Which Factors Capitalize into House Prices? A Bayesian Averaging Approach*

David Stadelmann†

February 25, 2009

Abstract

This paper investigates the robustness of 31 community specific explanatory variables for house prices in the Swiss metropolitan area of Zurich using Bayesian Model Averaging. The main variables which capitalize with a high posterior probability are location specific real estate characteristics, municipal taxes and expenditures for culture, health and social services. Demographic as well as other socioeconomic controls seem to be of minor importance. The analysis suggests a minimal list of variables that may be included in any estimation for capitalization of community specific characteristics in the context of a metropolitan area in a highly developed country.

Key words: Capitalization, House Price Hedonic, Taxes, Model Uncertainity, Bayesian Model Averaging.

JEL classification: R21, C11, H40.

***Acknowledgements:* I am grateful to Philippe Deschamps and Volker Grossmann for detailed suggestions on an earlier draft. I also benefited from inspiring discussions with Reiner Eichenberger and motivating comments by David M. Brasington and Oded Palmon. Finally, I also thank the seminar participants at the Department of Mathematics and the Department of Economics at the University of Fribourg, especially Olivier Dürr, Michael Funk, Manuela La Fauci, and Mark Schelker. The Zürcher Kantonalbank (Patrick Schellenbauer and Marco Salvi) kindly provided a house price dataset. Any remaining errors are entirely my own.

†University of Fribourg, Departement of Economics, Bureau F410, Bd. de Pérolles 90, 1700 Fribourg (Switzerland), +41 (026) 300 93 82, david.stadelmann@unifr.ch

1 Introduction

In an urban area with diverse and competing local governments, the price of housing depends on tax-expenditure packages, local school, demographic, socioeconomic and other locaction specific real estate characteristics of a community. In theory we expect that communities with attractive characteristics, for example, short distance to the center, access to public transport facilities, low taxes, high quality public services, neighborhood schools etc. will have higher property values than communities without these attributes.

Following the pioneering work of Oates (1969), a number of studies concerning capitalization of the most diverse variables have been performed in the last three decades.[1] Altogether many variables that influence property values have been identified. Indeed, as Fischel (2001) mentions, "Everything seems to be capitalized".

The basic methodology of most of these studies consists of running cross section regressions including the main variable of interest and a number of other controls. The question arises, however, which independent control variables to include in the regressions. If, for example, a particular toxic waste site's influence on house values is analyzed (see Ketkar 1992), this variable of interest is obviously included in the regression. But which other controls should be integrated in the estimations? Theory often produces a rather long "laundry list" of possible controls. Various theoretical arguments can be put forward for different variables. At the same time it is easy to find arguments against an inclusion in a regression. This tends to result in some degrees of freedom concerning the control variables, potentially introducing some considerable bias.

This paper employs model averaging in order to systematically analyze which community specific variables are of high importance when it comes to house prices. More precisely, we examine the robustness of independent variables in regressions for capital-

[1] Here is a selective list: Oates (1973), Pollakowski (1973), Edel and Sclar (1974), Reinhard (1981), Yinger, Bloom, Börsch-Supan, and Ladd (1988), Palmon and Smith (1998) consider the capitalization of taxes and/or public goods; Ketkar (1992) and Kiel (1995) analyze the influence of hazardous waste sites on property values; Hughes Jr. and Sirmans (1992) look at the effects of traffic on house prices; Portney (1981) shows that local air pollution reduces home values; Rohe and Stewart (1996) concentrate on the effects of homeownership and neighborhood stability; Jud and Watts (1981), Brasington (1999), Figlio and Lucas (2004), Reback (2005) analyze the effects of school specific characteristics on house prices.

ization of location specific real estate characteristics, local taxes, fiscal variables, school specific controls, median incomes, demographic and other socioeconomic characteristics on the price of a comparable property over a set of communities. Recent articles on house price hedonic such as Bajari and Kahn (2005) or Ekeland, Heckman, and Nesheim (2004) emphasize the issue of functional form instead of variable selection. Here, we focus on the latter as to the best of our knowledge no previous model of housing demand has systematically treated this problem.

We introduce Bayesian Model Averaging (BMA) to address the problem of model uncertainty inherent in the selection of control variables. BMA constructs estimates by "averaging" OLS coefficients across different models.[2] The weights to individual regressions stem from the Bayesian Information Criterion (see Raftery, Madigan, and Hoeting 1997). BMA is a way of taking account of uncertainty concerning the model's variables as it does an exhaustive search over the whole model space. The Bayesian approach is feasible and has been applied to various problems in economics by other authors: Fernandez, Ley, and Steel (2001b), Sala-i-Martin, Doppelhofer, and Miller (2004), Eicher, Papageorgiou, and Roehn (2007), attempt to identify the main determinants of economic growth. Raftery, Madigan, and Hoeting (1997), and Hoeting, Madigan, Raftery, and Volinsky (1999) give various other examples and mention possible applications. The interpretation of the estimates from BMA is straightforward as we can calculate conditional means and standard deviations which can be interpreted similarly to standard OLS coefficients and standard errors.

For this study a comprehensive dataset of 31 variables has been created that may affect the price of comparable single family houses in the Swiss metropolitan area of Zurich. Swiss municipalities are a particularly good study laboratory for the capitalization of municipal public goods and characteristics because of their high autonomy with regard to taxation.[3] Furthermore, the metropolitan area of Zurich is an ideal case

[2]Evidently, each combination of regressors represents a unique statistical model. Therefore, the number of combinations of possible models can be huge and usually follows some exponential function of the regressors. Classical statistics does not bring us any further to a solution of this problem. Accepting the fact that we do not know which model is really "true" we face substantial uncertainty in the choice of different models. Taking account of uncertainty is done by looking at the whole model space and attaching probabilities to its elements.

[3]Every year Swiss communities fix a communal income tax multiplier on the cantonal base tax. The

because of its high heterogeneity between municipalities with respect to location specific characteristics as well as demographic and socioeconomic variables. For the context of a highly developed and industrialized metropolitan area, the analysis suggests a minimal list of variables that might be included for other empirical analysis concerning the capitalization of community specific variables.

The remainder of this paper is organized as follows: Section 2 gives a brief overview concerning the methodology behind Bayesian statistics in general and BMA in particular. It also takes a look at the data and explains which variables are generally used in the literature on capitalization. Section 3 provides some baseline results that highlight the importance of certain regressors and show at the same time that a large number of explanatory variables have only little influence on property values. We then compare posterior mean of coefficients conditional on inclusion of the BMA algorithm with the setting having the highest posterior probability and with an OLS setting including all variables. Finally we perform an iterated Bayesian Model Averaging analysis for variable selection under different specifications which are common in other capitalization studies. Section 4 offers some concluding remarks.

2 Estimation

2.1 Methodological Issues of BMA

The empirical strategy is to address the issue of model uncertainty with respect to variable selection in regressions by employing Bayesian Model Averaging (BMA) as suggested by Raftery (1995). Here, we restrict ourselves to address the main intuitions behind the BMA methodology. For detailed introductions to model averaging and particularly BMA see Hoeting, Madigan, Raftery, and Volinsky (1999) and Raftery, Madigan, and Hoeting (1997).[4]

Broadly speaking, BMA is simple Bayesian statistics applied when there is model

differences in this tax multiplier are pronounced and continuously debated in public.

[4] Durlauf, Johnson, and Temple (2005) mention technical challenges with model averaging in economic applications and treat the choice of prior structure exhaustively.

uncertainty. In classical statistics a parameter has a true but unknown value. Such a parameter does not have a density because it is not random. As generally known, Bayesian statistics expresses all uncertainty in parameters and models in terms of probability. Moreover, the exploration of the joint posterior distribution gives a complete picture of the parameter uncertainty, which simply cannot be achieved via the classical approach. Inferences are made by applying the basic rules of probability calculus.[5]

Let y be the dependent variable, house prices, in n municipalities. The dependent variable is explained by an intercept α and $k-1$ independent predictor variables grouped in a design matrix $[\mathbf{1},X]$ of dimension $n \times k$. As we consider linear regression models only but are unsure about the k explanatory variables to include there are $K := 2^k$ possible statistical models, $M_1, ..., M_j, ..., M_K$. We denote X_j as a submatrix of X with dimension $n \times k_j$ and β_j is a column vector of regression coefficients with k_j elements. Model M_j with the regressors α and X_j can so be written in the standard form as

$$y = X_j \beta_j + \sigma \varepsilon, \tag{1}$$

where σ is a scale parameter and ε a random error term that follows an n-dimensional normal distribution.[6]

Denote $D = [y, \mathbf{1}, X]$ of dimension $n \times (k+1)$ as the full information we have got. Suppose we want to make inference about an unknown parameter of interest, say Δ, which has common interpretation across all models. A direct application of the law of total probabilities (see Durrett 2005 for a proof) implies that the posterior distribution of Δ given D is the weighted posterior distribution of that quantity under each model M_j

$$P(\Delta|D) = \sum_{j=1}^{K} P(\Delta|D, M_j) P(M_j|D). \tag{2}$$

The weights are given by the posterior model probabilities $P(M_j|D)$ which indicate the probability that M_j is the correct model given D. If we consider Δ as a single

[5]For additional advantages of Bayesian statistics and econometrics see Bolstad (2007).

[6]It is well know that for large sample sizes normality is not necessary by a central limit theorem (see Greene 2003 and Durrett 2005 for the respective Central Limit Theorems)

coefficient of a regression variable, the marginal posterior probability of including this variable is the sum of the posterior probabilities $P(\Delta|D, M_j)$ of the models which include this regressor, weighted by the posterior model probabilities $P(M_j|D)$. Again, with simple statistical theory ("Bayes rule") the posterior model probability of M_j is given by

$$P(M_j|D) = \frac{L(D|M_j)P(M_j)}{\sum_{l=1}^{K} L(D|M_l)p(M_l)}$$

where $P(M_j)$ is the prior model probability and $L(D|M_j)$ is the marginal (or integrated) likelihood of model M_j over all unknown parameters.[7] The marginal likelihood is a high dimensional integral and given by

$$\begin{aligned} L(D|M_j) &= \int_{\mathbb{R}^{k+1}} P(D|\alpha, \beta_j, \sigma, M_j)P(\beta_j|M_j)P(\alpha, \sigma)d\alpha d\beta_j d\sigma \\ &= \int_{\mathbb{R}^{k+1}} (\text{likelihood } \times \text{ prior})d\alpha d\beta_j d\sigma. \end{aligned}$$

Here, $P(D|\alpha, \beta_j, \sigma, M_j)$ is the likelihood to observe D given model M_j using the submatrix X_j corresponding to (1). $P(\beta_j|M_j)$ and $P(\alpha, \sigma)$ are the relevant priors. Yet again, when we are interested in a model parameter, say β_i, we get by (2) with $\Delta = \beta_i$ a model-averaged Bayesian "point" estimator

$$E(\beta_i|D) = \sum_{j=1}^{K} \tilde{\beta}_i^{(j)} P(M_j|D), \tag{3}$$

where the posterior mean $\tilde{\beta}_i^{(j)}$ of β_i is approximated by the corresponding maximum likelihood estimator $\hat{\beta}_i^{(j)}$ (see Raftery 1995). The same logic can be applied to obtain a posterior standard deviation. Clearly, we can also estimate the posterior probability that a certain variable has a nonzero coefficient and is therefore included in the regression. This probability is then called the posterior inclusion probability.

As we lack knowledge on the probability distributions of the models $P(M_j)$ it is straightforward to assume a uniform distribution and that the regressors are independent

[7] $L(D|M_j)$ is a probability measure itself.

of each other. This implies a prior probability of any model is

$$P(M_j) = \frac{1}{K} = \frac{1}{2^k}.$$

Each model is therefore equally likely. The probability of including any regressor then equals 1/2. We follow in the choice of the prior structure the common practice in the literature (see for example Raftery 1995, Raftery, Madigan, and Hoeting 1997, Hoeting, Madigan, Raftery, and Volinsky 1999 and Fernandez et al. (2001b, 2001a)).[8]

An additional implementation problem with BMA is that the number of models increases at an exponential rate 2^k in the number of regressor. Using $k = 32$ regressors (including the constant α) in the estimations corresponds to more than four billion different models. Efficient samplers have to be used to avoid exhaustive sampling. Our method follows Raftery (1995) who suggests the "Regression by Leaps And Bounds Algorithm" of Furnival and Wilson (1974). The main result of this algorithm is a reduction of several orders of magnitude in the number of operations required to find the best subsets. This is achieved by searching for the best subsets of possible variables for predicting the dependent variable. When applied to linear regressions a simple and fast computation of the Bayesian Information Criterion (BIC) is used.[9]

2.2 Data and Expected Capitalization Effects

For our purposes we use a panel dataset of 171 municipalities from 1998 to 2004 on the Swiss metropolitan area of Zurich. The Canton of Zurich is situated in the eastern Swiss midland. It extends from the river Rhein in the north to the beginning of the Alps in the south. With approximately 1.3 million inhabitants, Zurich is the most populous of all 26 Swiss cantons. Furthermore, it is one of the most densely populated areas in Europe. The City of Zurich itself is the center of the canton and the biggest

[8] Bayesian inference has been controversial because it uses the prior distribution, $P(M_j)$, which is subjectively determined by the user. However, it was shown that in large samples this has little importance. Furthermore, the chosen prior is safe and robust and may even be used as a benchmark. Sala-i-Martin, Doppelhofer, and Miller (2004) argue in favor of greater weights for smaller models.

[9] Other sampling algorithms include the "Markov Chain Monte Carlo Composition (MC3)" used by Fernandez et al. (2001b, 2001a) or the "Coinflip sampler" used by Sala-i-Martin, Doppelhofer, and Miller (2004).

urban agglomeration in Switzerland. Over a million people either work or live there.[10] The canton's 171 municipalities are very autonomous: Swiss municipalities have the possibility to levy income taxes via a municipal tax rate (collection rate) fixed by the municipality itself. Each community fixes every year a tax multiplier on the cantonal base tax which varies from canton to canton. Moreover, they may undertake a variety of local decisions concerning public expenditures without prior approval of the canton nor the federation. The Swiss legislation mostly addresses decision-making procedures and information the local governments must provide as decision support for voters. Communities have a large sovereignty of choice within certain minimum standards, regarding class size or other minimum requirements for public service production. This reduces, to some extent, the drawbacks of controlling correctly for public good provision by different proxies which can be problem according to Palmon and Smith (1998). Concerning population structure (population size, density, fraction of foreigners, etc.) and locational characteristics (distance to the center, distance to green spaces, distance to shops, pollution, etc.) the canton's municipalities are highly diverse. Further heterogeneity is added by the "Zürichsee", a 88.66 square kilometer large lake in the canton. Following the introduction of a harmonized public accounting system for bookkeeping and budgeting, reliable and consistent municipal financial data are available for all communities in our dataset. Consequently, the Canton of Zurich and its municipalities represents an ideal laboratory in order to identify the most important capitalization factors for house prices in an urban metropolitan area.

The dependent variable is the price of a standardized and comparable single family house for each community. The standardized house has five rooms, two wash rooms, 450 m2 garden area, 750 m3 volume, it is an end-terrace house, conveniently situated in the municipality, and has one garage space. The data was obtained from the Cantonal Bank of Zurich, the largest real estate banks in the canton, which evaluates houses by the sales comparison approach based on actual transactions. By looking at a comparable house for each municipality we can focus on differentials between communities as house

[10] For further information see the Statistical Office of the Canton of Zurich on http://www.statistik.admin.ch.

characteristics can be neglected.[11]

The independent variables were obtained from the Statistisches Amt des Kantons Zürich (Statistical Office of the Canton of Zurich), the Bildungsdirektion des Kantons Zürich (Secretary for Education of the Canton of Zurich), the GEFIS Finanzstatistik des Kantons Zürich (Financial Statistics of the Canton of Zurich), and from the Zürcher Kantonalbank (Cantonal Bank of Zurich). The variables include an array of different location specific characteristics, tax and expenditure information, school specific controls, median incomes, as well as demographic and other socioeconomic characteristics of the community. All independent variables, their sources, and a number of descriptive statistics are given in Table 1.

< Table 1 here >

The dataset contains observations for all 171 municipalities from the year 1998 to 2004. In the analysis we do not include the city of Zurich and Winterthur because as opposed to the other municipalities they are clearly considered as cities and have a different structure: Zurich and Winterthur have each a number of separate districts which form the cities. These districts differ in important aspects such as median incomes, unemployment rates, the fraction of foreigners etc. but they have the same tax multiplier and profit from the same public expenditures. Consequently, the effect of diverse fiscal variables cannot be measured for each district. Furthermore, Zurich is the center of the canton and we would like to control for the distance to the center in order to treat mobility issues in the analysis.[12] Finally, the two cities are large with respect the rest of the municipalities in the canton.[13]

[11] Capitalization studies such as Stull and Stull (1991), Palmon and Smith (1998) or Brasington (2001) look at heterogeneous houses and consequently have to control for housing characteristics such as the age of the house, number of rooms, the size of the house etc. Studies such as Oates (1969), Oates (1973), Ketkar (1992) or Reback (2005) use the median or average value of a house in a district.

[12] We performed a number of robustness tests (available on demand) including the city of Zurich and Winterthur. Our main insights do not change when we include these additional observations.

[13] Polinsky and Shavel (1976) show that using cross-section regressions to analyze the effect of amenities on house values is valid when the communities are considered "small" and there is mobility within and among them. The City of Zurich and Winterthur had an average number of 337262 inhabitants and 89757 inhabitants over the years 1998 to 2004. Whereas the average number for the other 169 municipalities were approximately 4700 inhabitants. The reduced samle of communities studied here is likely to approximate these conditions fairly well.

The literature on capitalization of single-family housing markets shows capitalization of a large number of different location specific variables. Still, most authors are interested in the effect of a specific variable on house prices and only include a limited subset of other controls in their studies, although some variables not considered were found as important determinants of house prices in other analyses. Here, we look at a large array of possible factors simultaneously and motivate the variables by referring to the existing literature. These variables are then used in BMA in order to identify the most important factors.

First we consider a number of location specific variables. We capture the possibility of living near a lake by taking the number of hectare of the lake which can be seen from the single-family house. Furthermore, we control for the percentage of hectare with south and west exposition of the house with respect to landmarks like mountains near the community. Both measures are expected to have a positive impact on house prices. Since Oates (1969) the inclusion of a measure for the distance to the main center is common (see Palmon and Smith 1998, Epple and Sieg 1999 or Brasington 2001 as further examples).[14] By including the distance to green spaces, we try to take account of the land in a community devoted to recreation. Similar variables are often considered in other capitalization studies. Stull and Stull (1991), for example, include the area devoted to industry in a community as an inverse measure for the living conditions in a community. Controlling for the distance to the center is not necessarily the same as controlling for the distance to commercial areas. Indeed, Lafferty and Frech III (1978) show that a wide dispersion of commercial developments reduces home values in Boston suburbs. Consequently, we include the distance to shopping centers as an additional measure of control. Estimating the effects of environmental damage is common in the capitalization literature: Portney (1981) looked at the effects of localized air pollution on home values in Pittsburgh. Ketkar (1992) and Kiel (1995) analyze the negative impacts of hazardous waste sites. For the estimations we include a control for environmental damage as NO2 particles in microgram per cubic meter of air. Usually in Europe and especially in Switzerland public transport plays a more important role than in the United

[14] Still not all older studies control for the distance to the center (see Edel and Sclar 1974 for instance)

States. We consequently include the distance to the next public transport facility and an indicator for access to fast trains in the estimations as is also done by Ketkar (1992). Note that these measures also represent an approximation of traffic on neighborhood streets which reduces property values according to a study by Hughes Jr. and Sirmans (1992).

Next we turn our attention to local fiscal variables that influence house prices. Clearly, the tax rate is the most common variable analyzed since Oates (1969). In the Swiss context, though, we focus on local income taxes instead of property taxes which are a common topic in the United States.[15] In a more recent study Palmon and Smith 1998 specifically focus on the estimation of property tax capitalization. Still, the effect of local income taxes was analyzed also for metropolitan areas in the United States by Stull and Stull (1991). They find a negative and significant impact. Moreover, we include an array of expenditure variables. Aggregated expenditures approximate public goods provision as mentioned by Oates (1973) and other authors but it is likely that a functional division of expenditures better reflects their real effect on house values. Cultural expenditures in a community reflect public good provision of cultural facilities. On the other hand expenditures for administration do not directly reflect public goods but high expenditures in this category might indicate waste of resources by the local government. Consequently, we expect that expenditures for administration capitalize negatively, where as expenditures for culture, health, security, social well being and for traffic facilities capitalize positively but with different coefficient values. Moreover, we take account of the municipal debt by constructing a theoretical number of years it takes a community to fully pay back its debts. This variable is expected to capitalize negatively as higher debts mean higher taxes in the future.[16]

The US literature on capitalization treats the effects of school characteristics, school distance, test scores or ethnic composition on house prices extensively. School characteristics are analyzed in some more recent studies by Black (1999), Figlio and Lucas (2004) or Reback (2005). The role of private schools is, for example, analyzed in Brasington

[15] Compared to income tax revenue, property tax revenue is small for Swiss communities.

[16] Stadelmann and Eichenberger (2008) discuss in detail the effect of local public debts and assets on house prices.

2000. The modern capitalization literature usually does not include communal expenditures for education in the regressions when educational quality serves as a pure control variable. Brasington 1999 as well as others use proficiency test passage rate in order to measure school quality and employ it as a control variable in their estimations. Because of stringent minimum standards in Swiss schools and an imposed and standardized curriculum the difference between Swiss public schools is rather limited. Furthermore, private schools are less common.[17] Still, we account for different measures of school characteristics in our study which might have an influence on house values. First we look at the average distance to the next school in meters which we expect to have a negative sign. Although minimum class sizes are imposed, there is some minor variation in class sizes and we include this variable as an additional control for school quality. Furthermore, living in a community which has an own grammar school may be more attractive as pupils do not depend on public transport and are closer to their homes. We consequently include a dummy variable that measures whether a community has a grammar school. Finally, the school organization in the canton of Zurich is different to other European municipal school structures. In most European countries, primary schools are usually managed by the communities themselves. In Zurich, schools can either be managed by the municipality itself or a special school community. The school communities can be seen as functional entities that overlap a number of different municipalities, usually geographical neighbors. The school communities have autonomous budgets and their accounts are often not consolidated in the political community because of the overlapping structure. Consequently, we do not have a consistent and reliable measure for education expenditures to include as an additional control. Instead, we have constructed a dummy variable which equals 0 if community and school are apart and takes the value of 1 if the community itself manages its schools. A priori, the impact of this variable is unclear. Frey and Eichenberger (2002) argue that a functional, overlapping and competing organization of federalism is likely to be more efficient but a number of other Swiss economists and bureaucrats argue in favor of centralization.[18]

[17] For more information on the Swiss schooling system in general and on schools in the Canton of Zurich in particular see http://www.bildungsdirektion.zh.ch.

[18] For an example see Regierungsrats des Kantons Zürich (2007). .

The median income or the income of a house owner is almost a standard variable used in the empirical capitalization literature. Again, Oates (1969) used this measure as well as most of the more recent articles on capitalization.

Moreover, we include a number of demographic characteristics in our estimation. These characteristics include the size of the population itself, the population density, the fraction of young people and the fraction of elderly. Most of these variables can be found in different empirical studies. They reflect general neighborhood and socioeconomic influences on house prices. Similar variables were used, for example, by Rohe and Stewart (1996). As further demographic controls, we include the fraction of foreigners which reflects the ethnic composition of the community. As Zurich is also the home of many highly educated expatriates the sign of this variable is unclear. The number of citizenship changes reflects communal openness as well as changes in the population structure. The fraction of persons employed in the third sector, the unemployment rate and the fraction of commuters in the workforce approximate working possibilities in the community. The unemployment rate itself may also serve as a measure for crime rates.[19]

Finally, we also include a variable which takes account of overall investments in living space per capita in the community. Such a type of variable captures municipal development and attractiveness (see Reback 2005 who also controls for a measure of the total market value derived from new construction and improvements).

3 Results

Now we are ready to conduct our BMA estimation. For every variable in the dataset we calculate the posterior inclusion probability as discussed in the last section. Figure 1 shows the posterior densities of the coefficients of 12 variables which are commonly used in the literature when analyzing the relation between housing prices, fiscal variables and location specific factors: the distance to center, NO2 pollution levels, the distance to public transport facilities, tax rates, expenditures for culture, expenditures for health, expenditures for social well-being, the distance to next school, median incomes, popula-

[19] Taylor (1995) mentions that high neighborhood crime rates reduce home values in Baltimore.

tion densities, the unemployment rate, and the fraction of commuters.

< **Figure 1 here** >

As mentioned by Sala-i-Martin, Doppelhofer, and Miller (2004) the marginal posterior distribution for each coefficient is a mixture-t distribution. Consequently such a distribution can take almost any form and we approximate it here by a finite mixture of normal densities which are scaled so that the height of the curve equals the posterior inclusion probability.[20] Indeed, we see that a number of variables such as the distance to the center or the level of pollution, expenditures for health and median income have a long tail distribution. The rest of the variables seems to follow a posterior distribution which is close to a normal one. The discrete mass at density zero gives the probability that a certain variable is not included in the model. Clearly, the inclusion probabilities for the distance to public transport facilities, the distance to the next school as well as the unemployment rate are fairly low. We also remark that for all variables apart from the distance to public transport facilities, the distance to schools and the unemployment rate the number of included control variables in the regression does not matter as far as the sign is concerned, i.e., all these variables always have the expected sign apart from the edges of the distributions.

We now present our numerical baseline results by analyzing simultaneously the 32 variables (including a constant) in the BMA framework. Subsequently, we take a closer look at the estimates by comparing the "Best Model" resulting from the BMA estimation procedure with the BMA posterior mean conditional on inclusion and an OLS estimation which includes all exogenous variables of our dataset. Furthermore, we implement an iterated BMA algorithm for variable selection which distinguishes itself from the baseline estimation by the fact that in each estimated model only a maximum number of explanatory variables are allowed. Finally, we present iterated BMA results for each year of our panel dataset separately.[21]

[20] We are using the statistical software R with the package BMA to obtain our computational and graphical results (http://www.r-project.org).

[21] Due to computational complexity when combining the issue of functional form with variable selection by BMA this paper focusses on the latter. Ekeland, Heckman, and Nesheim (2004) use non-parametric

To address possible problems of multicollinearity, numerical stability, and computational precision we calculate a "condition number" κ as often proposed by numerical mathematicians (see Schwarz and Köckler 2004, pages 51-54).[22] Generally speaking a very high condition number leads to problems with numerical estimations because a high number of floating point precision is lost. The condition number of all our estimated models is usually around $\kappa = 10^6$. Therefore, we are neither likely to have problems with precision nor with correlation between explanatory variables as the BMA algorithm itself chooses the variables having the highest effect on the Bayesian Information Criterion.[23]

3.1 Baseline results

< Table 2 here >

Table 2 presents the main results for the 32 variables. Note that all regressions used in the model averaging contain year fixed effects. Column (1) reports the posterior inclusion probability of a variable in the capitalization regression. Column (8) gives the ranking according to the posterior inclusion probability of the variable under consideration. For example median income is ranked first with an inclusion probability of 99.9 % whereas the distance to green spaces is ranked 6th with an inclusion probability of 22.9 %. The posterior inclusion probability can be interpreted as a goodness-of-fit measure of models including a specific variable versus models not including that variable. There is no clear consensus in the BMA literature about the threshold when a variable should be considered as effective. Raftery (1995) suggests that effectiveness is only reached when the posterior inclusion probability exceeds 50 %. This means that the posterior inclusion probability must be higher than the prior probability. There are 14 variables for which

methods to estimate a house price hedonic but does not focus on variable selection. BMA tests with semi-log forms lead to almost the same selection of controls as the linear form. Results for these additional tests are available on request.

[22] The condition number is defined as the ratio of the largest to the smallest non-zero singular value of the matrix.

[23] The machine accuracy of the statistical software R is approximately 16 floating point operations and R works with the perfectly stable QR algorithm when solving least squares problems. We also looked at the partial pair wise correlations of all variables. Only very few cases have partial pairwise correlations of over 0.5. There is no specific threshold value for the condition number. Still, as $\kappa = 10^6$ only approximately 7 floating point operations of 16 available are lost. Numerical accuracy is no problem in this case.

the posterior inclusion probability increases with respect to the prior. These variables are the constant, the controls for lake view, south-west exposition, the distance to the center, the distance to the next shop and the level of air pollution as far as major location specific characteristics are concerned. The most important fiscal variables are the tax rate, expenditures for culture and health as well as for social well-being. In the Swiss case it seems that school specific controls are of minor importance which is likely to be the case because of overall good public education.[24] As the curriculum is fixed, families cannot improve the education of their children by moving a specific community. Still, at least the average distance to the next school matters to some extent with an inclusion probability of 9.2 %. As supported by most other capitalization analyses starting with Oates (1969), Pollakowski (1973), Oates (1973) and a large follow up literature, the median income of a community is highly important. As far as demographic and socioeconomic characteristics are concerned only the population density, the fraction of elderly in the community and the fraction of commuters seem to satisfy the criterion of a posterior inclusion probability above 50 %.

Columns (2) and (4) show the posterior conditional mean and the posterior conditional standard deviations. Conditional, in this case, means that the variable is included in the model. The unconditional mean is calculated according to equation (3) and represents the weighted average of all OLS estimates including those where the variable in question is not included. The unconditional posterior mean is the product of the conditional mean and the inclusion probability. A similar approach may be applied for the conditional standard deviations (see Sala-i-Martin, Doppelhofer, and Miller 2004 or Raftery, Madigan, and Hoeting 1997). There is no notion of a single point estimate in Bayesian econometrics but only densities are known. Nevertheless, the posterior mean conditional on inclusion and posterior standard deviation conditional on inclusion have a straightforward interpretation: A researcher having the prior inclusion probability of 100 % for a specific variable and a 50 % inclusion probability for all remaining ones can interpret the coefficient of interest as the posterior conditional mean of Table 2.[25]

[24] Most likely this is also the case for the majority of countries in the European Union.

[25] The posterior conditional means and posterior conditional standard deviations from BMA can also be compared with coefficient and standard error estimates from standard OLS not taking account of

Moreover, we calculate an impact measure in column (3). This measure represents the influence in Swiss Francs on house values of a one percent increase of the mean of the respective independent variable with respect the posterior mean conditional on inclusion.

Looking at the estimated coefficients and standard errors in columns (2) and (4) as well as their impacts in Swiss francs we find that our expectations concerning the sign of the different variables are largely fulfilled. Indeed, all location specific variables have the expected sign as far as the posterior mean conditional on inclusion is considered and they are all significant apart from the distance to the next transport facility. Concerning the fiscal variables the tax rate capitalizes negatively and significantly with an impact of 963.32 Swiss Francs for a 1 % change in the mean tax multiplier. Expenditures for administration seem to capitalize positively but the posterior conditional mean is insignificant. The same is the case for expenditures for local security purposes. The measure for municipal debt repayment capacity capitalizes negatively and is marginally significant. Its posterior inclusion probability is with 4 % rather low but larger than the inclusion probabilities of administrative and security expenditures. A 1 % increase in the mean number of years for theoretical debt repayment reduces house values by 31.02 Swiss Francs. All other expenditure variables capitalize positively and significantly. For our schooling controls we find the expected signs: the distance to the next school is negative and marginally significant, the class size is negative but insignificant, living in a community with a grammar school has a positive but insignificant effect like the control for the school community has. Finally, we turn to the demographic and socioeconomic variables. The population size itself is insignificant but the density is highly significant and positive, indicating that densely populated areas show higher house values. The fraction of young capitalizes negatively and insignificantly but the fraction of elderly capitalizes positively and significantly. The main explanation of this finding is probably the fact that in the Canton of Zurich mainly the elderly have property and the higher the fraction of homeowners in a comunity the higher the property values according to Rohe and Stewart (1996). The fraction of foreigners has a positive impact because a large number of foreigners in the Canton of Zurich are highly educated expatriates.

model uncertainty.

Changes in citizenship, employment in the third sector and the unemployment rate have all the expected signs but are insignificant. Working possibilities in the community are important and a high fraction of commuters reduces house values significantly as expected. Finally, investments in living space which capture municipal development capitalize positively and marginally significant.

In column (5) we perform a Wilcoxon signed-rank test for the sign of the posterior mean conditional on inclusion (see Wilcoxon 1945). In the averaging procedure different models are estimated. In each of these models the sign of the variable under consideration is taken. It might be the case, for example, that the distance to transport facilities capitalizes positively in one model but negatively in another with different control variables. We test for this possibility and present the resulting p-values in column (4), i.e., we test whether the coefficients of the diverse models have the same sign as the reported posterior conditional mean. Apart from having a low inclusion probability a researcher cannot be sure about the sign of the variables used in his study. Depending on the variables included in the regression, the sign of the variables with a high p-value for the Wilcoxon test can either be positive or negative. For instance, in all models the coefficients for the lake view variable is positive and consequently the p-value equals 0.000. For the distance to public transport facilities, we find positive and negative coefficients over the whole model space. Still, at a level of 0.0591 the hypothesis that the distribution of the coefficients' signs is symmetric can be rejected. Most of our variables have p-values lower than 5 %. For the variables distance to public transport facilities, the theoretical debt repayment in years and changes in citizenship we find significance levels between 5 and 10 %. The signs of the variables expenditures for administration, expenditures for security, class size, own grammar school and employment in the third sector are unclear as the Wilcoxon test returns a p-value above 0.1 when the whole model space is considered.

Finally, we give a graphical representation of our results in Figure 2.

< **Figure 2 here** >

This plot is an immediate summary of the whole BMA output: Each row corresponds to a variable. In the columns the different models are represented. The width of the different columns corresponds proportionally to the model's posterior probability. Moreover, we represent negative coefficients with a darker color and positive ones with a lighter one.

3.2 Comparison of Coefficient Estimates

The Wilcoxon signed-rank test indicates that for a number of variables different specifications of the regression model return diverse coefficient signs. Now, we are interested whether the sizes of the coefficients are statistically different too. When estimating, for example, the impact of schools on house values such as in Jud and Watts (1981) the sign of a variable is of interest but also its impact on house prices. Consequently, the question shall be answered if variable and model selection influences the coefficients' sizes of the variables significantly.

We identify a "Best Model" as the model with the highest posterior probability over the whole model space. This model does not necessarily include all variables. Table 3 compares the "Best Model's" coefficients with an OLS specification including all 32 variables and the posterior mean conditional on inclusion. Moreover we calculate the standard deviation of the coefficients of the different models used in the averaging and estimate a bootstrap bias of these coefficients.

< Table 3 here >

Columns (1) and (2) give the coefficients and the standard errors of the "Best Model". Indeed, the highest posterior model probability is achieved when not all variables are included. Variables with a posterior inclusion probability below 50 % were dropped. In columns (3) and (4) we present OLS coefficient estimates and standard errors for house prices including all 32 variables of our dataset.

Turning to columns (5) to (7) gives an impression of the differences between the coefficients of three settings. We calculate the percentage differences between the coefficients of the "Best Model" versus the OLS setting in column (5), the "Best Model" versus the posterior mean conditional on inclusion in column (6) and finally the OLS setting versus the posterior mean conditional on inclusion in column (7).[26] Comparing the "Best Model" and the OLS estimates, there is no observable pattern as far as the differences of the coefficients are concerned. In ten cases the coefficient of the "Best Model" is higher than in the OLS setting. In five cases the OLS coefficient lies above the coefficient of the "Best Model". Especially, the relative difference for the variables distance to next shopping center, access to fast trains, and the population density are pronounced. Comparing the "Best Model" with the posterior mean conditional on inclusion in column (6), we notice that the absolute percentage values are lower, indicating that the conditional mean is closer to the "Best Model" than the simple OLS estimation with all control variables. Finally, we compare the OLS setting against the posterior conditional mean. Here all coefficients can be compared. Especially coefficients with a low posterior inclusion probability show very high differences in this setting. Interpreting the impact of coefficients in monetary terms should therefore be done cautiously especially if the true specification of the model is uncertain.

Column (8) gives a measure for the distribution of the coefficients over the different models estimated. When comparing the calculated standard deviation for each coefficient over all models with the posterior mean conditional on inclusion of Table 2 we find that the estimates with a high posterior inclusion probability show a comparatively small standard deviation. Coefficients with a low posterior inclusion probability tend to be distributed more widely around their mean.

Finally, we bootstrap the mean of each coefficient in all models 10000 times and compare it with the posterior mean conditional on inclusion resulting in an attempt to estimate of the bias of the posterior mean of Colum (1) of Table 2. For 16 values out of 32 the bias of the posterior mean conditional on inclusion is below 1 %. For ten variables the bias is between 1 and 10 %. Again variables with a low posterior inclusion probability tend to have higher biases. The distance to public transport facilities' posterior mean

[26] We do not consider the standard deviations but simply compare the coefficients themselves. This is motivated by the observation that the monetary impacts of different coefficient values on house prices are often discussed for policy consequences and in such cases the impact without a confidence interval is given.

conditional on inclusion shows (when compared with the bootstrapping results) a bias of approximately -10 %. Expenditures for administration have a negative bias of more than 40 %. The grammar school dummy and the population size as well as the fraction of persons employed in the third sector are biased between 10 and 20 % in absolute values.

3.3 Iterated BMA for Variable Selection and Robustness

Most capitalization studies only use a limited number of control variables. Clearly, the number of possible exogenous variables does not only depend on data availability but also on other issues such as problems with multi-collinearity and specific settings of the analysis: Palmon and Smith (1998), for example, do not need to take account of proxies for public goods as all the Texas municipal utility districts they analyze provide the same public services; other authors have to include additional controls as their houses are not comparable. Looking at the highest number of community specific control variables for a number of older and more recent studies shows the following picture: Oates (1969) includes seven controls, Pollakowski (1973) includes eight controls, Edel and Sclar (1974) include five controls, Stull and Stull (1991) include 15 municipal controls, and Brasington (2002) includes 13 municipal controls.

Next, we perform an iterated model averaging procedure with an exogenously determined number of maximum control variables to select from. Iterated BMA makes repeated calls to the Bayesian Model Averaging algorithm using the specified maximum number of control variables in every estimated model. After each iteration of BMA only those variables having a posterior inclusion probability greater than 5 % are used for the next step. Clearly, the lower the maximum number of control variables the higher the number of iterations performed.[27] The result of the iterated BMA procedure is a list of variables which were selected from the 32 possible controls (including the constant). Table 4 presents the results for three scenarios common in the literature.

< Table 4 here >

[27] If, for instance, the maximum number of controls to select is set to 32, no variable will be dropped and the iteration finishes after one step.

In column (1) we allow a maximum of ten variables to be selected. Indeed, especially the literature before 1990 mostly included a number of less than ten controls probably due to computational difficulties at the time. In such a setting mainly controls for locational characteristics remain in the estimation. Furthermore, the tax rate, expenditures for culture and health as well as the population density and the fraction of elderly are selected. All other variables drop out at a certain point during the iteration, meaning that at least in one out of 22 BMA iterations their posterior inclusion probability was below 5 %. A researcher interested in the effect of a certain policy variable on house prices, in a similar setting as presented here, should therefore try to find a set of location specific characteristics and at least some controls for fiscal characteristics of the community as well as the population density.[28] School and population specific variables turn out to be comparatively less important considering our data. In colum (2) we indicate the number of times a variable was included in a model. Most variables were dropped out directly after the first iteration.

In columns (3) and (4) we repeat the same exercise for a maximum of 15 controls to be selected. Again, except for the distance to green spaces and to public transport facilities all location dependent characteristics are selected by the algorithm. Expenditures for social well being, the median income of the community as well as the fraction of commuters are selected too. Lastly, year dummies become important in this setting.

Finally, we look at the selected variables when at most 20 controls are allowed to be chosen from the whole dataset. This represents a setting that is common in today's urban literature. Apparently, location specific characteristics remain important and only the control for the distance to the next public transport facility is dropped. Concerning fiscal variables the measure for the debt repayment capacity is selected. The distance to schools becomes important too but concerning demographic and socioeconomic characteristics no additional variables are selected nor dropped. Finally, the variable capturing investments in living space and consequently general attractiveness is selected. We remark, that the total number of iteration is low as 20 variables can be quickly selected from 32 possible

[28] Location specific characteristics and the selected fiscal variables are even more important than year fixed effects which were dropped during the iteration.

in only three iterations.

It is also worthwhile to note that most capitalization studies do not consider a long time horizon. Usually, only cross sectional data is used (Edel and Sclar 1974 are an exception). More recent papers such as, for example, Brasington (2002) analyze one year only but control for the quarter when the property transactions took place. To account for the fact that cross sectional observations instead of panel datasets are commonly used, we perform an iterated BMA procedure with a maximum number of 15 variables for each year from 1998 to 2004 separately. Additionally, we also include the maximum posterior probability of inclusion of all iterated BMA estimations and present the results in Table 5.

< **Table 5 here** >

The general picture concerning variables to be included when comparing Table 4 and Table 5 remains the same. Factors indicating the location of a house apart from the distance to public transport facilities remain highly important. Concerning fiscal variables, we see that expenditures for administration, security and the debt repayment measure drop out of the algorithm for all considered year. Expenditures for traffic facilities are dropped for all years except one. For school characteristics only the distance to the next school is selected for the year 1999 and the class size is selected for the year 2002. The median income is selected for all years and its maximum posterior inclusion probability is always at 99.9 %. As in Table 4 demographic and socioeconomic variables are mostly dropped during the iterating because their posterior inclusion probability is at least in one iteration below 5 %. Exceptions are the variabes population density which is selected four out of seven possible times, the fraction of elderly which is always selected, the fraction of foreigners which is selected four times, and the fraction of commuters which is always selected. The control for changes in citizenship remains in the estimation for the year 2000 but its maximum posterior inclusion probability is only 19.4 %. Finally, the control for living space investments is dropped in all subsamples apart of the year 2001.

4 Concluding Remarks

In this paper we took a fresh look at the capitalization of a number of location specific, fiscal, school, income, demographic and other socioeconomic variables for communities in a metropolitan area. There is a potentially large theoretical number of communal variables influencing house prices. Our analysis suggests that Bayesian Model Averaging (BMA) is a powerful way for researchers to analyze different factors of capitalization and make statements about their relative importance. The BMA method introduces many other improvements with respect to simple OLS regressions: For example, it is fully justified on Bayesian grounds and highlights the importance of particular variables via a posterior inclusion probability. We advocate a more formal treatment of model uncertainty in the light of a large choice of possible controls when estimating house prices. In addition, the BMA methodology provides clear, precise and easily interpretable results. It immediately allows posterior and predictive inference.

We use a rich dataset concerning communal characteristics of 171 municipalities in the Swiss metropolitan area of Zurich over the years 1998 to 2004. Due to the federal structure of Switzerland and the large autonomy of its municipalities as far as communal income taxes are concerned, Switzerland is an ideal case for the study of the capitalization of diverse community specific variables.

Our main results suggest that housing prices are primarily influenced by location specific factors such as the distance to the center, south-west exposition, the distance to local shopping centers, air pollution levels and access to fast public transport. Furthermore, fiscal variables and especially taxes play an important role. Expenditures for culture, health and social well-being clearly seem to determine house prices whereas expenditures for administration, public security and traffic have comparatively low posterior inclusion probabilities. Regarding the influence of schools, we find that in the Swiss context mainly the average distance to the next school matters to some extent. Class sizes, having and own grammar school in the community and living in a community with a separate school community seems to be of less essential. This might be due to the fact that private schools play a small role in education in Switzerland. Moreover, and in line with the capitalization literature, we find important effects of communal median incomes on property values. Most demographic and other socioeconomic characteristics seem of minor importance. Only the density, the fraction of elderly and the fraction of commuters seem to systematically influence house prices when the whole possible model space is considered. Comparisons with an OLS estimation including all variables in our dataset shows that the coefficients of the model with the highest posterior probability of the BMA algorithm are significantly different from the OLS coefficients. The most important variables identified in our baseline estimations also remain of high importance when performing an iterated BMA procedure with a maximum limit of controls for each specification.[29]

This analysis reveals scope for extending the research on capitalization further. Here we have established a set of main variables that may be considered in other capitalization studies in a similar context.

References

Bajari, P. and M. E. Kahn (2005). Estimating housing demand with an application to explaining racial segregation in cities. *Journal of Business & Economic Statistics 23*(1), 20–33.

Black, S. E. (1999). Do better schools matter? Parental valuation of elementary education. *Quarterly Journal of Economics 114*(2), 577–599.

Bolstad, W. M. (2007). *Introduction to Bayesian Statistics* (2 ed.). Wiley & Sons, Hoboken (NJ).

Brasington, D. M. (1999). Which measures of school quality does the housing market value? *Journal of Real Estate Research 18*(3), 395–414.

Brasington, D. M. (2000). Demand and supply of public school quality in metropolitan areas: The role of private schools. *Journal of Regional Science 40*(3), 583–605.

[29] Most capitalization studies only include a fraction of the number of community specific variables that are used in our setting.

Brasington, D. M. (2001). Capitalization and community size. *Journal of Urban Economics 50*(3), 385–395.

Brasington, D. M. (2002). Edge versus center: finding common ground in the capitalization debate. *Journal of Urban Economics 52*(3), 524–541.

Durlauf, S. N., P. A. Johnson, and J. R. W. Temple (2005). Growth econometrics. In P. Aghion and S. N. Durlauf (Eds.), *Handbook of Economic Growth*. North Holland, Amsterdam.

Durrett, R. (2005). *Probability - Theory and Examples* (3 ed.). Duxbury Advanced Series, Brooks/Cole, London.

Edel, M. and E. Sclar (1974). Taxes, spending, and property values: Supply adjustment in a Tiebout-Oates model. *Journal of Political Economy 82*(5), 941–954.

Eicher, T. S., C. Papageorgiou, and O. Roehn (2007). Unraveling the fortunes of the fortunate: An Iterative Bayesian Model Averaging (IBMA) approach. *Journal of Macroeconomics 29*(3), 494–514.

Ekeland, I., J. J. Heckman, and L. Nesheim (2004). Identification and estimation of hedonic models. *Journal of Political Economy 112*(1), S60–S109.

Epple, D. and H. Sieg (1999). Estimating equilibrium models of local jurisdictions. *Journal of Political Economy 107*(4), 645–681.

Fernandez, C., E. Ley, and M. F. Steel (2001a). Benchmark priors for Bayesian Model Averaging. *Journal of Econometrics 100*(2), 381–427.

Fernandez, C., E. Ley, and M. F. Steel (2001b). Model uncertainty in cross-country growth regressions. *Journal of Applied Econometrics 16*(5), 563–576.

Figlio, D. N. and M. E. Lucas (2004). What's in a grade? School report cards and the housing market. *American Economic Review 94*(3), 591–604.

Fischel, W. A. (2001). *The Homevoter Hypothesis: How Home Values Influence Local Government Taxation, School Finance, and Land-Use Policies*. Harvard University Press, Cambridge (MA).

Frey, B. S. and R. Eichenberger (2002). *The New Democratic Federalism for Europe: Functional, Overlapping, and Competing Jurisdictions*. Edward Elgar Publishing Limited, Cheltenham.

Furnival, G. M. and R. W. Wilson (1974). Regressions by leaps and bounds. *Technometrics 16*(4), 499–511.

Greene, W. H. (2003). *Econometric Analysis* (5 ed.). Pearson Education, Prentice Hall, Upper Saddle River (NJ).

Hoeting, J. A., D. Madigan, A. E. Raftery, and C. T. Volinsky (1999). Bayesian Model Averaging: A tutorial (with discussion). *Statistical Science 14*(4), 382–417.

Hughes Jr., W. T. and C. F. Sirmans (1992). Traffic externalities and single family house prices. *Journal of Regional Science 32*(4), 487–500.

Jud, G. D. and J. M. Watts (1981). Schools and housing values. *Land Economics 57*(3), 459–470.

Ketkar, K. (1992). Hazardous waste sites and property values in the State of New Jersey. *Applied Economics 24*(6), 647–59.

Kiel, K. A. (1995). Measuring the impact of the discovery and cleaning of identified hazardous waste sites on house values. *Land Economics 71*(4), 428–435.

Lafferty, R. N. and H. E. Frech III (1978). Community environment and the market value of single-family homes: The effect of the dispersion of land uses. *Journal of Law and Economics 21*(2), 381–394.

Oates, W. E. (1969). The effects of property taxes and local public spending on property values: An empirical study of tax capitalization and the Tiebout hypothesis. *Journal of Political Economy 77*(6), 957–971.

Oates, W. E. (1973). The effects of property taxes and local public spending on property values: A reply and yet further results. *Journal of Political Economy 81*(4), 1004–1008.

Palmon, O. and B. A. Smith (1998). New evidence on property tax capitalization. *Journal of Political Economy 106*(5), 1099–1111.

Polinsky, M. A. and S. Shavel (1976). Amenities and property values in a model of an urban area. *Journal of Public Economics 5*(1-2), 119–129.

Pollakowski, H. O. (1973). The effects of property taxes and local public spending on property values: A comment and further results. *Journal of Political Economy 81*(4), 994–1003.

Portney, P. R. (1981). Housing prices, health effects, and valuing reductions in risk of death. *Journal of Environmental Economics and Management 8*(1), 72–78.

Raftery, A. E. (1995). Bayesian model selection in social research. *Sociological Methodology 25*, 111–163.

Raftery, A. E., D. Madigan, and J. A. Hoeting (1997). Bayesian model averaging for linear regression models. *Journal of the American Statistical Association 92*(437), 179–191.

Reback, R. (2005). House prices and the provision of local public services: Capitalization under school choice programs. *Journal of Urban Economics 57*(2), 275–301.

Regierungsrats des Kantons Zürich (2007, November). Vernehmlassungsentwurf: Leitsätze für eine Reform der Gemeindestrukturen.

Reinhard, R. M. (1981). Estimating property tax capitalization: A further comment. *Journal of Political Economy 89*(6), 1251–1260.

Rohe, W. M. and L. S. Stewart (1996). Homeownership and neighborhood stability. *Housing Policy Debate 7*(1), 37–81.

Sala-i-Martin, X., G. Doppelhofer, and R. I. Miller (2004). Determinants of long-term growth: A Bayesian Averaging of Classical Estimates (BACE) approach. *American Economic Review 94*(4), 813–835.

Schwarz, H.-R. and N. Köckler (2004). *Numerische Mathematik* (5 ed.). Teubner Verlag, Wiesbaden.

Stadelmann, D. and R. Eichenberger (2008, January). The Ricardian equivalence versus debt capitalization in local public finance. Available at SSRN: http://ssrn.com/abstract=1088261.

Stull, W. J. and J. C. Stull (1991). Capitalization of local income taxes. *Journal of Urban Economics 29*(2), 182–190.

Taylor, R. B. (1995). The impact of crime on communities. *Annals of the American Academy of Political and Social Science 539*, 28–45.

Wilcoxon, F. (1945). Individual comparisons by ranking methods. *Biometrics Bulletin 1*(6), 80–83.

Yinger, J., H. S. Bloom, A. Börsch-Supan, and H. F. Ladd (1988). *Property Taxes and House Values: The Theory and Estimation of Intrajuristictional Property Tax Capitalization.* Academic Press, Bosten (MA).

Table 1
Data Description and Sources

Variable	*Description and source*	*Median*	*Mean*	*S.d.*
HousePrice	Price in Swiss Francs of standardized and comparable single family house. Cantonal Bank of Zurich.	788600	805900	133813
Lakeview	View on lake in number of hectare. Cantonal Bank of Zurich and Statistical Office of the Canton of Zurich (GIS system).	9.424	359.50	871.53
SWExposition	Percentage of hectare with south and west exposition. Cantonal Bank of Zurich.	0.404	0.429	0.275
DistCenter	Average time in minutes to Zurich main station. Cantonal Bank of Zurich and Statistical Office of the Canton of Zurich (GIS system).	26.700	26.700	8.493
DistGreen	Average distance to next park (green space) in meter. Cantonal Bank of Zurich.	33.270	40.290	27.554
DistShop	Average distance to shopping center in meter. Cantonal Bank of Zurich.	963.10	1200.00	674.99
NO2Pollution	Environmental damage as NO2 in microgram per cubic meter. Cantonal Bank of Zurich.	17.020	17.820	4.125
DistTransport	Average distance to next public transport facility in meter. Cantonal Bank of Zurich and Statistical Office of the Canton of Zurich (GIS system).	506.10	554.10	209.95
AccessFasttrain	Indicator for access to fast train (S-Bahn) as a fraction of the population. Statistical Office of the Canton of Zurich (GIS system).	27.000	32.320	32.754
TaxRate	Mean income tax rate (without churches). Statistical Office of the Canton of Zurich.	119.00	113.80	14.882
ExpAdmin	Expenditure for administration in Swiss Francs per capita. GEFIS Financial Statistics and Statistical Office of the Canton of Zurich.	354.00	373.00	139.92
ExpCulture	Expenditure for culture in Swiss Francs per capita. GEFIS Financial Statistics and Statistical Office of the Canton of Zurich.	78.000	93.540	63.555
ExpHealth	Expenditure for health in Swiss Francs per capita. GEFIS Financial Statistics and Statistical Office of the Canton of Zurich.	138.00	152.50	83.37
ExpSecurity	Expenditure for public security in Swiss Francs per capita. GEFIS Financial Statistics and Statistical Office of the Canton of Zurich.	149.00	153.60	45.50
ExpSocial	Expenditure for social well-being in Swiss Francs per capita. GEFIS Financial Statistics and Statistical Office of the Canton of Zurich.	265.00	296.20	158.81
ExpTraffic	Expenditure for traffic facilities and improvements in Swiss Francs per capita. GEFIS Financial Statistics and Statistical Office of the Canton of Zurich.	165.00	180.70	77.246
DebtRepay	Number of (theoretical) years for full debt repayment using tax revenues only (total debts divided by total tax revenues). GEFIS Financial Statistics and Statistical Office of the Canton of Zurich.	1.361	1.684	1.167
DistSchool	Average distance to schools in meter. Cantonal Bank of Zurich and Statistical Office of the Canton of Zurich (GIS system).	851.20	860.20	219.29
ClassSize	Average class size in primary school. Secretary for Education of the Canton of Zurich.	20.300	19.920	1.823

GrammarSchool	Identification whether community has a grammar school (value=1) or not (value=0). Secretary for Education of the Canton of Zurich.	0.000	-	0.170
NoSchoolComm	Identification whether the school is managed by the community itself (value=1) or a separate school community (value=0). Secretary for Education of the Canton of Zurich.	0.000	-	0.399
MedianIncome	Median income to tax of natural persons. Statistical Office of the Canton of Zurich.	46550	47310	5759.94
Pop1000	Population in 1000. Statistical Office of the Canton of Zurich.	3.104	4.726	4.944
Density	Population per square kilometer. Statistical Office of the Canton of Zurich.	405.90	601.00	598.49
Young	Fraction of population under 20 years. Statistical Office of the Canton of Zurich.	24.600	24.340	3.289
Elderly	Fraction of population over 65 years. Statistical Office of the Canton of Zurich.	12.300	12.560	2.981
Foreigners	Fraction of foreigners. Statistical Office of the Canton of Zurich.	12.100	13.290	7.579
CitizenshipChange	Changes in citizenship per 1000 population. Statistical Office of the Canton of Zurich.	3.300	3.744	3.027
Employed3sector	Fraction of labor force employed in third sector. Statistical Office of the Canton of Zurich.	64.900	65.200	12.907
Unemployment	Unemployment rate. Statistical Office of the Canton of Zurich.	2.000	2.234	1.237
Commuters	Fraction of commuters outgoing over labor force in community. Statistical Office of the Canton of Zurich.	0.699	0.690	0.067
LivingSpace	Investments in construction of living space in Swiss Francs per capita.	2938.00	3411.00	2432.60

Source: as mentioned in table
The median, mean and standard deviations are based on 1176 observations which are 168 municipalities from 1998 to 2004.

Table 2
Baseline Estimates for all 32 Variables (Including a Constant) on House Prices

Variable	*Posterior inclusion probability* *(1)*	*Posterior mean conditional on inclusion* *(2)*	*Impact** *(3)*	*Posterior s.e. conditional on inclusion* *(4)*	*Sign-test p-value* *(5)*	*Rank* *(6)*
Intercept	99.9	792353.9743		50424.6548	0.0000	1
Lakeview	99.9	34.8615	*125.32*	1.8653	0.0000	1
SWExposition	99.9	66316.5508	*284.24*	5263.2882	0.0000	1
DistCenter	99.9	-6154.2196	*-1643.44*	291.5937	0.0000	1
DistGreen	22.9	263.7982	*106.27*	119.5568	0.0000	6
DistShop	99.9	-10.1479	*-121.73*	2.2731	0.0000	1
NO2Pollution	99.9	-6631.1794	*-1181.66*	528.1966	0.0000	1
DistTransport	1.7	-6.5113	*-36.08*	7.8124	0.0591	13
AccessFasttrain	41.9	122.9670	*39.74*	47.3538	0.0000	4
TaxRate	99.9	-846.6231	*-963.32*	164.0857	0.0000	1
ExpAdmin	0.9	0.9792	*3.65*	10.8727	0.7893	21
ExpCulture	99.9	151.6901	*141.89*	29.8019	0.0000	1
ExpHealth	99.9	172.4338	*262.98*	23.0922	0.0000	1
ExpSecurity	0.9	10.2976	*15.82*	30.6931	0.1814	21
ExpSocial	99.9	84.0808	*249.06*	15.2329	0.0000	1
ExpTraffic	24.1	46.8725	*84.68*	20.7331	0.0000	5
DebtRepay	4.0	-1895.8472	*-31.92*	1401.6146	0.0059	10
DistSchool	9.2	-12.2076	*-105.00*	6.8065	0.0005	9
ClassSize	1.3	-549.8044	*-109.51*	782.3564	0.1003	16
GrammarSchool	1.1	2672.8920	*0.796*	8159.7119	0.2012	18
NoSchoolComm	1.6	2660.4071	*5.27*	3670.0103	0.0360	14
MedianIncome	99.9	7.0199	*3321.10*	0.4696	0.0000	1
Pop1000	0.9	0.1979	*0.009*	449.6415	1.0000	21
Density	92.6	17.0653	*102.55*	4.6770	0.0000	3
Young	2.0	-678.4537	*-165.16*	725.2237	0.0360	12
Elderly	99.9	4705.5229	*590.80*	726.2605	0.0000	1
Foreigners	14.4	677.8009	*90.08*	342.0352	0.0000	8
CitizenshipChange	1.3	280.9330	*10.51*	555.1850	0.0591	16
Employed3sector	1.1	11.9868	*7.81*	115.4348	0.5839	18
Unemployment	2.1	-2061.5701	*-46.05*	2433.5506	0.0143	11
Commuters	97.3	-111357	*-768.86*	29497.9247	0.0000	2
LivingSpace	15.1	1.1652	*39.74*	0.5646	0.0001	7
YearDummies	YES	YES	*YES*	YES	YES	YES

Source: own calculations
* The impact in Swiss Francs of a one percent increase of the mean of the respective independent variable on property prices.
The left-hand-side variable in all regressions is the price of a standardized and comparable single family house from 1998 to 2004 across 168 municipalities. The conditional mean and standard deviation are conditional on inclusion of the variable in the model. The sign-test is a Wilcoxon signed-rank test for the sign of the variable over all models. The p-value of the sign tests indicates whether the coefficient is on the same side zero as its posterior mean conditional on inclusion. The final column ranks all variables according to their posterior inclusion probability.

Table 3
Comparison of Estimates

Variable	*Coefficient of "Best Model" (1)*	*S.e. of "Best Model" (2)*	*OLS coefficient (3)*	*OLS s.e. (4)*	*"Best Model" vs OLS (5)*	*"Best Model" vs Conditional Mean (6)*	*OLS vs Conditional Mean (7)*	*Standard deviation of BMA coefficients (8)*	*Bootstrap bias of BMA coefficients (9)*
Intercept	805889.2984	43776.6194	829386.5675	62230.8640	-2.833%	1.708%	4.674%	31535.5638	-0.0063
Lakeview	34.9022	1.8060	34.9563	1.9100	-0.155%	0.117%	0.272%	0.6409	0.0027
SWExposition	67176.6900	5197.7384	67713.9831	5333.8298	-0.793%	1.297%	2.107%	947.4662	-0.0005
DistCenter	-6228.6240	274.3047	-6077.1458	333.4986	2.493%	1.209%	-1.252%	106.9665	-0.0039
DistGreen	dropped	dropped	109.7077	115.7414			-58.412%	71.6521	0.0013
DistShop	-10.0781	2.1931	-8.5594	2.3867	17.744%	-0.687%	-15.654%	0.5958	-0.0091
NO2Pollution	-6483.2650	487.7269	-6720.7681	553.0155	-3.534%	-2.231%	1.351%	214.4235	0.0056
DistTransport	dropped	dropped	0.0743	8.7314			-98.860%	3.1831	-0.1056
AccessFasttrain	126.9133	46.0425	84.9171	53.0112	49.456%	3.209%	-30.943%	11.9882	-0.0239
TaxRate	-887.6430	160.2777	-815.5980	169.1499	8.833%	4.845%	-3.665%	36.1971	-0.0117
ExpAdmin	dropped	dropped	-1.0787	11.8331			10.157%	3.2048	-0.4352
ExpCulture	149.3989	28.4580	128.3108	30.7769	16.435%	-1.510%	-15.413%	10.3389	-0.0059
ExpHealth	173.6083	22.7330	166.3394	23.3675	4.370%	0.681%	-3.534%	4.5530	-0.0002
ExpSecurity	dropped	dropped	-1.1704	32.4977			-88.634%	4.2264	0.0745
ExpSocial	81.7746	14.1376	79.0842	15.5572	3.402%	-2.743%	-5.943%	6.9591	0.0067
ExpTraffic	dropped	dropped	47.2833	21.4383			0.876%	4.5946	-0.0035
DebtRepay	dropped	dropped	-1448.0537	1470.2642			-23.620%	322.3567	-0.0286
DistSchool	dropped	dropped	-13.6432	7.8427			11.760%	0.9908	0.0094
ClassSize	dropped	dropped	-236.9147	833.4043			-56.909%	172.0593	-0.0232
GrammarSchool	dropped	dropped	4634.4298	8796.6272			73.386%	1995.8132	-0.1034
NoSchoolComm	dropped	dropped	1719.5900	3794.2335			-35.364%	927.8125	0.0428
MedianIncome	7.0083	0.4568	6.7744	0.4999	3.452%	-0.166%	-3.498%	0.1245	-0.0023
Pop1000	dropped	dropped	-355.8924	528.1132			179759.547%	131.4370	-0.1573
Density	16.7343	4.1032	14.2865	4.9197	17.134%	-1.939%	-16.283%	2.3375	-0.0309
Young	dropped	dropped	-377.9673	775.6294			-44.290%	128.9070	0.0015
Elderly	4566.3231	630.2024	4136.1894	801.0484	10.399%	-2.958%	-12.099%	417.7218	0.0107

Foreigners	dropped	dropped	875.8644	425.8275			29.221%	192.6884	0.0566
CitizenshipChange	dropped	dropped	-306.4083	614.2434			9.068%	139.8194	0.0159
Employed3sector	dropped	dropped	36.0288	118.4404			200.572%	26.1020	-0.1878
Unemployment	dropped	dropped	-3376.0115	2245.8921			63.759%	1522.4662	0.0758
Commuters	-116099.3508	27216.9895	-117933.8889	32455.1216	-1.556%	4.258%	5.905%	12138.4584	-0.0134
LivingSpace	dropped	dropped	0.9057	0.5754			-22.273%	0.0975	-0.0300
YearDummies	YES	YES	YES	YES					

Source: own calculations

The left-hand-side variable in all regressions is the price of a standardized and comparable single family house from 1998 to 2004 across 168 municipalities. The coefficients and the standard errors of the "Best Model" represent the coefficients and standard errors of the model with the highest posterior probability over the whole model space. The "Best Model" does not necessarily include all variables. The OLS coefficient and standard error represent the results of simple OLS using all 32 variables. The "Best Model's" coefficients are compared in percentage terms with the OLS model and the conditional mean of Table 2(1). The standard deviation of the BMA coefficients represents a measure of the distribution of coefficients over the whole model space. The bootstrap bias is represents a comparison of the conditional mean of Table 2(1) and the bootstrapped mean of the coefficients for all models estimated. The mean of the coefficients is bootstrapped 10000 times. The bias is the percent difference of the conditional mean and the bootstrap results.

Table 4
Iterated BMA for Variable Selection

Variable	*10 variables to select (1)*	*Number of times in model during iteration for 10 variables (2)*	*15 variables to select (3)*	*Number of times in model during iteration for 15 variables (4)*	*20 variables to select (5)*	*Number of times in model during iteration for 20 variables (6)*
Intercept	selected	22	selected	12	selected	3
Lakeview	selected	22	selected	12	selected	3
SWExposition	selected	22	selected	12	selected	3
DistCenter	selected	22	selected	12	selected	3
DistGreen	dropped	4	dropped	4	selected	3
DistShop	selected	22	selected	12	selected	3
NO2Pollution	selected	22	selected	12	selected	3
DistTransport	dropped	2	dropped	2	dropped	1
AccessFasttrain	selected	22	selected	12	selected	3
TaxRate	selected	22	selected	12	selected	3
ExpAdmin	dropped	1	dropped	1	dropped	1
ExpCulture	selected	22	selected	12	selected	3
ExpHealth	selected	21	selected	12	selected	3
ExpSecurity	dropped	1	dropped	1	dropped	1
ExpSocial	dropped	10	selected	12	selected	3
ExpTraffic	dropped	1	dropped	12	dropped	1
DebtRepay	dropped	1	dropped	1	selected	3
DistSchool	dropped	1	dropped	5	selected	3
ClassSize	dropped	1	dropped	1	dropped	2
GrammarSchool	dropped	1	dropped	1	dropped	1
NoSchoolComm	dropped	1	dropped	1	dropped	1
MedianIncome	dropped	1	selected	10	selected	3
Pop1000	dropped	1	dropped	1	dropped	1
Density	selected	10	selected	9	selected	2
Young	dropped	1	dropped	2	dropped	1
Elderly	dropped	1	selected	8	selected	2
Foreigners	dropped	1	dropped	4	dropped	1
CitizenshipChange	dropped	1	dropped	1	dropped	1
Employed3sector	dropped	1	dropped	1	dropped	1
Unemployment	dropped	1	dropped	2	dropped	1
Commuters	dropped	1	selected	3	selected	1
LivingSpace	dropped	1	dropped	1	selected	1
YearDummies	dropped	1	selected	1	selected	1

Source: own calculations
The left-hand-side variable in all regressions is the price of a standardized and comparable single family house from 1998 to 2004 across 168 municipalities. Iterated BMA works by making repeated calls to the Bayesian model averaging procedure, iterating through the variables. After each call to the Bayesian model averaging procedure only those variables which have posterior probability greater than 5 %. The maximum number of variables entering each BMA estimation is 10, 15 and 20 for columns (1), (3) and (5) respectively. The number of times in model during iteration indicates how often a variable was included during all iterations. The constant is always included.

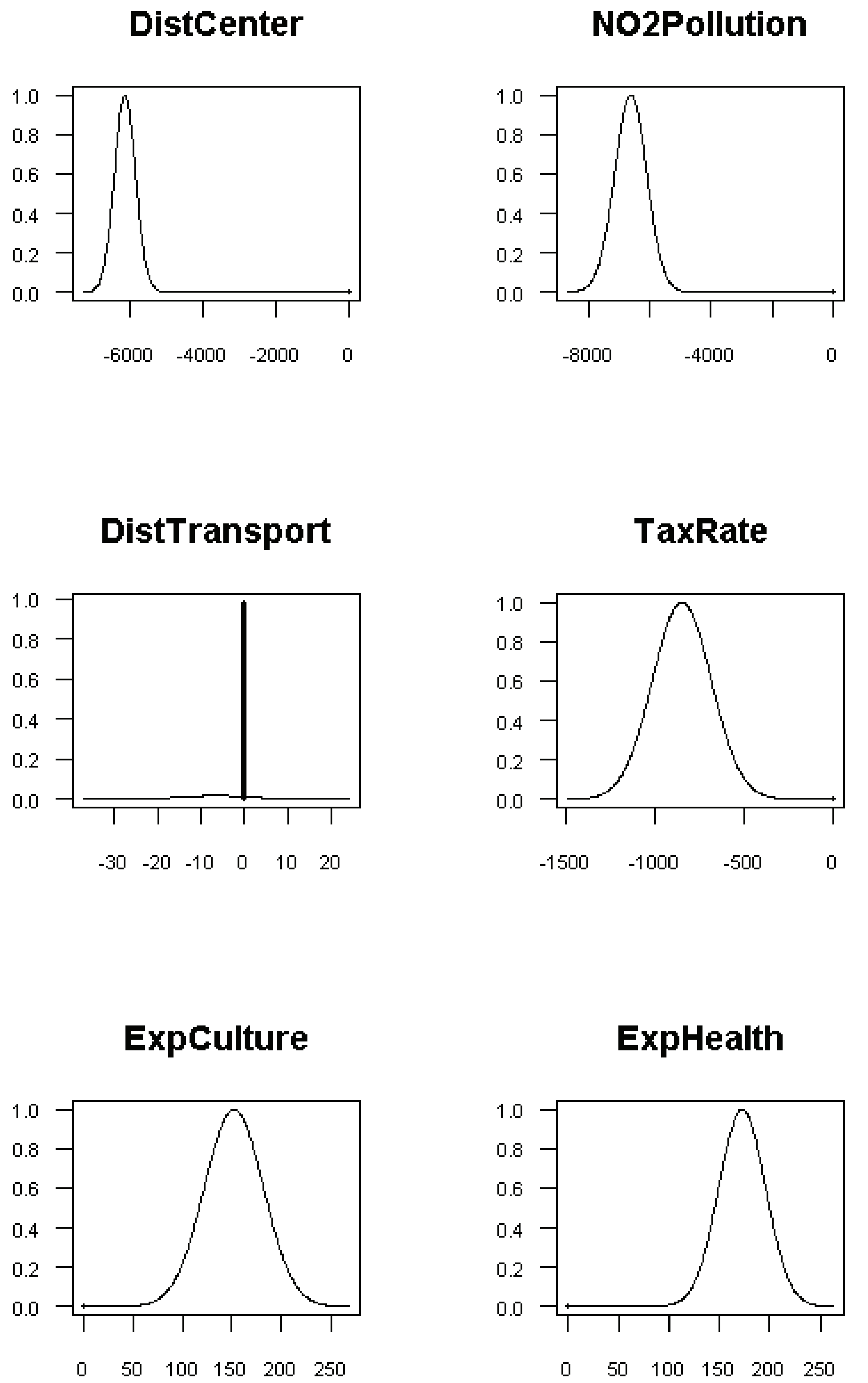
Table 5
Iterated BMA and Maximum Posterior Inclusion Probability for Different Subsets

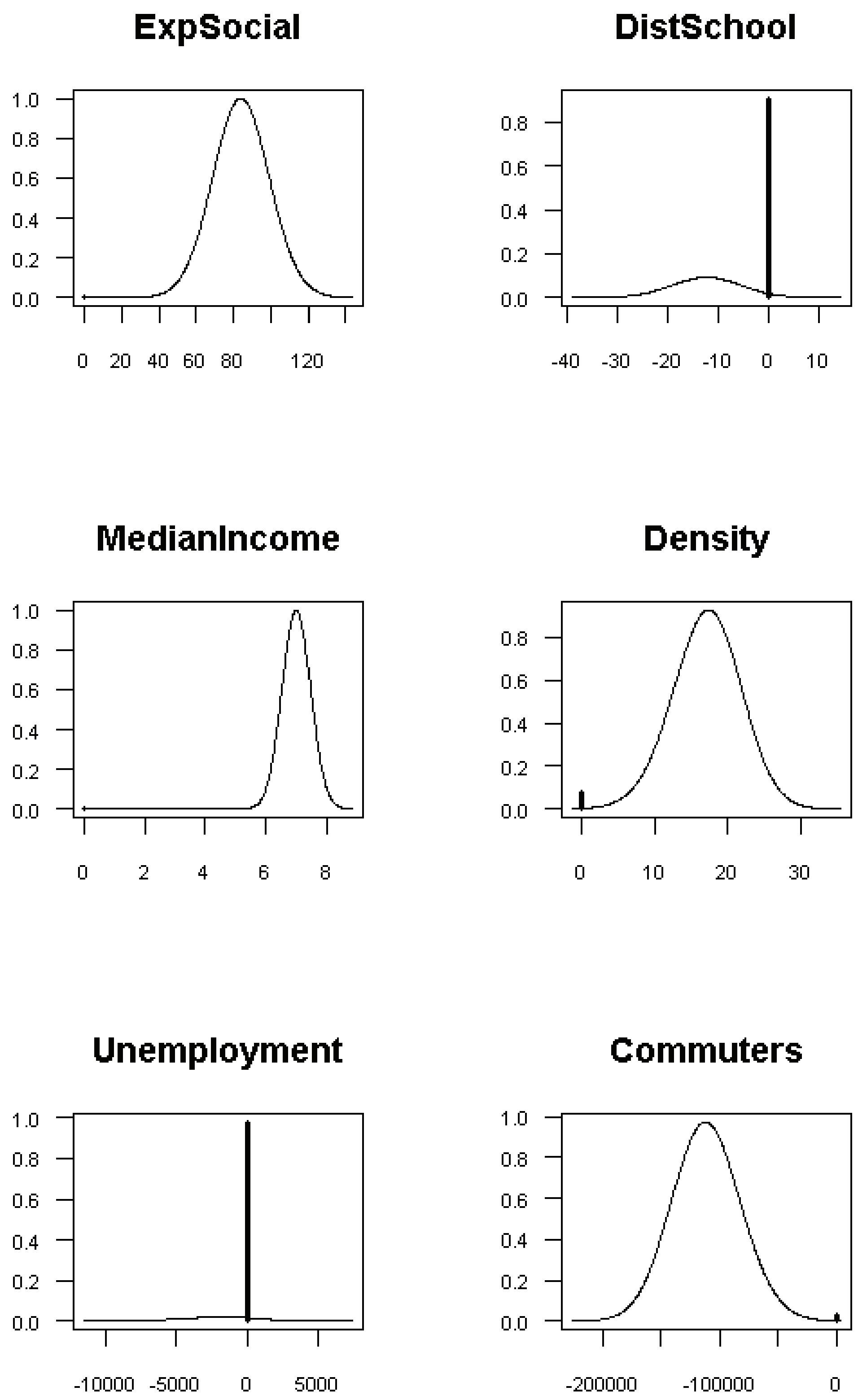
Variable	*iBMA selection 1998*	*iBMA selection 1999*	*iBMA selection 2000*	*iBMA selection 2001*	*iBMA selection 2002*	*iBMA selection 2003*	*iBMA selection 2004*
Intercept	99.9 selected	99.9 selected	99.9 selected	99.9 selected	99.9 selected	99.9 selected	99.9 selected
Lakeview	99.9 selected	99.9 selected	99.9 selected	99.9 selected	99.9 selected	99.9 selected	99.9 selected
SWExposition	99.9 selected	99.9 selected	99.9 selected	99.9 selected	99.9 selected	99.9 selected	99.9 selected
DistCenter	99.9 selected	99.9 selected	99.9 selected	99.9 selected	99.9 selected	99.9 selected	99.9 selected
DistGreen	77.2 selected	19.2 dropped	70.9 selected	4.5 dropped	55.3 selected	28.4 selected	45.3 selected
DistShop	99.9 dropped	90.2 dropped	84.5 dropped	95.6 selected	92.7 dropped	99.9 selected	91.4 selected
NO2Pollution	99.9 selected	99.9 selected	99.9 selected	99.9 selected	99.9 selected	99.9 selected	99.9 selected
DistTransport	6.0 dropped	8.4 dropped	12.6 dropped	5.7 dropped	5.0 dropped	2.1 dropped	3.2 dropped
AccessFasttrain	48.3 dropped	67.1 selected	60.1 dropped	42.9 selected	14.2 dropped	21.8 selected	58.7 dropped
TaxRate	99.9 selected	99.9 selected	99.9 selected	99.9 selected	99.9 dropped	99.9 selected	99.9 selected
ExpAdmin	13.9 dropped	2.5 dropped	3.9 dropped	3.8 dropped	6.7 dropped	6.8 dropped	2.5 dropped
ExpCulture	99.9 selected	99.9 selected	99.9 selected	99.9 selected	98.3 selected	99.9 selected	99.9 selected
ExpHealth	99.9 selected	48.2 dropped	59.0 dropped	99.9 selected	99.9 selected	99.9 selected	99.9 selected
ExpSecurity	4.2 dropped	2.2 dropped	3.5 dropped	3.7 dropped	4.9 dropped	18.0 dropped	3.1 dropped
ExpSocial	99.9 selected	99.9 selected	99.9 selected	99.9 selected	99.9 selected	99.9 selected	99.9 selected
ExpTraffic	69.3 dropped	31.7 dropped	42.4 dropped	8.2 dropped	61.3 selected	9.9 dropped	2.3 dropped
DebtRepay	3.8 dropped	1.9 dropped	3.5 dropped	10.8 dropped	6.2 dropped	4.7 dropped	2.0 dropped
DistSchool	8.8 dropped	31.3 selected	23.6 dropped	8.7 dropped	10.9 dropped	3.3 dropped	10.6 dropped
ClassSize	8.3 dropped	1.8 dropped	4.6 dropped	4.3 dropped	20.4 selected	7.4 dropped	2.0 dropped
GrammarSchool	3.8 dropped	1.3 dropped	2.1 dropped	5.8 dropped	3.9 dropped	2.1 dropped	8.2 dropped
NoSchoolComm	3.9 dropped	2.5 dropped	2.1 dropped	6.1 dropped	5.1 dropped	4.2 dropped	3.1 dropped
MedianIncome	99.9 selected	99.9 selected	99.9 selected	99.9 selected	99.9 selected	99.9 selected	99.9 selected
Pop1000	6.1 dropped	2.0 dropped	4.5 dropped	5.3 dropped	3.0 dropped	3.8 dropped	4.9 dropped
Density	39.9 selected	33.2 selected	36.9 selected	15.4 dropped	6.2 dropped	4.0 dropped	13.8 selected
Young	12.0 dropped	37.6 dropped	57.5 dropped	7.6 dropped	6.1 dropped	4.3 dropped	4.7 dropped
Elderly	99.9 selected	99.9 selected	99.9 selected	76.1 selected	73.0 selected	33.3 selected	98.3 selected

Foreigners	68.6 selected	31.6 selected	23.4 selected	30.1 selected	26.1 dropped	4.3 dropped	4.3 dropped
CitizenshipChange	1.6 dropped	3.9 dropped	19.4 selected	3.1 dropped	9.2 dropped	2.0 dropped	10.0 dropped
Employed3sector	2.4 dropped	3.3 dropped	1.1 dropped	1.0 dropped	1.3 dropped	2.4 dropped	2.1 dropped
Unemployment	7.4 dropped	4.0 dropped	13.6 dropped	2.5 dropped	4.2 dropped	2.1 dropped	6.8 dropped
Commuters	16.1 selected	5.1 selected	32.8 selected	47.7 selected	56.8 selected	34.9 selected	33.6 selected
LivingSpace	1.5 dropped	1.3 dropped	1.4 dropped	1.1 dropped	47.1 selected	4.0 dropped	5.0 dropped

Source: own calculations
The left-hand-side variable in all regressions is the price of a standardized and comparable single family house from 1998 to 2004. Iterated BMA works by making repeated calls to the Bayesian model averaging procedure, iterating through the variables. After each call to the Bayesian model averaging procedure only those variables which have posterior probability greater than 5 %. The maximum number of variables entering each BMA iteration is 15. Each column gives the maximum posterior inclusion probability and whether the variable was dropped or not (note that a variable can have a maximum 99.9 % inclusion probability but in one estimation an inclusion probability of less than 5 % and consequently the variable is dropped).

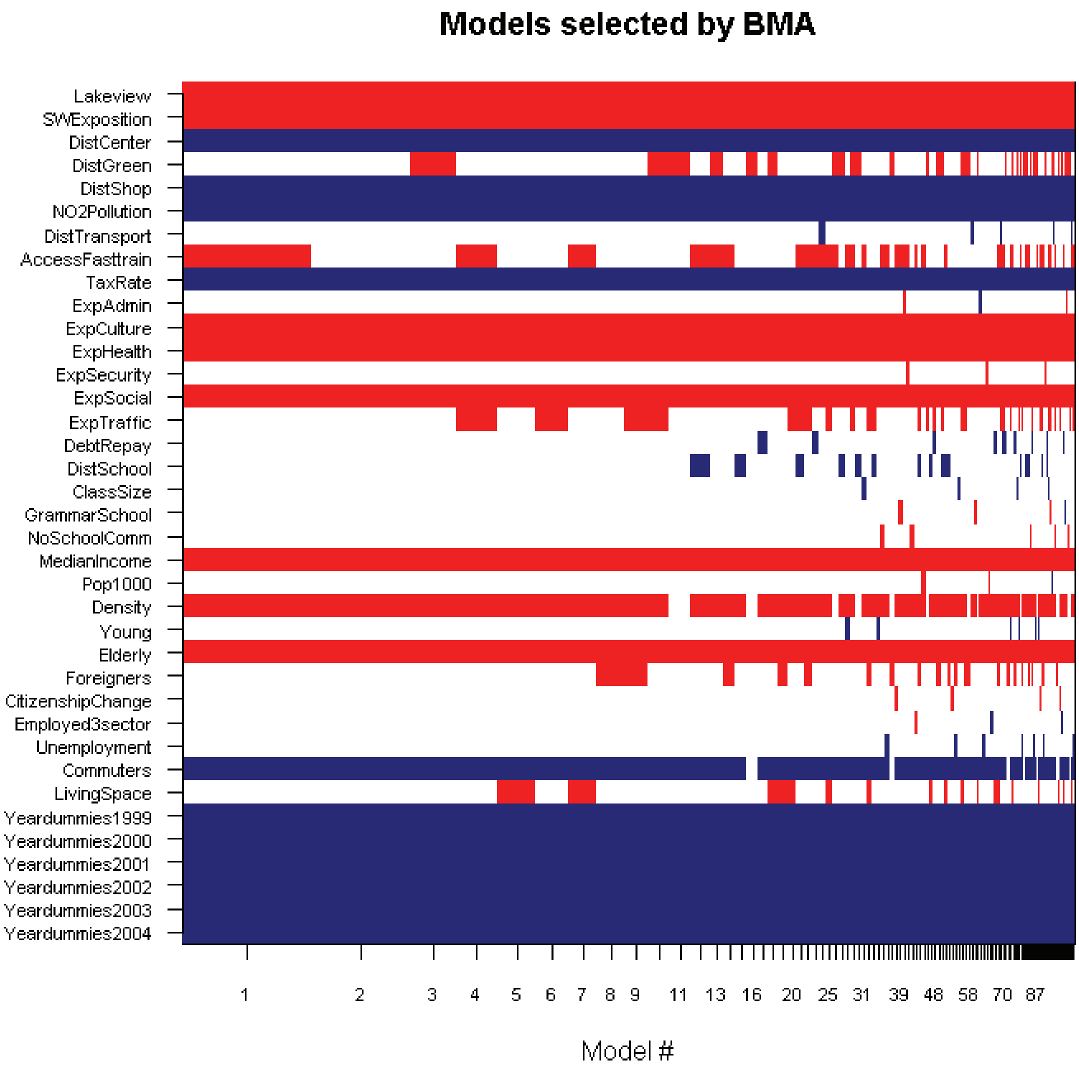
Figure 1
Posterior Distribution of Coefficients of Selected Variables





Source: own representation.
The figure shows the density function approximated by a finite mixture of normal densities of the coefficient estimates of the most common variables used in capitalization studies. The height of the density curve equals the posterior probability that the variable is included in the model. The discrete mass at density zero gives that probability that the variable is not included in the model.

Figure 2
BMA Best Model Selection for House Prices



Source: own representation.
Each row corresponds to one variable. The columns represent different models. The width of the columns corresponds proportionally to the model's posterior probability. Negative coefficients are highlighted with a darker color and positive ones with a lighter one. White spaces indicate that the variable was not included in the model considered.

Supplementary Material

(not intended for publication)

All results presented in the paper use a linear form. As the paper focuses on variable selection not every possible form can be tested. Table A1 gives additional robustness tests for a semi-logarithmic form, i.e. house prices are expressed in logs, with iterated BMA for variable selection as done in Table 4 of the article. There are only minor changes as far as variable selection is concerned (changes are highlighted).

Table A1

Iterated BMA for Variable Selection with Semi-logarithmic and Logarithmic Form

Variable	*10 variables to select (1)*	*Number of times in model during iteration for 10 variables (2)*	*15 variables to select (3)*	*Number of times in model during iteration for 15 variables (4)*	*20 variables to select (5)*	*Number of times in model during iteration for 20 variables (6)*
Intercept	selected	22	selected	17	selected	4
Lakeview	selected	22	selected	17	selected	4
SWExposition	selected	22	selected	17	selected	4
DistCenter	selected	22	selected	17	selected	4
DistGreen	dropped	13	selected	17	selected	4
DistShop	selected	22	selected	17	selected	4
NO2Pollution	selected	22	selected	17	selected	4
DistTransport	dropped	2	dropped	2	dropped	2
AccessFasttrain	selected	22	dropped	14	selected	4
TaxRate	selected	22	selected	17	selected	4
ExpAdmin	dropped	1	dropped	7	selected	4
ExpCulture	selected	22	selected	17	selected	4
ExpHealth	selected	21	selected	17	selected	4
ExpSecurity	dropped	1	dropped	1	dropped	1
ExpSocial	dropped	1	selected	17	selected	4
ExpTraffic	dropped	1	dropped	10	dropped	2
DebtRepay	dropped	1	dropped	6	dropped	1
DistSchool	dropped	1	selected	16	selected	4
ClassSize	dropped	1	dropped	1	dropped	1
GrammarSchool	dropped	1	dropped	1	dropped	1
NoSchoolComm	dropped	1	dropped	1	dropped	1
MedianIncome	dropped	1	selected	12	selected	4
Pop1000	dropped	1	dropped	3	dropped	2
Density	selected	10	dropped	1	dropped	1
Young	dropped	1	dropped	7	selected	3
Elderly	dropped	1	selected	8	selected	3
Foreigners	dropped	1	dropped	1	dropped	1
CitizenshipChang	dropped	1	dropped	1	dropped	1
Employed3sector	dropped	1	dropped	1	dropped	1
Unemployment	dropped	1	dropped	4	dropped	2
Commuters	dropped	1	selected	3	selected	2

LivingSpace	dropped	1	dropped	1	dropped	1
YearDummies	dropped	1	selected	1	selected	1

Source: own calculations

The left-hand-side variable in all regressions is the logarithm price of a standardized and comparable single family house from 1998 to 2004 across 168 municipalities. The explanatory variables are as in Table 4. The value in parenthesis indicates the number of times the variable was used in the model during iteration. Iterated BMA works by making repeated calls to the Bayesian model averaging procedure, iterating through the variables. After each call to the Bayesian model averaging procedure only those variables which have posterior probability greater than 5 %. The maximum number of variables entering each BMA estimation is 10, 15 and 20 for columns. The constant is always included.