Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Department of Management, Technology and Economics
Chair of Entrepreneurial Risks

Master Thesis

Analysis of real estate price dynamics in Switzerland using a fundamental factor model

Student:	Delong Lu
Tutor:	Prof. Dr. Didier Sornette
Supervisors:	Dr. Dorsa Sanadgol
	Diego Ardila
Date of Submission:	March 8^{th} 2016

(This page left blank.)

ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Declaration of originality

The signed declaration of originality is a component of every semester paper, Bachelor's thesis, Master's thesis and any other degree paper undertaken during the course of studies, including the respective electronic versions.

Lecturers may also require a declaration of originality for other written papers compiled for their courses.

I hereby confirm that I am the sole author of the written work here enclosed and that I have compiled it in my own words. Parts excepted are corrections of form and content by the supervisor.

Title of work (in block letters):

ANALYSIS OF REAL ESTATE PRICE DYNAMICS IN SWITZERLAND USING A FUNDAMENTAL FACTOR MODEL

Authored by (in block letters):
For papers written by groups the names of all authors are required.

Name(s):	First name(s):
LU	DE LONG

With my signature I confirm that
- I have committed none of the forms of plagiarism described in the 'Citation etiquette' information sheet.
- I have documented all methods, data and processes truthfully.
- I have not manipulated any data.
- I have mentioned all persons who were significant facilitators of the work.

I am aware that the work may be screened electronically for plagiarism.

Place, date	Signature(s)
Zürich, 2016/02/08	DeLong

For papers written by groups the names of all authors are required. Their signatures collectively guarantee the entire content of the written paper.

Analysis of real estate price dynamics in Switzerland using a fundamental factor model

Student
Delong Lu

Tutor
Prof. Dr. Didier Sornette

Supervisors
Dr. Dorsa Sanadgol
Diego Ardila

Abstract

Real estate markets represent an integral part of today's economy which is inextricably linked to other areas. With recovering and rising real estate prices since the 2008/09 financial crisis, the potential overheating of real estate markets is gaining attention from regulatory bodies, industry and research alike.

To analyze the relationship between real estate prices and economic fundamentals in Switzerland, we apply the inverted demand approach and model apartment prices as a function of population, income, housing stock and interest rates using quarterly panel data. The data includes the cantons of Geneva and Zurich, at the cantonal and district level, and covers the time span 2000q4–20012q4.

We apply three different methods to estimate the models (dynamic fixed effect, mean group, pooled mean group) and test for homogeneity of long run coefficients over different regions. Given the regional discrepancies within Switzerland, this gives an indication whether aggregate models sufficiently capture individual region dynamics.

We find limited evidence for poolability at the district level and continue by reporting individual district level model fits as well as estimating coefficients over varying timeframes as a first approximation of the fundamental relationship over time.

Contents

List of Figures

List of Tables

1 Introduction

The housing market is a fundamental component of today's society and economy. Not only does housing fulfill an individual's basic need for shelter, the housing market is also tightly integrated into the economy such that volatility in it can have far reaching consequences, as demonstrated by the preparatory role of the US housing market in 2006/07 [9], and the high rate of foreclosures and price movements thereafter.

Since 2007, real house prices globally and in OECD countries have managed to recover and in some areas even increased considerably [24]. The latter is especially noticeable in big metropolitan areas, which in Switzerland applies to Zurich and Geneva; Zurich for instance boasts the highest real house price levels in its history[23]. This steady increase has brought the possibility of a real estate bubble back into the spotlight for market participants. The potential overheating of the real estate market is also highlighted by the Swiss National Bank (SNB) in its annual Financial Stability Report, where it is designated as one of the key risks which it is monitoring [22].

Research and industry alike have been tackling this potential issue. In addition to the SNB, UBS and Credit Suisse, two of Switzerland's major financial institutions, publish periodic reports on the health of the real estate market. In its most recent publication of the UBS bubble index, the report finds that Zurich has the highest price-to-rent multiple among global cities and Geneva's price-to-rent multiple remains elevated with strained affordability [23]. Overall, the sources indicate that Swiss residential property market is overvalued and at risk, but cannot (yet) be considered as in speculative property price bubble [22, 23]. Research conducted at the Chair of Entrepreneurial Risks includes a biannual report on the Swiss real estate market in cooperation with comparis.ch. The report presents the results of applying a log periodic power law singularity (LPPLS) analysis and, in its most recent volume, shows several districts as potential hotspots for the formation of bubbles(see[5]).

Besides feeling "the pulse of the market", participants are interested in determining factors driving its development. For example, [22, 24] cite various structural characteristics of the market as factors for the high level of imbalances in Switzerland, including the low mortgage rates, high population growth driven by migration, as well as short supply due to spatial planning rigidities, environmental standards, taxations issues and tenancy law,

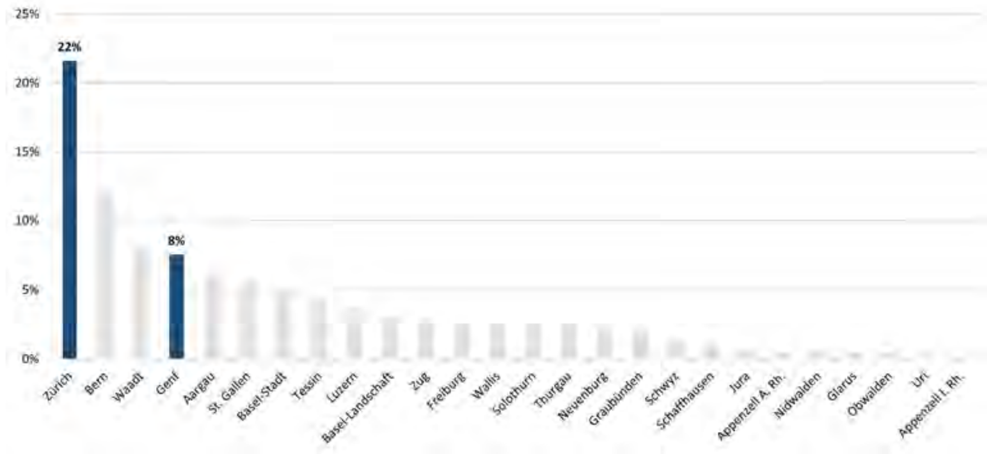
In line with this interest of analyzing the fundamental drivers of housing prices, this study aims to analyze the relationship of the Swiss apartment prices with fundamental economic variables. Specifically, given the regional disparities in Switzerland, as for example measured by GDP contribution of the different Swiss Cantons (see figure 1), we intend to test whether there is information lost by aggregating the analysis on a cantonal level (using data from the cantons Geneva and Zurich) as opposed to a lower level analysis which looks at each districts separately (using all districts of Zurich). In addition to testing aggregation, we estimate the coefficients over varying timeframes and examine the evolution of said coefficients over time.

In summary, we cannot reject the pooling assumption at the cantonal level. However, this conclusion is questioned as results at the district level reject the pooling assumption within the Canton of Zurich. Given that we reject the pooling assumption for districts, we report individual district model fits. Finally, when looking at the coefficient estimates over different timeframes, initial evidence does not suggest a stable long run relationship.

The following sections are structured as follows: In section 2 we give

Figure 1: Swiss GDP contributions by canton 2013

Total Swiss GDP 2013:
CHF 635bn ≈ USD 710bn (2013q4)



a brief overview of different approaches to modeling real estate prices in the academic literature. We continue in section 3 with the description of the methodology and model selected for this study. In section 4 and 5 we describe our data sample and their results respectively. Finally, we give a brief discussion in section 6 and conclude in section 7.

2 Review of existing Literature

Numerous studies have been conducted which aim to analyze the relationship of the real estate market with economic fundamentals, applying different methods and pursuing different hypotheses. Two of the most common approaches used in modeling the relationship are the price-to-rent ratio and the inverse demand approach. Additionally, recent work at the Chair of Entrepreneurial Risk at the ETH Zürich has examined the real estate markets using a log periodic power law singularity (LPPLS) model.

In the following subsections we shall give a brief overview of each of these approaches.

2.1 Existing work using a price to rent approach

The price-to-rent approach, in its most basic form, can be interpreted as an application of the concept of present value (PV) from the fields of Finance and Economics, where the present value of an asset is determined by its future cash flow. The PV is represented as $\mathrm{PV} = \frac{C_t}{(1+r)}$, where C_t represent the cash flows in period t and r_t represents the return a similar investment may have earned at the same time (either risk free or of a similar asset), at which C_t is discounted.

Parallel to PV calculations, price-to-rent states that the present value of a real estate asset should be the determined by the sum of its future rents minus certain costs, assuming conditions of an ideal market and under no arbitrage condition.

For example, [10] defined the price to rent model as

$$HP = \frac{\mathrm{RENT}}{r + \delta + t - \Delta \mathrm{hp}^e/\mathrm{hp}},$$

which can be rewritten as

$$\frac{\mathrm{HP}}{\mathrm{RENT}} = \frac{1}{r + \delta + t - \Delta \mathrm{hp}^e/\mathrm{hp}} =: \frac{1}{\mathrm{RUSER}} \tag{1}$$

and shortened to

$$\mathrm{HPRENT} = \frac{1}{\mathrm{RUSER}} \Leftrightarrow ln(\mathrm{HPRENT}) = -ln(\mathrm{RUSER}) \tag{2}$$

In equation (1), RENT stands for (future) rent payments, HP stands for house price, RUSER stands for real user costs which consists of the real after-tax interest rate r, the depreciation rate δ, the property tax t and the expected rate of house price change $\Delta \mathrm{hp}^e/\mathrm{hp}$. For convenience, the price-to-rent ratio is represented by HPRENT, which can be further simplified to equation (2) when working with log variables.

In their work [10], the authors investigate the impact of tightening credit constraints on buyers for the United States, which represents a deviation from the ideal market assumption. Specifically, they hypothesize that out of all groups, first time buyers are impacted the most by credit constraints. The authors test their hypothesis by including the variable of first time buyer loan to value ratios (LTV), as a proxy for credit constraints, into their model, which they formulate as an error correction model. Their results indicate that the long run relationship of HPRENT exists as described 1 and that it is further augmented by including the LTV ratio as a factor. In the long run, inclusion of the LTV ratio "show[s] stronger evidence of cointegration [...] than for models without the ratio", while in the short run inclusion of the ratio results in faster error correction adjustment speeds[10].

Earlier work by Mikhed & Zemcik [18], starts with a price to rent framework and links it to the inverted demand approach for the US housing market. Two data-sets are used, one at the national level, the other at the level of U.S. Metropolitan Statistical Areas. The authors look for evidence of a bubble in house prices by testing for unit roots in the relationship using panel data. Theory suggests that, given stationary demand and supply factors, a unit root in the price would decouple the price from the fundamental factors and indicate an unstable bubble process. The study concludes that the evidence does not support a cointegration relationship, implying that "house prices do not reflect movements in fundamental factors" and as such there is a bubble in the observed prices given his data time frame.

A different line of inquiry is pursued by Gallin [13]. In addition to estimating the price-to-rent ratios, the essay investigates whether prices follow rents or rents follow prices once there is a deviation from the estimated relationship. First, the author establishes a long run relationship between prices and rents by estimating an error correction relationship based on collected data of the U.S. market between 1970q1 and 2005q4. Subsequently, he uses these estimations to run simulations of hypothesized data generation processes which represent scenarios where either "prices do all the correcting" or "rents do all the correcting". His results reject the case where rents do all the adjusting, and provides evidence in favor of the hypothesis where house prices adjusting back to rent levels.

The price-to-rent ratio approach provides a straightforward pricing mechanism and can be related to similar measures in other fields of work (e.g. Present Value, or even the parallel to P/E ratios).

However, price-to-rent ratio modeling can have its shortcomings. Unlike financial assets, real estate assets are generally not easily substituted due to location and quality. As such, including adjustments based on hedonic factors can prove meaningful. Equation (1) also implies that given an ideal market, an agent should have no preference over owning vs renting a property[11, 18]. Finally, the price-to-rent approach usually does not feature any mechanism to adjust for market structures, such as regulations and preferences. As the authors of [11] note in their review, European markets may have more regulation when compared to the U.S. equivalent, suggesting that the price-to-rent may not be as well suited for the European case. Additionally, the Swiss real estate market may represent a structural exception in international comparison, with only 37.5% of all households owning their home at the end of 2013[1], compared to 65.2% in the U.S. [7, 8].

2.2 Existing work using the inverted demand approach

An alternate way of modeling housing prices starts with the economic principle of supply and demand. In theory, given a market in equilibrium, prices clear supply and demand. One can therefore invert the equation to back out the prices as a function of supply and demand variables.

A generic form of this approach can be states as follows (similar to [26]): We define demand D_t as

$$D_t = \alpha P_t + \beta' \theta_t + u_{dt},$$

Where P_t represent house prices, θ is a $(k_d \times 1)$ vector of demand shifters (e.g. income, population, etc.), u_{dt} is white noise and $\alpha < 0, \beta$ represent coefficients for each term.

Similarly, we can define supply S_t as

$$S_t = \lambda P_t + \psi' \eta_t + u_{st},$$

where $\eta_t, (k_s \times 1)$ represent a set of supply shifters, u_{st} also white noise and $\lambda > 0, \psi$ representing coefficients for each term.

In an equilibrium state, $S_t = D_t$ and solving for P_t yields

$$P_t = \frac{\psi'}{\alpha - \lambda}\epsilon_t - \frac{\beta'}{\alpha - \lambda}\theta_t + u_t \tag{3}$$

with $u_t = \dfrac{1}{\alpha - \lambda}(u_{st} - u_{dt})$.

[1]steadily rising tendency, with ownership rates at 28.5% in 1970. Again, there are large regional discrepancies, e.g. Canton of Jura (54.8%) vs. Canton of Geneva (17.9%

If the detailed supply and demand equation coefficients are not of interest, one might further simplify the equation above by restating (3) as

$$P_t = \gamma_s \epsilon_t - \gamma_d \theta_t + u_t \tag{4}$$

Previous work has used (4) as a generic starting point and further detailed out the model by choosing different sets of fundamentals in ϵ_t, θ_t, by including different dynamics through lag terms, and/or by reformulating the functional form depending on the research hypothesis. Often, research targets the questions whether 1) real estate prices are determined by fundamentals in the long run and, if they are, 2) how that relationship can be quantified.

Cointegration as long run relationship One notable concept which is frequently featured is the concept of cointegration. Engle & Granger originally defined it as follows in [12]: *Given x_t a vector of economic variables at time t, then the elements x_t are cointegrated of order d,b if there exists a vector $\alpha (\neq 0)$ such that $z_t = \alpha' x_t \sim I(d-b), b > 0$, with the vector α called the cointegration vector.*[2]

In practice, the case of $d = b = 1$ is often applied, in which case $x_t \sim I(1), z_t \sim I(0)$. For example, given two economic variables x_{1t}, x_{2t}, then $z_t = \alpha_1 x_{1t} + \alpha_2 x_{2t}$ is stationary with $E(z_t) = 0$.

Testing for cointegration therefore revolves around testing z_t for stationarity and the work proceeds to list seven possible testing methodologies[12].

Therefore, two variables which are cointegrated have a well-defined long run relationship while in each period, there may be deviations from said relationship with $z_t \neq 0$.

[2] Recent works have also expanded on the original definition of cointegration, e.g. Zhou introduces a methodology to test for nonlinear cointegration in [26]

Error correction as the representation of long run relationships Given a stable long run relationship as defined by cointegration with possible short run deviations, one modeling approach is the error correction model. In essence if y_t is cointegrated with a linear combination of $\boldsymbol{X_t} = (x_{1t}, x_{2t}, ..., x_{Nt})$, and given certain conditions captured in the Granger Representation theorem (see [12]), the relationship can be modeled as follows:[3]:

$$\Delta y_t = \mu + \alpha(y_{t-1} - \boldsymbol{\beta X_{t-1}}) + \sum_j \gamma_{dy} \Delta y_{t-j} + \sum_k \boldsymbol{\gamma_{dX}} \Delta \boldsymbol{X_{t-k}} \tag{5}$$

Where $\Delta X_t = X_t - X_{t-1}$ is the difference operator, and $\alpha(y_{t-1} - \beta X_t)$ is the error correction term. The error correction term captures any difference in observed y compared to its long run relationship $\boldsymbol{\beta X}$ and causes Δy to adjust by a factor of α, hence α is also called the "speed of adjustment". If a process exhibits a stable long run relationship, we expect $-1 < \alpha < 0$, which causes any deviations in the long run relationship to be corrected over time. Conversely, if $\alpha > 0$, any difference would be amplified over time and cause y and $\boldsymbol{\beta X}$ to diverge; $\alpha < -1$ would cause Δy to "over adjust".

Examples of applications of the inverted demand approach include [11, 3].

Duca et al. [11] also examine the impact of LTV on first time homebuyers using an inverted demand approach. Similar to [10], this work finds a long run relationship between housing price and fundamentals, which is further augmented by including LTV ratios. Results also suggest that credit constraints, as measured by LTV, represents one of the main factors driving the boom and bust from 2000 to 2007.

[3]We refer to [12, 6, 16] for a detailed discussion of the origin and background theory of error correction models

Recent work by Anundsen & Heboll [3] has also looked at the US housing market using an inverted demand approach using a sample from 1980q1 to 2010q2. In addition to estimating a long run relationship, the authors test whether regions can be pooled when conducting the analysis. This is achieved using a pooled mean group estimator, first defined in [19], which imposes a homogeneous long run relationship while allowing for short term discrepancies individual to each region.

2.3 Application of the LPPLS model in the real estate market

The Chair of Entrepreneurial Risk of ETH Zurich publishes a biannual risk analysis of the real estate market in Switzerland, with the most recent one published in 2015q2 [5].

The purpose of the study is "to provide buyers and sellers in the Swiss real estate market with critical information on price dynamics in every Swiss district". The analysis is based on applying a model which combines the log period power law singularity analysis(LPPLS) with a diffusion index to diagnose the risk of real estate overvaluations in Switzerland, the methodology for which is described in [4, 5].

The analysis is conducted using data collected from comparis.ch and comprises about 1'403'000 residential properties for sale between 2005q1 and 2015q2. The data covers both, houses and apartments and their respective asking price (as opposed to final transaction price).

The most recent diagnostic results in several findings. First, the study identifies the districts with highest median asking price per square meter for apartments and houses between 2007q1 and 2015q2. In this period, the districts Zug and Meilen are consistently ranked among the top three, for both categories apartment and houses. Second, a classification of the districts with respect to their status as diagnosed by the LPPLS model is conducted.

The LPPLS analysis classifies the status of a district regarding its risk of a real estate bubble in descending order as follow: “Critical”, “To watch”, “To monitor” and “Regime change”. The status “Critical” indicates a strong bubble signal with imminent regime change ahead, “To watch” is a less strong signal and “To monitor” occurs only after a “Critical” or “To watch” state has been obtained but not enough data points are present to confirm a change of regime. Lastly, the status “Regime change” occurs only after the status “To monitor” has been obtained and with the confirmation of the latest data points. For the detailed model construction and fitting we refer to the original documents [4, 5].

In summary, the current issue (2015q2) of the LPPLS diagnostics does not identify a bubble signal at the national level, with a of total, 11 districts categorized as either “To watch”(5) or “To monitor”(6). Previous issues’ “Critical” classifications for certain districts have since subsided [4].

Overall, given the moderate warning of the LPPLS analysis and other important economic factors such as low interest rate, unstable European geopolitics and slow European economy recovery, the report suggests the Swiss real estate market will remain stable with only moderate adjustments.

3 Methodology and Model Description

The aim of this work is to conduct a fundamental analysis of the Swiss housing market. Due to our belief of regional heterogeneity in the housing

[4]previous “Critical” classifications include: District Baden (Aargau, report 2013q2), District Bülach (Canton Zurich, reports 2013q2-q4) and Dielsdorf (Canton of Zurich, report 2013q2)

market, we test the validity of imposing homogeneous model coefficients (aggregating) at different regional levels as part of this analysis. Additionally, we examine the fundamental relationship over different timeframes.

Therefore, our methodology can be structured in five steps which are described in more detail in the following subsections: we begin by selecting the fundamental factors to be considered (3.1), and by choosing an appropriate modeling approach (3.2). With model and input selected, we continue by fitting the model parameters to the available data (3.3). Once the model is properly specified, we test the poolability assumption (3.4) and finish with a look at the evolution of the different model estimates over different timeframes (3.5).

3.1 Fundamental factor selection

Economic theory suggests a relationship between supply & demand variables and house prices. Previous works have used a varying combination of variables and operationalizations, chosen based on economic theory or research hypothesis, including for example: rent, mortgage rate, tax, depreciation and credit constraint (Loan-to-value) [11]; rent, income, population, mortgage volume and rates, housing stock and construction cost [18]; income, mortgage, construction cost [26]; and one specific example for the Swiss market included GDP, wage, population, depreciation and price appreciation, construction costs and interest rates [21] [5].

A notable exception to the aforementioned examples is the approach used by Ardila et. al in [4]. The authors select a pool of 90 potential input variables for the US market and 27 variables for the Swiss market. This selec-

[5]The aim of [21] is to examine the relationship of fundamentals and real estate investment behavior, as opposed to price development, using a stock-flow-model;

tion is then narrowed down using sparse partial least squares. The full set of variables can be categorized into variables measuring output & income, employment, housing availability, price levels, money and credit, bond and exchange rates as well as stock market indices.

Our criteria for choosing the fundamental factors, given the colorful palette of variables in the literature are: (i) variables which together represent both demand and supply side, (ii) variables recurring throughout literature, (iii) limiting ourselves to a few number of fundamental variables for which data is available.

Our final selection includes population and income for the demand side, as well as housing stock for the supply side. Additionally, we include interest rates as a high level proxy for mortgage rates and financing cost for construction for the demand and supply side respectively. As an auxiliary variable, CPI data was included to adjust income for inflation.

3.2 Selecting an appropriate model functional form

We believed there was potential regional difference in the behavior of real estate prices at the district level in the Switzerland due to the differences in economic activity across regions. This characteristic is suggested by the difference in GDP contribution per canton (figure 1), where the majority of Swiss GDP can be attributed to less than half a dozen of regions.

Given this suspected regional difference, we settled on the same approach similar to [3], with the goal of estimating a model which allows for short run heterogeneity (say, due to different economic shocks and transient differences) while allowing for a fundamental driven relationship in the long run.

3.2.1 Model Equation

The model we fitted can be described by the following equation [3]:

$$\Delta y_{i,t} = \mu_i + \alpha_i(y_{i,t-1} - \boldsymbol{\beta}_i \boldsymbol{X}_{i,t-1}) + \sum_{s=1}^{p-1} \gamma_{i,s} \Delta y_{i,t-s} + \sum_{s=0}^{q-1} \boldsymbol{\lambda}_{i,s} \Delta \boldsymbol{X}_{i,t-s} \quad (6)$$

Where i represent the different regions, t represents each time period, μ_i is a region specific, time constant intercept, $y_{i,t-1}$ is the house price of region i in time period $t-1$, $\boldsymbol{X}$ is the matrix of explanatory variables, $\Delta y_{i,t-s}$ and $\Delta \boldsymbol{X}_{i,t-s}$ represent the first difference of y and X in region i, lagged by s periods. The lag lengths for the dependent and independent variables are specified by p and q respectively; a $p = q = 5$ specification corresponds to 4 lag terms (quarters) in the differences of each variable.

The coefficients $\boldsymbol{\beta_i}$ in equation (6) can be interpreted as long run coefficients of the fundamental factors in $\mathbf{X}$, α can be interpreted as an adjustment coefficient measuring the speed of adjustment with which prices correct towards (or away form) the long run relationship. If a stable long run relationship exists between y_i and $\boldsymbol{X_i}$, we expect $\alpha_i < 0$, which causes Δy to correct towards equilibrium. Alternatively, an $\alpha \geq 0$ amplifies any difference between the relationship of y_i and $\boldsymbol{\beta_i X_i}$.

The coefficients $\gamma_i, \boldsymbol{\lambda_i}$ can be interpreted as the effect of short run fluctuations of the dependent and independent variables respectively.

An alternate representation using

$$\boldsymbol{\theta}_i = -\alpha_i \boldsymbol{\beta_i}$$

will also be used in this work, which would yield for equation (6)

$$\Delta y_{i,t} = \mu_i + \alpha_i y_{i,t-1} + \boldsymbol{\theta}_i \boldsymbol{X}_{i,t-1} + \sum_{s=1}^{p-1} \gamma_{i,s} \Delta y_{i,t-s} + \sum_{s=0}^{q-1} \boldsymbol{\lambda}_{i,s} \Delta \boldsymbol{X}_{i,t-s}$$

3.3 Fitting the model to the data

We start with a base case where we include all four factors ($k = 4$), and four lags in differences for the dependent and independent variables ($P = Q = 5$), corresponding to a lag length of one year.

3.3.1 Lag length selection

First we estimate individual region specific models at the least aggregated level (district level) with varying lag lengths with $P, Q \in \{5, 4, 3, 2\}$, and compare the resulting models using log likelihood ratios, the Akaike Information Criterion (AIC), the Bayesian information criterion (BIC) and AIC's small sample corrected version (AICc).

The panel Log Likelihood (pLL) We implemented the panel log likelihood as one of the criteria to compare the models with different lag lengths. At this point, we are solely interested in testing model lag length specification (independent of aggregation assumptions), and therefore treat each region as independent. As such, we fit 6 separately for each region's data and calculate the individual likelihoods. Finally, to measure the fit of the model over all regions, we sum each regions likelihood, which can be formulated as:

$$\begin{aligned} &\mathrm{pLL}(\mu_i, \alpha_i, \boldsymbol{\beta_i}, \gamma_i, \boldsymbol{\lambda_i}, \sigma_i^2) = \\ &\qquad -\frac{T}{2}\sum_{i=1}^{N} \ln 2\pi\sigma_i^2 - \frac{1}{2}\sum_{i=1}^{N}\frac{1}{\sigma_i^2}(\Delta \boldsymbol{y_i} - \widehat{\boldsymbol{\Delta y_i}})'(\Delta \boldsymbol{y_i} - \boldsymbol{\Delta}\hat{\boldsymbol{y}}_{\boldsymbol{i}}) \end{aligned} \tag{7}$$

where $\widehat{\boldsymbol{\Delta y_i}}$ represents the estimates of $\boldsymbol{\Delta y}$ from each individual regional fit of equation (6), N is the number of regions and σ^2 is the variance of the estimation error.

The log likelihood ratio used to compare two nested models is then:

$$LLR = 2log(\sqrt{\mathcal{LL}}_{ur}) - 2log(\sqrt{\mathcal{LL}}_{r}) \tag{8}$$

, where ur, r stand for the unrestricted and restricted model. As by Wilk's theorem [25], the LLR is approximately χ^2 distributed and as such can be used for testing the rejection of the null hypothesis (restricted model).

Information criteria AIC, BIC and AICc The tree information criteria published in [2], [20] and [15], provide three means to evaluate the fit between different nested models. They are defined as

$$AIC = -2log(\mathcal{L}) + 2k \tag{9}$$

$$BIC = -2log(\mathcal{L}) + k\ log(n) \tag{10}$$

$$AICc = -2log(\mathcal{L}) + \frac{2k(k+1)}{n-k-1} \tag{11}$$

, where $\mathcal{L}$ represents the likelihood of a model fit, k the number of independent linear parameters and n the number of available observations in the sample. The criteria allow for comparison between different models, however they do not provide information on the fit of a model. Selecting the model with the best fit equals minimizing the respective criteria.

As can be seen from their final definitions[6], they similarly capture the fit of a model based on its likelihood. Where the differ is how they adjust for

[6] for the nuances in their derivation, we refer to the original publications; AIC [2], BIC [20] and AICc [15]

sample size and number of parameters in the second term. While the original AIC adjusts for the number of parameters and disregards sample size, BIC and AICc both also include a terms for the sample size. AICc's extra term $\frac{2k(k+1)}{n-k-1}$ gives the strongest penalty for additional parameters and fewer observations, and as the authors in [15] have shown, it is preferable over AIC and BIC in small sample sizes.

3.3.2 Confirming factor selection

In a second step, we compare models (also at the lowest aggregation level) where factors are removed, one at a time. This was conducted in order to assess whether some of the selected parameters do not contribute additional information. The resulting models were again compared using the same statistics, log likelihood ratio, AIC, BIC and AICc.

After these two steps a final model is selected with appropriate lag length and factor selection.

3.4 Testing for coefficient homogeneity on different aggregation levels

Once the model lag lengths and factors have been selected, we test for long run coefficient homogeneity across different aggregation levels using likelihood ratio tests. Our hypotheses can be stated as:

$$\begin{aligned} H_0 &: \boldsymbol{\beta}_i = \boldsymbol{\beta},\ \forall i \\ H_A &: \boldsymbol{\beta}_i \neq \boldsymbol{\beta},\ \forall i \end{aligned} \tag{12}$$

To test the hypothesis, we fit the model (6) separately using three estimators: the dynamic fixed effects (DFE) estimator, the group mean(GM) estimator and the pooled group mean estimator(PMG).[7].

Out of the three, DFE and GM impose a high degree of homogeneity for the coefficients: with DFE, all coefficients are assumed homogeneous over all regions, with only the intercept μ_i allowed to vary ($\alpha_i = \alpha, \boldsymbol{\beta_i} = \boldsymbol{\beta}, \gamma_i = \gamma, \boldsymbol{\lambda}_i = \boldsymbol{\lambda}$
$\forall i$). The GM estimator fits (6) for each region separately and calculates the mean of each coefficient over all regions, including the intercept.

Finally, as introduced by Pesaran et al. in [19] and applied in [3], the PMG represents a hybrid case. The PMG estimator imposes homogeneity restrictions on the long run coefficients $\boldsymbol{\beta_i} = \boldsymbol{\beta}$, while the adjustment coefficient α, short run coefficients γ_i, λ_i and intercepts μ_i are allowed to vary over regions. The long run coefficients $\boldsymbol{\beta}$ and adjustment coefficients α_i are obtained through maximum likelihood estimation of a concentrated (profile) likelihood in which is described below in (13).

The concentrated panel log likelihood For estimating the homogeneous long run coefficients with the PMG estimator, we implemented the concentrated (profile-) likelihood function (13), as introduced in [19][8].

$$\begin{aligned} \mathrm{cpLL}(\boldsymbol{\varphi}) = & -\frac{T}{2}\sum_{i=1}^{N} \ln 2\pi\sigma_i^2 \\ & -\frac{1}{2}\sum_{i=1}^{N} \frac{1}{\sigma_i^2}(\Delta \boldsymbol{y}_i - \alpha_i \boldsymbol{\xi}_i(\boldsymbol{\beta}))' \boldsymbol{H}_i (\Delta \boldsymbol{y}_i - \alpha_i \boldsymbol{\xi}_i(\boldsymbol{\beta})) \end{aligned} \tag{13}$$

where

$$\boldsymbol{\xi_i(\beta)} = \boldsymbol{y_{i,-1}} - \boldsymbol{X_{i,-1}\beta}$$

[7]DFE & GM calculations performed with the `plm` package version 1.4-0 in `R` version 3.2.2

[8]Difference to original: for $\boldsymbol{\xi}$ our model uses $\boldsymbol{X_{i,-1}}$ instead of $\boldsymbol{X_i}$, this lets previous period's discrepancy in the long run relationship determine this period's change in prices

represents the difference in y as predicted by the long run relationship, $\boldsymbol{H_i} = \boldsymbol{I_T} - \boldsymbol{W_i}(\boldsymbol{W_i'}\boldsymbol{W_i})^{-1}\boldsymbol{W_i'}$, $\boldsymbol{\varphi} = (\boldsymbol{\theta'}, \boldsymbol{\alpha'}, \boldsymbol{\sigma'})$, $\boldsymbol{\alpha} = (\alpha_1, ..., \alpha_N)'$, $\boldsymbol{\sigma} = (\sigma_1^2, ..., \sigma_N^2)'$ and $\boldsymbol{I_T}$ is the $(T \times T)$ identity matrix. Contrary to a full likelihood function, the profile likelihood function takes only the adjustment coefficient α_i and the restricted long run coefficients $\boldsymbol{\beta_i} = \boldsymbol{beta}$ along with the variances σ^2 as input. Overall, this serves the purpose of restricting homogeneity in $\boldsymbol{\beta}$ while allowing the short run coefficients $\gamma_i, \boldsymbol{\lambda_i}$ to vary per region.

Based on the results of the homogeneity test, we estimate models at the most appropriate aggregation level.

3.5 Comparing model estimations over different timeframes

Theory suggests that markets in equilibrium are governed by a set of long run relationships which are stable over time, with possible deviations in the short run due to external shocks and structural rigidities.

We expect a stable long run relationship to be reflected as stable long run coefficient estimates for $\boldsymbol{\beta}$, independent of the timeframe used. For this purpose, we examine the evolution of estimated coefficients by varying the endpoint of our timeframe, akin to gaining new data every quarter.

4 Sample description

4.1 Available data

Our dataset comprises macroeconomic time series at the national, cantonal, and district level. Cantonal data covers Geneva and Zurich, whereas municipal data contains time series for the 12 districts within the Canton of Zurich. Variables included are population levels, nominal income level, housing stock, interest rates, CPI and the hedonic price index as described in [14].

The raw data spans different time periods and has mixed frequencies (monthly, quarterly, annually) – for an overview see table 3. In order to conduct a quarterly analysis, cubic spline interpolation was used for time series of lower frequency to interpolate missing the data points. The final date for the beginning and end of the sample were selected in order to ensure a balanced (panel) dataset.

The final working dataset covers the 2000q4-2012q4 period, for a total of 49 observations per region per variable. For reference, the most recent cross section for 2012q4 is shown in table 1.

Following [14, 4], we treat the Canton of Geneva as a canton and as a district. Given Geneva's population and size (see table 2), this is a sensible assumption.

In the following subsections, we describe the available data for each variable in more detail with an example plot.

4.1.1 Population data

Population data is readily available for all regional levels. The national level, aggregated series was available as annual values for the years 1981 to 2013, except for 1983 which was missing.

For the canton of Zurich and its individual districts, the data consists of annual values for the period 1962 to 2014.

Population data for Geneva represented annual levels between the years

Table 1: Panel data sample for 2012q4

In decreasing order of price index

	price_hedo_apt	pop	incr_median	hstock	X3mCHFL
Bezirk Zürich	146.200	379915	428.307	211942	0.012
Bezirk Meilen	144.740	97927	547.641	47854	0.012
Bezirk Bülach	142.740	135709	468.085	63551	0.012
Canton de Genève	140.550	470512	781.721	221880	0.012
Bezirk Affoltern	138.960	49384	493.987	22173	0.012
Bezirk Horgen	137.200	118462	491.212	58006	0.012
Kanton Zürich	133.713	1406083	458.834	693922	0.012
Bezirk Uster	132.290	122694	486.586	57532	0.012
Bezirk Dietikon	131.680	83590	448.659	39845	0.012
Bezirk Winterthur	131.460	158001	438.483	74638	0.012
Schweiz	129.956	8123721	411.656	1670054	0.012
Bezirk Pfäffikon	124.060	57269	462.535	26376	0.012
Bezirk Hinwil	122.480	90616	433.858	41280	0.012
Bezirk Dielsdorf	119.030	82516	470.860	37248	0.012

Variables	
price_hedo_apt	Hedonic price index for apartments as given by [14]
pop	Population numbers
incr_median	Median real income
hstock	Housing stock (# dwellings in districts/canton; # bldgs. at national level)
X3mCHFL	3 month CHF LIBOR

1989 and 2014.

4.1.2 Income level

The income data used at the national level is the nominal median *annual* nominal income, *categorized into four different employment groups*[10]. For this work, the mean value of the four categories was used. The timeframe available spans 1991 to 2014, reported annually.

[10] Categories are: self-employed, employee, family business and apprentice

Table 2: Comparison of Canton of Zurich vs Canton of Geneva

As of 2014q4	Canton of Zurich	Canton of Geneva
Geographic size [km^2]	1728.96	282.49
Population	1'443'436	482'545
Districts	12	–
Municipalities	169	45

The income data used within the Canton of Zurich is the taxable median *annual* nominal income. The available time span includes the years 1999 to 2012.

For the Canton of Geneva, the data represents the median *monthly* nominal income. The data was available between 2000 and 2012. All income levels are reported in Swiss Francs

The final values used in modeling are the *real* incomes adjusted for inflation as calculated by dividing each time series by the national level CPI[11].

4.1.3 Interest rate data

For interest rates the monthly 3 months CHF Libor rate at the end of each quarter is used *for all regional levels* similar to Anundsen [3].

[11]While cantonal level cpi was available, they showed high correlation to the national level aggregation

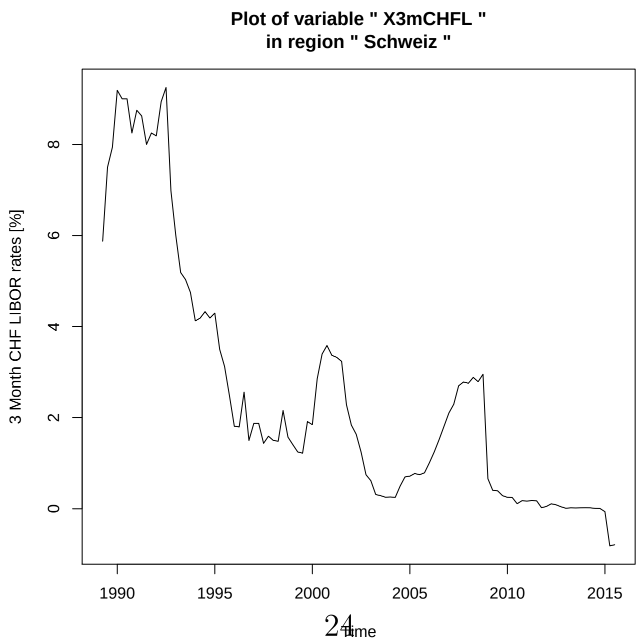
Table 3: Overview of available variables and their data sets for this study

Variable	**Region**	**Timeframe**	**Frequency**	**Source**	**Comment**
Hedonic Price index	All:	2000q4–2015q2	quarterly	[14]	–
Population level	Nat:	1982–2014	annual	FSO	1983 missing
	Zhr:	1962–2014	annual	ZSA	–
	Gnv:	1989–2014	annual	OCSTAT	–
Income level (median)	Nat:	1991–2014	annual	FSO	by 4 employment categories[9]
	Zhr:	1999–2012	annual	ZSA	annual inc
	Gnv:	2000–2012	annual	OCSTAT	monthly inc
CPI	All:	2000.05–2015.05	monthly	FSO	indexed to May 2000
Interest rates	All:	1989.03–2015.06	monthly	SNB	*
Housing stock	Nat:	1990, 2000, 2009–2013	mixed	FSO	–
	Zhr:	1990–2012	annual	ZSA	–
	Gnv:	1998–2014	annual	OCSTAT	–

Nat	: National level data
Zhr	: Zurich cantonal and district data
Gnv	: Geneva cantonal data
FSO	: Swiss Federal Statistics Office
ZSA	: Statistical Bureau of Canton of Zurich (Statistisches Amt)
OCSTAT	: Statistical office of Canton of Geneva (Office cantonale de la statistique)
SNB	: Swiss National Bank
*	rates assumed identical across all districts as 3 month CHF LIBOR

Figure 2: Population time series for district of Zurich and national level aggregation

Red diamond markers indicate original time series data points which were used to interpolate quarterly values; Black line is final quarterly series with interpolated values; regarding year index: 2000q1 = 2000.25, ..., 2000q4 = 2001



Time index: 2000q1 = 2000.25, ..., 2000q4 = 2001

Figure 4: Interest rates time series (3 month CHF Libor) used for all districts

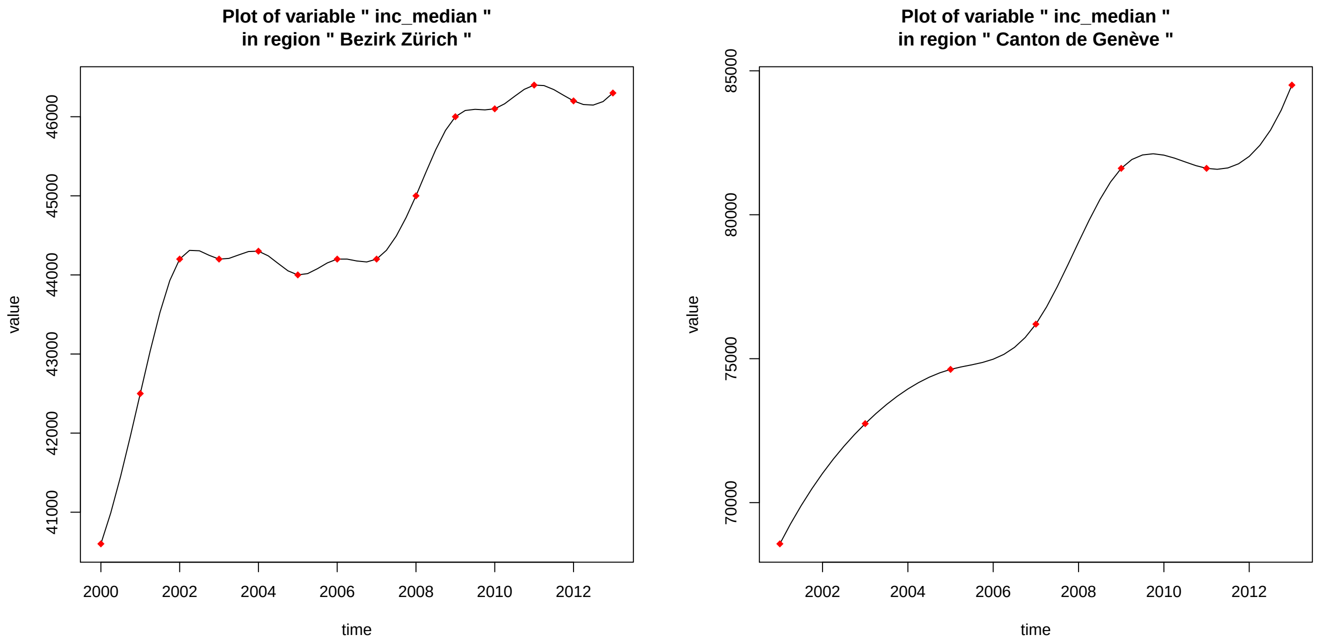


Figure 3: Population time series for district of Zurich and national level aggregation

Red diamond markers indicate original time series data points which were used to interpolate quarterly values; Black line is final quarterly series with interpolated values; regarding year index: 2000q1 = 2000.25, ..., 2000q4 = 2001

4.1.4 Housing stock data

The data points for the national level aggregation are infrequent. They represents the number of *buildings* with residential use[12] at the end of the year, with data points available for the years 1990, 2000, and 2009 to 2013 (7 data points in total); this represents the most sparsely available series.

For Zurich, the data represents the number of *dwellings* from 1990 to 2013, reported annually by the Zurich Cantonal statistics office.

Similar for Geneva, the data also represents the number of *dwellings*,

[12]including different types such as single family homes, apartment buildings, and buildings with partial residential use

between 1998–2014.

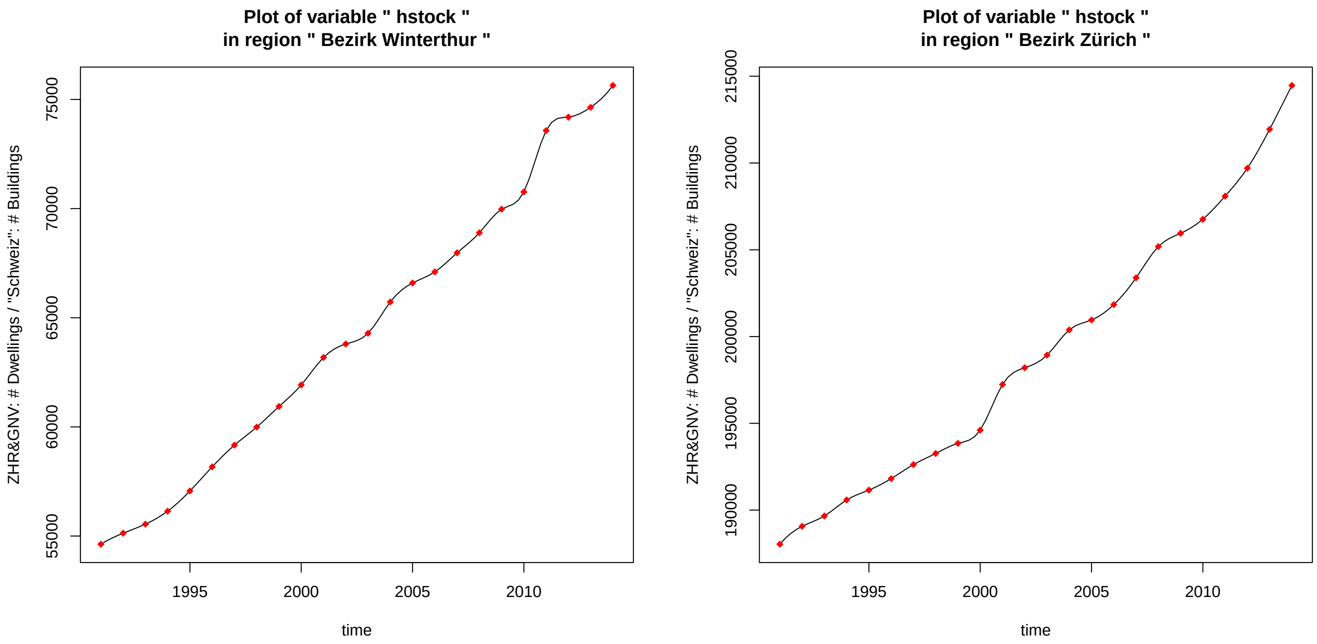


Figure 5: Housing stock time series for district of Zurich and national level aggregation

Red diamond markers indicate original time series data points which were used to interpolate quarterly values; Black line is final quarterly series with interpolated values

4.1.5 Real estate price data

We capture real estate prices for apartments using the hedonic price index as constructed in [14].

The index represents the estimated price of an 'average' apartment in each district as a function of apartment size (m^2) and a number of hedonic variables; hedonic features included are: construction year, surrounding area type (e.g. city center, suburb, rural, etc.), number of bathrooms, number of garages available, neighborhood 'quality' and building quality. An 'average apartment' is constructed from all measured values for the hedonic factors and then its price evolution fitted over time. For more detailed information we refer to the contents of [14].

The final index is reported quarterly between 2000q4 – 2015q2, normalized to 2007q1 and represent the prices for apartments.

They are available for 72 district, including Canton of Geneva as a district, and the national level aggregation represented by a weighted average of all districts.

By using a hedonic index (as compared to a traditional price-only index) we hope to capture some of the qualitative differences of the underlying assets.

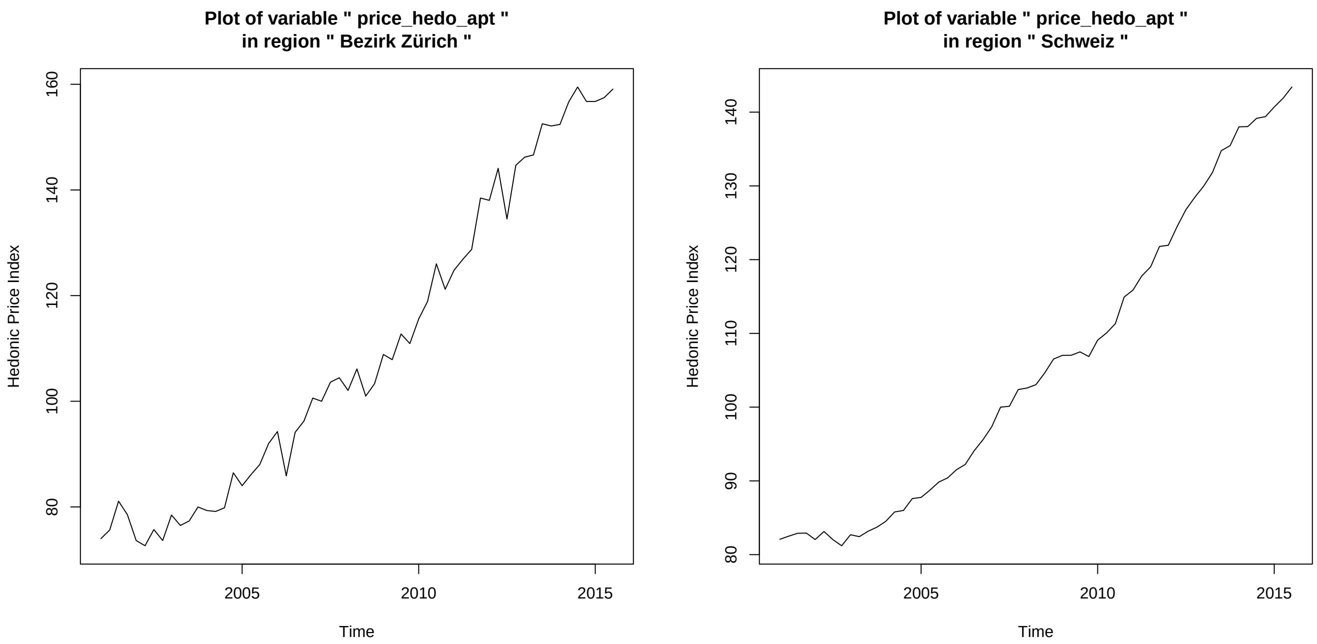


Figure 6: Hedonic price index time series for district of Zurich and national level aggregation

4.2 Unit root testing for panel data

Unit root processes are strongly dependent and highly persistent. As such, inference statistics of any estimated (level) linear model including a unit root process is highly sensitive to violations of the classic linear model assumptions[13].

We use the Augmented Dickey Fuller test on each time series to test for unit root, using the final time frame 2000q4-2012q4. The test regression is formulated as

$$\Delta y_t = \gamma y_{t-1} + a_0 + \sum_{i=1}^{4} a_i \Delta y_{t-i}$$

where y_t represents the time series tested, with the hypotheses formulated as

$$H_0 : \gamma = 0 \text{ (unit root present)}$$
$$H_A : \gamma < 0$$

The tests include a drift a_0, and a maximum lag length of 4, with the final lag length determined via the AIC criterion. The results, displayed in table 4, do not reject the null of unit root processes in (almost) any of the time series. The only exception is the district of Bülach, which is only marginally rejected at the 0.05 significance level. Surprisingly, we cannot reject a unit root for interest rate in the given time period which is contrary to expectation[14].

We specify a max lag length of 4, in order to allow each unit root regres-

[13]assumptions with u_t denoting estimation errors: no perfect collinearity; zero conditional mean $E(u_t|\boldsymbol{X}) = 0$, which also implies strict exogeneity for X; homoscedasticity $Var(u_t|\boldsymbol{X}) = \sigma^2\ \forall t$; no serial correlation $Cor(u_t, u_s|\boldsymbol{X}) = 0,\ \forall t \neq s$

[14]we note that similar observations have been encountered, such as by Lai in [17], however we postpone an excursion into macroeconomics for later projects

sion to include one non-interpolated data point for the time series which are interpolated.

Given that the data sample in each region is relatively short (49 observations), AIC will allow more parameters than a criteria adjusting for small sample size (say AICc). As such, we give the ADF regression more chances to retain a higher number of lags, and as such, favor the case where at least one non interpolated data point enters the regression.

Table 4: ADF Test results

ADF Test t-statistics: We reject a unit root only for the income time series in Bezirk Bülach at the 5% level

Region	price_hedo_apt	pop	incr_median	hstock	X3mCHFL
Bezirk Affoltern	2.596	0.130	-1.459	0.398	-1.602
Bezirk Bülach	3.080	0.540	-3.035**	-0.633	-1.602
Bezirk Dielsdorf	0.594	0.974	-2.446	0.749	-1.602
Bezirk Hinwil	3.036	0.532	-1.804	1.325	-1.602
Bezirk Horgen	1.824	1.470	-2.605*	1.417	-1.602
Bezirk Meilen	1.697	-0.724	-1.584	-1.171	-1.602
Bezirk Pfäffikon	1.483	1.699	-1.718	0.985	-1.602
Bezirk Uster	2.832	0.240	-2.847*	-0.793	-1.602
Bezirk Winterthur	1.684	1.097	-2.536	0.004	-1.602
Bezirk Dietikon	0.954	0.723	-2.774*	1.095	-1.602
Bezirk Zürich	2.583	0.256	-2.242	1.795	-1.602
Canton de Genève	-0.900	-0.522	0.290	0.607	-1.602
Schweiz	1.317	-0.053	-1.684	-0.791	-1.602

Signif. codes: 0.001 '***' 0.01 '**' 0.05 '*' 0.1 ' ' 1

ADF critical values with drift and 49 observations

critical level	1%	5%	10%
critical value	-3.58	-2.93	-2.60

Variables

price_hedo_apt	Hedonic price index for apartments as given by [14]
pop	Population numbers
incr_median	Median real income (nominal income divided by national cpi)
hstock	Housing stock (# dwellings in districts/canton; # bldgs at national level)
X3mCHFL	3 month CHF LIBOR

5 Results

5.1 Model specification

5.1.1 Lag length and confirming factor selection

Starting with the base case ($P = Q = 5$; all four factors population, income, housing stock, rates), we estimated coefficients using the three methods as described in section 3.3. Resulting individual region model fits show signs of overfitting, with R^2 values as high as 0.90 (for example, see figure 7, p.32). For reference, Abraham and Hendershott in their earlier study of the US housing market [1] find R^2 in the range of $0.5 - 0.6$.

This overfit likely results from the large number of parameters relative to the number of observations per region – with (P,Q) = (5,5), we have 31 parameters to be estimated per region while the number of observations per region is limited to 49. The effective number of observations is further reduced to 44, due to the lags and differences in the variables.

Lag length selection When comparing models with shorter lag lengths (P,Q $\in \{5, 4, 3, 2\}$), AIC and BIC suggest keeping the maximum possible lag length of base case, despite the overfit (see table 5, p. 33). In contrast, AICc favors the two cases of $(P, Q) = (3, 2)$ and $(P, Q) = (2, 2)$, with their respective AICc values being almost identical[15]. However, when using likelihood ratios to compare different lag lengths, any restricted model results in a $p << 0.001$ and would therefore be rejected.

[15] for easy readability we report $\Delta AICc$ which shows the improvement of the restricted model AICc over the unrestricted base case

Figure 7: Initial model fit results: Actual vs DFE vs GM estimations (P = 5, Q = 5)

Signs of overfitting at the individual district level; extremely close fit of individual estimation(red) to the actual data(black) for the selected parameters; cause is likely large number of parameters (31) vs effective number of observations per district(44)

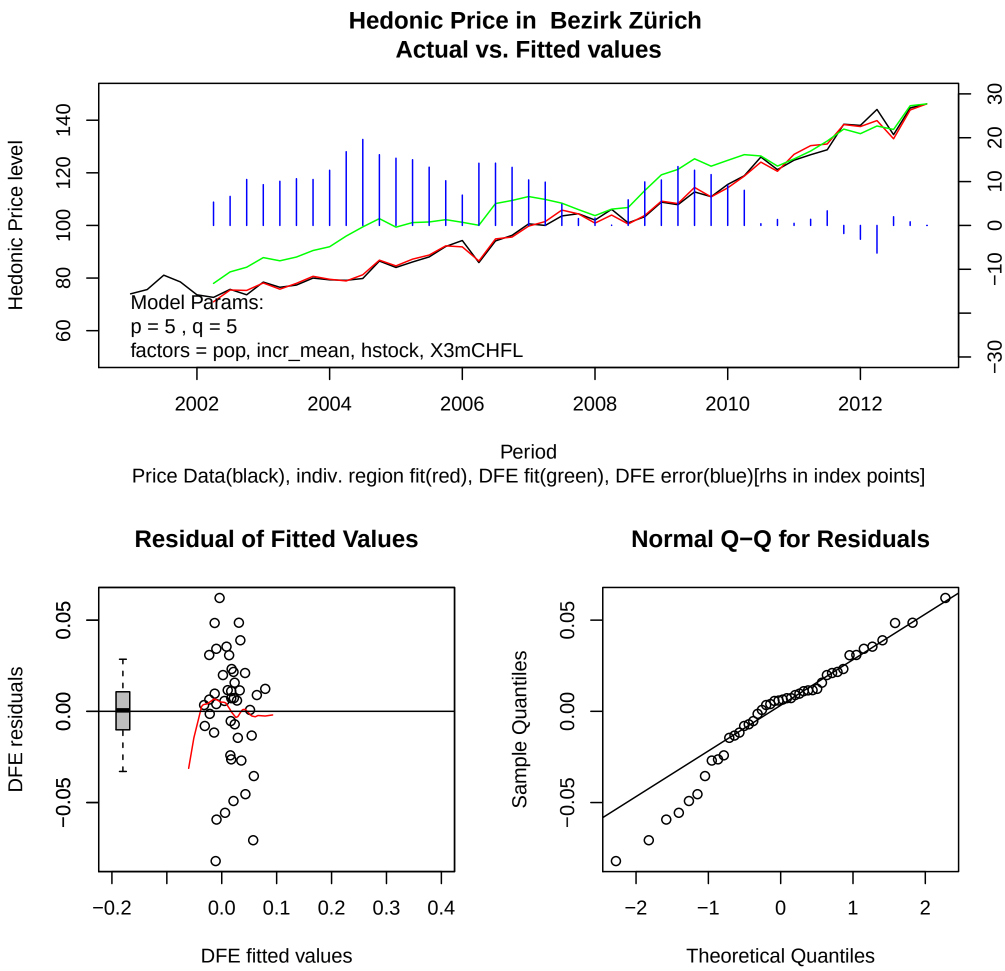


Table 5: Model selection criteria for lag lengths

The unrestricted (ur) model is t P = Q = 5; highlighted value is best fit according to AICc; smaller ΔAIC = better fit of restricted (r) model vs ur.

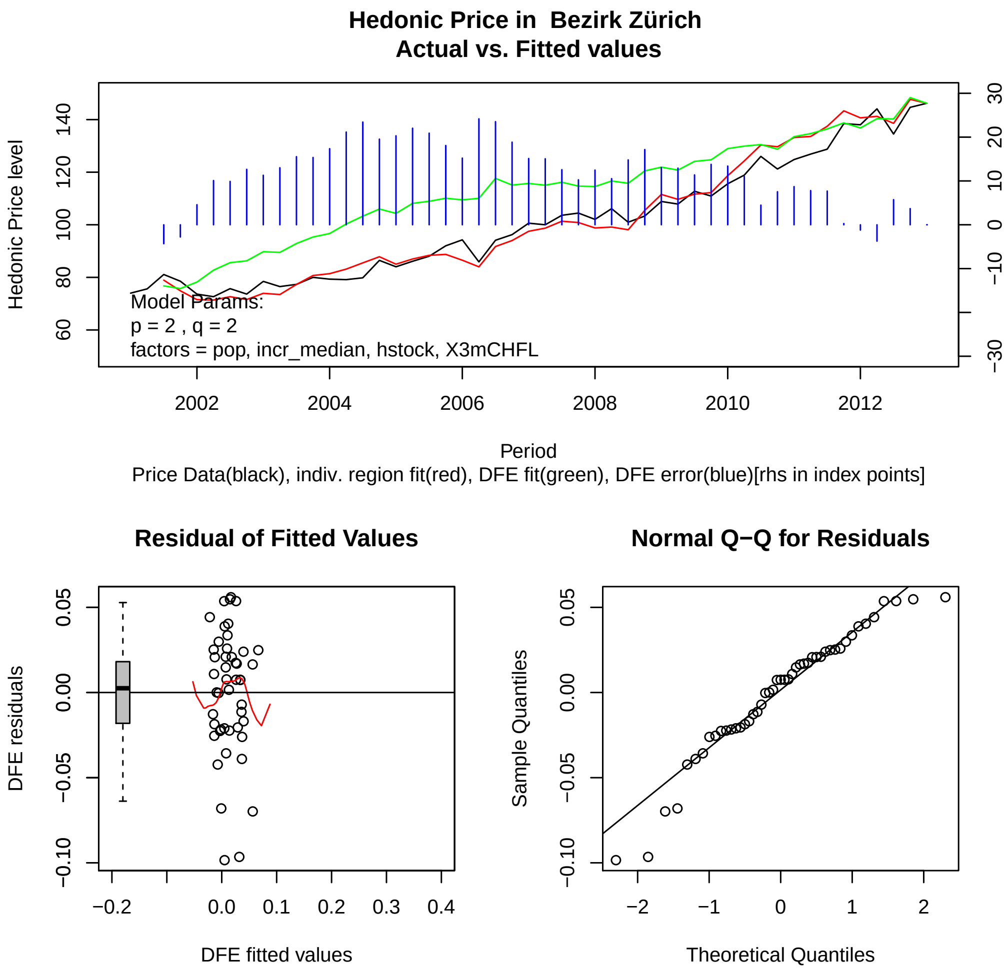
	P	Q	LL.ur	LL.r	LL.ratio	chisqdf	pval	n.obs
1	5	5	1614.46					484
2	5	4	1614.46	1466.89	295.13	44	0.00	484
3	5	3	1614.46	1405.14	418.63	88	0.00	484
4	5	2	1614.46	1343.59	541.74	132	0.00	484
5	4	5	1614.46	1582.29	64.34	11	0.00	484
6	4	4	1614.46	1429.84	369.24	55	0.00	495
7	4	3	1614.46	1378.53	471.86	99	0.00	495
8	4	2	1614.46	1315.91	597.10	143	0.00	495
9	3	5	1614.46	1551.02	126.87	22	0.00	484
10	3	4	1614.46	1409.66	409.59	66	0.00	495
11	3	3	1614.46	1347.58	533.76	110	0.00	506
12	3	2	1614.46	1280.99	666.93	154	0.00	506
13	2	5	1614.46	1527.68	173.56	33	0.00	484
14	2	4	1614.46	1387.09	454.73	77	0.00	495
15	2	3	1614.46	1328.06	572.79	121	0.00	506
16	2	2	1614.46	1250.29	728.34	165	0.00	517

	P	Q	k.ur	k.r	dBIC	dAIC	dAICc
1	5	5	31				
2	5	4	31	27	128.62	207.13	-572.04
3	5	3	31	23	85.62	242.63	-968.83
4	5	2	31	19	42.23	277.74	-1192.60
5	4	5	31	30	22.71	42.34	-202.48
6	4	4	31	26	167.53	259.24	-701.43
7	4	3	31	22	102.66	273.86	-1038.81
8	4	2	31	18	60.41	311.10	-1218.19
9	3	5	31	29	43.61	82.87	-368.66
10	3	4	31	25	166.01	277.59	-788.45
11	3	3	31	21	127.77	313.76	-1081.41
12	3	2	31	17	92.48	358.93	-1219.31
13	2	5	31	28	48.69	107.56	-520.17
14	2	4	31	24	169.28	300.73	-857.94
15	2	3	31	20	124.68	330.79	-1118.28
16	2	2	31	16	115.56	398.34	**-1220.86**

pLL : panel log Likelihood
k : number of parameters in model to be estimated
ΔXIC : smaller is better; calculated as XIC.r - XIC.ur;
regions = 11 (ZH excl. Andelfingen)

Figure 8: Reduced lag length model fit as recommended by AICc: Actual vs DFE vs GM estimations (P = 2, Q = 2)

GM individual region estimation(red) compared to the actual data(black) for the selected parameters still shows extremely close fit, with 22 parameters and 45 effective observations



Given the results described above, we decide for the case of $(P, Q) = (2, 2)$. Even though the likelihood ratios reject any model with shortened lag lengths, we note the high overfit at the individual district. As such, we gave precedence to the different information criteria during lag length selection. Within the three information criteria, we followed the recommendation of

AICc, as our sample is relatively small.

Confirming factor selection Testing for the relevancy of the selected factors yielded results summarized in table 6 (p. 36).

For the likelihood ratio tests, the pval of any model ($P, Q \in \{5, 4, 3, 2\}$) with a factor removed is $p << 0.001$ and thus the restricted model is always rejected.

Similarly, the AIC values of the unrestricted model are, in all but one case, consistency smaller than the values of any restricted model. Therefore the unrestricted model (factor included) is always preferable. The one exception is interest rates, for which the restricted model is slightly favored by AIC.

BIC and AICc on the other hand yield stronger results. At our selected lag length of (P,Q) = (2,2), BIC and AICc both favor the removal of each factor.

Ultimately, we reason to keep all four factors in the model due to the inconclusiveness of the likelihood ratio test, the relative weak evidence for removing a factor and theory suggesting a relationship with supply and demand.

Table 6: Model selection criteria for factors

The unrestricted (ur) model is the model including all factors at P, Q; the restricted (r) model has the factor in rmfactor entirely removed (long and short run effects)

Note: for spacing, test results for P,Q = {3,4} have been omitted; highlighted values represent AICc recommendations given our model; smaller ΔAIC = better fit of restricted (r) model vs ur; AICc suggests removing single factors (in order of significance) interest rates, income, housing stock and population

	P	Q	rmfactor	LL.ur	LL.r	LL.ratio	chisqdf	pval	n.obs
17	2	2	pop	1250.29	1197.63	105.32	33	0.00	517
18	2	2	incr_median	1250.29	1204.64	91.29	33	0.00	517
19	2	2	hstock	1250.29	1203.74	93.08	33	0.00	517
20	2	2	X3mCHFL	1250.29	1219.99	60.58	33	0.00	517
21	5	5	pop	1614.46	1407.72	413.47	66	0.00	484
24	2	2	pop	1250.29	1197.63	105.32	33	0.00	517
25	5	5	incr_median	1614.46	1473.51	281.90	66	0.00	484
28	2	2	incr_median	1250.29	1204.64	91.29	33	0.00	517
29	5	5	hstock	1614.46	1394.16	440.59	66	0.00	484
32	2	2	hstock	1250.29	1203.74	93.08	33	0.00	517
33	5	5	X3mCHFL	1614.46	1451.46	326	66	0.00	484
36	2	2	X3mCHFL	1250.29	1219.99	60.58	33	0.00	517

	P	Q	rmfactor	k.ur	k.r	dBIC	dAIC	dAICc
17	2	2	pop	16	13	-21.73	39.32	**-38.81**
18	2	2	incr_median	16	13	-35.77	25.29	**-52.85**
19	2	2	hstock	16	13	-33.97	27.08	**-51.05**
20	2	2	X3mCHFL	16	13	-66.47	-5.42	**-83.55**
21	5	5	pop	31	25	163.71	281.47	-742.75
24	2	2	pop	16	13	-21.73	39.32	-38.81
25	5	5	incr_median	31	25	32.14	149.90	-874.33
28	2	2	incr_median	16	13	-35.77	25.29	-52.85
29	5	5	hstock	31	25	190.83	308.59	-715.64
32	2	2	hstock	16	13	-33.97	27.08	-51.05
33	5	5	X3mCHFL	31	25	76.24	194	-830.22
36	2	2	X3mCHFL	16	13	-66.47	-5.42	-83.55

rmfactor : factor removed for testing
pLL : panel log Likelihood
k : number of parameters in model to be estimated
ΔXIC : smaller is better; calculated as XIC.r - XIC.ur;
regions = 11 (ZH excl. Andelfingen)

5.2 Testing for coefficient homogeneity

5.2.1 Testing pooling assumption at the cantonal level

We start the poolability test with the high level aggregation and test whether the two cantons (Zurich, Geneva) can be pooled for an aggregate model.

To test for homogeneity, we estimate the model with all three estimation methods (DFE, MG, PMG) and compare the results using the likelihood ratio test. The results of the poolability test are reported in table 7.

Out of the three estimation methods, PMG estimation results show the highest concentrated panel likelihood, followed by DFE results and GM results.

When testing the poolability hypothesis ($H_0 : \boldsymbol{\beta_i} = \boldsymbol{\beta}\ \forall i; H_A : \boldsymbol{\beta_i} \neq \boldsymbol{\beta}\ \forall i$), the resulting p-value $= 0.433$ indicates that we cannot reject the null hypothesis.

The resulting long run coefficient estimates seem to indicate that population, housing stock and rates have a positive relationship with apartment prices while income has a negative coefficient, inconsistent with the theory.

Table 7: PMG estimation results for testing cantonal Level aggregations with P = 2, Q = 2

Estimated LR coefficients

	cpLL	β_{pop}	β_{inc}	β_{hstock}	β_{rates}	α
DFE	266.487	27.600	-8.997	-21.998	0.137	-0.057
GM	139.364	-0.455	-0.455	9.423	0.032	-0.552
PMG	274.634	0.347	-1.415	16.530	0.039	

LL ratio results

	ur.cpLL	r.pmg.cpLL	LL.ratio	chisqdf	pval
1	276.538	274.634	3.807	4	0.433

cpLL	concentrated panel log Likelihood
ur.xxx	unrestricted model, with region specific $\alpha_i, \beta_i, \gamma_i, \lambda_i$
r.pmg.xxx	restricted model with pooled long run coefficient, $\beta_i = \beta \forall i$
	PMG estimates have region specific α_i, found in appendix A.3

5.2.2 Testing pooling assumption at the district level

While the evidence supports poolability at the cantonal level, this may mask intra-cantonal heterogeneities at the district level. Hence, we continue our poolability test and apply the same methodology for all districts within the Canton of Zurich[16].

The estimation results for the district level using different estimators are reported in table 8 (p.39), with region specific adjustment coefficients α_i for the PMG estimation available in the appendix A.3 due to spacing.

Coefficient estimates indicate a positive relationship of prices with population, housing stock and rates, and a negative one with income. Similar to the cantonal level pooling, this is inconsistent with theory. However, contrary to the cantonal level poolability test, the PMG estimation (which has

[16] except district Andelfingen, for which the hedonic price index was not available; this leaves us with 11 districts

a higher likelihood ratio than GM and DFE) is rejected when compared to the individual, district specific models, with a likelihood ratio of 178.10 and a p-value << 0.001 (table 8, p.39).

Table 8: PMG estimation results for all districts of Zurich with P = 2, Q = 2

Estimated LR coefficients

	cpLL	β_{pop}	β_{inc}	β_{hstock}	β_{rates}	α
DFE	1125.069	5.873	-0.612	-1.329	0.009	-0.095
GM	491.215	-1.016	-1.016	1.171	0.012	-1.061
PMG	1170.099	2.632	-0.309	0.110	0.004	

LL ratio results

ur.cpLL	r.pmg.cpLL	LL.ratio	chisqdf	pval
1250.285	1170.099	160.373	40	0.000

cpLL concentrated panel log Likelihood
ur.xxx unrestricted model, with region specific $\alpha_i, \beta_i, \gamma_i, \lambda_i$
r.pmg.xxx restricted model with pooled long run coefficient, $\beta_i = \beta \forall i$
PMG estimates have region specific α_i, found in appendix A.3

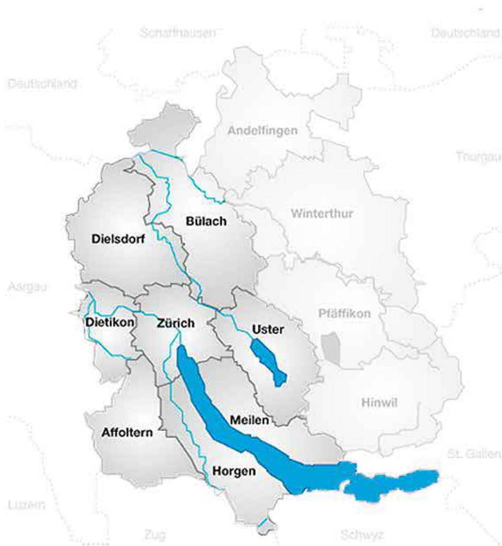
Further subdivision of the canton of Zurich We continue by further subdividing the pooling regions into three smaller geographical regions. For this, we select districts based on their proximity to Zurich's city center and split into three subsets: Zurich subregion I includes all districts bordering the city of Zurich, while Zurich subregion II consists of all the other districts (see figure 9, p. 41). Subregion III represents the districts bordering the Zurich lake, overall known for its affluent quarters and high end housing, including the district of Meilen, colloquially termed "Gold Coast".

The hypothesis behind this subdivision is that districts with a close proximity to Zurich district will have closer economic ties to the city center (e.g. more desirable for commuters, better transportation links etc.) while the other districts may be linked to other centers of economic activity, e.g. the next bigger city of Winterthur in the district of Winterthur.

We test the pooling assumption based on these smaller geographical regions and find results as shown in table 9 on page 42.

For subregion I & II, similar to the results for all districts, likelihood ratio test rejects the pooling assumption at the 5% level with $pval \approx 0.000$ for subregion I and $pval = 0.036$ for subregion II, albeit marginally.

Lastly, results indicate that we *cannot* reject the pooling assumption for subregion III at the 5% confidence level with a $pval = 0.094$. Coefficient estimates for subregion III indicate, again, a positive relationship with population and housing stock, a negative relationship with income and no relationship with interest rates. Adjustment coefficients range between $[-0.93, -0.13]$ (see appendix A.3).



Zurich subregion I

Zurich subregion II

Zurich subregion III

Figure 9: Subdivision of Zurich districts into three subregions

Original Image source: Canton of Zurich; `http://www.bezirke.zh.ch/`

Table 9: PMG estimation results for Zurich subregions with P = 2, Q = 2

Zurich Subregion I: Estimated LR coefficients

	cpLL	β_{pop}	β_{inc}	β_{hstock}	β_{rates}	α
DFE	823.002	7.294	-0.650	-1.921	0.022	-0.070
GM	343.283	-1.007	-1.007	1.291	0.013	-1.061
PMG	855.063	2.950	-0.400	-0.076	0.010	

Zurich Subregion I: LL ratio results

	ur.cpLL	r.pmg.cpLL	LL.ratio	chisqdf	pval
1	922.391	855.063	134.655	28	0.000

Zurich Subregion II: Estimated LR coefficients

	cpLL	β_{pop}	β_{inc}	β_{hstock}	β_{rates}	α
DFE	315.321	1.201	0.555	1.181	0.005	-0.606
GM	142.334	-1.041	-1.041	0.853	0.009	-1.160
PMG	319.648	2.177	0.736	0.469	0.008	

Zurich Subregion II: LL ratio results

	ur.cpLL	r.pmg.cpLL	LL.ratio	chisqdf	pval
1	327.894	319.648	16.492	8	0.036

Zurich Subregion III: Estimated LR coefficients

	cpLL	β_{pop}	β_{inc}	β_{hstock}	β_{rates}	α
DFE	321.129	5.179	-3.436	0.622	-0.008	-0.290
GM	252.573	-2.810	0.251	7.032	0.000	-1.015
PMG	331.420	4.131	-2.993	1.299	0.000	

Zurich Subregion III: LL ratio results

	ur.cpLL	r.pmg.cpLL	LL.ratio	chisqdf	pval
1	338.198	331.420	13.555	8	0.094

cpLL	concentrated panel log Likelihood
ur.x	unrestricted model, with region specific $\alpha_i, \beta_i, \gamma_i, \lambda_i$
r.pmg.x	restricted model with pooled long run coefficient, $\beta_i = \beta \forall i$
	PMG estimates have region specific α_i, found in appendix A.3

5.3 Estimating individual district models

Given that the district level pooling assumption is rejected when considering all districts within the Canton of Zurich, we report the individual district level fits of (6). A summary of the estimated long run coefficients can be found in table 10. Other short run parameter estimates can be found in appendix A.4 due to the spacing required to report all values.

As we would expect given the rejection of poolability in the previous sections, there seems to be no particular pattern in the sign, magnitude and significance of the long run coefficients ($\boldsymbol{\beta_i}$) across all districts.

Long run population coefficient is $\beta_{pop,i} > 0$ in 10 out of 14 regions considered, though only 4 estimates are significant at the 5% level in total; income is $\beta_{inc,i} > 0$ in 8 out of 14 regions, with 5 significant estimates in total; similarly $\beta_{hstock,i} > 0$ in 8 of 14 regions with 3 out of 14 significant in total[17].

It is interesting to note, all regions' interest estimates are consistently ≈ 0, albeit generally insignificant. This is consistent with the factor selection process where AICc suggested that the removal of interest rates would result in the least information loss (see table 6, 36).

Additionally, all regions exhibit a significant, negative adjustment coefficient ($\alpha_i < 0$), albeit with differing magnitude. This indicates that there is an error correction process in effect. The speed of adjustment varies, with the district of Meilen exhibiting the slowest correction ($\alpha_{Meilen} = -0.345$) and districts Dietikon, Hinwil and Horgen exhibiting slight "over-correction" ($\alpha_i < -1.100$).

[17] however, $\beta_{hstock,Geneva}$ seems unusually high compared to its peers.

Table 10: Overview of individual region model long run coefficient estimates

Expectation: theory suggests positive relationship with population & income; negative relationship with housing stock; error correction back to long run relationship with speed $-1 < \alpha < 0$

	α	β_{pop}	β_{inc}	β_{hstock}	β_{rates}	match expectation?
Bezirk Affoltern Estimate	-1.06***	-2.64	0.16	3.80.	-0.01	
SE	0.24	2.95	0.57	2.18	0.01	
Bezirk Bülach Estimate	-0.99***	-0.23	3.72**	3.65	0.04*	
SE	0.23	2.26	1.27	2.57	0.01	
Bezirk Dielsdorf Estimate	-1.06***	3.22*	0.44	-0.89	0.01	✓
SE	0.25	1.30	0.59	1.40	0.01	
Bezirk Dietikon Estimate	-1.39***	4.18	-0.72	-1.48	0.01	
SE	0.24	5.17	0.85	6.28	0.01	
Bezirk Hinwil Estimate	-1.16***	0.00	2.29.	1.96	0.02	
SE	0.26	2.30	1.24	1.86	0.01	
Bezirk Horgen Estimate	-1.40***	6.61	-3.63.	-1.14	0.01	
SE	0.28	4.77	1.87	3.74	0.02	
Bezirk Meilen Estimate	-0.35*	-6.78**	6.95***	9.21***	0.02	
SE	0.16	2.28	1.79	1.71	0.02	
Bezirk Pfäffikon Estimate	-0.96***	3.18	-0.45	-0.28	0.00	
SE	0.26	3.43	1.07	2.39	0.01	
Bezirk Uster Estimate	-0.95***	5.22***	-1.30	-2.21.	0.03**	
SE	0.23	1.39	0.87	1.19	0.01	
Bezirk Winterthur Estimate	-1.01***	1.92	1.10	0.87	0.01	
SE	0.26	3.10	0.95	3.03	0.01	
Bezirk Zürich Estimate	-0.87***	4.90***	-3.36**	-0.60	-0.00	
SE	0.20	1.27	1.06	2.65	0.01	
Canton de Genève Estimate	-0.55*	0.27	-1.37*	16.58**	0.04***	
SE	0.23	2.12	0.63	5.58	0.01	
Kanton Zürich Estimate	-0.36	1.87.	0.61	2.27*	0.03**	
SE	0.26	0.98	0.50	1.09	0.01	
Schweiz Estimate	-0.89***	4.55**	0.27*	0.09	0.03***	
SE	0.20	1.37	0.11	0.62	0.01	

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
SE: standard error; for βs: SE and significance represent estimates for $(-1) * \alpha\beta$

5.4 Coefficient estimates over time - Estimating model fits with varying timeframes

As a rough measure of judging whether the relationship between prices and the fundamental factors over time are stable, we estimate the relationship using varying timeframes, starting at the first observation 2000q4 and ending between 2008q4 and 2012q4. A stable long run relationship should in theory be represented by a constant long-run coefficient estimate, independent of the timeframe.

We estimate the varying timeframes over all districts in the canton of Zurich, including subregion I & II & III, using all methodologies (individual region fit, PMG, MG, DFE).

The results are summarized as sets of graphs displayed in appendix A.5.

In general, we can see that coefficient estimates vary depending on the timeframe used for all districts and estimation methods (none of the estimates are horizontal). If a fundamental relationship is captured by the specified model, we expect the estimations to remain constant over time. Judging by only the plots, we do not see any indication of a stable long run relationship, however, additional analysis quantitative analysis would be in order to draw a robust conclusion.

6 Discussion & Limitations

6.1 Limitations

A number of factors could skew the results obtained in this study.
First, the limitation due to the short time series for each panel ($T = 49$, 47 with lags included). In the introduction of the PMG estimate, [19] explicitly

states as a requirement that "T[number of observations per region] *must be large enough such that we can estimate the model for each group separately"*. While we are able to estimate the individual models, we have seen signs of overfitting for the individual models, likely due to short observations relative to the number of parameters.

In addition, this problem may be compounded by the fact that we interpolated quarterly series for three out of four factors (population, income and housing stock) where only annual observations were available, possibly introducing data which does not reflect reality as much as we would like. Finally, the choice of variables and their proxy may be imperfect; e.g. by choosing Zurich and Geneva as two examples, we captured two key housing markets, however given their similarity (socially comparatively wealthy, geographically characterized by lake proximity and with a strong focus on the service and financial industry) they may not be the best study candidates.

6.2 Discussion

Limited evidence in favor of poolability As the numeric results have shown, we do not reject pooling of long run coefficients at the cantonal level (between Zurich and Geneva). However, further analysis at the district level within the Canton of Zurich shows signs of heterogeneity and we are not able to reject pooled long run coefficients. This in turn seems to contradict the pooling assumption at the cantonal level; how can we pool different cantonal representations when there is no evidence in favor of homogeneous coefficients within the canton itself? Potentially, the similarity of Zurich and Geneva come to mind.

The first lesson one can draw is that homogeneity/poolability of long run coefficients should not be assumed as a given when modeling price dynamics.

Instead, given sufficient data availability, a disaggregated analysis should be preferred and poolability should be tested before moving with either continuing at a disaggregated level to analyze regional discrepancies, or with a higher aggregation level if evidence supports this approach.

Mixed factor coefficients at the district level Given that within the scope of this study there was no evidence in favor of pooling at the district level, we continued with estimating district specific models in the canton of Zurich. Here, the coefficients for the four fundamental factors showed mixed behavior.

The most consistent factor was interest rates, consistent with a near zero coefficient estimate, indicating that in our model, the three month CHF Libor rate plays a negligible effect. This contradicts theory, which suggests interest rates play a role both, on the demand as well as supply side by influencing mortgage availability and affordability, as well as investment cost for any housing related projects. This leaves us with the question, why does none of this seem to show up? One potential reason could be that by using a single interest rate variable to try and capture both supply and demand, we may have inadvertently suppressed the nuances of two credit markets, where mortgage rates available to buyers and investment rates available for investment efforts may differ. Potentially, a more targeted approach such as using LTV rations as in [10, 11] may have yielded more significant relationships.

No sign of stable coefficient estimates over time As the plots of section 5.4 have shown, there is no apparent evidence for a stable long run relationship. Not only do some region specific long run coefficients vary in amplitude over time, such as $\beta_{Pop,Dietikon}$, varying between a magnitude around 0.00 (2000q4-2009q4) to 10.00 (2000q4 - 2010q2), oftentimes the coefficients also change signs, such as for the district of Meilen with $\beta_{pop,Meilen} \in [-8.00, 7.50]$. Since the long run coefficient is estimated in log variables, the latter would suggest that a 1% increase in population would have caused a -8% or +7.5% change in prices respectively - a substantial difference in relationship. The question whether this is caused by a missing fundamental relationship in the give time period or due to the model not properly capturing the dynamics needs to be clarified in further detail. Additionally, the methodology for determining long run coefficient stability can be improved, given that we are only examining this relationship by evaluating coefficient evolution using plots of different timeframes. By sequentially shortening the timeframe over which we estimate the coefficients, we are additionally shortening an already short observation time period, thus exacerbating the ill effects of short observation periods.

6.2.1 Comparison of this model vs the LPPLS model

This work looks at 12 districts (when counting the Canton of Geneva) which are also covered in the most recent publication of the biannual Risk Analysis of the Real Estate Market in Switzerland [5]. Of the districts which are covered in both, [5] classifies the district of Bülach as to Watch meaning a bubble signal is being reported by the LPPLS analysis. Additionally, the districts of Pfäffikon, Hinwil, Uster, Horgen, and Dielsdorf are reported as "To Monitor" which indicates either that a post-bubble regime change is underway, or that future periods may continue building a bubble.

In regard to these districts, while our estimated model with the most data available (2000q4 - 20012q4) showed error correction terms as expected for an error correction process ($\alpha < 0$), none of these districts showed stable long run relationships over time. As discussed above, one of the reasons hereof could the lack of a fundamental relationship.

7 Conclusion

We set out to analyze the Swiss real estate price dynamics with a model driven by fundamental factors using quarterly panel data.

After evaluating existing works, we settled on modelling apartment prices as a function of population, income, housing stock and interest rates using the inverted demand approach.

The model we specify, allows for a long run stable (cointegrated) relationship between prices and the fundamentals, while allowing short run dynamics to deviate from the long run relationship. Specifically, the chosen approach (using pooled mean group estimation) allows us to test the hypothesis of homogeneous long run coefficients in different districts.

We gather panel data at the national level, at the cantonal level for Zurich and Geneva, as well as for the 12 districts within the Canton of Zurich.

The results suggest that a regionally pooled model cannot be assumed to capture the price dynamics properly, even for district within a canton, as shown by results in section 5.2.2. Further regional subdivision within the Canton of Zurich is necessary to find areas where the pooling assumption cannot be rejected. For the homogeneity hypothesis we therefore conclude that 1) in general, to obtain a feasible description for regional real estate market dynamics, it may be preferable to disaggregate the analysis to a district level, as coefficient homogeneity does not seem to hold per default at the cantonal level; 2) regional aggregation can be possible, however administrative borders such as cantons may not be a good aggregation criteria. It seems to be up to the investigator to disentangle potential regional relationships.

Following the evidence against poolability, we estimate individual regional models for each district. One conclusion we are able to draw is that interest rates (as captured by the three month CHF Libor) does not seem to contribute towards house price dynamics.

Finally, we look at the evolution of coefficients over time by estimating the model with different data timeframes, starting in 2000q4 and ending between 2008q4 and 2012q4. We find no evidence suggesting stable coefficient estimates over time, with results subject discussions in section 6.

Acknowledgements My gratitude to my supervisors Dorsa Sanadgol, Diego Ardila, and tutor Prof. Didier Sornette, for their continued input, understanding and support.

References

[1] Jesse M Abraham and Patric H Hendershott. Bubbles in Metropolitan Housing Markets. *Journal of Housing Resarch*, 7(June):191–207, 1994.

[2] Hirotugu Akaike. A New Look at the Statistical Model Identification. *Automatic Control, IEEE Transactions on*, 19(6):716–723, 1974.

[3] Andre K. Anundsen. *Regional US house price formation : One model fits all ?* 2014.

[4] Diego Ardila, Dorsa Sanadgol, and Didier Sornette. Identification and critical time forecasting of real estate bubbles in the U.S .A and Switzerland. pages 1–56, 2014.

[5] Diego Ardila, Dorsa Sanadgol, and Didier Sornette. Risk Analysis of the Real Estate Market in Switzerland (Diagnostic as of 2014-Q4). 1:1–17, 2015.

[6] A. Banerjee, J. J. Dolado, J. W. Galbraith, and D. F. Hendry. Cointegration, error correction, and the econometric analysis of non-stationary data. page 344, 1993.

[7] Bundesamt für Statistik BfS. Bau- und Wohnungswesen Panorama. Technical report, 2016.

[8] Robert R. Callis and Melissa Kresin. U.S. Census Bureau News: Residential Vacancies and Home Ownership in the Fourth Quarter 2015. 2016.

[9] James Crotty. Structural causes of the global financial crisis: A critical assessment of the 'new financial architecture'. *Cambridge Journal of Economics*, 33(4 SPEC. ISS.):563–580, 2009.

[10] John V. Duca, John Muellbauer, and Anthony Murphy. House Prices and Credit Constraints : Making Sense of the U . S . Experience. 121(March):533–551, 2011.

[11] John V Duca, John Muellbauer, and Anthony Murphy. Shifting Credit Standards and the Boom and Bust in U.S. House Prices. (0076), 2011.

[12] Robert F Engle and C W J Granger. Co-Integration and Error Correction : Representation , Estimation , and Testing Published by : The Econometric Society Stable URL : http://www.jstor.org/stable/1913236 . yet drift too far apart . Typically economic theory will propose forces which tend to. *Econometrica*, 55(2):251–276, 1987.

[13] Joshua Gallin. Real Estate The Long-Run Relationship Between House. pages 635–658, 2008.

[14] Ioannis Georgiadis. Bubble Analysis of the Swiss Real Estate Market using a hedonic index. 2015.

[15] Cm M Hurvich and Chih-Ling L Tsai. Regression and time series model selection in small samples. *Biometrika*, 76(2):297–307, 1989.

[16] Soren Johansen. Likelihood-Based Inference in Cointegrated Vector Autoregressive Models. *Econometric Theory*, 14(July 2015):517–524, 1995.

[17] Kon S. Lai. The puzzling unit root in the real interest rate and its inconsistency with intertemporal consumption behavior. *Journal of International Money and Finance*, 27(1):140–155, 2008.

[18] Vyacheslav Mikhed and Petr Zemčík. Do house prices reflect fundamentals? Aggregate and panel data evidence. *Journal of Housing Economics*, 18(2):140–149, 2009.

[19] M. Hashem Pesaran, Yongcheol Shin, and Ron P. Smith. Pooled Mean Group Estimation of Dynamic Heterogeneous Panels. (September 2015), 1999.

[20] Gideon Schwarz. Estimating The Dimension Of A Model. 1978.

[21] Elizabeth Steiner. Swiss National Bank Working Papers. 2010.

[22] Swiss National Bank. Financial stability report 2015. *Financial Stability Report*, pages 1–126, 2015.

[23] UBS. UBS Global Real Estat Bubble Index. 2015.

[24] Petar Vujanovic. 2015 OECD Economic Survey of Switzerland Policies to Tame the Housing Cycle. (December), 2015.

[25] S. S. Wilks. The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses. *The Annals of Mathematical Statistics*, 9(1):60–62, 1938.

[26] Jian Zhou. Testing for Cointegration between House Prices and Economic Fundamentals. *Real Estate Economics*, 38(4):599–632, 2010.

A Appendix

A.1 Variable definitions

Abbreviation	Variable	Source/Comment
price_hedo_apt	Hedonic price index for apartments	[14]
pop	Population level	FSO/ZSA/OCSTAT
cpi	Consumer Price Index as defined by "Landesindex der Konsumentenpreise"	FSO
incr_median	Median real income (nominal income divided by national cpi)	FSO/ZSA/OCSTAT
hstock	Housing stock (# dwellings in districts/canton; # bldgs at national level)	FSO/ZSA/OCSTAT
X3mCHFL	3 month CHF LIBOR	SNB

FSO : Swiss Federal Statistics Office
ZSA : Statistical Bureau of Canton of Zurich (Statistisches Amt)
OCSTAT : Statistical office of Canton of Geneva (Office cantonale de la statistique)
SNB : Swiss National Bank

A.2 Unit root test lag lengths selected by AIC

Table 11: ADF test number of lag terms selected by AIC

Region	price_hedo_apt	pop	incr_median	hstock	X3mCHFL
Bezirk Affoltern	4	4	4	4	1
Bezirk Bülach	3	4	4	3	1
Bezirk Dielsdorf	2	4	4	4	1
Bezirk Hinwil	3	4	4	4	1
Bezirk Horgen	2	4	4	4	1
Bezirk Meilen	1	4	4	4	1
Bezirk Pfäffikon	4	4	4	4	1
Bezirk Uster	3	4	4	3	1
Bezirk Winterthur	3	4	4	4	1
Bezirk Dietikon	3	4	4	4	1
Bezirk Zürich	4	4	4	4	1
Canton de Genève	3	3	4	3	1
Schweiz	3	3	3	4	1

Variables	
price_hedo_apt	Hedonic price index for apartments as given by [14]
pop	Population numbers
incr_median	Median real income (nominal income divided by national cpi)
hstock	Housing stock (# dwellings in districts/canton; # bldgs at national level)
X3mCHFL	3 month CHF LIBOR

A.3 Model estimation results: region specific adjustment coefficients α_i

Table 12: PMG estimation results for testing cantonal Level aggregations with P = 2, Q = 2

Estimated LR coefficients

	cpLL	β_{pop}	β_{inc}	β_{hstock}	β_{rates}	α
DFE	266.487	27.600	-8.997	-21.998	0.137	-0.057
GM	139.364	-0.455	-0.455	9.423	0.032	-0.552
PMG	274.634	0.347	-1.415	16.530	0.039	

	ur.cpLL	r.pmg.cpLL	LL.ratio	chisqdf	pval
1	276.538	274.634	3.807	4	0.433

Individual region specific adjustment coefficients

	Canton de Genève	Kanton Zürich
DFE	-0.057	-0.057
GM	-0.552	-0.359
PMG	-0.542	-0.017

cpLL	concentrated panel log Likelihood
ur.xxx	unrestricted model, with region specific $\alpha_i, \beta_i, \gamma_i, \lambda_i$
r.pmg.xxx	restricted model with pooled long run coefficient, $\beta_i = \beta \forall i$
	PMG estimates have region specific α_i, found in appendix A.3

Table 13: PMG estimation results for all districts with P = 2, Q = 2

	cpLL	β_{pop}	β_{inc}	β_{hstock}	β_{rates}	α
DFE	1125.069	5.873	-0.612	-1.329	0.009	-0.095
GM	491.215	-1.016	-1.016	1.171	0.012	-1.061
PMG	1170.099	2.632	-0.309	0.110	0.004	

ur.cpLL	r.pmg.cpLL	LL.ratio	chisqdf	pval
1250.285	1170.099	160.373	40	0.000

Individual region specific adjustment coefficients

	Bezirk Affoltern	Bezirk Bülach	Bezirk Dielsdorf	Bezirk Hinwil
DFE	-0.095	-0.095	-0.095	-0.095
GM	-1.061	-0.992	-1.016	-1.016
PMG	-0.877	-0.228	-0.317	-0.347

	Bezirk Horgen	Bezirk Meilen	Bezirk Pfäffikon	Bezirk Uster
DFE	-0.095	-0.095	-0.095	-0.095
GM	-1.016	-1.016	-1.016	-1.016
PMG	-0.024	0.022	-0.952	-0.650

	Bezirk Winterthur	Bezirk Dietikon	Bezirk Zürich
DFE	-0.095	-0.095	-0.095
GM	-1.016	-1.016	-1.016
PMG	-0.605	-1.350	-0.023

pLL	panel log Likelihood
cpLL	contentrated panel log Likelihood
ur.x	unrestricted model, with region specific $\alpha_i, \beta_i, \gamma_i, \lambda_i$
r.pmg.x	restricted model with pooled long run coefficient, $\beta_i = \beta \forall i$

Table 14: PMG estimation results for Zurich subregion I with P = 2, Q = 2

	cpLL	β_{pop}	β_{inc}	β_{hstock}	β_{rates}	α
DFE	823.002	7.294	-0.650	-1.921	0.022	-0.070
GM	343.283	-1.007	-1.007	1.291	0.013	-1.061
PMG	855.063	2.950	-0.400	-0.076	0.010	

	ur.cpLL	r.pmg.cpLL	LL.ratio	chisqdf	pval
1	922.391	855.063	134.655	28	0.000

Individual region specific adjustment coefficients

	Bezirk Affoltern	Bezirk Bülach	Bezirk Dielsdorf	Bezirk Horgen
DFE	-0.070	-0.070	-0.070	-0.070
GM	-1.061	-0.992	-1.007	-1.007
PMG	-0.776	-0.203	-0.296	-0.021

	Bezirk Meilen	Bezirk Uster	Bezirk Dietikon	Bezirk Zürich
DFE	-0.070	-0.070	-0.070	-0.070
GM	-1.007	-1.007	-1.007	-1.007
PMG	0.020	-0.653	-1.389	-0.032

pLL	panel log Likelihood
cpLL	contentrated panel log Likelihood
ur.x	unrestricted model, with region specific $\alpha_i, \beta_i, \gamma_i, \lambda_i$
r.pmg.x	restricted model with pooled long run coefficient, $\beta_i = \beta \forall i$
	PMG estimates have region specific α_i, found in appendix

Table 15: PMG estimation results for Zurich subregion II with P = 2, Q = 2

	cpLL	β_{pop}	β_{inc}	β_{hstock}	β_{rates}	α
DFE	315.321	1.201	0.555	1.181	0.005	-0.606
GM	142.334	-1.041	-1.041	0.853	0.009	-1.160
PMG	319.648	2.177	0.736	0.469	0.008	

	ur.cpLL	r.pmg.cpLL	LL.ratio	chisqdf	pval
1	327.894	319.648	16.492	8	0.036

Individual region specific adjustment coefficients

	Bezirk Hinwil	Bezirk Pfäffikon	Bezirk Winterthur
DFE	-0.606	-0.606	-0.606
GM	-1.160	-0.956	-1.041
PMG	-0.450	-0.964	-0.999

pLL	panel log Likelihood
cpLL	contentrated panel log Likelihood
ur.x	unrestricted model, with region specific $\alpha_i, \beta_i, \gamma_i, \lambda_i$
r.pmg.x	restricted model with pooled long run coefficient, $\beta_i = \beta \forall i$

Table 16: PMG estimation results for Zurich subregion III with P = 2, Q = 2

	cpLL	β_{pop}	β_{inc}	β_{hstock}	β_{rates}	α
DFE	321.129	5.179	-3.436	0.622	-0.008	-0.290
GM	252.573	-2.810	0.251	7.032	0.000	-1.015
PMG	331.420	4.131	-2.993	1.299	0.000	

	ur.cpLL	r.pmg.cpLL	LL.ratio	chisqdf	pval
1	338.198	331.420	13.555	8	0.094

Individual region specific adjustment coefficients

	Bezirk Horgen	Bezirk Meilen	Bezirk Zürich
DFE	-0.290	-0.290	-0.290
GM	-1.015	-0.259	-0.781
PMG	-0.930	-0.131	-0.805

pLL	panel log Likelihood
cpLL	contentrated panel log Likelihood
ur.x	unrestricted model, with region specific $\alpha_i, \beta_i, \gamma_i, \lambda_i$
r.pmg.x	restricted model with pooled long run coefficient, $\beta_i = \beta \forall i$

A.4 Individual region model fits

Table 17: Individual region model fit: Bezirk Affoltern

Multiple R^2: 0.687 ; Adjusted R^2: 0.55

	Estimate	[Std. Error]$\star$	[t-value]$\star$	[Pr($>|t|$)]$\star$	[Signif.]$\star$
(Intercept)	-5.910	10.866	-0.544	0.590	
α	-1.061	0.241	-4.413	0.000	***
β_{pop}	-2.636	2.948	-0.894	0.378	
β_{inc}	0.162	0.575	0.281	0.780	
β_{hstock}	3.795	2.184	1.738	0.092	.
β_{rates}	-0.010	0.010	-1.041	0.306	
lag(diff(log(price_hedo_apt)), 1)	-0.136	0.160	-0.847	0.403	
lag(diff(log(pop)), 0)	25.305	19.194	1.318	0.197	
lag(diff(log(pop)), 1)	-11.223	18.763	-0.598	0.554	
lag(diff(log(incr_median)), 0)	-1.454	1.105	-1.315	0.198	
lag(diff(log(incr_median)), 1)	-1.800	1.119	-1.608	0.118	
lag(diff(log(hstock)), 0)	14.049	8.972	1.566	0.127	
lag(diff(log(hstock)), 1)	-21.535	9.319	-2.311	0.027	*
lag(diff(X3mCHFL), 0)	-0.034	0.021	-1.640	0.111	
lag(diff(X3mCHFL), 1)	-0.037	0.019	-1.959	0.059	.

$\star$	for long run coefficients β; Std.Error, t-value, Pr(), represent estimates for $(-1) * \alpha\beta$
Signif. codes:	0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Table 18: Individual region model fit: Bezirk Bülach

Multiple R^2: 0.574 ; Adjusted R^2: 0.387

	Estimate	[Std. Error]⋆	[t-value]⋆	[Pr(> $\|t\|$)]⋆	[Signif.]⋆
(Intercept)	-55.091	14.026	-3.928	0.000	***
α	-0.992	0.229	-4.323	0.000	***
β_{pop}	-0.230	2.263	-0.102	0.920	
β_{inc}	3.723	1.274	2.922	0.006	**
β_{hstock}	3.653	2.568	1.423	0.165	
β_{rates}	0.039	0.014	2.710	0.011	*
lag(diff(log(price_hedo_apt)), 1)	0.083	0.169	0.493	0.625	
lag(diff(log(pop)), 0)	-24.620	32.270	-0.763	0.451	
lag(diff(log(pop)), 1)	0.217	31.211	0.007	0.994	
lag(diff(log(incr_median)), 0)	0.203	0.896	0.226	0.822	
lag(diff(log(incr_median)), 1)	-1.720	1.056	-1.628	0.113	
lag(diff(log(hstock)), 0)	3.048	11.564	0.264	0.794	
lag(diff(log(hstock)), 1)	1.678	9.642	0.174	0.863	
lag(diff(X3mCHFL), 0)	-0.015	0.015	-1.007	0.321	
lag(diff(X3mCHFL), 1)	-0.013	0.015	-0.820	0.418	

⋆ for long run coefficients β; Std.Error, t-value, Pr(), represent estimates for $(-1) * \alpha\beta$

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Table 19: Individual region model fit: Bezirk Dielsdorf

Multiple R^2: 0.668 ; Adjusted R^2: 0.523

	Estimate	[Std. Error]$\star$	[t-value]$\star$	[Pr(> $\|t\|$)]$\star$	[Signif.]$\star$
(Intercept)	-26.227	7.375	-3.556	0.001	**
α	-1.055	0.251	-4.210	0.000	***
β_{pop}	3.215	1.300	2.473	0.019	*
β_{inc}	0.443	0.587	0.755	0.456	
β_{hstock}	-0.894	1.136	-0.788	0.437	
β_{rates}	0.012	0.009	1.270	0.213	
lag(diff(log(price_hedo_apt)), 1)	-0.117	0.151	-0.775	0.444	
lag(diff(log(pop)), 0)	18.329	13.243	1.384	0.176	
lag(diff(log(pop)), 1)	-30.607	13.886	-2.204	0.035	*
lag(diff(log(incr_median)), 0)	-1.426	1.032	-1.382	0.177	
lag(diff(log(incr_median)), 1)	-1.424	1.011	-1.409	0.169	
lag(diff(log(hstock)), 0)	-16.565	9.052	-1.830	0.077	.
lag(diff(log(hstock)), 1)	18.774	8.134	2.308	0.028	*
lag(diff(X3mCHFL), 0)	-0.009	0.016	-0.561	0.579	
lag(diff(X3mCHFL), 1)	-0.009	0.017	-0.521	0.606	

$\star$	for long run coefficients β; Std.Error, t-value, Pr(), represent estimates for $(-1) * \alpha\beta$
Signif. codes:	0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Table 20: Individual region model fit: Bezirk Dietikon

Multiple R^2: 0.663 ; Adjusted R^2: 0.516

	Estimate	[Std. Error]⋆	[t-value]⋆	[Pr(> $\|t\|$)]⋆	[Signif.]⋆
(Intercept)	-31.155	15.650	-1.991	0.055	.
α	-1.392	0.237	-5.876	0.000	***
β_{pop}	4.183	5.175	0.808	0.425	
β_{inc}	-0.718	0.854	-0.841	0.407	
β_{hstock}	-1.482	6.283	-0.236	0.815	
β_{rates}	0.010	0.012	0.856	0.398	
lag(diff(log(price_hedo_apt)), 1)	0.285	0.159	1.785	0.084	.
lag(diff(log(pop)), 0)	52.787	34.705	1.521	0.138	
lag(diff(log(pop)), 1)	-68.795	30.836	-2.231	0.033	*
lag(diff(log(incr_median)), 0)	0.913	1.062	0.860	0.396	
lag(diff(log(incr_median)), 1)	1.005	1.172	0.857	0.398	
lag(diff(log(hstock)), 0)	-32.466	29.250	-1.110	0.275	
lag(diff(log(hstock)), 1)	40.045	25.571	1.566	0.127	
lag(diff(X3mCHFL), 0)	0.018	0.018	0.992	0.329	
lag(diff(X3mCHFL), 1)	-0.032	0.019	-1.713	0.096	.

⋆	for long run coefficients β; Std.Error, t-value, Pr(), represent estimates for $(-1) * \alpha\beta$
Signif. codes:	0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Table 21: Individual region model fit: Bezirk Hinwil

Multiple R^2: 0.645 ; Adjusted R^2: 0.489

	Estimate	[Std. Error]$\star$	[t-value]$\star$	[Pr($>|t|$)]$\star$	[Signif.]$\star$
(Intercept)	-34.745	10.284	-3.379	0.002	**
α	-1.160	0.264	-4.391	0.000	***
β_{pop}	0.003	2.300	0.001	0.999	
β_{inc}	2.293	1.237	1.854	0.073	.
β_{hstock}	1.961	1.861	1.054	0.300	
β_{rates}	0.016	0.012	1.376	0.178	
lag(diff(log(price_hedo_apt)), 1)	0.044	0.179	0.244	0.809	
lag(diff(log(pop)), 0)	9.078	20.105	0.452	0.655	
lag(diff(log(pop)), 1)	-21.221	22.567	-0.940	0.354	
lag(diff(log(incr_median)), 0)	1.666	0.864	1.929	0.063	.
lag(diff(log(incr_median)), 1)	1.096	0.934	1.174	0.249	
lag(diff(log(hstock)), 0)	7.808	7.106	1.099	0.280	
lag(diff(log(hstock)), 1)	-6.826	6.678	-1.022	0.314	
lag(diff(X3mCHFL), 0)	0.009	0.016	0.538	0.595	
lag(diff(X3mCHFL), 1)	0.016	0.016	0.983	0.333	

$\star$ for long run coefficients β; Std.Error, t-value, Pr(), represent estimates for $(-1) * \alpha\beta$

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Table 22: Individual region model fit: Bezirk Horgen

Multiple R^2: 0.682 ; Adjusted R^2: 0.544

	Estimate	[Std. Error]⋆	[t-value]⋆	[Pr(> $\|t\|$)]⋆	[Signif.]⋆
(Intercept)	-51.997	15.366	-3.384	0.002	**
α	-1.395	0.279	-5.010	0.000	***
β_{pop}	6.613	4.769	1.387	0.175	
β_{inc}	-3.629	1.868	-1.943	0.061	.
β_{hstock}	-1.145	3.744	-0.306	0.762	
β_{rates}	0.008	0.015	0.533	0.598	
lag(diff(log(price_hedo_apt)), 1)	0.106	0.168	0.630	0.533	
lag(diff(log(pop)), 0)	13.895	19.862	0.700	0.489	
lag(diff(log(pop)), 1)	2.551	23.103	0.110	0.913	
lag(diff(log(incr_median)), 0)	-1.821	1.111	-1.640	0.111	
lag(diff(log(incr_median)), 1)	0.398	1.349	0.295	0.770	
lag(diff(log(hstock)), 0)	3.312	15.547	0.213	0.833	
lag(diff(log(hstock)), 1)	-38.146	14.723	-2.591	0.014	*
lag(diff(X3mCHFL), 0)	0.002	0.021	0.073	0.942	
lag(diff(X3mCHFL), 1)	-0.007	0.020	-0.361	0.720	

⋆ for long run coefficients β; Std.Error, t-value, Pr(), represent estimates for $(-1) * \alpha\beta$

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Table 23: Individual region model fit: Bezirk Meilen

Multiple R^2: 0.607 ; Adjusted R^2: 0.436

	Estimate	[Std. Error]$\star$	[t-value]$\star$	[Pr(> $\|t\|$)]$\star$	[Signif.]$\star$
(Intercept)	-20.728	13.702	-1.513	0.140	
α	-0.345	0.162	-2.128	0.041	*
β_{pop}	-6.781	2.284	-2.969	0.006	**
β_{inc}	6.949	1.791	3.879	0.000	***
β_{hstock}	9.211	1.712	5.380	0.000	***
β_{rates}	0.025	0.016	1.545	0.132	
lag(diff(log(price_hedo_apt)), 1)	-0.436	0.162	-2.687	0.011	*
lag(diff(log(pop)), 0)	0.758	36.286	0.021	0.983	
lag(diff(log(pop)), 1)	-5.687	39.833	-0.143	0.887	
lag(diff(log(incr_median)), 0)	-0.450	1.049	-0.429	0.671	
lag(diff(log(incr_median)), 1)	-0.707	1.049	-0.674	0.505	
lag(diff(log(hstock)), 0)	-4.567	15.157	-0.301	0.765	
lag(diff(log(hstock)), 1)	0.920	14.885	0.062	0.951	
lag(diff(X3mCHFL), 0)	0.006	0.018	0.364	0.718	
lag(diff(X3mCHFL), 1)	-0.008	0.016	-0.530	0.600	

$\star$ for long run coefficients β; Std.Error, t-value, Pr(), represent estimates for $(-1) * \alpha\beta$

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Table 24: Individual region model fit: Bezirk Pfäffikon

Multiple R^2: 0.493 ; Adjusted R^2: 0.271

	Estimate	[Std. Error]⋆	[t-value]⋆	[Pr(> $\|t\|$)]⋆	[Signif.]⋆
(Intercept)	-23.405	15.046	-1.556	0.130	
α	-0.956	0.260	-3.680	0.001	***
β_{pop}	3.178	3.431	0.926	0.361	
β_{inc}	-0.452	1.071	-0.422	0.676	
β_{hstock}	-0.276	2.392	-0.116	0.909	
β_{rates}	0.000	0.011	0.004	0.997	
lag(diff(log(price_hedo_apt)), 1)	0.062	0.191	0.323	0.749	
lag(diff(log(pop)), 0)	38.460	27.879	1.380	0.177	
lag(diff(log(pop)), 1)	-22.927	28.593	-0.802	0.429	
lag(diff(log(incr_median)), 0)	-0.064	1.285	-0.050	0.960	
lag(diff(log(incr_median)), 1)	0.723	1.329	0.544	0.590	
lag(diff(log(hstock)), 0)	-14.971	14.108	-1.061	0.297	
lag(diff(log(hstock)), 1)	14.593	12.695	1.150	0.259	
lag(diff(X3mCHFL), 0)	0.005	0.021	0.243	0.810	
lag(diff(X3mCHFL), 1)	0.003	0.019	0.153	0.879	

⋆ for long run coefficients β; Std.Error, t-value, Pr(), represent estimates for $(-1) * \alpha\beta$

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Table 25: Individual region model fit: Bezirk Uster

Multiple R^2: 0.56 ; Adjusted R^2: 0.368

	Estimate	[Std. Error]$\star$	[t-value]$\star$	[Pr($>|t|$)]$\star$	[Signif.]$\star$
(Intercept)	-22.774	8.834	-2.578	0.015	*
α	-0.950	0.231	-4.108	0.000	***
β_{pop}	5.215	1.385	3.765	0.001	***
β_{inc}	-1.301	0.872	-1.492	0.145	
β_{hstock}	-2.209	1.187	-1.862	0.072	.
β_{rates}	0.026	0.009	2.806	0.008	**
lag(diff(log(price_hedo_apt)), 1)	0.073	0.165	0.441	0.662	
lag(diff(log(pop)), 0)	-5.890	13.566	-0.434	0.667	
lag(diff(log(pop)), 1)	-15.000	13.500	-1.111	0.275	
lag(diff(log(incr_median)), 0)	0.600	0.810	0.741	0.464	
lag(diff(log(incr_median)), 1)	1.714	0.666	2.573	0.015	*
lag(diff(log(hstock)), 0)	0.208	9.342	0.022	0.982	
lag(diff(log(hstock)), 1)	7.703	8.288	0.929	0.360	
lag(diff(X3mCHFL), 0)	0.022	0.011	1.993	0.055	.
lag(diff(X3mCHFL), 1)	0.005	0.012	0.401	0.691	

$\star$	for long run coefficients β; Std.Error, t-value, Pr(), represent estimates for $(-1) * \alpha\beta$
Signif. codes:	0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Table 26: Individual region model fit: Bezirk Winterthur

Multiple R^2: 0.566 ; Adjusted R^2: 0.376

	Estimate	[Std. Error]$\star$	[t-value]$\star$	[Pr($>\|t\|$)]$\star$	[Signif.]$\star$
(Intercept)	-34.718	11.049	-3.142	0.004	**
α	-1.006	0.260	-3.870	0.001	***
β_{pop}	1.918	3.105	0.618	0.541	
β_{inc}	1.102	0.950	1.160	0.255	
β_{hstock}	0.873	3.032	0.288	0.775	
β_{rates}	0.012	0.013	0.894	0.378	
lag(diff(log(price_hedo_apt)), 1)	-0.040	0.178	-0.224	0.824	
lag(diff(log(pop)), 0)	3.487	23.270	0.150	0.882	
lag(diff(log(pop)), 1)	-19.600	22.912	-0.855	0.399	
lag(diff(log(incr_median)), 0)	1.094	0.929	1.177	0.248	
lag(diff(log(incr_median)), 1)	-0.014	1.117	-0.013	0.990	
lag(diff(log(hstock)), 0)	0.458	11.688	0.039	0.969	
lag(diff(log(hstock)), 1)	1.755	10.492	0.167	0.868	
lag(diff(X3mCHFL), 0)	-0.002	0.021	-0.079	0.938	
lag(diff(X3mCHFL), 1)	-0.000	0.021	-0.023	0.982	

$\star$	for long run coefficients β; Std.Error, t-value, Pr(), represent estimates for $(-1) * \alpha\beta$
Signif. codes:	0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Table 27: Individual region model fit: Bezirk Zürich

Multiple R^2: 0.579 ; Adjusted R^2: 0.395

	Estimate	[Std. Error]⋆	[t-value]⋆	[Pr(> $\|t\|$)]⋆	[Signif.]⋆
(Intercept)	-26.207	28.050	-0.934	0.357	
α	-0.865	0.200	-4.322	0.000	***
β_{pop}	4.901	1.266	3.872	0.001	***
β_{inc}	-3.363	1.058	-3.178	0.003	**
β_{hstock}	-0.604	2.648	-0.228	0.821	
β_{rates}	-0.004	0.012	-0.319	0.752	
lag(diff(log(price_hedo_apt)), 1)	-0.099	0.159	-0.626	0.536	
lag(diff(log(pop)), 0)	24.400	8.641	2.824	0.008	**
lag(diff(log(pop)), 1)	-25.711	9.003	-2.856	0.007	**
lag(diff(log(incr_median)), 0)	0.230	1.257	0.183	0.856	
lag(diff(log(incr_median)), 1)	1.650	1.147	1.438	0.160	
lag(diff(log(hstock)), 0)	22.676	24.160	0.939	0.355	
lag(diff(log(hstock)), 1)	-1.820	26.523	-0.069	0.946	
lag(diff(X3mCHFL), 0)	0.006	0.023	0.263	0.794	
lag(diff(X3mCHFL), 1)	0.012	0.019	0.621	0.539	

⋆ for long run coefficients β; Std.Error, t-value, Pr(), represent estimates for $(-1) * \alpha\beta$

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Table 28: Individual region model fit: Canton de Genève

Multiple R^2: 0.676 ; Adjusted R^2: 0.534

	Estimate	[Std. Error]⋆	[t-value]⋆	[Pr(> $\|t\|$)]⋆	[Signif.]⋆
(Intercept)	-106.686	48.168	-2.215	0.034	*
α	-0.552	0.227	-2.431	0.021	*
β_{pop}	0.272	2.116	0.129	0.899	
β_{inc}	-1.367	0.631	-2.167	0.038	*
β_{hstock}	16.579	5.577	2.973	0.006	**
β_{rates}	0.038	0.009	4.051	0.000	***
lag(diff(log(price_hedo_apt)), 1)	-0.417	0.160	-2.599	0.014	*
lag(diff(log(pop)), 0)	1.743	14.595	0.119	0.906	
lag(diff(log(pop)), 1)	2.349	14.315	0.164	0.871	
lag(diff(log(incr_median)), 0)	-1.304	0.861	-1.515	0.140	
lag(diff(log(incr_median)), 1)	-1.194	0.822	-1.453	0.156	
lag(diff(log(hstock)), 0)	53.961	29.257	1.844	0.074	.
lag(diff(log(hstock)), 1)	-80.843	33.285	-2.429	0.021	*
lag(diff(X3mCHFL), 0)	0.019	0.014	1.352	0.186	
lag(diff(X3mCHFL), 1)	-0.002	0.015	-0.131	0.897	

⋆ for long run coefficients β; Std.Error, t-value, Pr(), represent estimates for $(-1) * \alpha\beta$

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Table 29: Individual region model fit: Kanton Zürich

Multiple R^2: 0.501 ; Adjusted R^2: 0.283

	Estimate	[Std. Error]$\star$	[t-value]$\star$	[Pr($>|t|$)]$\star$	[Signif.]$\star$
(Intercept)	-19.987	14.226	-1.405	0.170	
α	-0.359	0.259	-1.384	0.176	
β_{pop}	1.868	0.985	1.897	0.067	.
β_{inc}	0.612	0.498	1.228	0.228	
β_{hstock}	2.267	1.088	2.084	0.045	*
β_{rates}	0.026	0.009	2.842	0.008	**
lag(diff(log(price_hedo_apt)), 1)	-0.371	0.206	-1.800	0.081	.
lag(diff(log(pop)), 0)	4.240	8.765	0.484	0.632	
lag(diff(log(pop)), 1)	-12.028	9.674	-1.243	0.223	
lag(diff(log(incr_median)), 0)	-0.339	0.524	-0.646	0.523	
lag(diff(log(incr_median)), 1)	-0.353	0.475	-0.744	0.463	
lag(diff(log(hstock)), 0)	-2.677	6.224	-0.430	0.670	
lag(diff(log(hstock)), 1)	5.422	5.251	1.032	0.310	
lag(diff(X3mCHFL), 0)	-0.004	0.009	-0.488	0.629	
lag(diff(X3mCHFL), 1)	-0.007	0.008	-0.919	0.365	

$\star$ for long run coefficients β; Std.Error, t-value, Pr(), represent estimates for $(-1) * \alpha\beta$

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Table 30: Individual region model fit: Schweiz

Multiple R^2: 0.659 ; Adjusted R^2: 0.51

	Estimate	[Std. Error]$\star$	[t-value]$\star$	[Pr($>|t|$)]$\star$	[Signif.]$\star$
(Intercept)	-62.920	15.670	-4.015	0.000	***
α	-0.893	0.204	-4.370	0.000	***
β_{pop}	4.552	1.367	3.330	0.002	**
β_{inc}	0.268	0.106	2.518	0.017	*
β_{hstock}	0.088	0.620	0.142	0.888	
β_{rates}	0.027	0.006	4.527	0.000	***
lag(diff(log(price_hedo_apt)), 1)	-0.047	0.150	-0.312	0.757	
lag(diff(log(pop)), 0)	-7.241	6.144	-1.179	0.247	
lag(diff(log(pop)), 1)	3.918	7.220	0.543	0.591	
lag(diff(log(incr_median)), 0)	0.164	0.132	1.243	0.223	
lag(diff(log(incr_median)), 1)	0.105	0.144	0.728	0.472	
lag(diff(log(hstock)), 0)	39.792	36.106	1.102	0.279	
lag(diff(log(hstock)), 1)	-17.056	33.268	-0.513	0.612	
lag(diff(X3mCHFL), 0)	0.011	0.004	2.522	0.017	*
lag(diff(X3mCHFL), 1)	-0.004	0.005	-0.823	0.417	

$\star$ for long run coefficients β; Std.Error, t-value, Pr(), represent estimates for $(-1) * \alpha\beta$

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

A.5 Long Run coefficient estimations over time

A.5.1 Canton of Zurich all districts

Figure 10: Long Run coefficient estimations over time
Canton of Zurich all districts

Timeframe: 2000q4 - varying end time

Time representation:
$2000q1 = 2000.\mathit{25}, \ldots , 2000q4 = 200\mathit{1.00}$

A.5.2 Canton of Zurich Subregion I

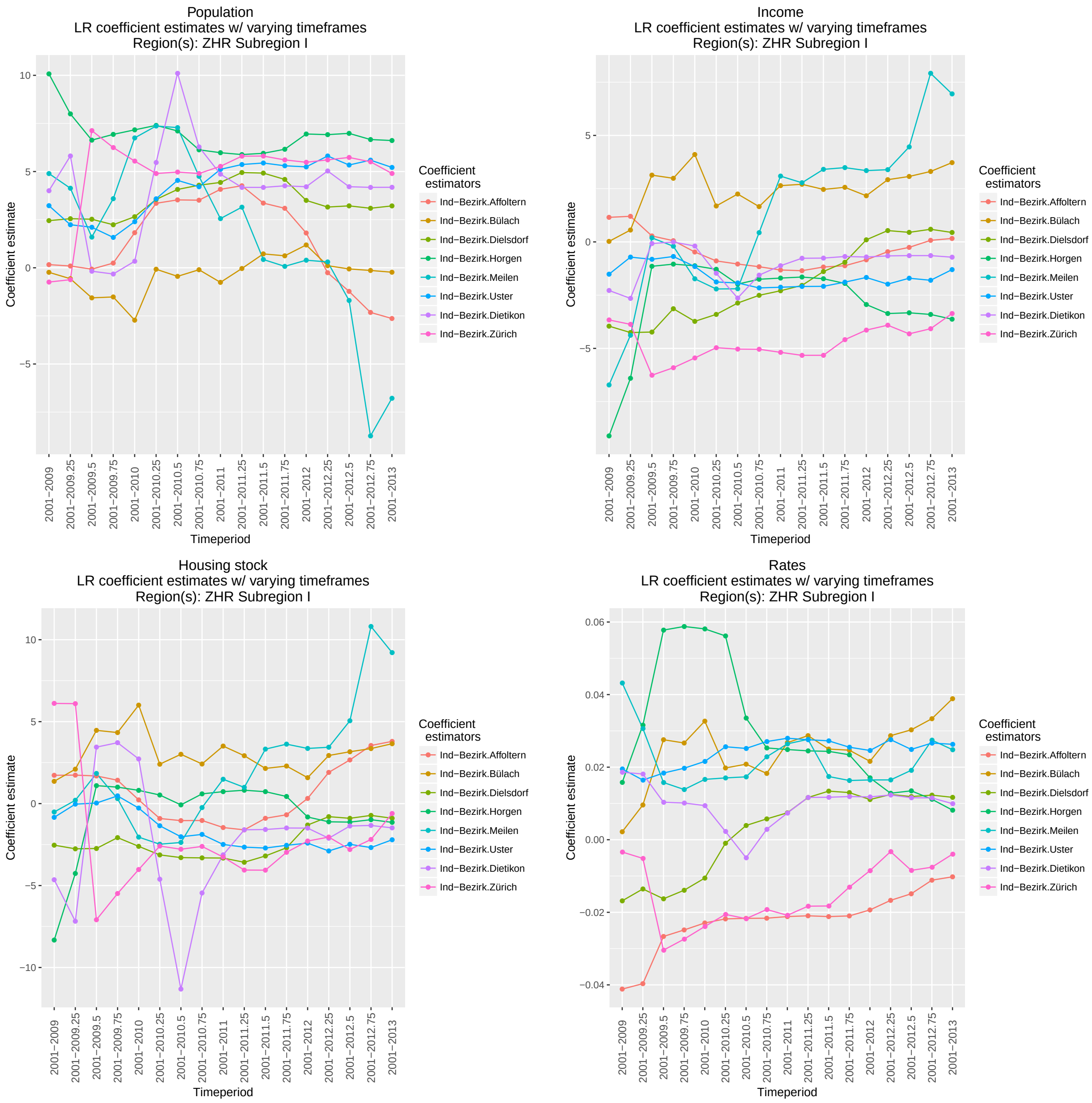
Figure 11: Long Run coefficient estimations over time
Canton of Zurich subregion I

Timeframe: 2000q4 - varying end time

Time representation:
2000q1 = 2000.*25*, ... , 2000q4 = 200*1.00*

Figure 12: Long Run coefficient estimations over time
Canton of Zurich subregion I (region specific est)

Timeframe: 2000q4 - varying end time



Time representation:
2000q1 = 2000.*25*, ... , 2000q4 = 200*1.00*

A.5.3 Canton of Zurich Subregion II

Figure 13: Long Run coefficient estimations over time
Canton of Zurich subregion II

Timeframe: 2000q4 - varying end time

Time representation:
2000q1 = 2000.*25*, ... , 2000q4 = 200*1.00*

Figure 14: Long Run coefficient estimations over time
Canton of Zurich subregion II (region specific est)

Timeframe: 2000q4 - varying end time

Time representation:
2000q1 = 2000.*25*, ... , 2000q4 = 200*1.00*