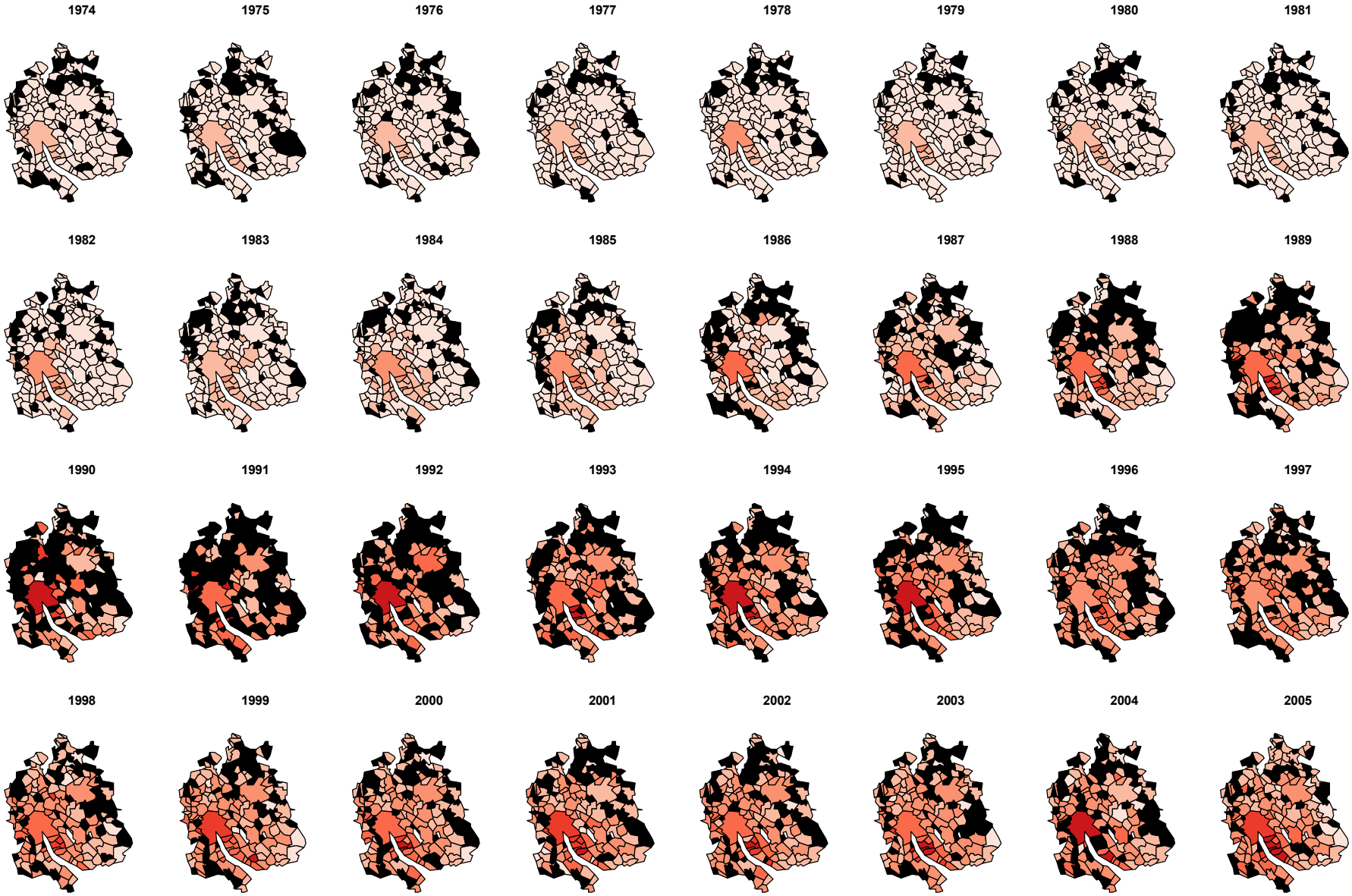
Modelling the Spatio-temporal Development of Building Land Prices in the Canton of Zurich: A Mixed-Effects Approach

Starting Point

- For a variety of reasons (scientific, personal, legal, fiscal etc.) lots of people are interested in actual but also historical „mean prices" paid for residential building land plots in the municipalities of the Canton of Zurich.
- There is lots of data available: Since 1974 there exists in the Canton of Zurich a more or less complete data base of all real estate transactions notified with the authorities, among which there are no less than
- ~ 58 000 Transactions concerning undeveloped land plots for residential building (unbebautes Wohnbauland)
- There are, however a few problems….

Missing values and unmotivated price jumps...



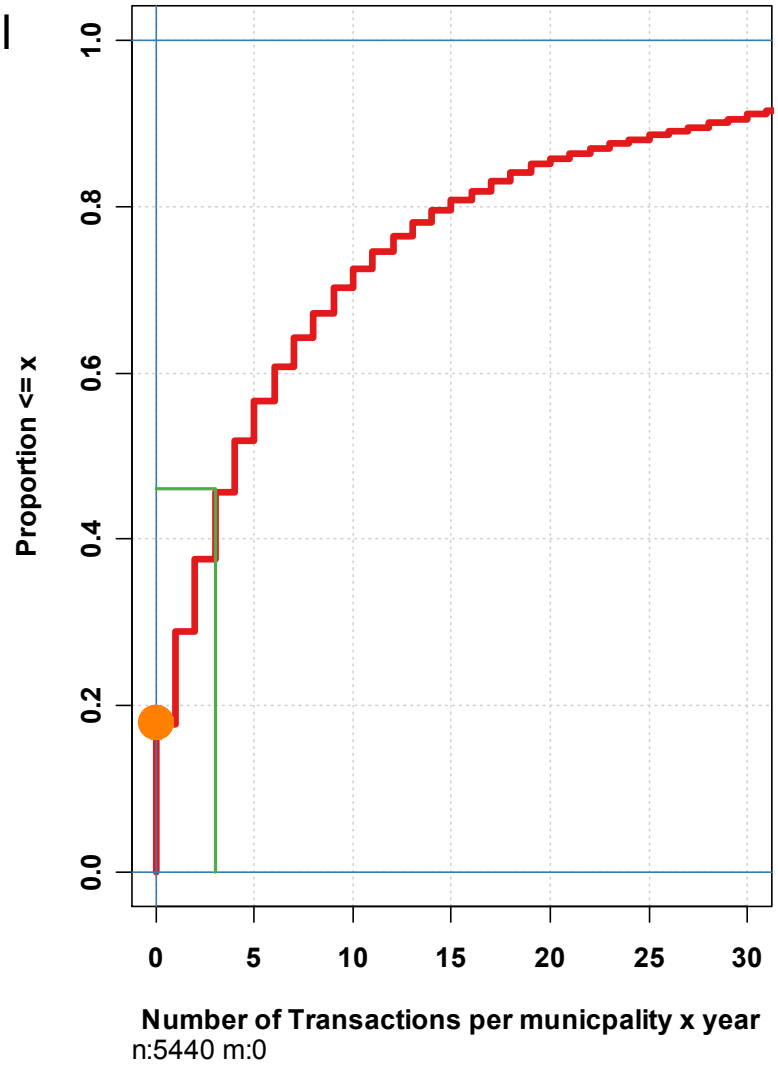
....implied in the spatio-temporal sparsity of the data

- For no less than **18%** of the 5472 potential combinations of (171) municipalities X years (32; 1974–2005) no transactions took place at all.
- For a further **20%** of these grid points mean prices can‘t be published for reasons of data protection (number of transactions <3)

➔ **For a total of 38% of the grid-points no mean price can be published**.

- And the n of Transactions is small for much of the other grid-points, so means are of limited value (Small-area Problem!)

➔ **High temporal variability of communal prices from year to year (dependent on the actual mix of micro conditions, perhaps also on particular supply-demand constellations)**

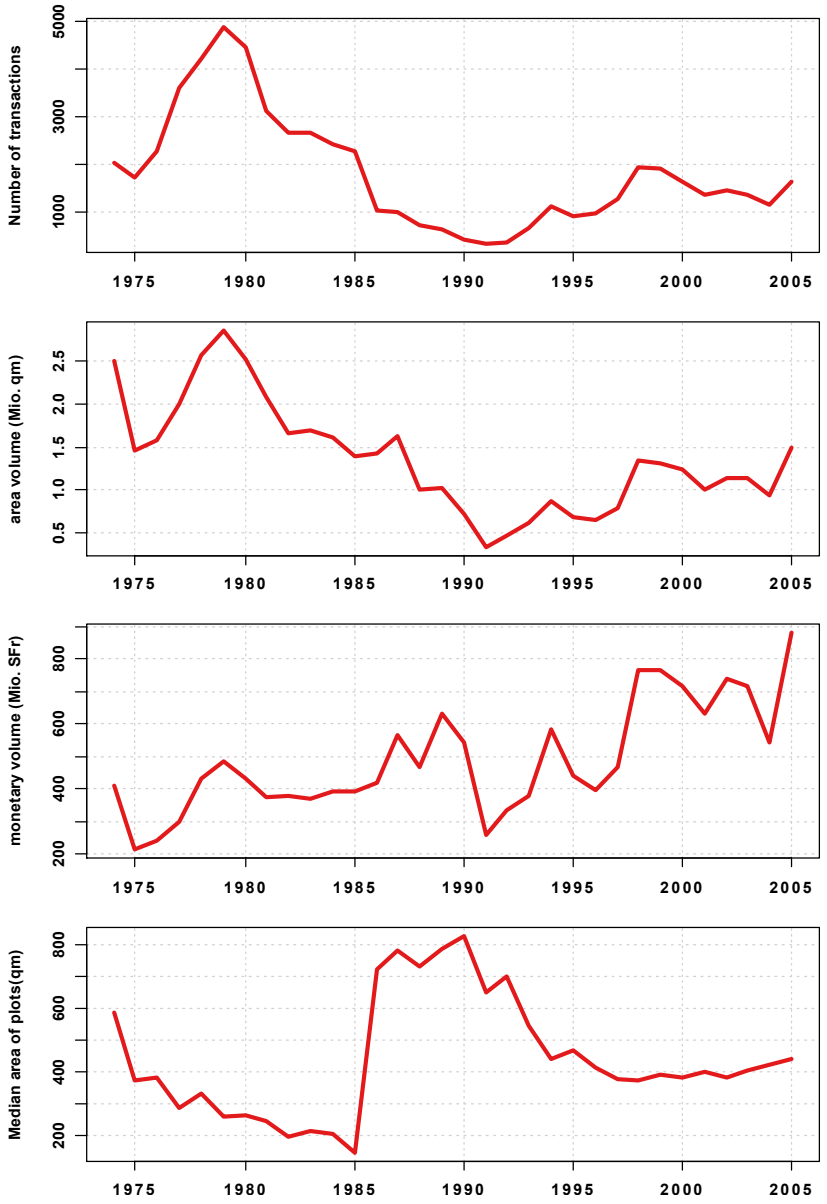


What do we want?

- Estimation of mean prices for all missing year x municipality gridpoints in a methodically transparent and consistent framework
- Smoothing of the temporal development of mean prices; we want to extract the „signal“ and eliminate the „noise“ created by particular mixes of micro conditions of traded land plots, especially if the number of transactions is small.
- Optimal use of all available information about transactions.
- At an acceptable Cost: No back-collection of historical data about the micro-locations of traded land plots – eg. exposition ,distance to infrastructure like schools, train-stations etc. –: This is out of the question.
- Which means: we have to do with covariates relating to the municipal level or above: a detailed **hedonic model** of building land-prices, including data on micro-conditions is not feasible - at least for the more distant past.

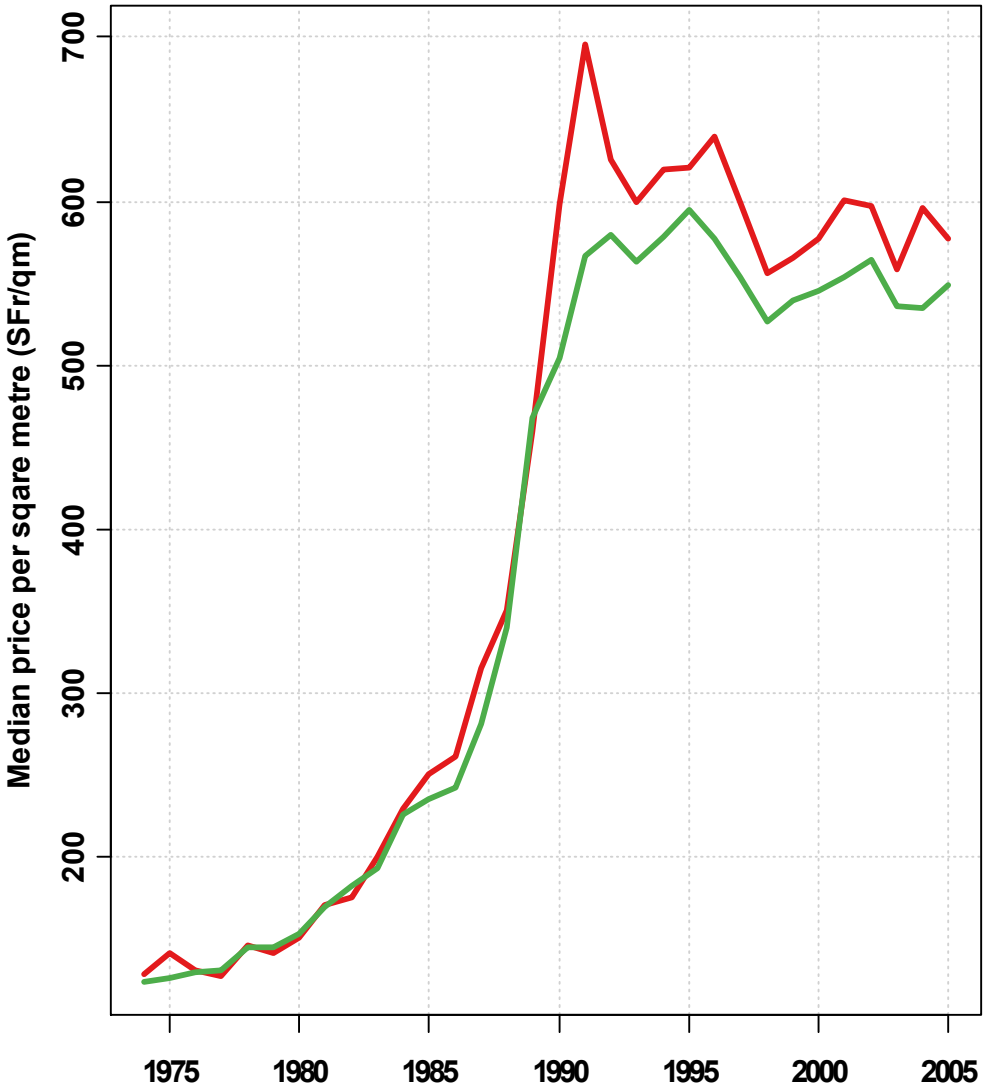
Some indicators of the market in land plots 1974-2005 I

- Number of transactions ...
- ... volume traded (in qm)....
- .... Transaction volume (in SFr).....
- ... Median surface area of land plots traded



Price development 1974-2005 II

- ...the cantonal median price per sqare metre of land (unweighted)
- and the geometric mean, which is a quite robust measure for the central tendency in the case of left-skewed data, which can be „normalized“ by taking the log.
- Caveat: We don‘t want to explain this general development of prices. It is endogenous to our model. We are only interested in the divergent developments in the municipalites of the canton!



Land Prices develop differently in the municipalities of the Canton

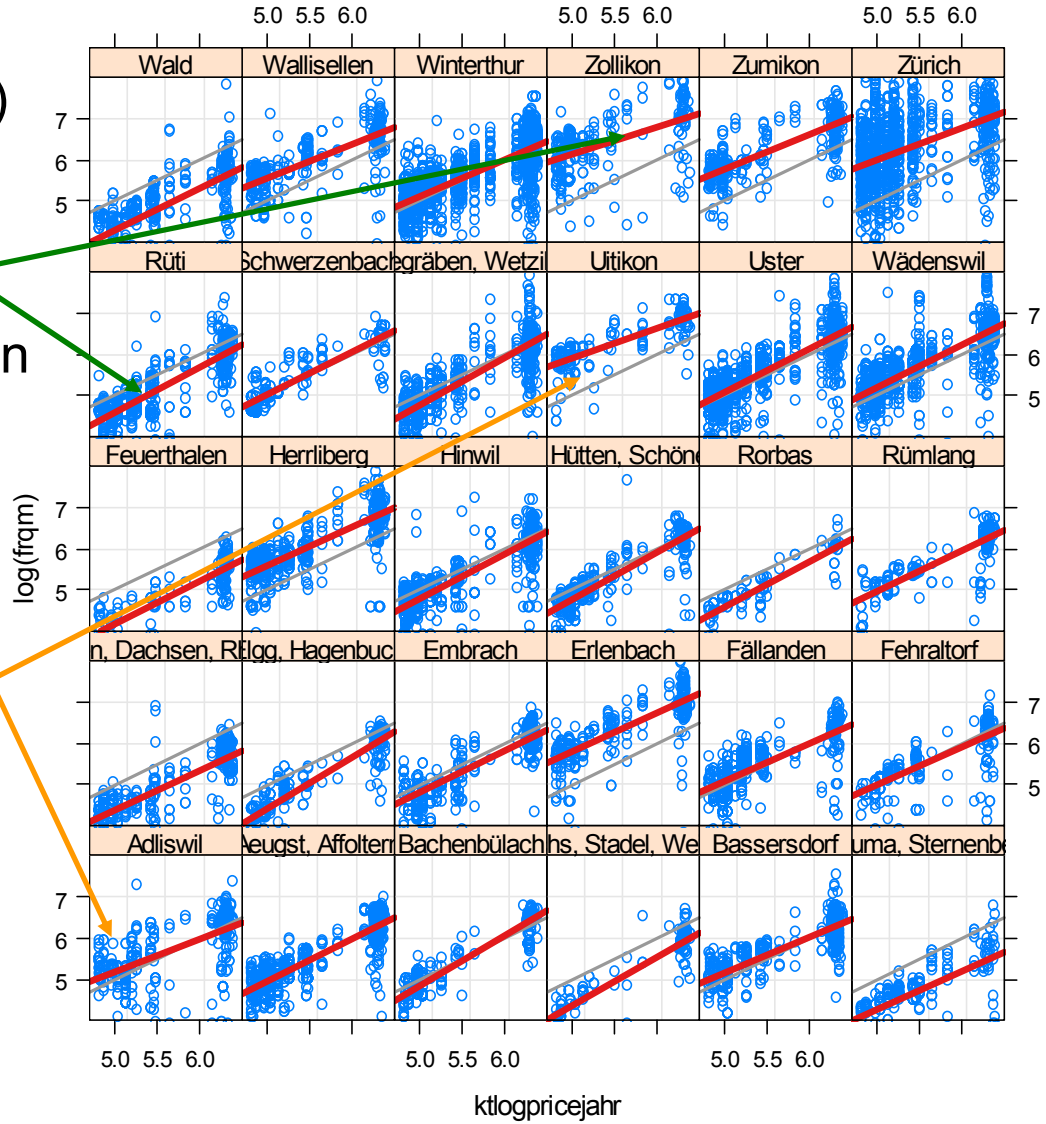
- $\log(frqm)=\beta_0+\beta_1\log(ktpricejahr)$
- β_0 (Intercepts): differ between the municipalities

➔ general price levels as compared to the cantonal mean differ (High on the „Gold-Coast“, low in periphereal regions)

- β_1 (Slopes) differ too:

➔ municipal prices „react“ differently to the general development, elasticities vary (log~log=Δ%~Δ%)

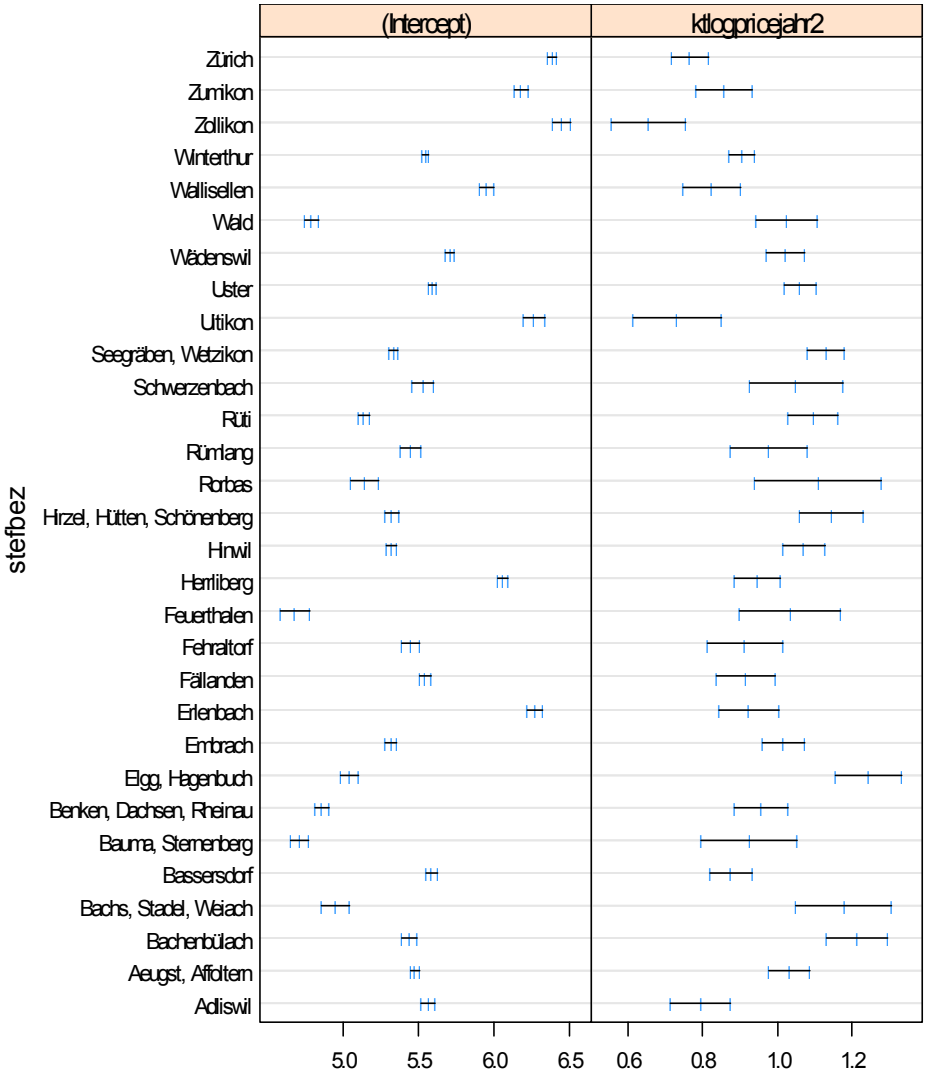
- ~ the relationship is more or less linear



Fixed effects Modelling municipality-wise

- A somewhat more schematic view of the group-wise Parameters and their confidence intervals
- May be a solution: a collection of fixed effects models for all municipalites and variables!
- Or we could wrap this up into a single model with complex interactions

$$\log(frqm) = \beta_0 + \beta_1(frqmzh)$$



The problem with fixed effects – and a solution

- A fixed-effects-model with interactions becomes complex very rapidly especially if the basic model incorporates many variables (e. g. municipal tax rates etc.) parameters tend to proliferate.
- Municipality-specific parameter estimates are often based on a very small number of cases (transactions) and are therefore unreliable: this is especially true for the Slopes, the β_1

➔ Solution: **Mixed-effects model**

- Basic principle: we estimate overall fixed-effect-parameters for the independent variables. They represent something like mean effects of a variable.
- The variability of the parameters between the municipalities is modelled by a distribution – often a normal distribution with the corresponding fixed effect as a mean and a standard deviation dependent on the variability of the municipality-specific estimates of the parameter in question.

A simple mixed-effects model

- **Fixed effects**:
- $\log(frqm)=\beta_0+\beta_1\log(frqmz_h)+\beta_2 steuerfuss$

➔ log transaction prices per square metre depend on a cantonal price index (exogenous!) and the municipal tax rate.

- **Random effects**:
- The intercept as well as the slopes of the explanatory variables *log(ktpricejahr)* and *steuerfuss (tax rate)* are allowed to vary between municipalites; we assume the parameter distribution to be normal.

➔ Transaction prices in the municipalites are allowed to have different levels and may also differ in their elasticity towards the general trend in land prices and changes in the municipal tax rates.

- A model specified this way explains about 71% of the variability of log(*frqm*) (pseudo-R^2; correlation^2 between actual transaction prices and model values). The lion‘s share of the explanatory power of the model is, unsurprisingly, due to the explanatory power of the cantonal building land price index over time.

Implications and advantages of this modelling strategy

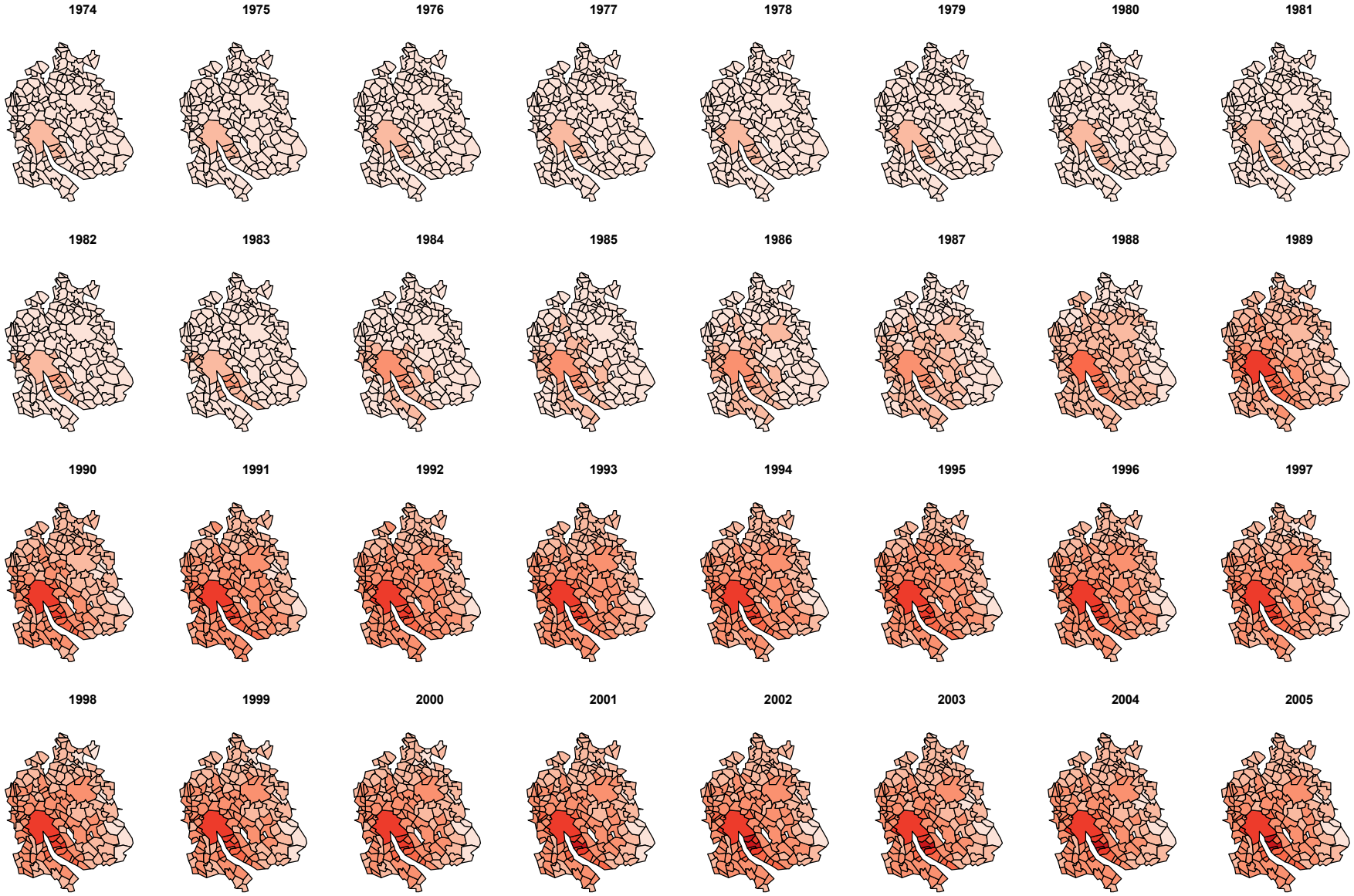
- The normal distribution of the random effects parameters implies that municipal parameter values far from the (fixed effect) mean (in the tails of the distribution), are rather improbable.
- The random-effects parameters of the explanatory variables (and the intercept) of a particular municipality are (ceteris paribus) relatively close to the mean:

➔ If price information (the number of transactions) is sparse.

➔ If the variability of this within-municipality price information is high.

- In difference to fixed-effects model with interactions, the estimates of the municipal parameters of the explanatory variables depend not only on the information from within this particular spatial unit.
- Municipality-specific parameter values are more or less strongly influenced by the „general situation“, if information within the unit is sparse or ambigous. This principle of "borrowing strength" stabilises estimates and implies a smoothing of the estimates for mean values.

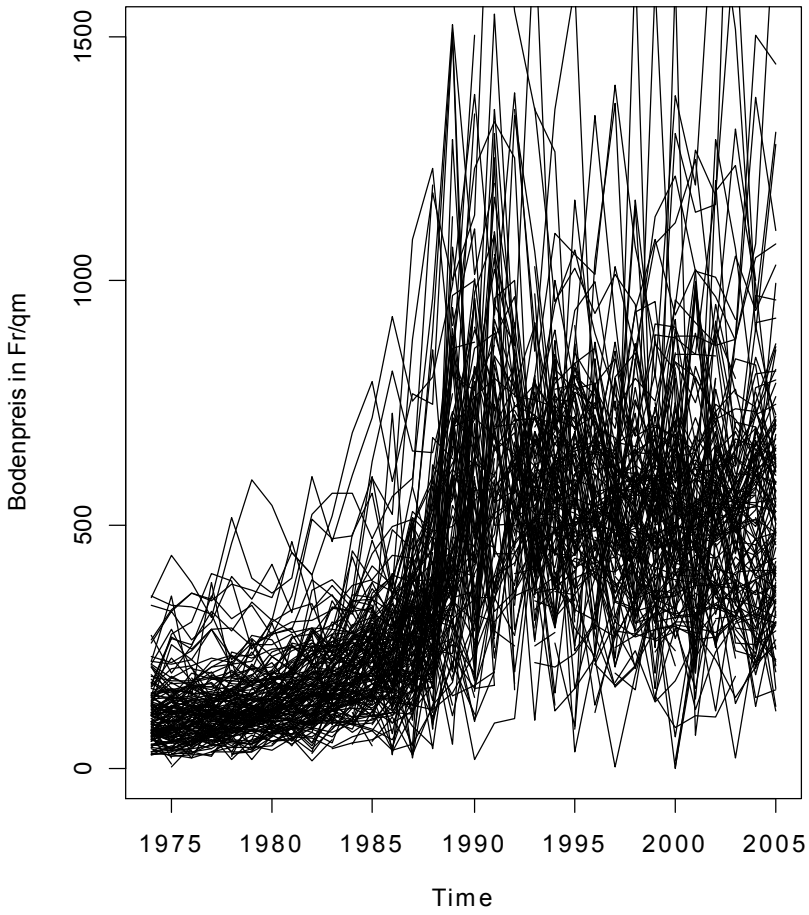
The smoothing effect of the model I



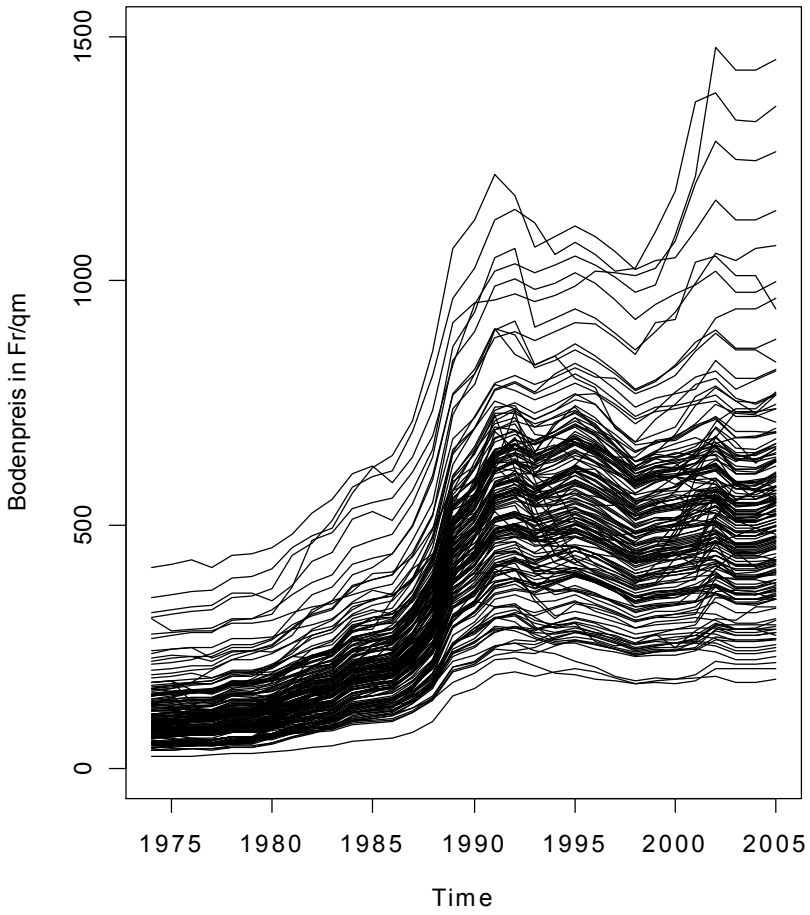
Smoothing of the communal mean price trajectories

•A different view of the temporal aspect:

"raw" mean prices



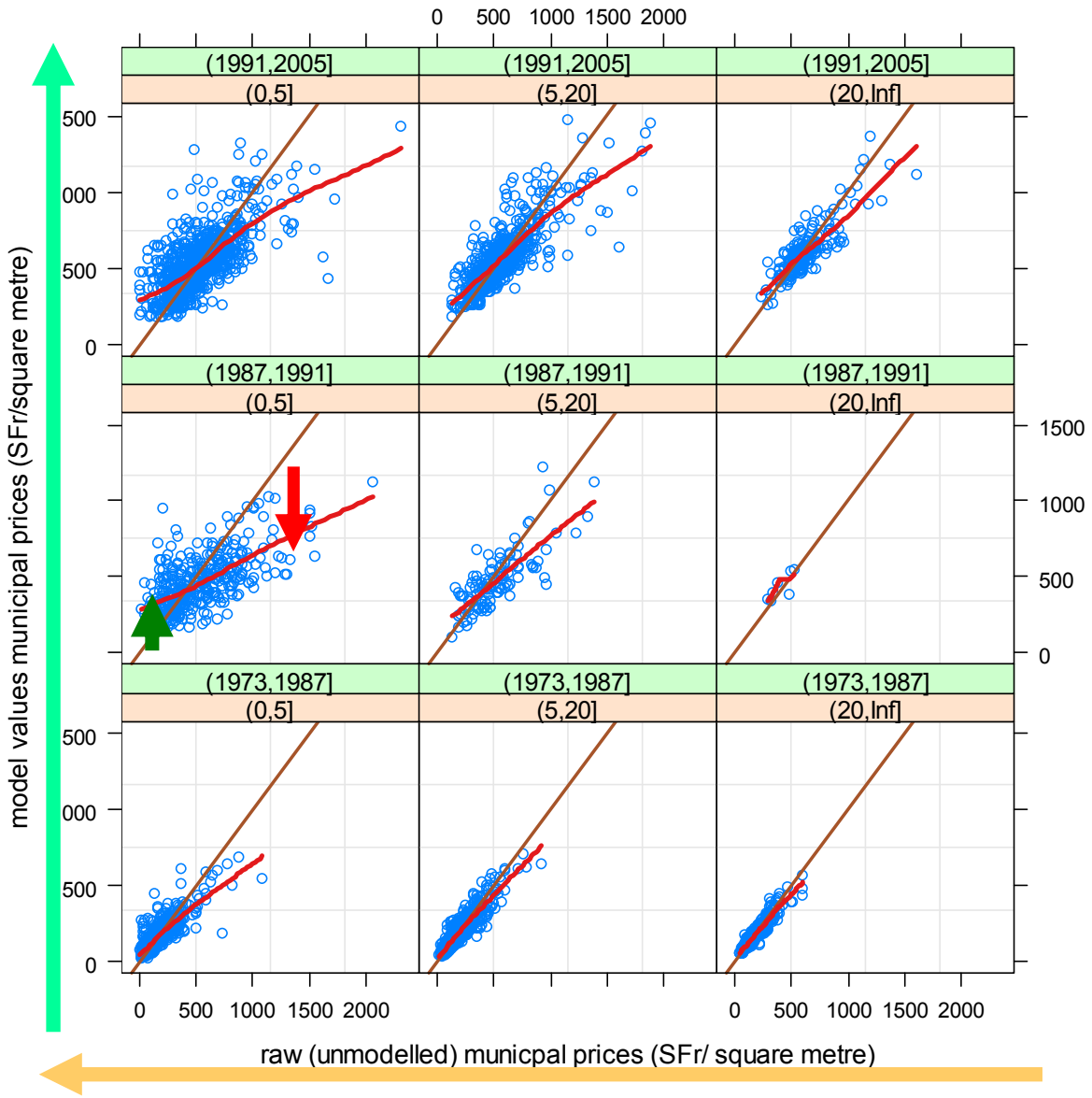
modelled mean prices



The logic of the smoothing

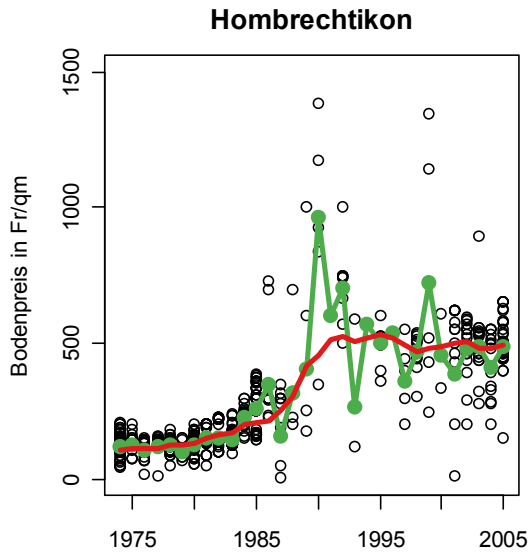
- A small number of trans-actions per year and municipality and a large amount of variabilty:
- Implies a shifting of model estimates of the mean land price in a munici-pality towards the can-tonal mean (the fixed effect): In these cases:
- If the raw mean is „too high“ model estimates are lower.
- And vice versa

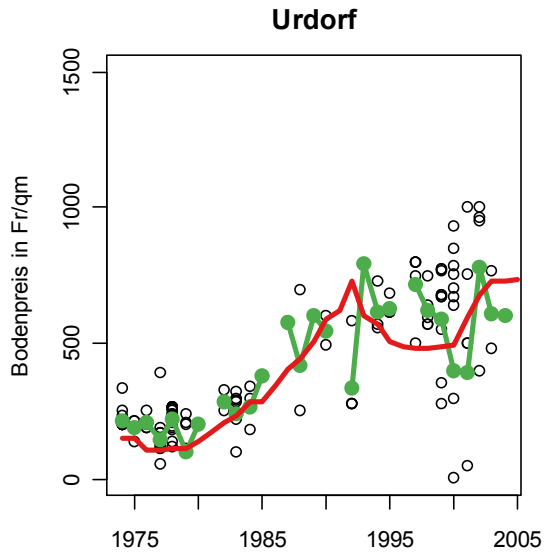
➔ That‘s what "borrowing strength„ means

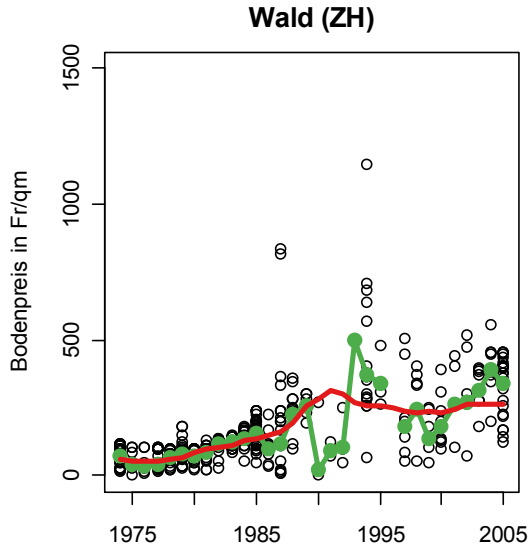


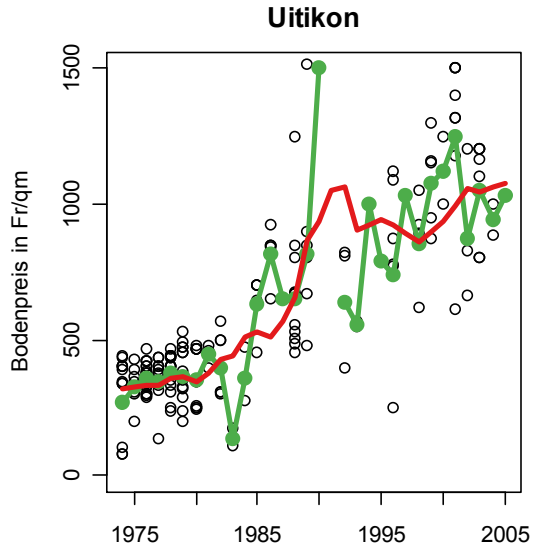
Some examples ....

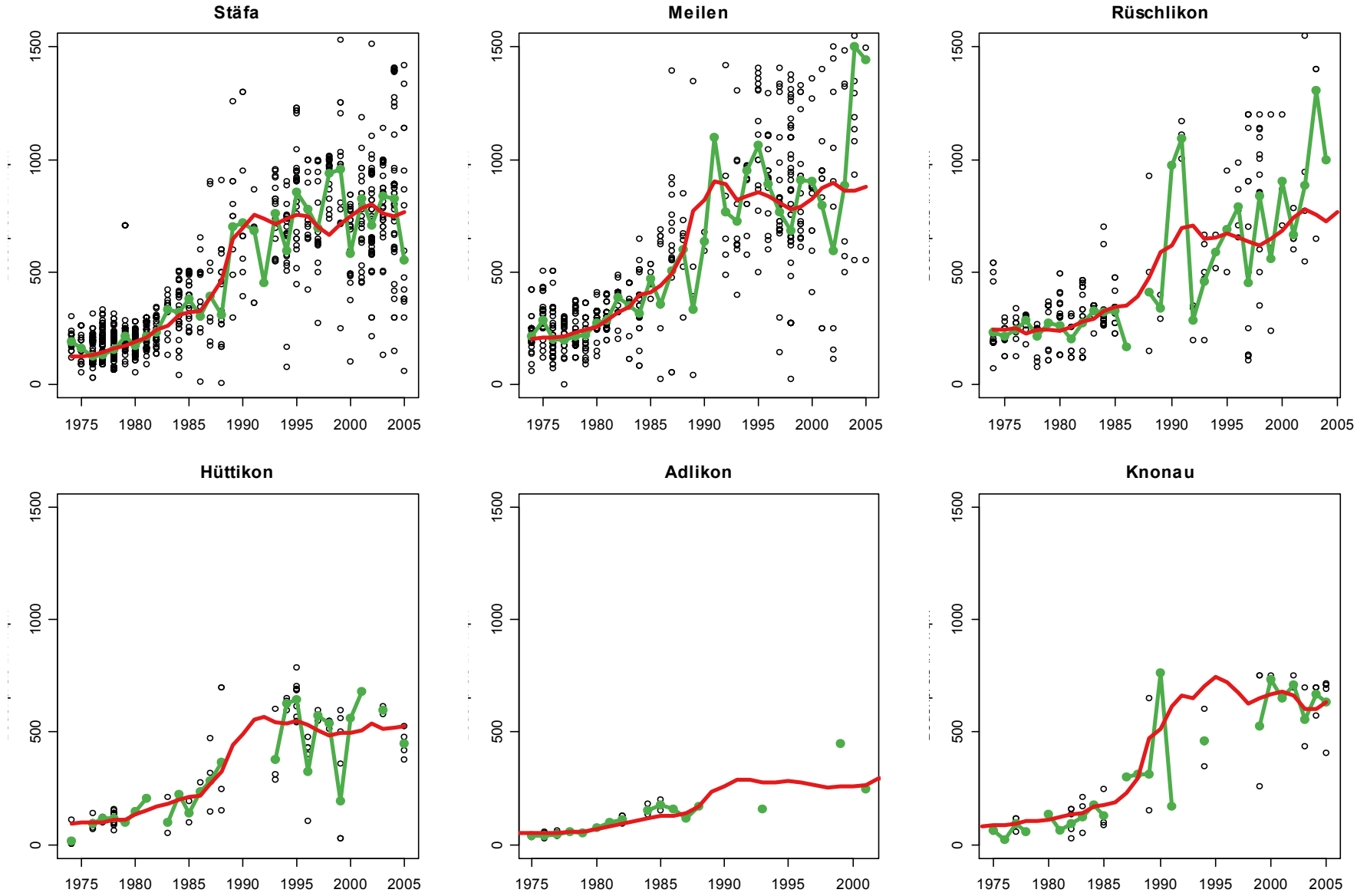
- Model estimates
- (Geometric) mean values
- Actual transactions



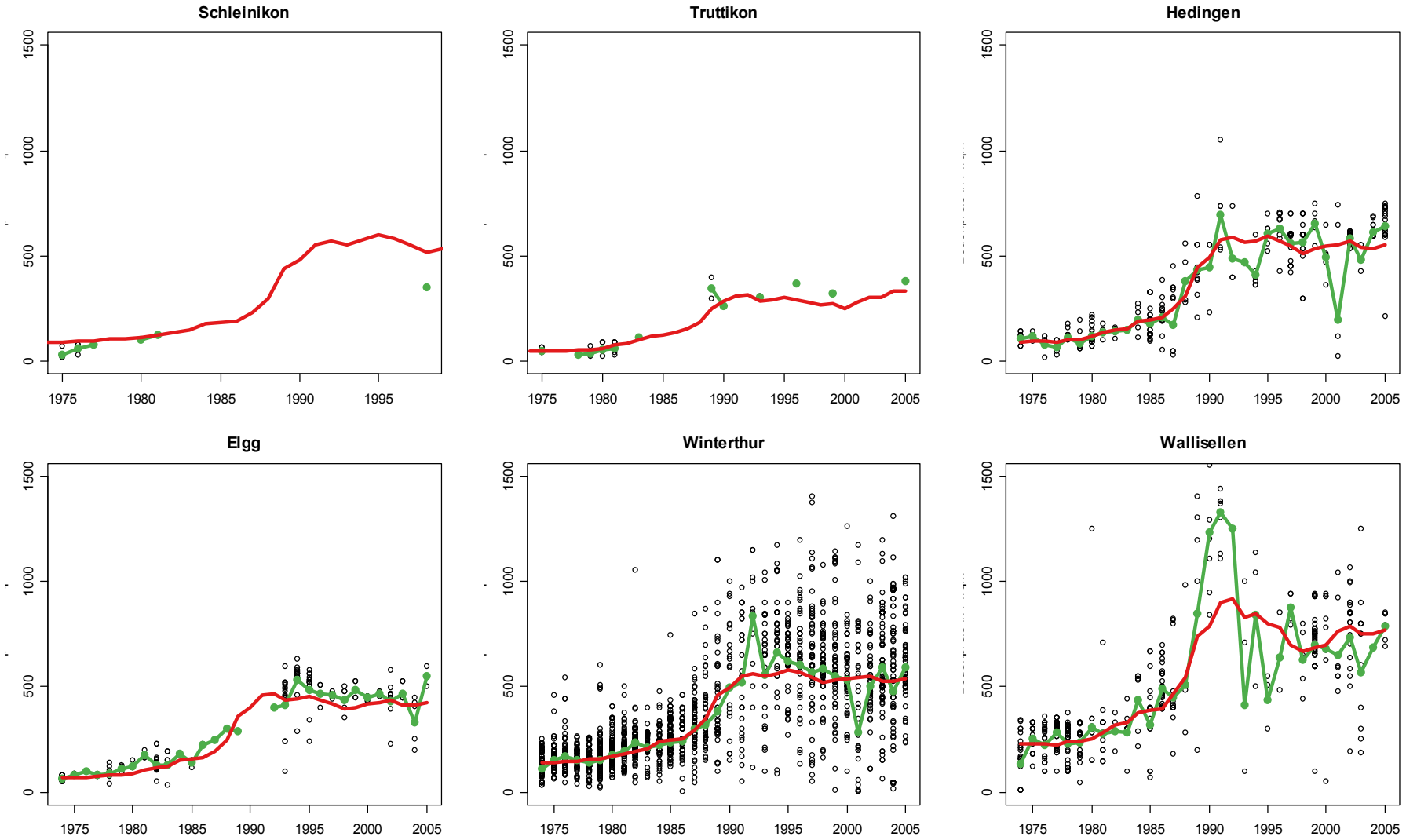







Stäfa
Meilen
Rüschlikon
Hüttikon
Adlikon
Knonau
0
500
1000
1500
1975
1980
1985
1990
1995
2000
2005

And some more ...



Summary - Perspectives

- Mixed-Effects modelling als a flexible but nevertheless parametric strategy for this problem, which is at its core a small area estimation problem
- Estimates are based on a single model, formally implementing some kind of "common sense" – which is what statistics is - or should be - all about in the first place.
- The model actually has an explanatory value, and could be used for analytical purposes.
- The future:

➔ mixed-effects quantile regression: a L1-norm-strategy would allow us to model medians and quantiles directly, without having to transform the dependent variable – which is always difficult to explain to lay people.

➔ Endogenization of the general trend in land prices would allow for a much more powerful model.

Thanks for your attention

For further information, do not hesitate to contact me:

Dr. Peter Moser
Forschungsleiter
Statistisches Amt des Kantons Zürich
Bleicherweg 5
8090 Zürich

peter.moser@statistik.ji.zh.ch