and a Multivariate Normal prior $\boldsymbol{\theta}$. These conjugate prior distributions made the posterior distributions simpler and enabled us generating MCMC samples using Gibbs sampling. As alternatives, one may wish to use other prior distributions as they fit instead of the provided priors. However, the use

of alternative prior distributions may complicate the computational process especially when the posterior distributions are not in closed form. In such a situation, one may need to use Metropolis or Metropolis-Hastings algorithm instead of Gibbs sampling for σ^2 and $\boldsymbol{\theta}$. On the other hand, the applications of open-source MCMC softwares, such as JAGS [13], WinBUGS [29], or Stan [7] may appear handy to improve computational issues.

The proposed methodology of this paper might appear useful to model other macroeconomic variables such as gross domestic product, economic growth, inflation, and unemployment rates of a country to detect potential threshold effects of some events defined over time using Eq. 1 after careful exploratory analysis of the data using scatterplot such as in Fig. 7 and autocorrelation function plot as in Fig. 1. To understand the effects of financial crisis on international asset diversification in real estate market, via structural breaks in increased volatility during a high turbulent period, the readers are encouraged to read the article by Gerlach et al. [20]. Furthermore, to detect structural breaks using autoregressive model with exogenous inputs (ARX) and GARCH model for the conditional variances, via volatility during a high turbulent period, we encourage readers to read the article by Than-Thi et al. [41], which examined the impact of inflation and interest rates on housing price dynamics.

Declarations

Conflict of interest No authors declare any conflicts of interest.

References

1. Bai, J., Perron, P.: Estimating and testing linear models with multiple structural changes. Econometrica **66**(1), 47–78 (1998). http://www.jstor.org/stable/2998540
2. Banner, K.M., Irvine, K.M., Rodhouse, T.J.: The use of bayesian priors in ecology: the good, the bad and the not great. Methods Ecol. Evol. **11**(8), 882–889 (2020). https://doi.org/10.1111/2041-210X.13407. https://besjournals.onlinelibrary.wiley.com/doi/abs/10.1111/2041-210X.13407
3. Begiazi, K., Katsiampa, P.: Modelling UK house prices with structural breaks and conditional variance analysis. J. Real Estate Finance Econ. **58**(2), 290–309 (2019). https://doi.org/10.1007/s11146-018-9652-5
4. Brenden, T., Wang, L., Su, Z.: Quantitative identification of disturbance thresholds in support of aquatic resource management. Environ. Manage. **42**, 821–832 (2008). https://doi.org/10.1007/s00267-008-9150-2
5. Bucci, A., Skrami, E., Faragalli, A., Gesuita, R., Cameriere, R., Carle, F., Ferrante, L.: Segmented Bayesian calibration approach for estimating age in forensic science. Biom. J. **61**(6), 1575–1594 (2019). https://doi.org/10.1002/bimj.201900016. https://onlinelibrary.wiley.com/doi/abs/10.1002/bimj.201900016
6. Burda, C.: Priced out: Understanding the factors affecting home prices in the GTA, @2013 Royal Bank of Canada and the Pembina Institute (2013). https://www.pembina.org/reports/priced-out.pdf. Accessed 15 Feb 2021
7. Carpenter, B., Gelman, A., Hoffman, M.D., Lee, D., Goodrich, B., Betancourt, M., Brubaker, M., Guo, J., Li, P., Riddell, A.: Stan: a probabilistic programming language. J. Stat. Softw. **76**(1), 1–32 (2017). https://doi.org/10.18637/jss.v076.i01. https://www.jstatsoft.org/v076/i01
8. Chen, C., So, M.: On a threshold heteroscedastic model. Int. J. Forecast. **22**, 73–89 (2006). https://doi.org/10.1016/j.ijforecast.2005.08.001

9. Chien, M.S.: Structural breaks and the convergence of regional house prices. J. Real Estate Finance Econ **40**(1), 77–88 (2010). https://doi.org/10.1007/s11146-008-9138-y

10. Clark, W.A., Dieleman, F.M.: Households and housing: choice and outcomes in the housing market, 1st edn. Routledge (2012). https://doi.org/10.4324/9780203789773

11. CMHC: Examining escalating house prices in large Canadian metropolitan centres. Canada Mortgage and Housing Corporation (2018). https://www.cmhc-schl.gc.ca/en/data-and-research/publications-and-reports/examining-escalating-house-prices-in-large-canadian-metropolitan-centres

12. Denis, J.S.: What's driving Vancouver's soaring real estate market? Vancouver is Awesome (2016). https://www.vancouverisawesome.com/courier-archive/news/whats-driving-vancouvers-soaring-real-estate-market-3026617

13. Denwood, M.J.: runjags: An R package providing interface utilities, model templates, parallel computing methods and additional distributions for MCMC models in JAGS. J. Stat. Softw. **71**(9), 1–25 (2016). https://doi.org/10.18637/jss.v071.i09

14. Draper, N.R., Smith, H.: Applied Regression Analysis. Wiley Series in Probability and Statistics. Wiley, New York (2014). https://books.google.ca/books?id=uSReBAAAQBAJ

15. Evans, M., Rosenthal, J.S.: Probability and Statistics: The Science of Uncertainty, 2 edn. W.H. Freeman & Company (2009). http://www.utstat.toronto.edu/mikevans/jeffrosenthal/book.pdf

16. Fallis, G.: The demand for housing. In: Fallis, G. (ed.) Housing Economics, pp. 27–48. Butterworth-Heinemann (1985). https://doi.org/10.1016/B978-0-409-82940-2.50005-8. https://www.sciencedirect.com/science/article/pii/B9780409829402500058

17. Fallis, G.: Housing policy in Canada. In: Fallis, G. (ed.) Housing Economics, pp. 167–187. Butterworth-Heinemann (1985). https://doi.org/10.1016/B978-0-409-82940-2.50012-5. https://www.sciencedirect.com/science/article/pii/B9780409829402500125

18. Fallis, G.: The supply of housing. In: Fallis, G. (ed.) Housing Economics, pp. 49–72. Butterworth-Heinemann (1985). https://doi.org/10.1016/B978-0-409-82940-2.50006-X. https://www.sciencedirect.com/science/article/pii/B978040982940250006X

19. Ficetola, G.F., Denoël, M.: Ecological thresholds: An assessment of methods to identify abrupt changes in species-habitat relationships. Ecography **32**, 1075–1084 (2009). https://doi.org/10.1111/j.1600-0587.2009.05571.x

20. Gerlach, R., Wilson, P., Zurbruegg, R.: Structural breaks and diversification: the impact of the 1997 Asian financial crisis on the integration of Asia–Pacific real estate markets. J. Int. Money Finance **25**(6), 974–991 (2006). https://doi.org/10.1016/j.jimonfin.2006.07.002. https://www.sciencedirect.com/science/article/pii/S0261560606000489

21. Gerlach, R.H., Chen, C.W.S., Chan, N.Y.C.: Bayesian time-varying quantile forecasting for value-at-risk in financial markets. J. Bus. Econ. Stat. **29**(4), 481–492 (2011). https://www.sciencedirect.com/science/article/pii/S0261560606000489

22. Gordon, J.C.: Solving puzzles in the Canadian housing market: foreign ownership and de-coupling in Toronto and Vancouver. Housing Stud., 1–24 (2020). https://doi.org/10.1080/02673037.2020.1842340

23. Grigoryeva, I., Ley, D.: The price ripple effect in the Vancouver housing market. Urban Geogr. **40**(8), 1168–1190 (2019). https://doi.org/10.1080/02723638.2019.1567202

24. Hossain, B., Latif, E.: Determinants of housing price volatility in Canada: a dynamic analysis. Appl. Econ. **41**(27), 3521–3531 (2009). https://doi.org/10.1080/00036840701522861

25. King, R.: Why Vancouver's real estate prices are so crazy. Money Sense (2016). https://www.moneysense.ca/spend/real-estate/8-factors-that-really-mess-up-vancouvers-real-estate-prices/

26. Lee, J., Strazicich, M.C.: Minimum Lagrange multiplier unit root test with two structural breaks. Rev. Econ. Stat. **85**(4), 1082–1089 (2003). https://doi.org/10.1162/003465303772815961

27. Ley, D., Tutchener, J.: Immigration, globalisation and house prices in Canada's gateway cities. Housing Stud. **16**(2), 199–223 (2001). https://doi.org/10.1080/02673030120038483

28. Liang, Z., Qian, S.S., Wu, S., Chen, H., Liu, Y., Yu, Y., Yi, X.: Using Bayesian change point model to enhance understanding of the shifting nutrients-phytoplankton relationship. Ecol. Model. **393**, 120–126 (2019). https://doi.org/10.1016/j.ecolmodel.2018.12.008

29. Lunn, D.J., Thomas, A., Best, N., Spiegelhalter, D.: Winbugs—a Bayesian modelling framework: concepts, structure, and extensibility. Stat. Comput. **10**, 325–337 (2000). https://doi.org/10.1023/A:1008929526011

30. Mohammad-Djafari, A., Féron, O.: A Bayesian approach to change points detection in time series. Int. J. Imaging Syst. Technol. **16**, 215–221 (2006). https://doi.org/10.1002/ima.20080

31. Muggeo, V.M.R.: Estimating regression models with unknown break-points. Stat. Med. **22**(19), 3055–3071 (2003). https://doi.org/10.1002/sim.1545. https://onlinelibrary.wiley.com/doi/abs/10.1002/sim.1545

32. Muth, R., Goodman, A.: The Economics of Housing Markets. Contemporary Concepts in Physics. Harwood Academic Publishers (1989). https://books.google.ca/books?id=-HceAQAAIAAJ

33. Nistor, A., Reianu, D.: Determinants of housing prices: evidence from Ontario cities, 2001–2011. Int. J. Hous. Markets Anal. **11**(3), 541–556 (2018). https://doi.org/10.1108/IJHMA-08-2017-0078

34. Ouyang, Z., Lin, M., Chen, J., Fan, P., Qian, S.S., Park, H.: Improving estimates of built-up area from night time light across globally distributed cities through hierarchical modeling. Sci. Total Environ. **647**, 1266–1280 (2019). https://doi.org/10.1016/j.scitotenv.2018.08.015

35. Qian, S., King, R., Richardson, C.: Two statistical methods for the detection of environmental thresholds. Ecol. Model. **166**(1), 87–97 (2003). https://doi.org/10.1016/S0304-3800(03)00097-8

36. Ravazzolo, F., Paap, R., van Dijk, D., Franses, P.H.: Chapter 15 Bayesian model averaging in the presence of structural breaks. In: Rapach, D.E., Wohar, M.E. (eds.) Forecasting in the Presence of Structural Breaks and Model Uncertainty (Frontiers of Economics and Globalization, vol. 3), vol. 3. Emerald Group Publishing Limited, Bingley (2008). https://doi.org/10.1016/S1574-8715(07)00215-1

37. Ruggieri, E.: A Bayesian approach to detecting change points in climatic records. Int. J. Climatol. **33**(2), 520–528 (2013). https://doi.org/10.1002/joc.3447

38. Shea, E., Vecchione, M.: Quantification of ontogenetic discontinuities in three species of oegopsid squids using model II piecewise linear regression. Mar. Biol. **140**(5), 971–979 (2002). https://doi.org/10.1007/s00227-001-0772-7

39. Teräsvirta, T.: Specification, estimation, and evaluation of smooth transition autoregressive models. J. Am. Stat. Assoc. **89**(425), 208–218 (1994). http://www.jstor.org/stable/2291217

40. Teräsvirta, T., Yang, Y.: Specification, estimation and evaluation of vector smooth transition autoregressive models with applications. LIDAM Discussion Papers CORE 2014062, Université catholique de Louvain, Center for Operations Research and Econometrics (CORE) (2014). https://ideas.repec.org/p/cor/louvco/2014062.html

41. Than-Thi, H., Dong, M.C., Chen, C.W.S.: Bayesian modelling structural changes on housing price dynamics. In: Kreinovich, V., Sriboonchitta, S. (eds.) Structural Changes and their Econometric Modeling, pp. 83–104. Springer International Publishing, Cham (2019). https://doi.org/10.1007/978-3-030-04263-9_6

42. Tomal, J., Ciborowski, J.J.: Ecological models for estimating breakpoints and prediction intervals. Ecol. Evol. **13500–13517**, 13500–13517 (2020). https://doi.org/10.1002/ece3.6955

43. Tomal, J., Rahmati, S., Boroushaki, S., Jin, L., Ahmed, E.: The impact of COVID-19 on students' marks: a Bayesian hierarchical modeling approach. Metron **79**, 57–91 (2021). https://doi.org/10.1007/s40300-021-00200-1

44. Toms, J., Lesperance, M.: Piecewise regression: a tool for identifying ecological thresholds. Ecology **84**(8), 2034–2041 (2003). https://doi.org/10.1890/02-0472

45. Toms, J.D., Villard, M.A.: Threshold detection: matching statistical methodology to ecological questions and conservation planning objectives. Avian Conserv. Ecol. **10**(1), 2 (2015). https://doi.org/10.5751/ACE-00715-100102

46. Valsamis, E., Husband, H., Chan, G.: Segmented linear regression modelling of time-series of binary variables in healthcare. Comput. Math. Methods Med. (2019). https://doi.org/10.1155/2019/3478598

47. Vizek, M., Posedel, P.: Are house prices characterized by threshold effects? evidence from developed and post-transition countries. Czech J. Econ. Finance (Finance Uver) **61**, 584–600 (2011)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.