TI 2023-039/III
Tinbergen Institute Discussion Paper

Time-varying effects of housing attributes and economic environment on housing prices

Marina Friedrich1[1,3]
Yicong Lin[1,3]
Pavitram Ramdaras[1]
Sean Telg[1,3]
Bernhard van der Sluis[2]

1 Department of Econometrics and Data Science, Vrije Universiteit Amsterdam

2 Department of Econometrics, Erasmus University Rotterdam

3 Tinbergen Institute

Time-varying effects of housing attributes and economic environment on housing prices

Marina Friedrich[1,2], Yicong Lin*[1,2], Pavitram Ramdaras[1], Sean Telg[1,2] and Bernhard van der Sluis[3]

[1]Department of Econometrics and Data Science, Vrije Universiteit Amsterdam

[2]Tinbergen Institute

[3]Department of Econometrics, Erasmus University Rotterdam

June, 2023

Abstract

We propose a flexible framework that allows for the relationship between housing prices and their determinants to vary over time. Our model incorporates housing-specific characteristics and macroeconomic variables, while accounting for a gradual global trend that reflects the unobserved external environment. We estimate the trend and coefficient curves by local linear estimation and propose a bootstrap procedure for conducting inference. By employing monthly data from the Dutch housing market, covering 60 municipalities from 2006 to 2020, the proposed models show the capability to accurately describe the comovements of housing prices. Our results show strong statistical evidence of time variation in the effects of housing attributes and macroeconomic variables on prices throughout the entire sample period, revealing that the unemployment rate plays a crucial role between approximately 2012 and 2017. The extracted latent global trend reveals a significant influence of the economic environment and takes the shape of a leading indicator of the property market index. Moreover, we find that both the housing characteristics and the external environment explain comparably high proportions of the variation in housing prices, which stresses the importance of including both components in empirical analyses.

Keywords: housing prices, time-varying panels, nonparametric estimation, autoregressive wild bootstrap, simultaneous bands

JEL Classification: C14, C15, C23, R30

*Corresponding author: Department of Econometrics and Data Science, Vrije Universiteit Amsterdam, De Boelelaan 1105, 1081 HV, Amsterdam, the Netherlands. E-mail address: yc.lin@vu.nl.

1 Introduction

The housing market exerts a significant impact on the overall economy through various channels. It is therefore crucial for policymakers to gain insights into the dynamics of housing prices to ensure market stability. There is a large body of literature that investigates the relationship between housing prices and their determinants (see, e.g., Ortalo-Magne and Rady (2006), Adams and Füss (2010), Iacoviello and Neri (2010), Fuerst and Warren-Myers (2018), Chen et al. (2022), Møller et al. (2023)), but almost exclusively based on fixed parameter models. This assumption appears rather restrictive, especially during the past two decades, in which the European housing market has been exposed to multiple major global events such as the financial crisis (2007-2008), the migrant crisis (2015), and the Covid-19 pandemic (2019-2022).[1] Therefore, we propose a time-varying panel model that explains housing prices by both hedonic attributes and macroeconomic drivers in the presence of a latent global trend. A bootstrap procedure is proposed to construct confidence bands, following the methods in Friedrich et al. (2020) and Friedrich and Lin (2022). These bands allow one to infer whether relationships have remained stable over the considered time period. The strength of the model lies in its flexibility and interpretability. Specifically, it avoids imposing any specific functional form on the parameter curves, while still providing interpretable results that may facilitate policymaking. Despite its linear formulation in terms of explanatory variables, the model effectively captures the nonlinear dynamics of housing prices for each cross-sectional unit and exploits the co-moving behavior across the units.

To illustrate the points above, we display the logarithm of the real monthly housing price index in the Netherlands over the period 2006-2020 in Figure 1. The left panel shows the development of these time series for 60 different municipalities, and it is evident that they exhibit a strong co-movement. For almost all series, an overall trend of decreasing prices until approximately 2013 is followed by a sharp ongoing upward trend. In the right panel, the cross-sectional average of these prices is depicted, which shows an asymmetric V-shape. These dynamics may be partially explained by the occurrence of various major global events as mentioned earlier. Some clear fluctuations from the trend are visible, such as the major drops in housing prices around 2008 and 2012, and the stark decline at the start of 2020 that is immediately followed by a steep increase.

A possible strategy to incorporate these dynamics of housing prices into a model is by allowing for the presence of trend breaks. Nevertheless, this strategy is prone to misleading inference since both the number and location of breaks are typically unknown to the researcher and have to be consistently estimated.[2] Therefore, it is common to base the analysis on data with a lower

[1] Discussions on the impact of pandemics on urban housing markets over the short- and long-term can be found, for instance, in Francke and Korevaar (2021).

[2] In the presence of a deterministic trend, Yang (2017) finds that consistent estimates of trend break locations

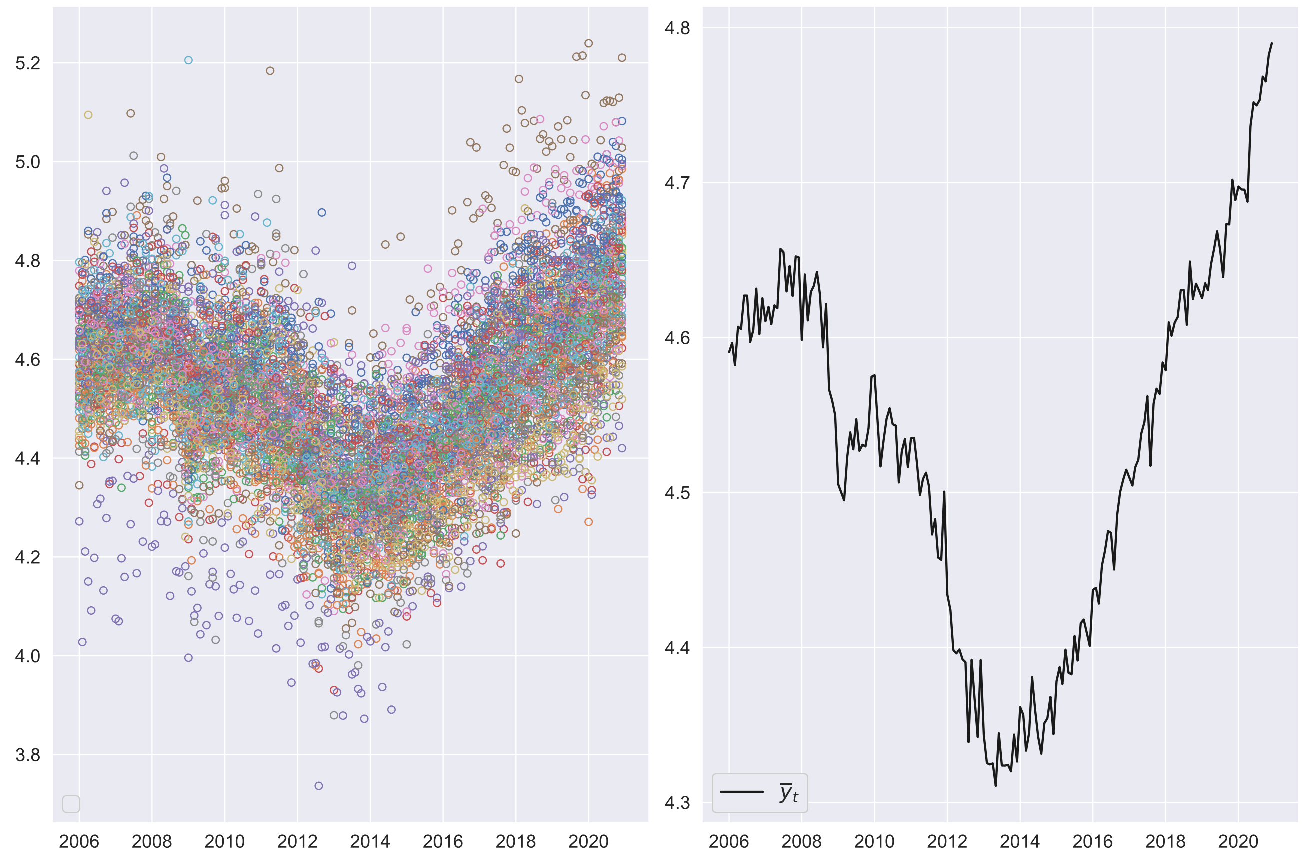


Figure 1: Housing prices y_{it} (circles) in the Netherlands deflated by CPI and adjusted to the log scale for 60 municipalities ranges from January 2006 until December 2020. The solid line, $\bar{y}_t$, on the right side, represents the cross-sectional averages.

frequency, such as annual averages. While this approach simplifies the modeling framework, it may come at the cost of losing a better understanding of price dynamics in the short run. Indeed, empirical implications can be more insightful to different agents if one can conduct inference and perform forecasts using the information at higher frequencies.

A growing literature has pointed out time-varying relations in the housing market (see e.g., Brown et al., 1997; Gelain and Lansing, 2014; Jordà et al., 2015), but to the best of our knowledge, most papers model individual time series separately and consequently ignore the changing environment shared by individuals. As shown in Figure 1, it appears more natural to model them jointly. Current joint studies include but are not limited to, spatial approaches (Holly et al., 2010; Baltagi et al., 2015; Paci et al., 2020), cointegration analyses (Mikhed and Zemcik, 2009), and structural break models (Boldea et al., 2020). However, existing analyses of housing prices either overlook the time-varying relations or allow for them but neglect common trending behaviors. Similar to the Dutch housing market, various studies have pointed out a range of common patterns in the housing markets in other countries. For instance, studies such as Fang et al. (2015) have found similar trends in housing prices among different tiers of cities in China. Additionally, Knoll et al.

can only be guaranteed if the number of breaks is correctly specified.

(2017) have examined annual house prices in 14 advanced economies since 1870 and discovered a hockey-stick pattern in the long-run paths of these prices, which has become one of the recognized stylized facts. Therefore, it is crucial to establish a joint modeling framework that incorporates the co-movements of cross-sectional units and allows for time-varying relationships.

To this end, we propose a flexible global-local panel model with time-varying coefficients, which allows for the decomposition of housing prices into *local* and *global* components. The local component consists of movements induced by hedonic attributes, such as observable house characteristics. The global component represents overall trending movements and consists of variables that are indicative of the macro environment as well as a latent global component. This unobserved factor encompasses the external environment and is able to absorb the influences of other key macroeconomic indicators that are possibly not available to the researcher (at the considered frequency), such as technological progress considered in Iacoviello and Neri (2010). By constructing confidence bands around the time-varying coefficients, we are able to investigate whether the relation between the regressors and the housing prices has remained stable over the considered time interval. In this way, the proposed model offers a powerful and unified framework for modeling housing prices, in which the relative importance of the local and global components can be assessed due to the explicit decomposition of the model.

In order to determine the importance of drivers, we consider two models in our analysis and we refer to the specification including all factors as the full model and the one excluding macro variables as the base model. We find that the simple base model adequately captures the asymmetric V-shape of the average housing prices displayed in Figure 1, except for 2012-2017, in which the average is overestimated. The inclusion of macro variables leads to three important findings. First, we observe that the coefficients of some local drivers appear constant over time when we control for the effects of the macro variables. This reveals that restricting the analysis to solely hedonic attributes and ignoring time variation might lead to misleading results. The importance of both components is also reflected by the finding that both components show similar proportions of explained variation. Second, the latent global component clearly changes patterns. Whereas it closely mimics the quadratic pattern of housing prices in the base model, it shows an upward trend in the full model. This change in global trend underlines the importance of a flexible specification, as the decomposition in local and global components might be misspecified when a deterministic linear or polynomial trend is imposed a priori. Finally, the asymmetric V-shaped pattern in Figure 1 is more accurately described during 2012-2017, largely due to the inclusion of the unemployment rate.

The paper is structured in the following way. Section 2 introduces the global-local panel model and outlines the estimation and inference procedures. Section 3 provides details about the data used in the analysis. The empirical results are presented in Section 4. Section 5 concludes the paper.

2 The modeling framework

Housing prices are determined by a large spectrum of quantities which we summarize in two main components: local and global. The local component is widely studied in the hedonic pricing literature. It encompasses heterogeneous attributes of a house such as lot size and the number of rooms, which cannot be bought and sold on a market separately. Whereas the local component aims at explaining the variation per house separately, as shown in Figure 1, there are also common fluctuations in housing prices and these are captured by the so-called global component. Houses are commonly subject to the same economic conditions, and this global factor, whether it is observable or latent, captures the primary economic and financial indicators.

2.1 The global-local panel model

The decomposition above can be assessed empirically using a linear panel regression model. In this model, both the macro and local components are quantified as linear combinations of multiple observed variables. Each of these observed variables is allowed to have a time-varying impact on housing prices. For cross-sectional units, say municipalities, $i = 1, \ldots, N$, and time $t = 1, \ldots, T$, our proposed model is as follows:

$$y_{it} = \alpha_i + g_t + \sum_{j=1}^{d_1} \beta_{t,j} x_{it,j} + \sum_{j=1}^{d_2} \gamma_{t,j} w_{t,j} + e_{it} = \alpha_i + g_t + \boldsymbol{x}_{it}' \boldsymbol{\beta}_t + \boldsymbol{w}_t' \boldsymbol{\gamma}_t + e_{it}, \tag{2.1}$$

where y_{it} represents the natural logarithm of housing prices, $d_1, d_2 \in \mathbb{Z}^+$ and $\boldsymbol{x}_{it} = (x_{it,1}, \ldots, x_{it,d_1})'$ stacks the individual-specific explanatory variables. An important characteristic of the model is that the explanatory variables are allowed to have bounded, deterministic trends (Chen et al., 2012; Chen and Huang, 2018). Such flexibility is crucial in our application as will be seen later in Figure 3. Moreover, the (unobservable) heterogeneous effects are captured by α_i which are allowed to correlate with $\boldsymbol{x}_{it}$, i.e., fixed effects.

Figure 1 shows that some global comovements exist in the housing market. As such, we use the vector $\boldsymbol{w}_t = (w_{t,1}, \ldots, w_{t,d_2})'$ for *observable* macroeconomic variables that possibly explain the global behavior, for instance, unemployment rate and mortgage rate. Moreover, g_t is an *unobserved*, representative component, driving the overall level of housing prices (y_{it}) that is not explained by $\boldsymbol{w}_t'\boldsymbol{\gamma}_t$, e.g., the external environment possibly due to policy changes. Finally, the time-varying coefficients are stacked in the vectors $\boldsymbol{\beta}_t = (\beta_{t,1}, \ldots, \beta_{t,d_1})'$ and $\boldsymbol{\gamma}_t = (\gamma_{t,1}, \ldots, \gamma_{t,d_2})'$. Model (2.1) separates the global trending effects $g_t + \boldsymbol{w}_t'\boldsymbol{\gamma}_t$ of housing prices from the local drivers $\boldsymbol{x}_{it}'\boldsymbol{\beta}_t$. Our primary interest lies in the estimation and inference of both the global component g_t and the time-varying coefficients $(\boldsymbol{\beta}_t, \boldsymbol{\gamma}_t)$.

2.2 Nonparametric estimation

As observed in Figure 1, global trends typically vary slowly over time. We therefore assume $g_t = g(t/T)$, where $g(\cdot) : [0,1] \to \mathbb{R}$ is a nonparametric, smooth function. Similarly, $\boldsymbol{\beta}_t = \boldsymbol{\beta}(t/T)$ and $\boldsymbol{\gamma}_t = \boldsymbol{\gamma}(t/T)$, where $\boldsymbol{\beta}(\cdot) = (\beta_1(\cdot), \ldots, \beta_{d_1}(\cdot))' : [0,1] \to \mathbb{R}^{d_1}$ and $\boldsymbol{\gamma}(\cdot) = (\gamma_1(\cdot), \ldots, \gamma_{d_2}(\cdot))' : [0,1] \to \mathbb{R}^{d_2}$ are vectors of unknown smooth functions, respectively. Alternatively, one can consider specifying $(g_t, \boldsymbol{\beta}_t, \boldsymbol{\gamma}_t)$ as latent processes in a state-space representation and estimate them using methods such as the Kalman filter or score-driven filter (Durbin and Koopman, 2012; Creal et al., 2013). Since it is not possible to separately identify g_t from the fixed effects α_i without imposing any condition, we enforce a commonly used identification condition $\sum_{i=1}^N \alpha_i = 0$, see, e.g., Eq. (1.3) in Chen et al. (2012).

To estimate the coefficient curves, we adapt the local linear dummy variable (LLDV) estimation proposed in Li et al. (2011). The main idea behind the LLDV method is that any smooth (twice continuously differentiable) function at any fixed $\tau \in (0,1)$ can be approximated using observations around τ in the spirit of the Taylor series.[3] More specifically, for any $\tau \in (0,1)$ and $\tau_t = t/T$,

$$
\begin{aligned}
&y_{it} \approx \alpha_i + \boldsymbol{z}_{it}(\tau)'\boldsymbol{\theta}(\tau) + e_{it}, \qquad \boldsymbol{z}_{it} = (1, \boldsymbol{x}_{it}', \boldsymbol{w}_t', \tau_t - \tau, (\tau_t - \tau)\boldsymbol{x}_{it}', (\tau_t - \tau)\boldsymbol{w}_t')', \\
&\boldsymbol{\theta}(\tau) = \left(g(\tau), \boldsymbol{\beta}(\tau)', \boldsymbol{\gamma}(\tau)', g^{(1)}(\tau), \boldsymbol{\beta}^{(1)}(\tau)', \boldsymbol{\gamma}^{(1)}(\tau)'\right)' \in \mathbb{R}^{2(1+d_1+d_2)},
\end{aligned}
\tag{2.2}
$$

where $g^{(1)}(\cdot)$, $\boldsymbol{\beta}^{(1)}(\cdot)$, and $\boldsymbol{\gamma}^{(1)}(\cdot)$, are the first-order derivatives of $g(\cdot)$, $\boldsymbol{\beta}(\cdot)$, and $\boldsymbol{\gamma}(\cdot)$, respectively. Then the LLDV estimator minimizes the following weighted quadratic loss function:

$$
\widehat{\boldsymbol{\theta}}(\tau) = \arg\min_{\boldsymbol{\theta}(\tau)} \sum_{i=1}^N \sum_{t=1}^T \left[y_{it} - \alpha_i - \boldsymbol{z}_{it}(\tau)'\boldsymbol{\theta}(\tau)\right]^2 K\left(\frac{\tau_t - \tau}{h}\right), \qquad \tau \in (0,1), \tag{2.3}
$$

where $K(\cdot)$ is a kernel function and $h \downarrow 0$ is a bandwidth.

The LLDV estimator $\widehat{\boldsymbol{\theta}}(\cdot)$ has a closed-form expression, and thus the minimization problem (2.3) is numerically straightforward. To save space, we provide the detailed steps in Supplemental Appendix B. Moreover, it is worth noting that nonparametric estimation is generally not sensitive

[3] We refer the interested reader to Cai (2007) and Friedrich and Lin (2022) for the univariate case.

to the kernel. Therefore, the common Epanechnikov kernel is adopted in our estimation. That is, we take $K(x) = 3/4(1-x^2)\mathbb{1}\{|x| \leq 1\}$, where $\mathbb{1}\{\cdot\}$ is an indicator function. For the selection of the bandwidth parameter h, we shall use a data-driven procedure which we describe in the next section.

2.3 Bandwidth selection

In practice, the bandwidth shall be carefully chosen. On the one hand, an overlarge bandwidth parameter h may cause oversmoothing, and thus large estimation bias, which can be problematic for inference. This is because a large value of h likely overlooks the local features of the functions, see Section 4.1 in Friedrich and Lin (2022) for further discussion. On the other hand, a relatively small bandwidth leads to a large variance in estimation. Therefore, the optimal bandwidth should result in the best tradeoff between bias and variance. Besides these potential problems, the presence of unknown fixed effects causes the conventional leave-one-out method in time series to fail in providing satisfactory results. As such, we adapt the leave-one-unit-out cross-validation (CV) method developed in Sun et al. (2009). This procedure removes a cross-sectional unit $\{(y_{it}, \boldsymbol{x}_{it}, \boldsymbol{w}_t)\}_{t=1}^{T}$, $i = 1, \ldots, N$, from the dataset each time, and use the remaining $(N-1)T$ observations to estimate $\boldsymbol{\theta}(\tau)$, denoted by

$$\widehat{\boldsymbol{\theta}}_{(-i)}(\tau) = \left(\hat{g}_{(-i)}(\tau), \widehat{\boldsymbol{\beta}}_{(-i)}(\tau)', \widehat{\boldsymbol{\gamma}}_{(-i)}(\tau)', \hat{g}^{(1)}_{(-i)}(\tau), \widehat{\boldsymbol{\beta}}^{(1)}_{(-i)}(\tau)', \widehat{\boldsymbol{\gamma}}^{(1)}_{(-i)}(\tau)'\right)'.$$

The optimal bandwidth is selected such that it minimizes a weighted squared loss function, i.e.,

$$\hat{h}_{\text{opt}} = \underset{h\in[h_L,h_U]}{\arg\min} \left\| \boldsymbol{M}_{\mathcal{D}} \left(\boldsymbol{y} - \widehat{\boldsymbol{g}}_{(-)} - \boldsymbol{B}_1\left(\boldsymbol{X}, \widehat{\boldsymbol{\beta}}_{(-)}\right) - \boldsymbol{B}_2\left(\boldsymbol{w}, \widehat{\boldsymbol{\gamma}}_{(-)}\right)\right)\right\|^2, \quad 0 < h_L < h_U < \infty, \tag{2.4}$$

where $\|\cdot\|$ is the Euclidean norm, $\boldsymbol{M}_{\mathcal{D}} = \boldsymbol{I}_{NT} - T^{-1}\boldsymbol{I}_N \otimes (\boldsymbol{\imath}_T\boldsymbol{\imath}_T')$ is a residual-maker matrix that eliminates the fixed effects α_i from Eq. (2.4), $\boldsymbol{I}_K$ stands for a $K \times K$ identity matrix, and $\boldsymbol{\imath}_T = (1, \ldots, 1)'$ is a T-dimensional vector of ones. Moreover, $\boldsymbol{y} = (y_{11}, \ldots, y_{1T}, y_{21}, \ldots, y_{2T}, \ldots, y_{N1}, \ldots, y_{NT})'$, $\widehat{\boldsymbol{g}}_{(-)} = \left(\hat{g}_{(-1)}(1/T), \ldots, \hat{g}_{(-1)}(T/T), \hat{g}_{(-2)}(1/T), \ldots, \hat{g}_{(-2)}(T/T), \ldots, \hat{g}_{(-N)}(1/T), \ldots, \hat{g}_{(-N)}(T/T)\right)'$,

$$\begin{aligned}\boldsymbol{B}_1\left(\boldsymbol{X}, \widehat{\boldsymbol{\beta}}_{(-)}\right) = \Big(&\boldsymbol{x}_{11}'\widehat{\boldsymbol{\beta}}_{(-1)}(1/T), \ldots, \boldsymbol{x}_{1T}'\widehat{\boldsymbol{\beta}}_{(-1)}(T/T), \boldsymbol{x}_{21}'\widehat{\boldsymbol{\beta}}_{(-2)}(1/T), \ldots, \boldsymbol{x}_{2T}'\widehat{\boldsymbol{\beta}}_{(-2)}(T/T),\\ &\boldsymbol{x}_{N1}'\widehat{\boldsymbol{\beta}}_{(-N)}(1/T), \ldots, \boldsymbol{x}_{NT}'\widehat{\boldsymbol{\beta}}_{(-N)}(T/T)\Big)',\\ \boldsymbol{B}_2\left(\boldsymbol{w}, \widehat{\boldsymbol{\gamma}}_{(-)}\right) = \big(&\boldsymbol{w}_1'\widehat{\boldsymbol{\gamma}}_{(-1)}(1/T), \ldots, \boldsymbol{w}_T'\widehat{\boldsymbol{\gamma}}_{(-1)}(T/T), \boldsymbol{w}_1'\widehat{\boldsymbol{\gamma}}_{(-2)}(1/T), \ldots, \boldsymbol{w}_T'\widehat{\boldsymbol{\gamma}}_{(-2)}(T/T),\\ &\boldsymbol{w}_1'\widehat{\boldsymbol{\gamma}}_{(-N)}(1/T), \ldots, \boldsymbol{w}_T'\widehat{\boldsymbol{\gamma}}_{(-N)}(T/T)\big)'.\end{aligned}$$

The second term within the norm in Eq. (2.4) can be viewed as a compact form for the difference between housing prices and the estimated sum of local and global components by repeatedly leaving out one cross-sectional unit.

We take $h_L = 0.1$ and $h_U = 0.6$ with a step size 0.01 in Eq. (2.4). As pointed out by Friedrich et al. (2020) and Friedrich and Lin (2022), conventional CV methods have to be applied with care since they often select local minima; visual inspection of coefficient estimates is often necessary. Our investigation shows that the leave-one-unit-out CV selects reasonable bandwidths in the current application (see Figure 13, Supplemental Appendix C).

2.4 Simulation-based inference

Pointwise confidence intervals and simultaneous confidence bands are used to quantify the estimation uncertainty around the coefficient curves. More specifically, pointwise intervals $I_{j,N,T,\alpha}(\tau)$ for $\beta_j(\tau)$, are constructed to satisfy

$$\liminf_{T\to\infty,\, N\to\infty} \mathbb{P}\Big(\beta_j(\tau) \in I_{j,N,T,\alpha}(\tau)\Big) \geq 1-\alpha, \qquad \tau \in (0,1). \tag{2.5}$$

Namely, $I_{j,N,T,\alpha}(\tau)$ are only statistically valid for a given time point $\tau \in (0,1)$. However, if one aims to determine the overall variation of the coefficient curves, it is not sufficient to use pointwise confidence intervals. For instance, statistical statements such as a coefficient curve remaining zero over time, or having an upward trend over a certain period, cannot be answered with pointwise intervals: simultaneous confidence bands have to be constructed for these purposes. More specifically, for a given set of time points G, simultaneous confidence bands $I^G_{j,N,T,\alpha}(\cdot)$ satisfy

$$\liminf_{T\to\infty,\, N\to\infty} \mathbb{P}\Big(\beta_j(\tau) \in I^G_{j,N,T,\alpha}(\tau),\ \forall \tau \in G\Big) \geq 1-\alpha. \tag{2.6}$$

Unfortunately, as shown in Chen et al. (2012) and Chen and Huang (2018), the pointwise asymptotic distribution of the LLDV estimator $\widehat{\boldsymbol{\theta}}(\cdot)$ relies on various nuisance parameters such as the second-order bias and long-run covariance matrices. The estimation of these nuisance parameters is highly challenging in nonparametric settings, complicating the use of asymptotic results to conduct inference (Friedrich and Lin, 2022). Moreover, a simultaneous construction of confidence bands based on asymptotic results is currently unavailable in nonparametric panel models.

To circumvent this problem, we consider an autoregressive wild bootstrap (AWB) method, initially proposed by Smeekes and Urbain (2014) for multivariate unit root tests. The AWB can potentially handle serial dependence and heteroskedasticity. The following procedure is a straightforward modification of Friedrich et al. (2020) and Friedrich and Lin (2022) for nonparametric time series models:

S1 Let $\widetilde{\alpha}_i$, $\widetilde{g}(\cdot)$, $\widetilde{\boldsymbol{\beta}}(\cdot)$, and $\widetilde{\boldsymbol{\gamma}}(\cdot)$ be the LLDV estimates described in Section 2.2, but using a larger bandwidth $\tilde{h} > h$. Obtain residuals

$$\widetilde{e}_{it} = y_{it} - \widetilde{\alpha}_i - \widetilde{g}\left(t/T\right) - \boldsymbol{x}_{it}'\widetilde{\boldsymbol{\beta}}(t/T) - \boldsymbol{w}_t'\widetilde{\boldsymbol{\gamma}}(t/T), \qquad i = 1, \ldots, N, \quad t = 1, \ldots, T.$$

S2 For $\gamma \in (0,1)$, generate scalar sequence $\nu_1^*, \ldots, \nu_T^*$ as i.i.d. $\mathcal{N}\left(0, 1-\gamma^2\right)$ and let $\xi_t^* = \gamma\xi_{t-1}^* + \nu_t^*$, $t = 2, \ldots, T$, where $\xi_1^* \sim \mathcal{N}(0,1)$.

S3 Calculate the bootstrap errors $e_{it}^* = \xi_t^*\widetilde{e}_{it}$ and generate the bootstrap observations by

$$y_{it}^* = \widetilde{\alpha}_i + \widetilde{g}\left(t/T\right) + \boldsymbol{x}_{it}'\widetilde{\boldsymbol{\beta}}(t/T) + \boldsymbol{w}_t'\widetilde{\boldsymbol{\gamma}}(t/T) + e_{it}^*, \quad i = 1, \ldots, N, \qquad t = 1, \ldots, T,$$

where $\widetilde{\alpha}_i$, $\widetilde{g}(\cdot)$, $\widetilde{\boldsymbol{\beta}}(\cdot)$, and $\widetilde{\boldsymbol{\gamma}}(\cdot)$, are the same estimates given in Step S1.

S4 Using $\left\{\left(y_{it}^*, \boldsymbol{x}_{it}, \boldsymbol{w}_t\right), i = 1, \ldots, N, t = 1, \ldots, T\right\}$, obtain the bootstrap LLDV estimates $\widehat{\alpha}_i^*$, $\widehat{g}^*(\cdot)$, $\widehat{\boldsymbol{\beta}}^*(\cdot)$, and $\widehat{\boldsymbol{\gamma}}^*(\cdot)$ with the same bandwidth h as used for the original estimates.

S5 Repeat Step S2 to Step S4 B times, and let

$$\widehat{q}_{j,\alpha}(\tau) = \inf\left\{u \in \mathbb{R}:\ \mathbb{P}^*\left(\widehat{\beta}_j^*(\tau) - \widetilde{\beta}_j(\tau) \leq u\right) \geq \alpha\right\}, \qquad j = 1, \ldots, d_1, \tag{2.7}$$

denote for the 100αth percentile of the B centered bootstrap statistics $\widehat{\beta}_j^*(\tau) - \widetilde{\beta}_j(\tau)$, similarly for $\widehat{g}^*(\tau) - \widetilde{g}(\tau)$ and $\widehat{\gamma}_j^*(\tau) - \widetilde{\gamma}_j(\tau)$ with $j = 1, \ldots, d_2$.[4] These bootstrap quantiles are then used to construct confidence bands as described below.

Some words on the implementation. In Step S1, an oversmoothed bandwidth is commonly used to produce a consistent estimate of the second-order bias as discussed in Friedrich and Lin (2022). We follow their suggestion to set $\tilde{h} = 2h^{5/9}$. Step S2 of the bootstrap procedure is to account for both serial dependence and heteroskedasticity. A new parameter γ is introduced for this purpose. This parameter has a similar interpretation as the block length in block bootstrap procedures. It can be considered a tradeoff between capturing more dependence and allowing for more variation in the bootstrap samples (Smeekes and Urbain, 2014). We adopt $\gamma = 0.2$ as suggested in Friedrich et al. (2020).[5] Finally, we take $B = 1,499$ in Step S5.

[4] The notation $\mathbb{P}^*$ denotes for the probability measure conditional on the samples.

[5] Other values of γ do not give qualitatively different results, including the rule of thumb $\gamma = \theta^{1/\ell}$, where $\theta = 0.01$

2.5 Constructing confidence intervals and bands

We only describe the procedures for constructing confidence bands for $\beta_j(\cdot)$, because the constructions for $g(\cdot)$ and $\gamma_j(\cdot)$ are similar. Based on the bootstrap quantiles, the pointwise confidence intervals that fulfill Eq. (2.5) can be constructed as follows:

$$I^{P*}_{j,N,T,\alpha}(\tau) = \left[\widehat{\beta}_j(\tau) - \widehat{q}_{j,1-\alpha/2}(\tau),\ \widehat{\beta}_j(\tau) - \widehat{q}_{j,\alpha/2}(\tau)\right], \qquad \tau \in (0,1), \tag{2.8}$$

where $1-\alpha$ is the confidence level, $j = 1\ldots,d_1$, and $\widehat{q}_{j,\alpha}(\tau)$ is defined in (2.7). For the simultaneous confidence bands to satisfy Eq. (2.6), we require additional steps:

S1 Compute the pointwise quantiles $\widehat{q}_{j,\alpha_p/2}(\tau), \widehat{q}_{j,1-\alpha_p/2}(\tau)$ by varying $\alpha_p \in [1/B, \alpha]$, for $\tau \in G$, $j = 1,\ldots,d_1$.

S2 Choose $\hat{\alpha}_s = \hat{\alpha}_s(\alpha)$ as

$$\hat{\alpha}_s = \underset{\alpha_p\in[1/B,\alpha]}{\arg\min} \left| \mathbb{P}^* \left(\widehat{q}_{j,\alpha_p/2}(\tau) \le \widehat{\beta}^*_j(\tau) - \widetilde{\beta}_j(\tau) \le \widehat{q}_{j,1-\alpha_p/2}(\tau),\ \forall \tau \in G \right) - (1-\alpha) \right|.$$

S3 Given $\hat{\alpha}_s$ above, construct the simultaneous confidence bands as

$$I^{G*}_{j,N,T,\hat{\alpha}_s}(\tau) = \left[\widehat{\beta}_j(\tau) - \widehat{q}_{j,1-\hat{\alpha}_s/2}(\tau),\ \widehat{\beta}_j(\tau) - \widehat{q}_{j,\hat{\alpha}_s/2}(\tau)\right], \qquad \tau \in G.$$

This three-step procedure is similar to the one described in Bühlmann (1998) and Friedrich and Lin (2022).[6] Note that we take $G = \{1/T, 2/T, \ldots, 1\}$ to construct the simultaneous bands.

3 Data

We study $N = 60$ municipalities in the Netherlands during the period January 2006 until December 2020, resulting in $T = 180$ monthly time points. We make use of both housing market and macroeconomic data which stem from different sources. The data on the attributes of houses and their corresponding transaction prices are provided by the Nederlandse Vereniging van Makelaars (NVM) – the Dutch association of real estate agents. The macroeconomic data are collected from the Federal Reserve Economic Data (FRED), the Dutch National Bank (DNB), Statistics Netherlands (CBS), and Yahoo Finance. We give more information on the exact variables in the remainder of this section.

and $\ell = 1.75T^{1/3}$.

[6]One can find the Python codes for the LLDV estimation and bootstrap confidence intervals/bands on https://yiconglin.com/.

3.1 Housing prices

In order to make the data set suitable for our research, we have carefully pre-processed the raw data obtained from NVM. More specifically, we exclude apartments and properties with abnormally high transaction prices from our analysis as we want to strictly confine ourselves to houses. Moreover, since a house is traded infrequently, analyzing housing prices for a panel of houses introduces sparsity issues. Therefore, it is standard practice to analyze pools of houses within a region such as a municipality (Adams and Füss, 2010). However, to ensure the credibility of the pooled housing prices, we restrict our attention to municipalities that have at least ten housing transactions each month. The pre-processing stage consists of more specific adjustments to the raw data and we provide more detailed information about this process in Supplemental Appendix A. The resulting housing price data have already been alluded to in Introduction and are displayed in Figure 1. It can clearly be seen that the housing prices appear to co-move over the time span considered with a change in the direction of the trend in the year 2013.

Table 1 (first row) provides descriptive statistics regarding housing prices over the complete sample period. We observe that the average housing price is around € 300,000, but that the dispersion is quite large as the standard deviation lies around € 190,000. Thus, the commonality between the housing prices is striking, but zooming out on the different municipalities appears to reveal that they behave quite differently from each other. Figure 2 confirms this conjecture. It displays the average nominal housing prices in the selected municipalities. We observe that, especially in the urban area called the Randstad (which covers the largest cities Amsterdam, Rotterdam, and The Hague), the mean price is relatively high compared to the rural areas (e.g. Friesland and Groningen in the North). More specifically, prices in Amsterdam (in yellow) are nearly twice as high as the overall mean, while in the Northern and Western parts of the Netherlands, average prices are close to the total average. As municipalities vary greatly in size, we expect that heterogeneity might be present quite pronouncedly. Since our model allows for municipality-specific intercepts, it can account for such effects.

Another interesting observation can be made regarding the relatively high standard deviation of housing prices in the sample. Instead of considering the average nominal housing price, a similar heatmap can be made based on the standard deviations, see Supplemental Appendix A. One can observe that some standard deviations in the West of the Netherlands are 2-3 times as large compared to the remaining regions. More generally, we find that the municipalities with higher average prices also have a larger standard deviation. This could result in the presence of heteroskedasticity, whenever we do not have regressors at our disposal that are able to capture this variation in housing prices. However, this does not pose any problems for our analysis, as the

	Mean	St.dev	Min	Max
Price (€)	302,251	191,215	11,000	10,000,000
Size (m^2)	132.18	42.86	28	538
Number of Rooms	4.97	1.29	1	25
Number of Floors	2.78	0.59	1	9
Parking Space Availability	0.38	0.49	0	1
Presence of a Garden	0.97	0.18	0	1
Construction Year: 1945-1959	0.06	0.23	0	1
Construction Year: 1960-1970	0.10	0.30	0	1
Construction Year: 1971-1980	0.13	0.34	0	1
Construction Year: 1981-1990	0.12	0.33	0	1
Construction Year: 1991-2000	0.14	0.35	0	1
Construction Year: 2000-2020	0.19	0.39	0	1

Table 1: Descriptive statistics, NVM data

inference we conduct is robust against a non-constant error variance.

With respect to the development of housing prices over time, we also observe some noteworthy patterns. Initially, prices increased mostly for urban areas, however, this tendency changed since 2019. For example, the Randstad region went from the highest relative price change between 2013 and 2019 to the lowest change between 2019 and 2021, while rural areas such as the provinces Drenthe and Flevoland experienced exactly the opposite development (Langenberg and Jonkers, 2022). To illustrate that our method is able to deal with such diversity, we include both urban and rural municipalities in our sample which may be very distinct both in terms of size and behavior.

3.2 Housing attributes

There is a large body of literature that relates housing prices to the attributes of houses (see e.g. Ekeland et al., 2004; de Groote et al., 2018; Dröes and Koster, 2021). These hedonic pricing models are based on the notion that housing characteristics adequately determine the worth of a house (and its corresponding market price). We refer to these regressors as the local component, because they are house-specific and may be valued differently per municipality. For example, it might be more difficult to obtain a house with a garden in urban areas than rural areas.

Table 1 highlights the housing attributes we consider in this paper and we summarize them in three different categories: capacity, environment, and building year. The first category includes the quantitative variables *Size*, *Number of Rooms*, and *Number of Floors*, as they indicate how spacious the houses are. We can observe that an average house in the sample covers 132 m^2, divided over three floors and five rooms. The second category consists of the binary variables

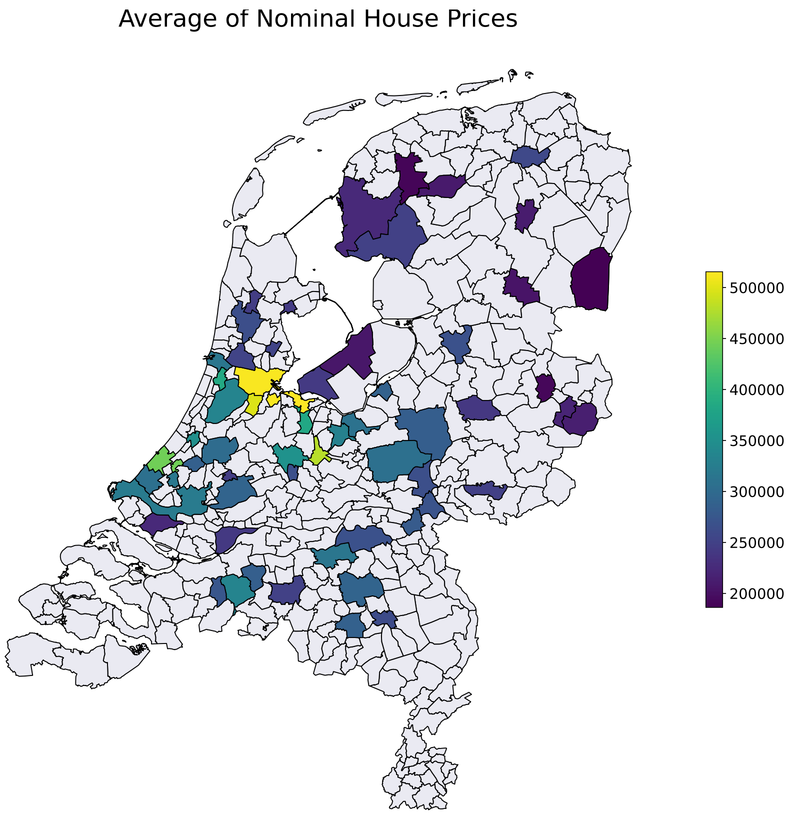


Figure 2: Heatmap with the average nominal house prices in the 60 municipalities from January 2006 until December 2020.

Parking Space Availability and *Presence of a Garden*, which cover facilities outside the direct living area. The mean of these variables can be interpreted as the proportion of houses in the sample that have a parking space or garden, respectively. For all properties, we see that only 39% includes parking space, while an overwhelming 97% has a garden. This last result can be explained by the fact that the definition of the garden is rather broad, as it typically refers to the presence of some type of outside area (e.g., a balcony). The third category consists of binary variables called *Construction Year* and specifies six different sub-periods over the years 1945 up to 2020. We observe that there is a relatively equal spread with approximately 10% − 15% of the houses being built in each period. A percentage of 26% of the houses in our sample are built before 1945.

In Figure 3, we display some of the hedonic attributes for the three major cities Amsterdam, The Hague, and Rotterdam. We select these regressors because they jointly exhibit typical dynamic behavior that can also be observed in the other variables. In the top-left panel, we see that the logarithm of the variable *Size* overall appears stable around a fixed value for these cities but the mean value lies much higher for The Hague compared to the other cities. The top-right panel shows that the size of the fluctuations in the *Presence of a Garden* differs per city over different sub-periods. In the first part of the sample, the behavior of the data appears very similar for

Figure 3: Time series plots of some (local) variables regarding housing characteristics for three major cities: Amsterdam, The Hague, and Rotterdam.

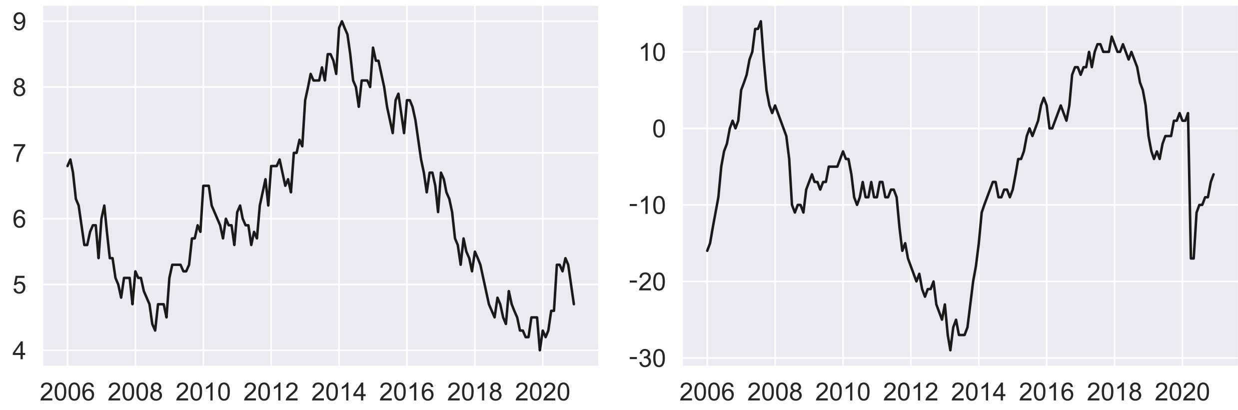
all cities. In contrast, during the second part of the sample, Amsterdam is much more volatile than before, and vice versa for the other two cities. This partially signals that the housing market in Amsterdam has overheated over the past years, as house buyers pay more for fewer facilities on average. The bottom-left panel shows that the presence of parking spaces has some changing trends, which might be attributable to changes in local parking policies. In particular, we see an upward trend from 2006 until 2011 followed by an overall downward trend until 2015. After a short increase, the series seems to stabilize around a fixed value. In the bottom-right panel, we see that the amount of houses built in the most recent time period (2000-2020) steadily increases over time. As indicated in Section 2, our model allows the inclusion of regressors that have deterministic trends. This trending effect is generally more noticeable in urban areas such as the major cities, as more new houses are built in their direct surroundings to combat the housing shortage.

3.3 Macroeconomic variables

Housing prices might not solely be affected by hedonic attributes but also by the overall economic climate and the willingness of consumers to make such a large purchase. A major example of such a situation is the period leading to the 2008 global financial crisis, in which housing prices were not driven by their attributes but rather treated as speculative objects, due to the favorable macroeconomic and financial climate. Therefore, we take into account several key macroeconomic indicators of the Netherlands as well as a sentiment variable that captures Dutch households' willingness to buy.

As key macroeconomic indicators, we collect data on inflation, interest rate, mortgage rate, unemployment rate, a measure of economic output, and a stock index from the Netherlands. As a measure of inflation, we use the Dutch CPI, which importantly excludes housing prices, as reported by the FRED. Inflation affects housing prices through multiple channels. Higher inflation leads to an overall higher cost of living and lower disposable income that can be used for saving or larger purchases, resulting in a lower demand for houses. It also leads to increased prices for goods that enter the production of houses which results in higher asking prices for newly built properties. Finally, there is an indirect link through raised interest rates by the central bank in response to inflation. To control for interest rates, we use the yield on 10-year Dutch government bonds from the same source. Alternatively, we consider the mortgage rate on loans in euros as reported by banks residing in the Netherlands. This data is retrieved from the DNB, as is the Dutch unemployment rate. A high unemployment rate often correlates with low economic growth and durable goods, such as properties, can experience major drops in demand. In addition, the labor market is directly linked to the housing market, as many jobs in the construction and real estate sector are dependent on the market's performance. Both arguments demonstrate why we expect a negative link between the unemployment rate and house prices. As a measure of economic activity, we cannot use GDP for our analysis, since the frequency of our data is monthly. Therefore, we use the Industrial Production Index (IPI) which is reported for the Netherlands by CBS and is given in nominal terms with 2015 as the base year. Finally, we download the Amsterdam Exchange Index (AEX) from Yahoo Finance which can be another indicator of economic activity and the economic climate. We expect a rise in economic activity and overall economic climate to have a positive effect on house prices. The exact sources with links to the above data are given in Supplemental Appendix A.

In line with Rouwendal and Longhi (2008), we additionally consider a regressor that reflects consumers' confidence and their expectations of the market. In the remainder of the paper, we will call this variable *Willingness to Buy*, which is in line with the label given by the data source (CBS). It is an indicator constructed from households' answers to three survey questions which are answered by approximately 1,000 households. Two questions require respondents to assess their own financial situation over the previous and upcoming 12 months. The third question asks



(a) Unemployment rate. **(b)** Willingness to buy.

Figure 4: The macroeconomic variables used in the full model.

them whether they believe that the current time is the right time to make large purchases.

To illustrate our data, the trajectories of the variables *Unemployment rate* and *Willingness to Buy* can be found in Figure 4. As expected, *Unemployment Rate* and housing prices mostly move in opposite directions. The main difference is that this macroeconomic variable progresses rather smoothly and does not display the same abrupt drops and surges as the housing prices. In contrast, the variable *Willingness to Buy* shows more extreme and erratic behavior. As it is based on the sentiments of households, it is probably able to capture extreme movements more adequately. As such, it might be able to explain the housing prices in times of booms and busts better. It should be noted that the variable partly follows a similar pattern as housing prices, but there are also clear differences. Whereas housing prices have a positive trend from 2013 until 2020, *Willingness to Buy* only shows this behavior until 2018 but is in a downward spiral after. Note that we do not seasonally adjust our data, as our model allows for regressors to have bounded, deterministic trends which capture possible seasonality.

As a final remark, we note that the effect of many macroeconomic and financial variables might be captured by the variable *Willingness to Buy*, as it represents how potential buyers are perceiving the economic climate. Information on variables such as interest rate, mortgage rate, and economic activity are all affecting the decision to make larger purchases.

4 Time-varying panel results

We present the estimation results for model (2.1) using the data as described in Section 3. We estimate two different versions of this model: a restricted version without macroeconomic variables, which we refer to as the *base* model, and the *full* model which includes the macroeconomic variables. We display the estimated global trend and the time-varying paths of slope coefficients.

Moreover, to assess the relative power of the hedonic and macroeconomic components, we employ a decomposition of variation, given by

$$\frac{\left|\Delta\widehat{\text{macro}}_t\right|}{\left|\Delta\widehat{\text{macro}}_t\right| + \left|\Delta\widehat{\text{local}}_t^{avg}\right|}, \tag{4.1}$$

where $\left|\Delta\widehat{\text{macro}}_t\right| := \left|\widehat{\text{macro}}_t - \widehat{\text{macro}}_{t-1}\right|$ measures the change of macro environments, and $\left|\Delta\widehat{\text{local}}_t^{avg}\right| := N^{-1}\sum_{i=1}^{N}\left|\Delta\widehat{\text{local}}_{i,t}\right|$ the average change of the local component. Specifically, $\widehat{\text{macro}}_t = \hat{g}_t + \boldsymbol{w}_t'\widehat{\boldsymbol{\gamma}}_t$, where in the base model we restrict $\boldsymbol{\gamma}_t = \mathbf{0}$, and thus $\widehat{\boldsymbol{\gamma}}_t = \mathbf{0}$, for all t. The variation decomposition can be interpreted as the percentage of variation explained by the macro component. Since we consider a monthly frequency, this percentage can be erratic. Therefore, we take quarterly and semi-annual averages, without overlapping months, to construct smoother measures. Throughout this section, the estimated coefficient paths are plotted as black solid lines, 95%-level pointwise confidence intervals are displayed in blue dotted lines and simultaneous confidence bands in red dashed lines.

4.1 Base model

We first present the results for the base model. The optimal bandwidth for this model, based on the leave-one-unit-out CV approach, is $\hat{h}_{\text{opt}} = 0.42$ (Figure 13, Supplemental Appendix C). Figures 5(a) - 5(d) show the effect of the hedonic attributes *Size*, *Number of Rooms*, *Number of Floors* and *Presence of Garden* on housing prices over time. The vertical axis shows the magnitude of the estimated coefficients and the horizontal axis captures the development over our sample period. The graphs can be interpreted as follows. At a given point in time, we can judge the significance of a specific attribute by focusing on the pointwise confidence intervals. When zero is not included, the attribute showed a significant effect on housing prices at this point in time. When we want to say something about the development of a coefficient over time, we look at the simultaneous confidence bands. For example, when a horizontal line at zero cannot be completely embedded in these bands, it shows that the hedonic attribute significantly affects housing prices (at least over some periods of time).

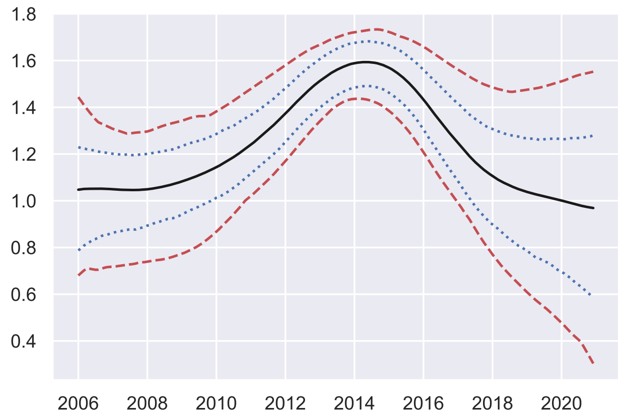
The results suggest that there is time variation in the relationships between housing prices and some hedonic attributes. The simultaneous confidence bands for *Number of Rooms* and *Number of Floors* support time-varying preferences towards these hedonic attributes in our sample period. In particular, in the recovery from the financial crisis in 2008, relationships start changing. However, these relationships seem to be slowly returning to their sign and size before the financial crisis of 2008. Though we observe a flat shape of the coefficient curve estimate of *Presence of Garden*, by zooming in, one can see its estimated coefficient is significantly different from zero around 2013, as suggested by the pointwise intervals. Similar conclusions hold for the coefficients of the other hedonic attributes shown in Supplemental Appendix C, such as the coefficients related to the construction years.

The estimated global trend in Figure 5(e) resembles an asymmetric V-shape, which corresponds to the shape we have seen in Figure 1. The global trend is significantly negative from 2010 to 2017, This shows that the trend reversal pattern of the housing prices cannot be explained by any of the included hedonic attributes. We note in Figure 5(b) that the estimated coefficient *Number of rooms* displays a similar pattern. It even turns significantly negative around 2014. This goes against our expectations as - controlling for size and number of floors of a house - the number of rooms should positively affect the price. This indicates that the base model is probably misspecified and that important predictors might be missing from our model. We address this by including macroeconomic variables in Section 4.2.

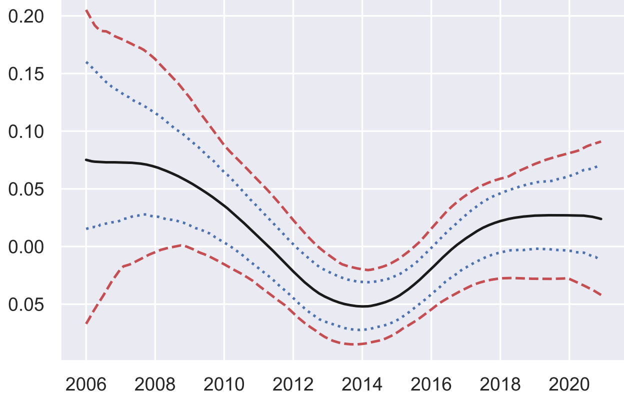
Until now, we have empirically assessed the local and global components separately. We can examine them relative to each other by computing the explained variation given in Eq. (4.1). The explained variation by the global trend is depicted in Figure 6. Interestingly, during the financial crash in 2007-2008, the hedonic attributes explain most variation in housing prices for the base model. The large size of explained variation for both components shows the importance of including both the local and global components in analyzing housing prices. Excluding either one of the two components corresponds to missing approximately 50% of the explanation of the time variation in housing prices.

4.2 Full model

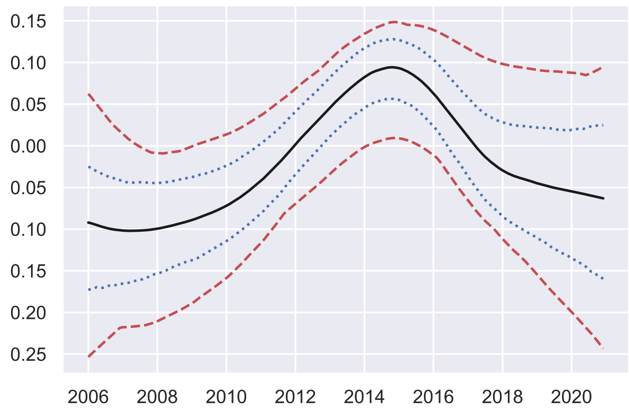
As we have seen in the previous section, the estimated global component $\hat{g}_t$ of the base model displays the same trend reversal pattern as the raw housing prices. By considering our full model, we can investigate whether some of this movement can be explained by specific macroeconomic variables. Out of the macroeconomic variables described in Section 3, the two variables that show a significant effect on housing prices are the *Unemployment Rate* and the indicator of *Willingness to Buy*. It might seem surprising that variables like the mortgage rate and inflation do not show a significant effect. This can potentially be explained by the fact that the information given by these variables is partly included in the factor *Willingness to Buy*. This survey-based indicator captures exactly the part of the more general information on the economic climate that is relevant for larger purchases such as property, rendering the additional information given by the other macroeconomic variables insignificant in explaining the development of housing prices. We further



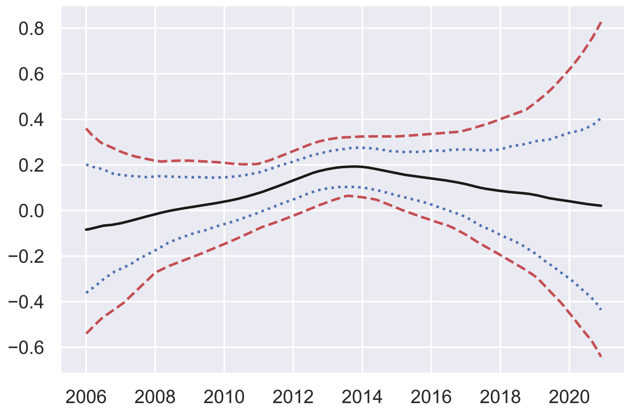
(a) Size.



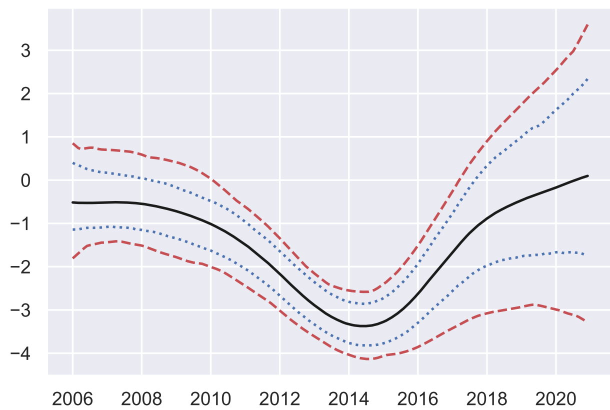
(b) Number of rooms.



(c) Number of floors.



(d) Presence of garden.



(e) Global trend.

Figure 5: Base model: the estimated coefficient paths (block solid lines) for number of rooms, number of floors, and global trend; the 95%-level pointwise intervals and simultaneous bands are displayed in blue dotted and red dashed lines, respectively.

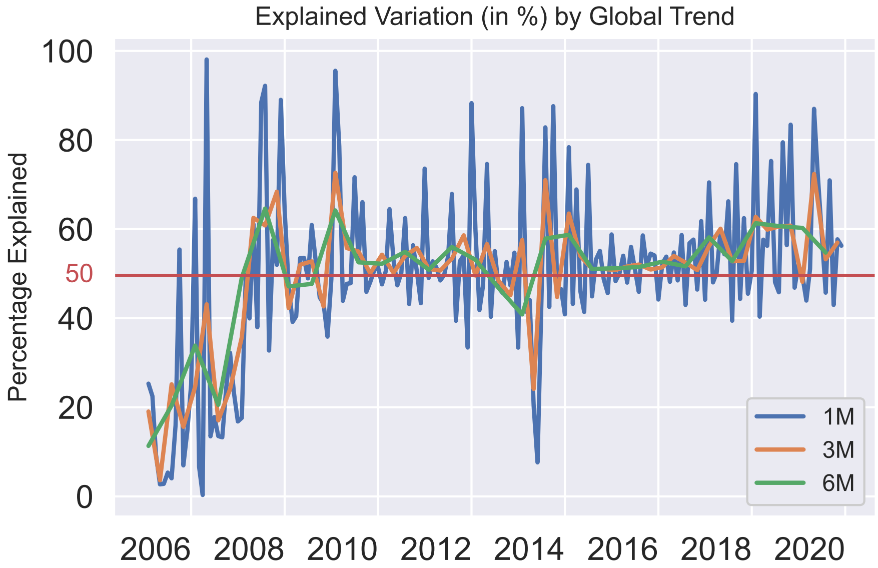
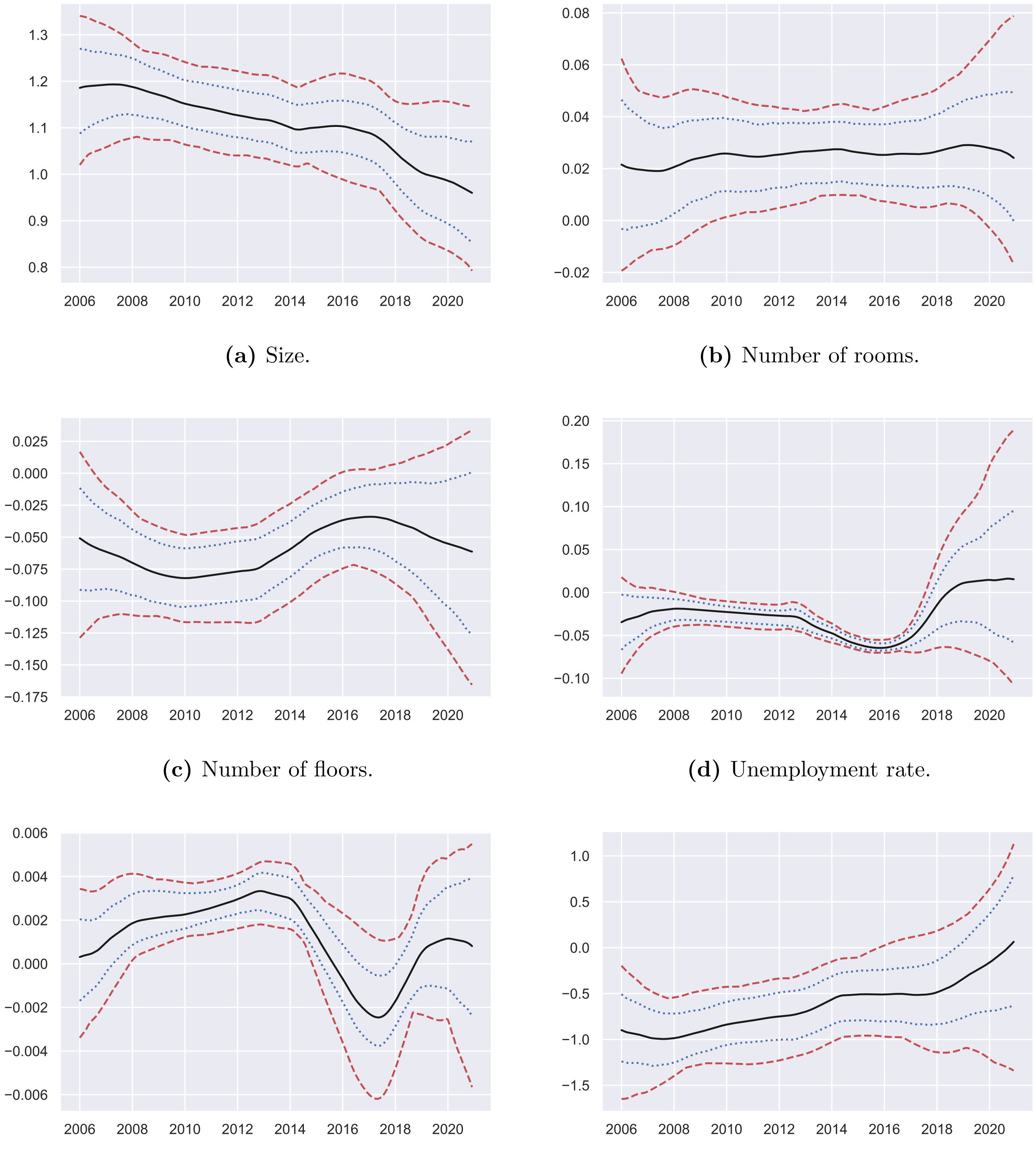


Figure 6: The percentage of housing price variation explained by the global trend in the base model. 'M' denotes the number of months averaged. For 1 month, no average is taken. The horizontal red line denotes the average monthly explained variation.

discuss this in Section 4.4.

Figure 7 displays the estimated coefficient paths for a selected number of variables and the global trend resulting from this new model. The model is estimated using the same bandwidth as the base model ($\hat{h}_{\text{opt}} = 0.42$) to keep the findings between the two models comparable. As mentioned, the macroeconomic variables that do not have any significant effect on housing prices in our model are not included here.

Adding the *Unemployment Rate* and *Willingness to Buy* greatly improves the results of the base model. For example, Figures 7(a)–7(c) plot the estimated coefficient paths of the three hedonic housing attributes which display significant time variation in the base model: *Size*, *Number of Rooms* and *Number of Floors*. In the full model, these three coefficient curves show a relatively flat shape (compared to the base model), indicating no or minor time variation occurs when adding macroeconomic variables. The predictor *Size* is significantly positive over the complete sample, but its importance diminishes over time - especially around the time that we observe a large acceleration in housing prices. This result could be interpreted as house buyers' willingness to accept a smaller house when the real estate market is overheated. The hedonic attribute *Number of Rooms* shows a significant positive effect on housing prices over most of the considered period in contrast to the partially negative effect from the base model. The positive effect is an expected result since even controlling for size, houses with more (bed)rooms will achieve a higher price on the market. Similarly, the attribute *Number of Floors* now shows a negative relation to housing prices. It indicates that, given the size and number of rooms of a house, more floors decrease



(a) Size.

(b) Number of rooms.

(c) Number of floors.

(d) Unemployment rate.

(e) Willingness to buy.

(f) Global trend.

Figure 7: Full model: estimated coefficient paths for number of rooms, number of floors, the global trend, unemployment rate, and willingness to buy.

the price. This is an expected result, as property in the Netherlands is expensive. A house with multiple floors will be built on a smaller piece of land and the property component of the house price will be lower compared to a house with the same size spread out over just one floor. Figure 7(d) shows that *Unemployment Rate* has a negative effect on housing prices which is significant over the period from 2009 to 2017, which confirms the finding in Vermeulen and van Ommeren (2009). As can be seen in Figure 7(e), *Willingness to Buy* has a positive effect on housing prices which is significant from 2008 to 2015. During the period of the sharp price increase, *Willingness to Buy* plays a smaller role, and the relationship becomes insignificant according to the simultaneous confidence bands.

The estimated global trend for this model is displayed in Figure 7(f). Compared to the base model, the form of the trend drastically changes. Instead of resembling the overall movement of housing prices as displayed in Figure 1, the trend is now mostly linear and upward trending. This suggests that a significant part of the variation in housing prices that can be explained by the included macro variables has been taken up by the global trend component and some of the hedonic attributes in the base model. Thus, restricting the analysis to the hedonic pricing model may result in a misspecified model and misleading conclusions. In addition, it stresses the importance of allowing for a flexible specification of the global trend in modeling housing prices. The shape of the global trend component changes from a quadratic to a linear trend in the full model. Including a deterministic quadratic trend component in the model would be in line with the overall price movement visible in Figure 1. However, this modeling choice could lead to a wrong decomposition into the global and local components. In our flexible model, we allow the data to determine the shape of the global trend without the need to specify the functional form in advance.

In Figure 6 the global trend becomes relatively more important from 2009 onward. At the end of the sample, there seems to be an increase in the explained variation by the global trend, suggesting that the effect of changes in the macro environment becomes more apparent in Dutch housing prices over time. However, we see in Figure 8 that, after we control for macroeconomic variables, the variation explained by the macro environment has a tendency to decrease after 2016. Additionally, the low percentage of explained variation at the beginning of the sample in Figure 6 has disappeared in Figure 8. It suggests that *Unemployment Rate* and *Willingness to Buy* are able to explain parts of the housing price variation that was neglected by the global trend in the base model.

Figure 9 shows the predicted housing price curves for the base model and the full model. We see that the hedonic attributes in the base model are able to capture the variation of the housing

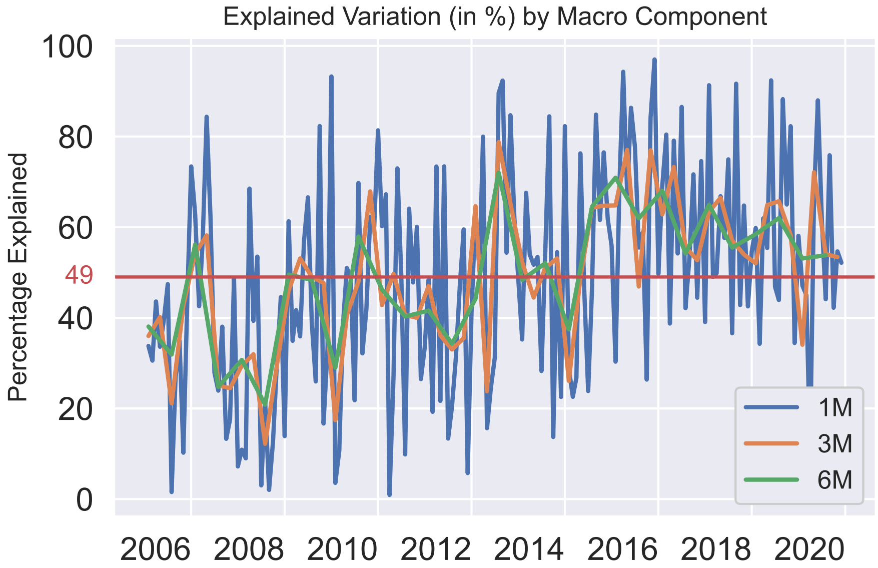


Figure 8: The percentage of housing price variation explained by the macro environment in the full model. 'M' denotes the number of months averaged. For 1 month, no average is taken. The red line denotes the average monthly explained variation by the macro environment.

prices throughout this sample, apart from the years 2013 to 2016. This result further supports our suspicion that some of the unexpected coefficient curves in the base model are simply due to model misspecification. When the macro variables are added, the full model does capture the variation of the housing prices also in this period. This result shows the importance of both components to predict housing prices.

Lastly, we investigate the role of *Unemployment Rate* by leaving it out of the full model. We see in Figure 10 that *Willingness to Buy* alone cannot accurately predict the housing prices from 2014 to 2016. Since the discrepancy between the predicted curve from the full model and the observed housing prices is not visible in Figure 9, it implies that *Unemployment Rate* is crucial for explaining housing prices.

4.3 What could the global trend capture?

The estimated global trend from the full model captures external factors that are not reflected by the macroeconomic measures in the full model. We see in Figure 8 that this external environment exhibits a linear trend. To understand this trend, we trace the global co-movement in $\hat{g}_t$ back to co-movements in the financial market. We measure the financial sub-market of real estate using a property market index. A property market index consists of stocks that are companies owning real estate property. According to the efficient market hypothesis, the value of the index should reflect all available information relevant to the investor, which makes it a suitable measure to compare with the estimated global trend. We take the GPR 250 index which consists of the 250

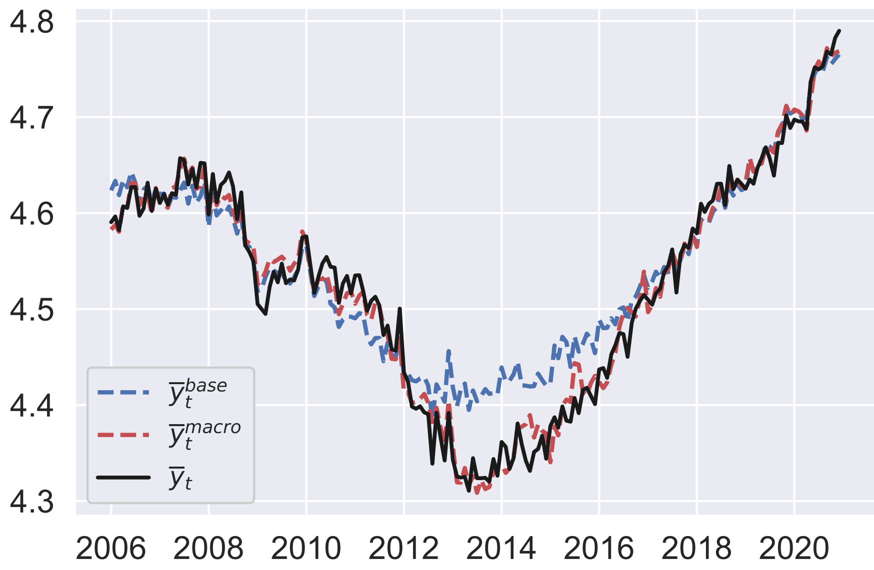


Figure 9: The solid lines represent the sample path cross-sectional averages, namely $\bar{y}_t = N^{-1}\sum_{i=1}^{N} y_{it}$. The dashed lines are the predicted path of $\bar{y}_t$ based on our models (Sections 4.1 - 4.2), i.e., $\bar{y}_t^j = \hat{g}_t^j + \left(N^{-1}\sum_{i=1}^{N} \boldsymbol{x}_{it}\right)' \widehat{\boldsymbol{\beta}}_t^j$, $j \in \{\text{base}, \text{full}\}$, where $\hat{g}_t^j$ and $\widehat{\boldsymbol{\beta}}_t^j$ are the corresponding LLDV estimates.

most liquid property stocks around the world, then take monthly averages to obtain a measure with the same frequency as the estimated global trend.[7] We see in Figure 11 that the shape of the estimated global trend is similar to the linear shape of the world property market index.

More specifically, the estimated global trend of the full model seems to be a leading indicator of the property market index. The slight decrease in the global trend in 2007 is followed by an increase in 2008. For the property market index, the drop occurs slightly later, in 2008, and has a similar increase starting in 2009. Both series seem to cool down around 2015. A sharp increase follows for both. The latter increase seems to occur simultaneously. The estimated global trend in the full model, however, does not show the sudden drop in the stock market index around 2020. This drop in the property market index may be intrinsic to the systemic risk of the stock market, which does not translate to housing prices. The outbreak of the Covid-19 pandemic in early 2020 induced risks to the financial market, such as increasing credit risk. These risks do not seem to apply to housing prices. Therefore, the increase we observe in the estimated global trend cannot be linked to the property market. The interpretation of the global trend requires the future attention of economic theorists.

[7]The data of GPR 250 index is publicly available at https://www.globalpropertyresearch.com/downloadable-index-data.

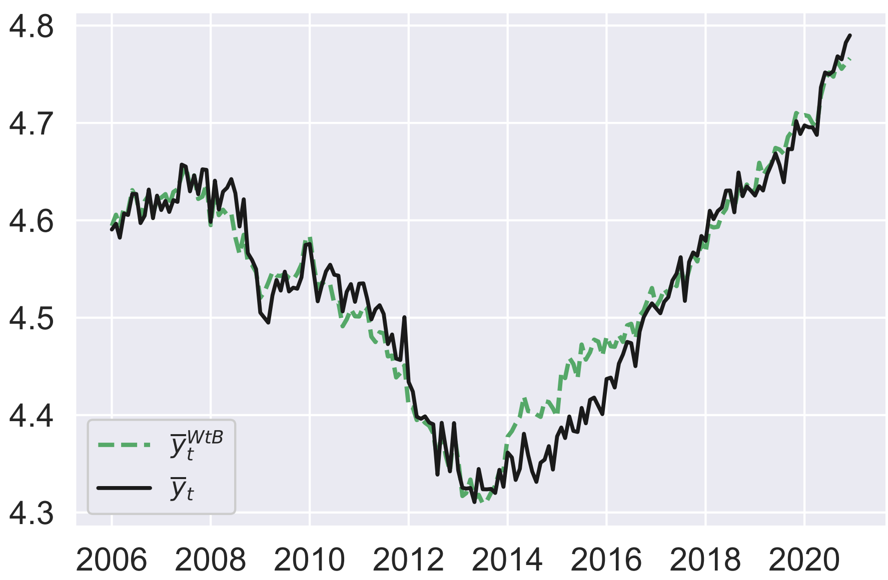


Figure 10: $\bar{y}_t^{WtB}$ is obtained using the variable *Willingness to Buy* only, i.e., excluding *Unemployment Rate* in the full model, see Figure 9 for further information.

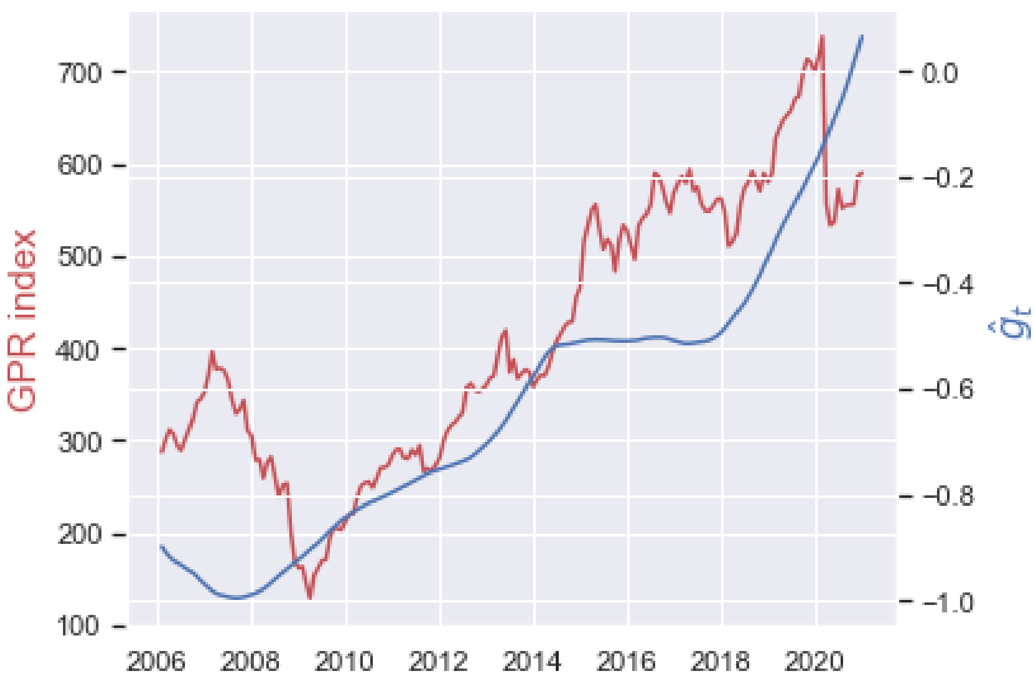


Figure 11: Estimated global trend and the GPR 250 Global Index.

4.4 Further discussions

There is extensive literature showing that housing prices are related to a large spectrum of factors, whose effects might change over time. This includes, but is not limited to, equity and stock prices (Kakes and End, 2004), consumers' confidence and their expectations of the housing market (Rouwendal and Longhi, 2008), financing conditions (Galati et al., 2011), GDP (Teulings, 2014), mortgage policies (Boelhouwer, 2017; Rouwendal and Petrat, 2022), parking policies (de Groote et al., 2018), environmental factors such as the presence of wind turbines and solar farms (Dröes and Koster, 2021). Motivated by the large set of potential predictors, we also consider a multitude of these variables in our preliminary analysis. To cover the effect of financing conditions on housing prices, we use the interest rate, measured as the 10-year maturity treasury government bonds of the Netherlands, and the Dutch mortgage rates. To account for the effect of stock prices on housing prices, we take the Amsterdam Exchange Index as an index of the Dutch stock market. Additionally, consumer confidence and a measure of the economic climate are considered. Since we are interested in studying the housing prices measured at a relatively high frequency (monthly), we are limited in our access to data. Given that monthly GDP data are not available, we use Industrial Production Index as a proxy of GDP.[8] However, we find no significant improvements in explaining the time variation of monthly housing prices when including the variables above in our model.

As argued in Section 4, we expect that part of the effect could be picked up by the *Willingness to Buy* variable. However, a few additional comments are in place. We hypothesize that there is possibly not an instantaneous but rather delayed effect between certain variables. For example, many house buyers take out a mortgage some months before the house is bought, which could explain that the effect of both the mortgage and interest rate is found to be insignificant. Following Ortalo-Magne and Rady (2006), a more informative measure could be the ability for young households to afford the down payment on a starter home. Moreover, macroeconomic variables such as economic activity and inflation might have a relatively weak short-run influence and changes rather materialize in the medium run. Lastly, the variable *Willingness to Buy* has to be treated with caution, as it could possibly be prone to some endogeneity concerns. It seems plausible to argue that consumers' willingness to buy has an effect on the housing prices as it affects the demand side directly. However, housing prices could also be an indicator of consumers' confidence in the overall economy which then directly affects their willingness to buy.

5 Conclusion

We have proposed a global-local panel model that allows for a time-varying relationship between housing prices and their price drivers. As potential price drivers, we include both hedonic housing characteristics, such as *Number of Rooms*, as well as macroeconomic variables, such as *Unemployment Rate*. The model is flexible as it can capture nonlinear movements of prices. It is nevertheless simple to interpret; it reveals how the prices can be jointly affected by housing attributes and the external economic environment. We have constructed local linear dummy

[8] See Supplemental Appendix A for the data description of the variables mentioned above.

variable estimates of the unknown coefficient curves and implemented an autoregressive wild bootstrap procedure to conduct inference.

By studying a dataset of the Dutch housing market with 60 municipalities from 2006 to 2020, we found that the model accurately captures the comovements of housing prices with an asymmetric V-shape. Moreover, we found significant time variation in the relationship between hedonic attributes as well as macroeconomic variables and housing prices. Next to the housing characteristics *Number of Rooms*, *Number of Floors* and some specific construction year indicators, we found that from the group of included macroeconomic variables, *Willingness to Buy* and *Unemployment Rate* are most important in predicting housing prices. We also saw that leaving out those macroeconomic indicators and thus, restricting the analysis to the hedonic characteristics, may lead to misleading conclusions.

Our findings stress the importance of allowing for a flexible model in which the functional form of coefficient and trend curves does not have to be determined in advance. Imposing a quadratic shape for the global trend, which constitutes a reasonable choice based on the overall V-shape of housing prices, would lead to a wrong global-local decomposition as the global trend we found in the full model has a nearly linear, upward-trending shape. We saw that this shape follows quite closely the leading indicator of the property market. Finally, we found that both the housing attributes and the economic environment explain a similar proportion of the housing price variation which shows that including both components is crucial in explaining housing prices.

For future research, it could be interesting to see whether empirical analyses benefit from the inclusion of other components. A possible direction involves the inclusion of spatial information to account for potential similarities between or even within municipalities. For example, Baltagi et al. (2015) find that spatial lags increase the explanatory power for the prices of flats in Paris. Even though the Netherlands does not have a historical spatial division as strong as Paris, there might still be information to exploit. Alternatively, municipalities could be divided into different categories: e.g., municipalities including the larger cities in the Netherlands (often with a large student population) might be grouped together as they are likely to behave differently from the remaining municipalities. Moreover, data specific to the municipalities, such as the crime rate, could be employed. It can be argued that the prices of houses are affected by the safety of their respective area. Thus, indicators of the overall perception of the neighborhood could be considered. Lastly, a more structural interpretation of the global trend is worth investigating.

Acknowledgements

We thank Hans Koster and Nederlandse Vereniging van Makelaars (Brainbay) for kindly providing the data. Earlier versions of this paper have been presented at the NESG 2023 (Rotterdam), the QFFE 2023 (Marseille), as well as seminars held at CREST and Lund University. We gratefully acknowledge the comments of the participants. We extend our thanks to Siem Jan Koopman for his insightful suggestions.

References

Adams, Z. and R. Füss (2010). Macroeconomic determinants of international housing markets. *Journal of Housing Economics 19*(1), 38–50.

Baltagi, B. H., G. Bresson, and J.-M. Etienne (2015). Hedonic housing prices in Paris: An unbalanced spatial lag pseudo-panel model with nested random effects. *Journal of Applied Econometrics 30*(3), 509–528.

Boelhouwer, P. (2017). The role of government and financial institutions during a housing market crisis: a case study of the netherlands. *International Journal of Housing Policy 17*(4), 591–602.

Boldea, O., B. Drepper, and Z. Gan (2020). Change point estimation in panel data with time-varying individual effects. *Journal of Applied Econometrics 35*(6), 712–727.

Brown, J. P., H. Song, and A. McGillivray (1997). Forecasting UK house prices: A time varying coefficient approach. *Economic Modelling 14*(4), 529–548. Modelling Residential Property Markets.

Bühlmann, P. (1998). Sieve bootstrap for smoothing in nonstationary time series. *The Annals of Statistics 26*(1), 48 – 83.

Cai, Z. (2007). Trending time-varying coefficient time series models with serially correlated errors. *Journal of Econometrics 136*(1), 163–188.

Chen, B. and L. Huang (2018). Nonparametric testing for smooth structural changes in panel data models. *Journal of Econometrics 202*(2), 245–267.

Chen, J., J. Gao, and D. Li (2012). Semiparametric trending panel data models with cross-sectional dependence. *Journal of Econometrics 171*(1), 71–85.

Chen, M., Y. Liu, D. Arribas-Bel, and A. Singleton (2022). Assessing the value of user-generated images of urban surroundings for house price estimation. *Landscape and Urban Planning 226*, 104486.

Creal, D., S. J. Koopman, and A. Lucas (2013). Generalized autoregressive score models with applications. *Journal of Applied Econometrics 28*(5), 777–795.

de Groote, J., J. van Ommeren, and H. R. A. Koster (2018). The impact of parking policy on house prices. *Journal of Transport Economics and Policy 52*(3), 267–282.

Dröes, M. I. and H. R. Koster (2021). Wind turbines, solar farms, and house prices. *Energy Policy 155*, 112327.

Durbin, J. and S. J. Koopman (2012). *Time Series Analysis by State Space Methods* (2nd ed.). Oxford University Press (Oxford).

Ekeland, I., J. Heckman, and L. Nesheim (2004). Identification and estimation of hedonic models. *Journal of Political Economy 112*(S1), S60–S109.

Fang, H., Q. Gu, W. Xiong, and L.-A. Zhou (2015). Demystifying the Chinese housing boom. Working Paper 21112, National Bureau of Economic Research.

Francke, M. and M. Korevaar (2021). Housing markets in a pandemic: Evidence from historical outbreaks. *Journal of Urban Economics 123*, 103333.

Friedrich, M., E. Beutner, H. Reuvers, S. Smeekes, J.-P. Urbain, W. Bader, B. Franco, B. Lejeune, and E. Mahieu (2020). A statistical analysis of time trends in atmospheric ethane. *Climatic Change 162*, 105–125.

Friedrich, M. and Y. Lin (2022). Sieve bootstrap inference for linear time-varying coefficient models. *Journal of Econometrics*. In press.

Friedrich, M., S. Smeekes, and J.-P. Urbain (2020). Autoregressive wild bootstrap inference for nonparametric trends. *Journal of Econometrics 214*(1), 81–109.

Fuerst, F. and G. Warren-Myers (2018). Does voluntary disclosure create a green lemon problem? energy-efficiency ratings and house prices. *Energy Economics 74*, 1–12.

Galati, G., F. Teppa, and R. Alessie (2011, March). Macro and micro drivers of house price dynamics: An application to dutch data. DNB Working Paper 288.

Gelain, P. and K. J. Lansing (2014). House prices, expectations, and time-varying fundamentals. *Journal of Empirical Finance 29*, 3–25.

Holly, S., M. H. Pesaran, and T. Yamagata (2010). A spatio-temporal model of house prices in the USA. *Journal of Econometrics 158*(1), 160–173. Twenty Years of Cointegration.

Iacoviello, M. and S. Neri (2010). Housing market spillovers: evidence from an estimated DSGE model. *American Economic Journal: Macroeconomics 2*(2), 125–164.

Jordà, O., M. Schularick, and A. M. Taylor (2015). Betting the house. *Journal of International Economics 96*, S2–S18. 37th Annual NBER International Seminar on Macroeconomics.

Kakes, J. and J. W. V. D. End (2004). Do stock prices affect house prices? evidence for the netherlands. *Applied Economics Letters 11*(12), 741–744.

Knoll, K., M. Schularick, and T. Steger (2017). No price like home: Global house prices, 1870-2012. *American Economic Review 107*(2), 331–53.

Langenberg, H. and W. Jonkers (2022). Achtergrond bij de huizenprijsstijgingen vanaf 2013. *CBS Internal Report*.

Li, D., J. Chen, and J. Gao (2011). Non-parametric time-varying coefficient panel data models with fixed effects. *The Econometrics Journal 14*(3), 387–408.

Mikhed, V. and P. Zemcik (2009). Do house prices reflect fundamentals? Aggregate and panel data evidence. *Journal of Housing Economics 18*(2), 140–149.

Møller, S. V., T. Pedersen, E. C. Montes Schütte, and A. Timmermann (2023). Search and predictability of prices in the housing market. *Management Science*. In press.

Ortalo-Magne, F. and S. Rady (2006). Housing market dynamics: On the contribution of income shocks and credit constraints. *The Review of Economic Studies 73*(2), 459–485.

Paci, L., A. E. Gelfand, Beamonte, and María Asunción , P. Gargallo, and M. Salvador (2020). Spatial hedonic modelling adjusted for preferential sampling. *Journal of the Royal Statistical Society: Series A (Statistics in Society) 183*(1), 169–192.

Rouwendal, J. and S. Longhi (2008). The effect of consumers' expectations in a booming housing market: Space-time patterns in the netherlands, 1999–2000. *Housing Studies 23*(2), 291–317.

Rouwendal, J. and A. Petrat (2022). Mortgage underwriting and house prices: Evidence from the 2011 change in the Dutch Code of Conduct for mortgage loans. *Real Estate Economics 50*(4), 1141–1159.

Smeekes, S. and J. Urbain (2014). A multivariate invariance principle for modified wild bootstrap methods with an application to unit root testing. Working Paper 008, Maastricht University, Graduate School of Business and Economics.

Sun, Y., R. J. Carroll, and D. Li (2009). Semiparametric estimation of fixed-effects panel data varying coefficient models. In *Nonparametric econometric methods*. Emerald Group Publishing Limited.

Teulings, C. N. (2014). Unemployment and house price crises: Lessons for fiscal policy from the Dutch recession. *IZA Journal of European Labor Studies 3*, 1–19.

Vermeulen, W. and J. van Ommeren (2009). Compensation of regional unemployment in housing markets. *Economica 76*(301), 71–88.

Yang, J. (2017). Consistency of trend break point estimator with underspecified break number. *Econometrics 5*(1).

Online supplemental appendix to:

Time-varying effects of housing attributes and economic environment on housing prices

A Data

A.1 Additional data sources

Our data are adopted from multiple (online) sources. The corresponding links are provided as follows.

(i) AEX: https://finance.yahoo.com/quote/%5EAEX/

(ii) CPI: https://fred.stlouisfed.org/series/NLDCPIALLMINMEI

(iii) Consumer confidence, economic climate and willingness to buy: https://opendata.cbs.nl/#/CBS/en/dataset/83693ENG/table

(iv) Interest rate: https://fred.stlouisfed.org/series/IRLTLT01NLM156N

(v) IPI: https://www.cbs.nl/en-gb/figures/detail/83838ENG

(vi) Mortgage rate: https://www.dnb.nl/en/statistics/data-search

(vii) Unemployment rate: https://opendata.cbs.nl/#/CBS/en/dataset/80590eng/table

A.2 Processing the data

The complete dataset from NVM is aggregated by municipality and month. As a result, the interpretation of binary variables changes to represent the percentage of total houses sold in each month that fall into "this interval" or have a private parking space. Monthly housing prices are indexed based on a pre-sample date and adjusted for inflation using the Dutch CPI before taking the logarithm. Indexing prices enables us to compare municipalities of different sizes, while deflating them allows for an analysis of price growth in real terms. Additionally, the variable Size is transformed into a logarithmic scale.

A.3 Standard deviations of the house prices

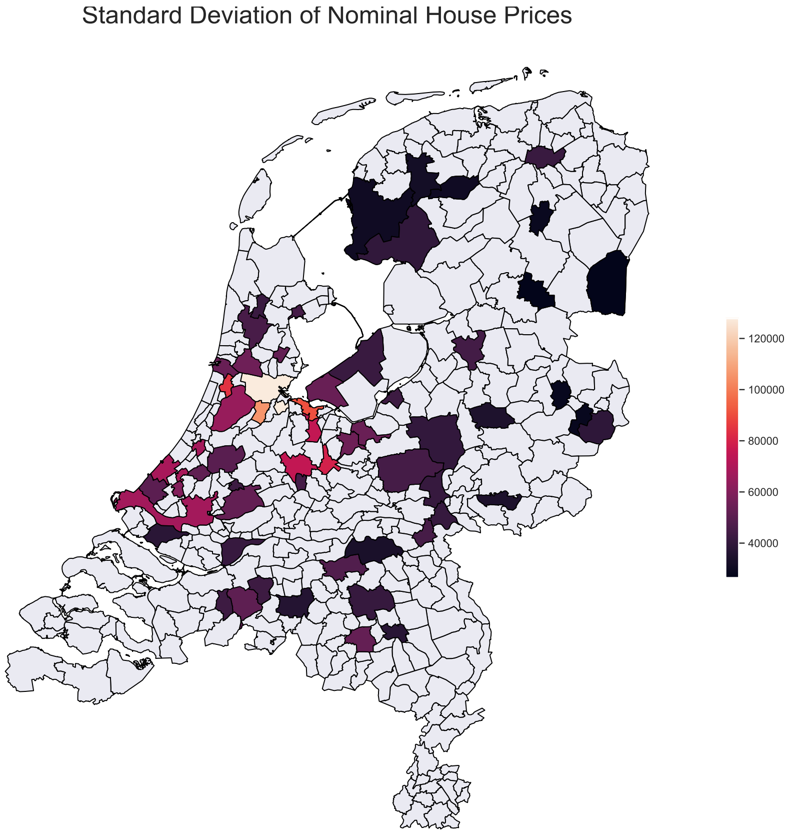


Figure 12: Heatmap of the standard deviations of the nominal house prices in the 60 municipalities from January 2006 until December 2020.

B Local linear dummy variable estimation

Further notation is required for illustration. Recall the approximation (2.2), and $\tau_t = t/T$. For any $\tau \in (0,1)$, Eq. (2.2) can be expressed in a compact form:

$$\boldsymbol{y}_i \approx \alpha_i \boldsymbol{\imath}_T + \boldsymbol{Z}_i(\tau)\boldsymbol{\theta}(\tau) + \boldsymbol{e}_i, \qquad \tau \in (0,1), \tag{B.1}$$

where $\boldsymbol{y}_i = (y_{i1}, \dots, y_{iT})'$, $\boldsymbol{\imath}_T = (1, \dots, 1)'$ is a T-dimensional vector of ones,

$$\boldsymbol{Z}_i(\tau) = \begin{pmatrix} \boldsymbol{z}_{i1}(\tau)' \\ \boldsymbol{z}_{i2}(\tau)' \\ \vdots \\ \boldsymbol{z}_{iT}(\tau)' \end{pmatrix} = \begin{pmatrix} 1 & \boldsymbol{x}_{i1}' & \boldsymbol{w}_1' & \frac{\tau_1-\tau}{h} & \frac{\tau_1-\tau}{h}\boldsymbol{x}_{i1}' & \frac{\tau_1-\tau}{h}\boldsymbol{w}_1' \\ 1 & \boldsymbol{x}_{i2}' & \boldsymbol{w}_2' & \frac{\tau_2-\tau}{h} & \frac{\tau_2-\tau}{h}\boldsymbol{x}_{i2}' & \frac{\tau_2-\tau}{h}\boldsymbol{w}_2' \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & \boldsymbol{x}_{iT}' & \boldsymbol{w}_T' & \frac{\tau_T-\tau}{h} & \frac{\tau_T-\tau}{h}\boldsymbol{x}_{iT}' & \frac{\tau_T-\tau}{h}\boldsymbol{w}_T' \end{pmatrix}, \tag{B.2}$$

and $\boldsymbol{e}_i = (e_{i1}, \dots, e_{iT})'$. For convenience, let $\boldsymbol{k}_h(\tau) = \left[K\left(\frac{\tau_1-\tau}{h}\right), \dots, K\left(\frac{\tau_T-\tau}{h}\right)\right]'$ be a T-dimensional vector, and $\boldsymbol{K}_h(\tau) = \text{diag}\left[\boldsymbol{k}_h(\tau)\right]$ a diagonal matrix with the elements of $\boldsymbol{k}_h(\tau)$ on the diagonal.

Moreover, for a vector $\boldsymbol{x} \in \mathbb{R}^n$ or a diagonal matrix $\boldsymbol{D} = \operatorname{diag}(d_1, \ldots, d_n) \in \mathbb{R}^{n \times n}$, let $\boldsymbol{x}^k = \left(x_j^k\right)$ and $\boldsymbol{D}^k = \operatorname{diag}\left(d_1^k, \ldots, d_n^k\right)$ take the power k element-wise.

Under the identification restriction $\sum_{i=1}^N \alpha_i = 0$ in Section 2.2, our LLDV estimator can be constructed using the following procedure.

S1 Project $\boldsymbol{K}_h^{1/2}(\tau)\boldsymbol{Z}_i(\tau)$ on $\boldsymbol{k}_h^{1/2}(\tau)\alpha_i$, $i = 1, \ldots, N$, and obtain the residuals $\widetilde{\boldsymbol{Z}}_i(\tau)$. That is,

$$\begin{pmatrix} \widetilde{\boldsymbol{Z}}_1(\tau) \\ \vdots \\ \widetilde{\boldsymbol{Z}}_N(\tau) \end{pmatrix} = \left(\boldsymbol{I}_N \otimes \boldsymbol{K}_h^{1/2}(\tau)\right) \begin{pmatrix} \boldsymbol{Z}_1(\tau) \\ \vdots \\ \boldsymbol{Z}_N(\tau) \end{pmatrix} - \left(\sum_{t=1}^T K\left(\frac{\tau_t - \tau}{h}\right)\right)^{-1} \left(\boldsymbol{I}_N \otimes \boldsymbol{k}_h^{1/2}(\tau)\boldsymbol{k}_h(\tau)'\right) \begin{pmatrix} \boldsymbol{Z}_1(\tau) - \bar{\boldsymbol{Z}}(\tau) \\ \vdots \\ \boldsymbol{Z}_N(\tau) - \bar{\boldsymbol{Z}}(\tau) \end{pmatrix},$$

where $\bar{\boldsymbol{Z}}(\tau) = N^{-1}\sum_{i=1}^N \boldsymbol{Z}_i(\tau)$.

S2 Project $\boldsymbol{K}_h^{1/2}(\tau)\boldsymbol{y}_i$ on $\boldsymbol{k}_h^{1/2}(\tau)\alpha_i$, $i = 1, \ldots, N$, and obtain the residuals $\widetilde{\boldsymbol{y}}_i$. That is,

$$\begin{pmatrix} \widetilde{\boldsymbol{y}}_1 \\ \vdots \\ \widetilde{\boldsymbol{y}}_N \end{pmatrix} = \left(\boldsymbol{I}_N \otimes \boldsymbol{K}_h^{1/2}(\tau)\right) \begin{pmatrix} \boldsymbol{y}_1 \\ \vdots \\ \boldsymbol{y}_N \end{pmatrix} - \left(\sum_{t=1}^T K\left(\frac{\tau_t - \tau}{h}\right)\right)^{-1} \left(\boldsymbol{I}_N \otimes \boldsymbol{k}_h^{1/2}(\tau)\boldsymbol{k}_h(\tau)'\right) \begin{pmatrix} \boldsymbol{y}_1 - \bar{\boldsymbol{y}} \\ \vdots \\ \boldsymbol{y}_N - \bar{\boldsymbol{y}} \end{pmatrix},$$

where $\bar{\boldsymbol{y}} = N^{-1}\sum_{i=1}^N \boldsymbol{y}_i$.

S3 Project $\widetilde{\boldsymbol{y}}_i$ on $\widetilde{\boldsymbol{Z}}_i(\tau)$, $i = 1, \ldots, n$, and obtain $\widehat{\boldsymbol{\theta}}(\tau)$ given by

$$\widehat{\boldsymbol{\theta}}(\tau) = \begin{pmatrix} \widehat{g}(\tau) \\ \widehat{\boldsymbol{\beta}}(\tau) \\ h\widehat{g^{(1)}}(\tau) \\ h\widehat{\boldsymbol{\beta}^{(1)}}(\tau) \\ h\widehat{\boldsymbol{\gamma}^{(1)}}(\tau) \end{pmatrix} = \left(\sum_{i=1}^N \widetilde{\boldsymbol{Z}}_i(\tau)'\widetilde{\boldsymbol{Z}}_i(\tau)\right)^{-1} \left(\sum_{i=1}^N \widetilde{\boldsymbol{Z}}_i(\tau)'\widetilde{\boldsymbol{y}}_i\right).$$

S4 Given $\widehat{\boldsymbol{\theta}}(\tau)$, we can obtain the estimates of fixed effects $\widehat{\boldsymbol{\alpha}}$. More specifically,

$$\widehat{\boldsymbol{\alpha}}(\tau) = \begin{pmatrix} \widehat{\alpha}_1(\tau) \\ \vdots \\ \widehat{\alpha}_N(\tau) \end{pmatrix} = \left(\sum_{t=1}^{T} K\left(\frac{\tau_t - \tau}{h} \right) \right)^{-1} (\boldsymbol{I}_N \otimes \boldsymbol{k}_h(\tau)') \begin{pmatrix} (\boldsymbol{y}_1 - \bar{\boldsymbol{y}}) - \left(\boldsymbol{Z}_1(\tau) - \bar{\boldsymbol{Z}}(\tau)\right) \widehat{\boldsymbol{\theta}}(\tau) \\ \vdots \\ (\boldsymbol{y}_N - \bar{\boldsymbol{y}}) - \left(\boldsymbol{Z}_N(\tau) - \bar{\boldsymbol{Z}}(\tau)\right) \widehat{\boldsymbol{\theta}}(\tau) \end{pmatrix}.$$

Define $\widehat{\boldsymbol{\alpha}} = T^{-1} \sum_{t=1}^{T} \widehat{\boldsymbol{\alpha}}(\tau_t)$.

C Further empirical results

This section provides additional empirical results.

C.1 Optimal bandwidth

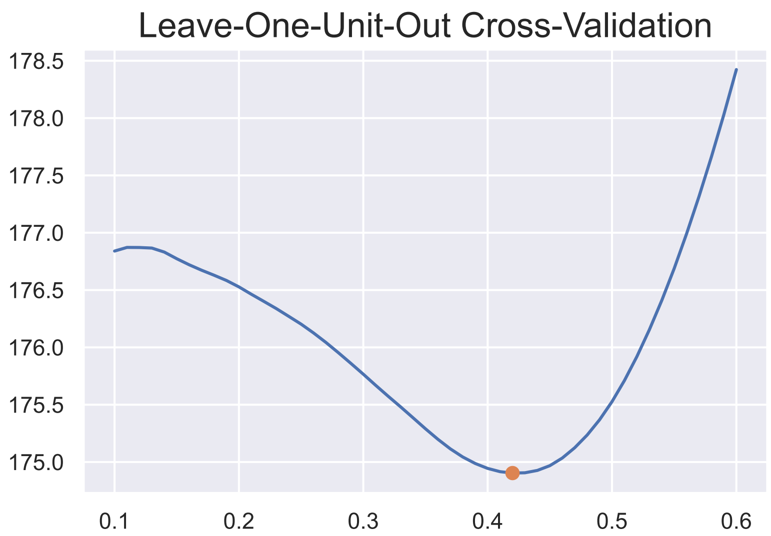


Figure 13: The selected bandwidth ($\hat{h}_{\text{opt}} = 0.42$) for the base model using the leave-one-unit-out cross-validation procedure in Section 2.3.

C.2 Base model: additional regression results

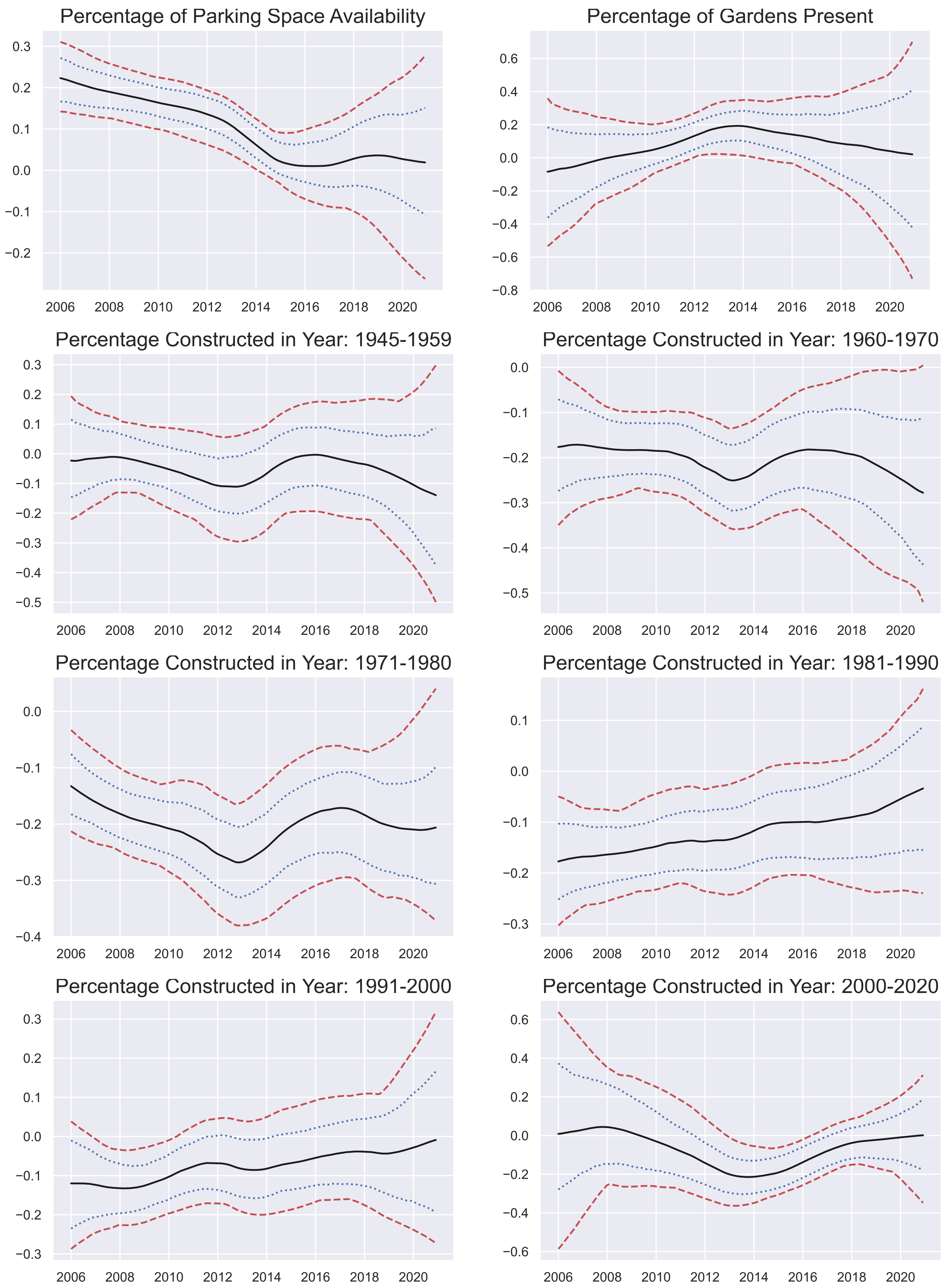
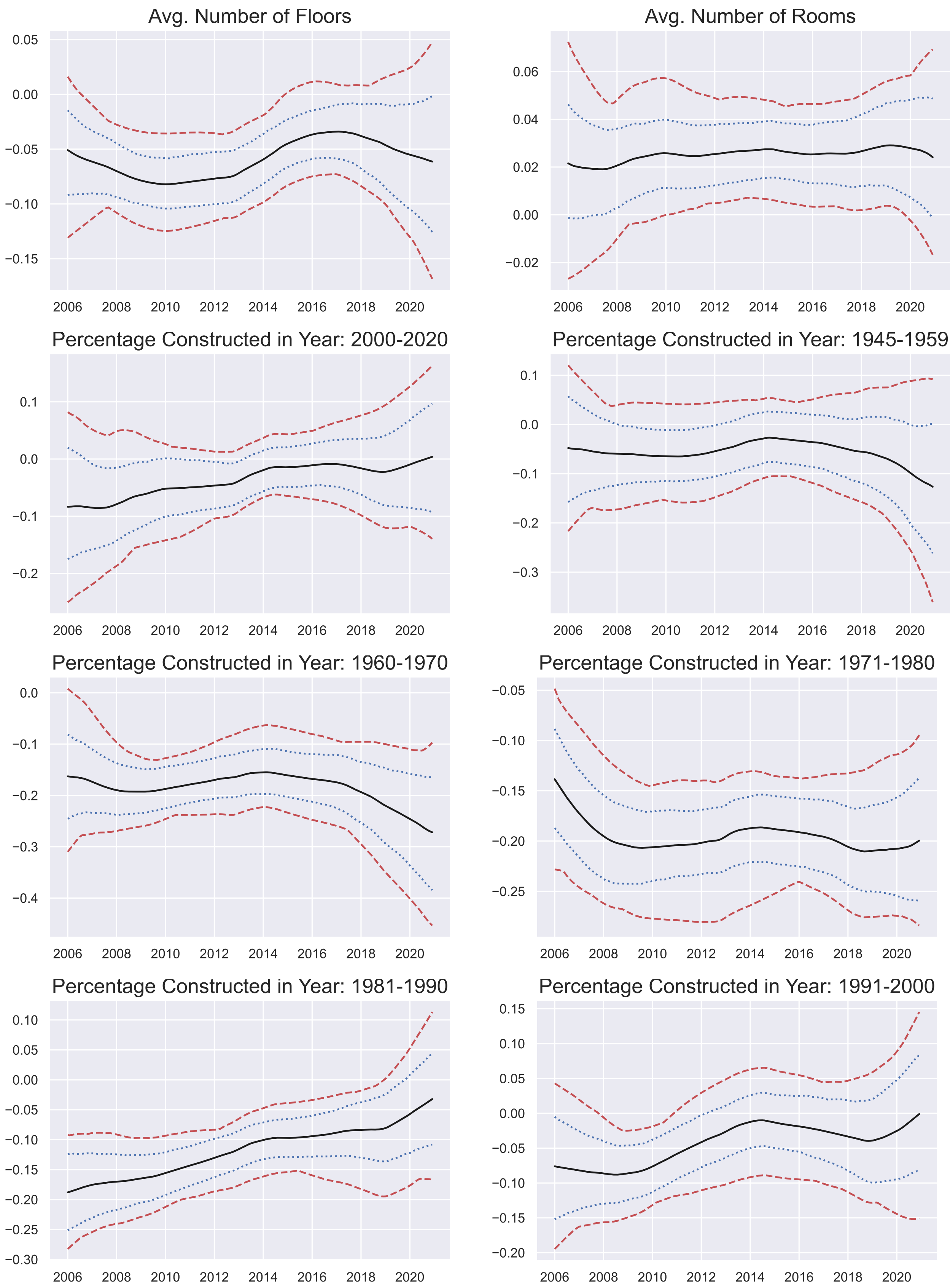


Figure 14: Base model: all estimated coefficient paths (black solid) and the corresponding 95%-level confidence intervals (blue dotted) and bands (red dashed).

C.3 Full model: additional regression results



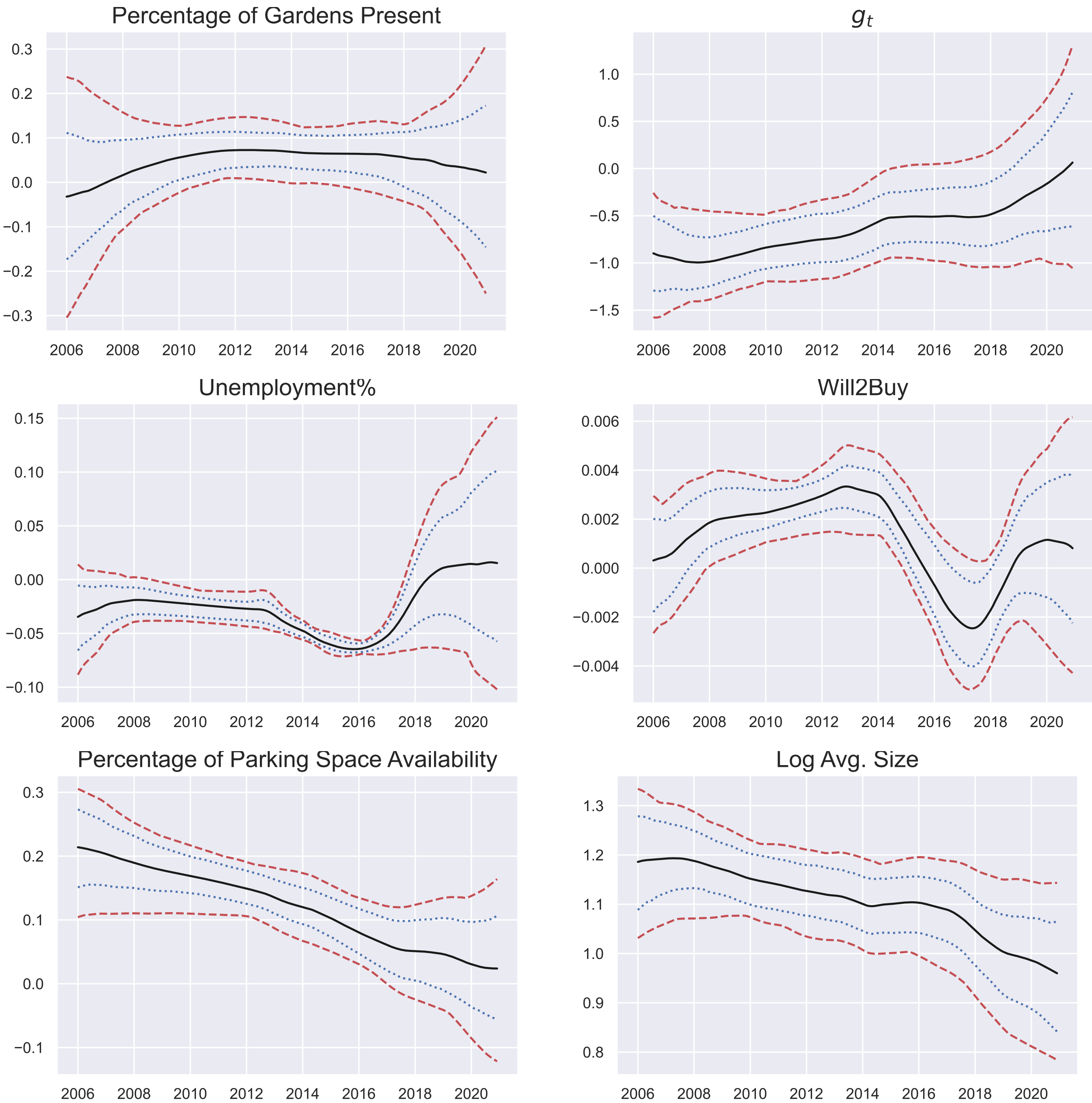


Figure 15: Full model: all estimated coefficient paths. (black solid) and the corresponding 95%-level confidence intervals (blue dotted) and bands (red dashed).