

Schulz, Rainer; Wersing, Martin; Werwatz, Axel

Working Paper

Automated valuation modelling: A specification exercise

SFB 649 Discussion Paper, No. 2013-046

Provided in Cooperation with:

Collaborative Research Center 649: Economic Risk, Humboldt University Berlin

Suggested Citation: Schulz, Rainer; Wersing, Martin; Werwatz, Axel (2013) : Automated valuation modelling: A specification exercise, SFB 649 Discussion Paper, No. 2013-046, Humboldt University of Berlin, Collaborative Research Center 649 - Economic Risk, Berlin

This Version is available at:

<https://hdl.handle.net/10419/91574>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

SFB 649 Discussion Paper 2013-046

Automated Valuation Modelling: A Specification Exercise

Rainer Schulz*
Martin Wersing*
Axel Werwatz**



*University of Aberdeen Business School, United Kingdom
**Technische Universität Berlin, Germany

This research was supported by the Deutsche
Forschungsgemeinschaft through the SFB 649 "Economic Risk".

<http://sfb649.wiwi.hu-berlin.de>
ISSN 1860-5664

SFB 649, Humboldt-Universität zu Berlin
Spandauer Straße 1, D-10178 Berlin



SFB 649 ECONOMIC RISK BERLIN

Automated Valuation Modelling: A Specification Exercise

Rainer Schulz, Martin Wersing, and Axel Werwatz*

September 2013

*Schulz and Wersing: University of Aberdeen Business School, Edward Wright Building, Dunbar Street, Aberdeen AB24 3QY, United Kingdom. Emails: r.schulz@abdn.ac.uk and martin.wersing@abdn.ac.uk. Werwatz: Technische Universität Berlin, Institut für Volkswirtschaftslehre und Wirtschaftsrecht, Straße des 17. Juni 135, 10623 Berlin, Germany, and Collaborative Research Center 649 *Economic Risk*, Humboldt-Universität zu Berlin. Email: axel.werwatz@tu-berlin.de.

Abstract

Market value predictions for residential properties are important for investment decisions and the risk management of households, banks, and real estate developers. The increased access to market data has spurred the development and application of Automated Valuation Models (AVMs), which can provide appraisals at low cost. We discuss the stages involved when developing an AVM. By reflecting on our experience with md*immo, an AVM from Berlin, Germany, our paper contributes to an area that has not received much attention in the academic literature. In addition to discussing the main stages of AVM development, we examine empirically the statistical model development and validation step. We find that automated outlier removal is important and that a log model performs best, but only if it accounts for the retransformation problem and heteroscedasticity.

Keywords: Hedonic regression, log transformation, predictive performance

1 Introduction

The market value of a residential property is the price one should expect in an arm's-length transaction between informed and willing buyers and sellers. This value depends on the property's structural and location characteristics, some of these can be observed easily, but others will require a full physical inspection. Professional valuers predict the market value of a property by taking all characteristics into account. Such a full appraisal should provide an accurate prediction of the market value, but is also time consuming and expensive. Automated Valuation Models (AVMs) use recorded transaction or listing information to fit a statistical model. Once the model is fitted, the market value of any common type of residential property can be predicted. Such automated appraisals might be less accurate than full appraisals, but are also less costly.

Market participants will be prepared to trade off predictive accuracy for cost in several applications. First, banks can use low cost appraisals from AVM services when underwriting further loan advances, home equity withdrawals, and remortgaging (Fitch Ratings Structured Finance, 2012). Bank risk managers may use an AVM as a cost-effective tool to monitor the collateral values underlying the bank's portfolio of mortgage loans. Such monitoring can be required by banking regulation. Second, AVM appraisals are also used when mortgage loans are pooled and securitised. Rating agencies request information about current loan-to-value ratios, as they relate to default probabilities and loss severity given default, and AVM appraisals contribute to such information (Moody's Investors Service, 2009, 2012). Third, aspiring property owners might want to get a feel for the expected price they have to pay for their dream house. Existing property owners might be interested in a snapshot appraisal, maybe to inform the decision to relocate. An appraisal from an AVM service will fit the bill. The availability of low cost appraisals should, therefore, increase the transparency of residential property markets. Fourth, government agencies might be interested in cost-effective appraisals for taxation, planning, and land use regulation (IAAO, 2003).

In Europe, several providers offer AVM services. Examples are Calnea, Hometrack, and RightmoveData in the UK, Calcasa in the Netherlands, and ImmobilienScout24 and md*immo in Germany. As most of these services are proprietary businesses, only limited information can be obtained about how a specific AVM is implemented. At the same time, there is no scarcity of academic papers examining house prices with

statistical models. This literature mostly focusses on the construction of house price indices or the estimation of marginal valuations of individual property characteristics (Hill et al., 1997; Palmquist, 1980, 1991). Few papers have examined the performance of statistical models when predicting property prices out-of-sample (Anglin and Gençay, 1996; Thibodeau, 2003; Bin, 2004; Case et al., 2004). These papers work with log prices and ignore the retransformation problem. Only Gençay and Yang (1996) examine predictive performance using property prices. Papers examining out-of-sample performance focus on finding the best statistical model for a given data set, but do not examine if and how the model found could be implemented as an AVM.

In our paper, we discuss the stages that are required to implement an AVM. At each stage, pragmatic choices have to be made. We exemplify this using `md*immo`, a not-for-profit AVM service for Berlin, with which we have been involved from its inception in 2002 (the service went online in 2006). Our paper is informed by the academic literature and also the pragmatic trade-offs inherent in developing and implementing an AVM. By reflecting on our experience, we contribute to an area that has not received much attention in the academic literature.

We focus, in particular, on the model development and validation stage of an AVM. We conduct the empirical analysis using single-family house transactions from Berlin and split the data into a sub-sample for model development and a sub-sample for model validation.

In the model development step, we specify the market value with a flexible linear parametric function. We consider regression specifications with either the price or the log price as dependent variable. We also examine the optimal length of the estimation window over which a model should be fitted. For the log model, we use different approaches to deal with the retransformation problem. In line with the practical purpose of an AVM, we are not interested in specifying regression models with good in-sample fit, but rather models with good out-of-sample performance. We, therefore, specify the different possible models using a measure related to out-of-sample predictive accuracy.

In the validation step, we assess the predictive performance of the different model specifications using Monte Carlo simulation (Haupt et al., 2010). The simulation results in a distribution of performance measures for many possible splits of the validation sample into sub-samples used for model fitting and sub-samples used for performance validation. The performance measures are those commonly used: mean error, median

error, mean absolute error, and mean squared error. These measures summarize individual appraisal errors, which measure deviations of actual transaction prices from appraisals predicted with a statistical model.

We obtain the following results. First, data cleaning is an important step of the implementation of an AVM. We find that outliers in the development sample can cause large appraisal errors in the validation sample. We are, therefore, quite generous with removing aberrant observations, which is defensible when developing an AVM. The aim of an AVM service is to provide good appraisals for ‘average’ properties. It is likely that properties with unusual observed characteristics are unusual in other aspects too. Such properties should not be appraised with an AVM, but require a physical inspection by an expert valuer.

Second, specifying a linear statistical model with the log price as dependent variable has the advantage that this partially corrects the inherent heteroscedasticity of property prices. However, it requires that log predictions are adjusted for the retransformation problem. Without such an adjustment, appraisals underestimate systematically. We examine two different adjustments and find that both improve the performance of the appraisals. The best appraisal performance results when the retransformation also considers the observed characteristics of the house for which the market value is predicted. Specifying the model with the price as dependent variable has the advantage that no retransformation adjustment is required. Individual appraisals from this specification can perform very poorly, however, and can lead to very high mean squared errors.

Third, we find that the estimation window for our preferred model should cover three years of data. A shorter estimation window would allow the coefficients of the statistical model to be more flexible, but fewer observations are available for fitting the model. A longer estimation window provides more observations, but assuming that the coefficients are constant over the longer period is too restrictive. Appraisals from a statistical model fitted to log prices and log predictions corrected with the adjustment suggested by Garderen (2001) produce the best out-of-sample performance.

We do not claim that our procedure of implementing an AVM is the only way of doing it. As our experience with the Berlin data shows, one has to be pragmatic when implementing an AVM. The existing literature has focussed exclusively on finding the best model for a given data set without asking if the model is practical for an AVM.

Practicality is an important aspect of an AVM and should not be ignored when

developing statistical models for market value predictions. We focus on parametric statistical models to keep the computing time during model development and validation at a reasonable level. A parametric model is also convenient in the implementation stage. Once characteristics of a subject property are provided, the computation of an appraisal is instant. Backtesting should become easier too.

The remainder of this paper is organized as follows. Section 2 discusses the different stages when developing an AVM. Section 3 presents the data we use for the model specification exercise. Section 4 explains the model development. Section 5 presents the performance assessment obtained from the validation step. Section 6 concludes. Details of the analysis are relegated to the Appendix.

2 Stages of AVM development

Developing and running an AVM service involves the following stages:

1. Establishing continuous access to reliable data
2. Model development and validation
3. Roll-out and service provision
4. Backtesting

The first stage of any AVM is establishing continuous access to data. In most cases, the data will be collected for purposes other than the AVM. Examples of such data include listing information from property websites, recorded transaction data from land title registers, data from syndicates of local solicitors, or information from banks acquired during the mortgage underwriting process. Most of this data is itself proprietary, and the owner of the data might be interested in setting up an AVM on its own. In addition to establishing access to data sources, it must be ensured that the data provision is reliable and the data current. There exists a clear trade-off between listing information and transaction data. The former are current, but are only asking prices, whereas the latter are the best indicator of market values, but data might become available, if at all, only with a delay. Depending on the intended coverage of an AVM, it might be required to establish contacts with several local data providers.

The second stage of an AVM starts with data cleaning procedures, which should become as automated as possible. Data cleaning is followed by the selection of variables

that are observed and relevant for the market value of a property. This requires full understanding of the respective market and knowledge of the data and their definitions. Variable selection can be based on statistical significance levels also. In such a case, the variable selection step and the model specification step overlap. In the model specification step, the suitable functional form for the market value function has to be established. Semi-parametric and spatial models provide much flexibility at this step, but have the disadvantage that appraisals are more complicated to compute.¹ For instance, a nonparametric function allows the estimation of a location value surface with great flexibility, but computing the location value for a requested prediction will then be either time consuming or reliant on interpolation. Market value prediction with an estimated parametric model is straightforward, because the functional form provides this interpolation. Once a suitable model (or a set of suitable models) has been established, the model has to be validated out-of-sample. This corresponds to a dry-run of the AVM before the roll-out. The validation step also helps to discriminate between models if several seem suitable during model development.

The third stage of an AVM relates to the technical implementation of the service. Often, the appraisals should be provided in real time on desktops of a institution and an efficient technical implementation is important for this purpose. The technical implementation becomes more complicated if the appraisals should be available online. Depending on the experience and knowledge of the prospective users, pragmatic choices have to be made regarding the information that can be requested when using the service. For instance, homeowners will know the street address of their house and also some structural conditions, but understanding categories of the state of repair might be already too complicated.² This also applies to the rendered appraisal information itself. The average user might not understand what a confidence interval is and clever ways have to be found to provide this information in an intuitive manner. For instance, AVM services could give a confidence ranking for each appraisal, which maps the standard error of the appraisal onto an ordinal scale (Moody's Investors Service, 2008).

¹For such models, see Yatchew (2003) and LeSage and Pace (2009).

²Such variables will be considered in the statistical model and predictions will be based on the most common category. The AVM could have a request mode for expert users, who understand the categories.

The fourth stage consists of backtesting the AVM once it is rolled out. This will be done by the service provider itself, but also by users of the appraisals, such as rating agencies. Backtesting implies that any remaining structure in appraisal errors should be detected and used to improve the statistical model. For instance, if appraisal errors in one local area have the tendency to be positive, then consideration should be given to allowing more flexibility in the statistical model or to fitting a separate model for this area.

3 Data

For the empirical analysis, we use single-family transactions from Berlin over the period 2000 to 2011. Observations from 2000-2005 are used for model specification (development step) and observations from 2006-2011 are used for model fitting and prediction (validation step). These two steps of the empirical analysis are detailed in Sections 4 and 5, respectively.

The data is from the transaction data base of Berlin's Committee of Valuation Experts (GAA, Gutachterausschuss für Grundstückswerte). The GAA is obliged and authorized by law to collect information on all real estate transactions occurring in Berlin. The GAA transaction data base is therefore an example of market information that is collected continuously and reliably, but not with the initial purpose of setting up an AVM. The data is, however, ideally suited for this purpose.

Table 1 contains summary statistics for the variables in the cleaned data set.

[Table 1 about here.]

To indicate the cross-sectional variation of house prices, we present the statistics for real prices, i.e., prices deflated with Berlin's CPI. In the empirical analysis below, we work with nominal prices and use time dummy coefficients to control explicitly for any variation in the general market trend. In addition to information on the transaction price and several structural characteristics of the buildings, we also know in which of Berlin's 96 sub-districts the house is located. We also observe an expert-based rating for each location. This ordinal rating is provided by Berlin's Senate Department for Urban Development and the Environment and uses four levels to summarize the quality of a particular street block. The experts consider the amount of natural amenities such as lakes and forests, the quality of existing buildings, and the access to public transport

and shopping facilities. Unusual features of the houses in Table 1 include physical aspects such as structural damage or flooding risk, and legal aspects such as rights of way or use for pipes or cables. Such easements are rather common. We also know about the structural conditions as assessed by a building surveyor and we have some information about the interior layout (attic and type of cellar). We do not observe, however, the quality of the fittings and other characteristics that can only be assessed during a physical inspection of the property.

To obtain the clean data set summarized in Table 1, we apply the following steps. First, we exclude all observations provided by the GAA that have misreported or missing values for relevant variables. This requires that values of individual variables are examined and understood. It also requires that we have a notion about which variables will be relevant. The first step of data cleaning is time consuming and cannot be automated. It results in a data set consisting of 19,553 arms-length transactions of single-family houses. As the boxplots in Figure 1 show, all continuous variables show substantial variation. There seems to be a number of observations with unusual (outlying) values, such as particularly large or old houses.

[Figure 1 about here.]

It could be that some or all of these unusual observations are perfectly normal once examined in detail. However, such examination would require an inspection and costly analysis. An AVM is intended for cost-effective appraisals of average properties, and it is, therefore, appropriate to use a mechanical and automated criterion to remove unusual observations.

In the second step of data cleaning, we use the robust Mahalanobis distance

$$d = \sqrt{(\mathbf{x}_i - \hat{\boldsymbol{\mu}})\hat{\boldsymbol{\Sigma}}^{-1}(\mathbf{x}_i - \hat{\boldsymbol{\mu}})'} \quad (1)$$

to detect outliers. The (1×3) vector \mathbf{x}_i contains the age, floor area, and lot area of a house. The (1×3) vector $\hat{\boldsymbol{\mu}}$ and the (3×3) matrix $\hat{\boldsymbol{\Sigma}}$ give the means of the characteristics and their covariance matrix estimated with all observations used in the development step (i.e., observations from the period 2000-2005), but using the robust estimators proposed by Rousseeuw (1985). This prevents that the outlier detection measure d is affected by outliers. Assuming that the three continuous characteristics in the population follow a normal distribution, d is the root of the sum of squared standard

normal variables and will be $\sqrt{\chi_3^2}$ distributed. We then have to decide on a confidence level. For the analysis to follow, we choose a confidence level of 99 percent (critical value 3.4) and remove all observations n for which $d_n > 3.4$. If the inequality holds, house n 's distance from the average characteristics is unusual under the null, because the probability for observing this distance is less than 1 percent. The observation is, removed.

To assess the role of the confidence level, we also conduct the whole analysis with the stricter confidence level of 95 percent (critical value 2.8) and without any outlier removal (setting the confidence level to 100 percent). The stricter confidence level leaves the qualitative results of our analysis unchanged. Without any automated outlier removal, however, the performance of the appraisals worsens substantially. Robustness analysis of this kind is important, because we want to find the most sensible cut-off point. We want to remove observations that are too far away from the average combination of characteristics, but we do not want to lose too many observations.

The clean data set has 18,444 observations, of which 8,429 are used in the development step and 10,015 in the validation step. The boxplots in Figure 2 show that the variation in the transaction price and the continuous house characteristics is reduced in the cleaned data set.

[Figure 2 about here.]

As Table 1 indicates, this still translates into substantial cross-sectional variation of real house prices and house characteristics, as one expects from a market with heterogeneous goods. The average characteristics between observations used in the development step and the validation step are nearly identical, which supports the notion that information on a cross-section of past transactions should be helpful in modelling prices in a current cross-section.

4 Model development

We start with the assumption that the price of a property

$$P = V(\mathbf{x})U \tag{2}$$

depends on property's market value $V(\mathbf{x})$ and transaction noise $U > 0$. The market value is a function of property's characteristics, which are collected in the vector \mathbf{x} . Full

knowledge of \mathbf{x} will require a physical inspection of the property. When implementing an AVM, only a subset of \mathbf{x} will be available. We do not distinguish this subset to keep the notation simple, but come back to this point below.

The transaction noise considers that in a particular deal either the buyer or the seller could be better informed and exploit this advantage. We assume, however, that the noise is independent of property’s characteristics and $E[U] = 1$. This implies that knowing \mathbf{x} *ex ante* does not help to predict the realization of U . It follows that the market value is the price we expect in an arm’s-length transaction between informed and willing parties.

An appraisal is then a prediction of $V(\mathbf{x})$ given a property’s characteristics. AVM providers usually use one of three methods to predict the market value: *i*) indexation, *ii*) weighted comparable sales, and *iii*) hedonic regression (Downie and Robson, 2007; Moody’s Investors Service, 2008; Nattagh, 2007).

Indexation uses the last existing full appraisal of the subject house and rolls it forward with a quality-controlled house price index. This requires that a previous appraisal exists and that cross-sectional variation in the housing stock and the valuation of individual characteristics can be neglected. There exists a large literature on house price index construction, including papers that examine the out-of-sample performance and papers that examine the adjustments that are necessary when log prices are used for constructing the index (Goh et al., 2012; Goetzmann, 1992; Goetzmann and Peng, 2002). We are not aware of any study that examined the performance of full appraisal indexation in an out-of-sample prediction exercise.³

Weighted comparable sales mimic the sales comparison approach, whereby recent transactions of similar houses are used to value the subject property. A very simple implementation would be to compute the average price of houses located nearby that happen to be transacted recently. More complicated weighting is conceivable.

Hedonic regression fits observed transaction prices on house characteristics. Once fitted, the estimated hedonic function can be used to predict a property’s market value given its characteristics. Effectively, hedonic regression weighs recent sales with respects to the similarity of the subject house and the transacted house.⁴ Hedonic regression is

³We cannot implement such a exercise, as we do not observe full appraisals for the transactions in our data set.

⁴With the vector of prices \mathbf{p} , the matrix \mathbf{X} of characteristics, and the characteristics \mathbf{x}_0 of the subject house, a fitted linear hedonic model predicts $\hat{p}_0 = \mathbf{x}_0\hat{\boldsymbol{\beta}} = \mathbf{x}_0(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{p} = \mathbf{w}_0\mathbf{p}$. Element w_n in the

therefore a variant of the sales comparison approach.

In our specification exercise, we consider two additive specifications of a hedonic regression model. For the first, we take the log of Eq. 2

$$\begin{aligned}\ln P &= \ln V(\mathbf{x}) + E[\ln U] + (\ln U - E[\ln U]) \\ p &= w(\mathbf{x}) + u ,\end{aligned}\tag{3}$$

where $w(\mathbf{x}) \equiv \ln V(\mathbf{x}) + E[\ln U]$ and $u \equiv \ln U - E[\ln U]$. It follows for the noise term that $E[u] = 0$. Eq. 3 states that the log price of a property equals the log market value function $w(\mathbf{x})$ plus some noise. On average, we expected the log price to be equal to the log market value. One could assume that all efforts should focus on finding $w(\mathbf{x})$. However, this assumption ignores the retransformation problem. Even if we knew $w(\mathbf{x})$ (and did not have to estimate it), market value predictions would be biased,

$$\exp\{w(\mathbf{x})\} = V(\mathbf{x}) \exp\{E[\ln U]\} < V(\mathbf{x}) ,\tag{4}$$

because $E[\ln U] < \ln E[U] = 0$ (Jensen's inequality). Using the first additive specification requires therefore that we either ignore this bias or that we adjust for it. We will consider both possibilities in the validation step.

For the second additive specification, we reformulate Eq. 2

$$\begin{aligned}P &= V(\mathbf{x}) + V(\mathbf{x})(U - 1) \\ P &= V(\mathbf{x}) + e\end{aligned}\tag{5}$$

with $E[e] = 0$ and $\text{Var}[e] = V(\mathbf{x})^2 \sigma_U^2$. This specification has the untransformed price as dependent variable and does not suffer from a retransformation problem. The regression will suffer, however, from heteroscedasticity. A correctly specified model will be consistent, but the estimator will not be efficient.

Economic theory does not provide much guidance about a particular form for the market value function. In the academic literature, the (log) market value function has been specified as parametric model (Palmquist, 1980; Case and Quigley, 1991; Hill et al., 1997; Thibodeau, 2003) and as semiparametric model (Anglin and Gencay, 1996; Clapp, 2004). Semiparametric models provide flexibility, but the technical implementation in an AVM will be complicated.

weight vector \mathbf{w}_0 is the larger (smaller) the more (less) similar the characteristics of the transacted house n and the subject house are. The importance of the observed price p_n for the prediction of \hat{p}_0 is proportional to w_n .

In our hedonic regressions, we use the flexible parametric model of Bunke et al. (1999)

$$y = \beta_0 + \sum_{c=1}^C \beta_c T_\lambda(x_c) + \sum_{c=1}^C \sum_{\tilde{c}=1}^C \beta_{c\tilde{c}} T_\lambda(x_c) T_\lambda(x_{\tilde{c}}) + \sum_{d=1}^D \gamma_d D_d + \varepsilon. \quad (6)$$

The dependent variable y is either the price or the log price of a house, x_c is a continuous characteristic (floor area, lot area, age), $T_\lambda(\cdot)$ is a Box-Cox type transformation function depending on the parameter λ , and D is an indicator for discrete characteristics, such as number of storeys. We also include time and sub-district dummies. The hedonic regression is linear in the coefficients β and γ , making market value predictions easy to compute once Eq. 6 has been estimated. The model is still flexible regarding the continuous explanatory variables and nests linear models that are commonly used. Appendix A.1 provides details on the transformation function.

In addition to finding the best transformations for the three continuous explanatory variables, we also want to establish the optimal length of the estimation window. The longer the estimation window, the more observations are available and the more precise the coefficient estimates should become. Coefficients might not be constant, however, in a growing housing market with an influx of buyers who have heterogenous tastes and a supply of new houses with contemporary specifications. In this case, a short estimation window will be advantageous.

We choose the best model and estimation window length simultaneously. To do so, we fit Eq. 6 to the development sample of six years, where the estimation windows cover either one, two, three or six years. This implies that the specification with the yearly estimation window is fitted six times, providing much flexibility regarding any variation in the coefficients. At the other extreme, the specification is fitted only once, covering the whole six years. If the coefficients are constant over time, then the long estimation window is better, if they vary a lot, then the yearly estimation window should be better. To assess the fit of the different transformation and estimation window combinations, we use the standardized cross-validation criterion (CVS). This criterion is similar to the coefficient of determination, R^2 , but uses errors from predictions that use all observation except the one that should be predicted, see Eq. A3 in the Appendix. This measure focusses, therefore, on the out-of-sample fit (Myers, 1989, 4.2).

[Table 2 about here.]

Table 2 reports for each of the four possible estimation windows the *CVS* and the best model. We see that the best model for log prices favours an estimation window of three years and the transformation parameters $\lambda^* = (0.5, 0.5, -2)$ for the floor area, lot area, and age, respectively. We also see that performance is very similar if the model is fitted only once over the development sample. If the price is the dependent variable, then the long estimation window is best. It also appears that λ^* varies more between the different estimation windows. In hedonic regression studies, common transformations for all explanatory variable are the linear and the log form, which correspond to $\lambda = (1, 1, 1)$ and $\lambda = (0, 0, 0)$. For our data, these are never the best transformations. The results in the two panels of Table 2 cannot be compared directly, because the scales of the dependent variables are different. We will compare the performance of the two best models in the validation step.

To assess if the estimates of the individual coefficients are also ‘plausible’, Table 3 reports OLS estimates of Eq.6 fitted to the second three-year estimation sample with the log price as dependent variable.

[Table 3 about here.]

The overall fit of the regression, as measured by the R^2 and the *CVS*, is good by usual standards. The estimated coefficients have reasonable signs and are mostly statistically significant at the usual levels. The interpretation of the estimated coefficients on the discrete indicator variables is straightforward. For instance, semi-detached (row) houses sell with a rebate of about 4.5 (6.5) percent compared to detached houses. In order to interpret the effect of the continuous characteristics on house prices, we need to evaluate Eq. 6 at all values of the characteristics of interest (keeping all other characteristics constant). The estimated coefficients for the transformed age variable, for instance, implies a decreasing rate of depreciation for older buildings. This is sensible in market with vintage houses of relatively high quality as in Berlin. The floor area and lot area, on the other hand, have an overall positive impact on the price of a house.

The estimated coefficients may increase confidence in a statistical model, even though we hold an agnostic view. First, the values of the estimated coefficients represent the interaction of supply and demand and have only a clear economic interpretation if one makes strong assumptions (for instance, fixed supply of characteristics). Second, we *know* that we do not observe all characteristics that are relevant for the market

value. Only a full inspection can reveal all relevant characteristics, but this is not the type of data used in an AVM. The estimates of the coefficients, therefore, could be biased due to omitted variables. However, when predicting market values, this bias will even out on average. Given the observed house characteristics, our model provides the best linear predictor under squared error loss (Cameron and Trivedi, 2005, 4.2.3). Predictive performance is all we are interested in when developing and implementing an AVM.

5 Validation

We validate and compare the performance of the different specified models with a Monte Carlo simulation. We randomly split the validation sample into a sub-sample with 80 percent of the observations. This sub-sample is used to fit the hedonic regression in Eq. 6 as specified in the development step. We fit the model for the best log price specification and the best price specification. If the model is fitted to log prices, we adjust the log prediction with one of two adjustment factors to deal with the retransformation problem. For comparison, we also predict market values without retransformation adjustment. Appendix A.2 explains the market value predictors (MVPs), three of which are based on the log price regression and one of which is based on the price regression. The MVPs provide appraisals for remaining 20 percent of the observations in the validation sample. We replicate this exercise 10,000 times. The Monte Carlo simulation prevents that our results are driven by a specific split of the validation sample and provide a fair assessment of the out-of-sample performance of the different MVPs.

To assess the performance, we compute for each replication r of the Monte Carlo simulation the appraisal errors

$$e_{n,r} = \frac{P_{n,r} - \text{MVP}_{n,r}}{\text{MVP}_{n,r}},$$

where n is an observation in the 20 percent validation sub-sample. This validation sub-sample has $N = 2003$ observations. $P_{n,r}$ is the actual transaction price of house n in the validation sub-sample. $\text{MVP}_{n,r}$ is one of the four market value predictors, evaluated at the characteristics $\mathbf{x}_{n,r}$. Negative errors imply that the appraisal value is larger than the actual transaction price, while positive errors imply the appraisal is

smaller than the price. An appraisal error of zero implies that the appraisal predicts the price perfectly. However, because of the transaction noise, prices are themselves only indicators of market values. Even knowing the market value function $V(\mathbf{x})$ would not prevent that non-zero valuation errors occur. But it is also clear that knowledge of $V(\mathbf{x})$ would produce the ‘smallest’ appraisal errors. For instance, knowing the actual market value function should lead to an average appraisal error of zero when many house prices are predicted. Given our different MVPs, we are looking, therefore, for the predictor that shows good performance when summary statistics of appraisal errors are compared.

We use the following summary performance measures. The mean prediction error (MPE)

$$\text{MPE}_r = \frac{1}{N} \sum_{n=1}^N e_{n,r} ,$$

which is the arithmetic average over all appraisal errors. The median prediction error (MDPE)

$$\text{MDPE}_r = \text{Med}\{e_{n,r}\}_{n=1}^N ,$$

which is the error in the middle of the appraisal error distribution. A MPE and MDPE of zero indicate that the respective MVP is correct, on average, and an unbiased estimator of the market value function. The MPE and the MDPE do not take the dispersion of the appraisal errors into account, however. A particular MVP could have a MPE (MDPE) close to zero, but could also exhibit large individual errors that lie far away from zero. The mean absolute prediction error (MAPE)

$$\text{MAPE}_r = \frac{1}{N} \sum_{n=1}^N |e_{n,r}|$$

and the mean squared prediction error (MSPE)

$$\text{MSPE}_r = \frac{1}{N} \sum_{n=1}^N e_{n,r}^2$$

take dispersion into account. Both are symmetric measures and weigh positive and negative errors of the same absolute magnitude equally.

Depending on the application for which AVM appraisals are requested, some of the above statistics will have more relevance than others when discriminating between different MVPs. For instance, if the appraisals should inform purchase decisions, then the MPE should be zero and the MSPE should be as small as possible. If the AVM is

used for risk management purposes, such as the evaluation of loss severity for a portfolio of mortgages, banks are likely to be concerned about large overvaluations and may even favour a MVP that has the tendency to err on the cautious side. It follows that, whereas the statistical assessment of different MVPs is clear, an economic assessment can be complicated because it requires an economic loss function for over-and undervaluations (Shiller and Weiss, 1999).

[Figures 3–6 about here.]

Figures 3–6 show the simulation-based cumulative distribution functions (CDF) of the four performance measures. The CDFs are for the three MVPs based on log price specification, but only MVP2 and MVP3 adjust for the retransformation problem. The vertical lines to the left and right of each distribution indicate the smallest and largest performance measure obtained for each predictor. Summary statistics are given in the first three rows of Table 4.

[Table 4 about here.]

The CDF of the MPEs in Figure 3 show that MVP1 is systematically downward biased. On average, the bias is about 2.8 percent of the appraisal, see Table 4. MVP2 and MVP3 consider the retransformation problem and thus perform better. A judgment with respect to the MDPEs is not as clear cut. While MVP1 gives slightly too low predictions of median prices, MVP1 and MVP2 overstate the median by 1.9 to 2.0 percent, see Table 4. Turning to the simulated distributions of the MAPE and MSPE, we find that the two adjusted MVPs always lie to the left of the unadjusted predictor. Thus, MVP2 and MVP3 first-order stochastically dominate the unadjusted MVP1. In summary, the results show that an unadjusted retransformation of predicted log price not only introduces a systematic downward bias, but also increases the dispersion of prediction errors. We favour MVP3, as it is slightly better than the appraisals using MVP2. The reason for this good performance is that MVP3's retransformation adjustment depends on the specific house characteristics. It therefore takes any remaining heteroscedasticity into account.

Figures 7-10 show the simulation-based CDFs of the four performance measures for MVP3 and MVP4. The latter MVP is calculated from a regression with the price as the dependent variable. Summary statistics for MVP4 are given in the last row of Table 4.

[Figures 7–10 about here.]

The CDFs clearly show that MVP4 performs very poorly in some instances. While for the majority of replications the performance measures are of a similar magnitude as for MVP3, appraisals from MVP4 can be very large or small (and even negative). Even though MVP4 avoids the retransformation problem, because it is based on a hedonic regression with the price as dependent variable, it still performs worse than MVPs based on log predictions.

6 Conclusion

Predicting market values of residential properties is important for investment decisions and the risk management of households, banks, and real estate developers. The increased access to electronically stored information has spurred the development and implementation of AVMs, which provide appraisals at low cost. The key challenge of AVM development is finding a workable statistical approach that uses the data and provides reasonably reliable appraisals. A large body of academic literature has developed concerning the statistical modelling of real estate prices. This literature offers a variety of approaches, but does not aim at prediction or, if it does, eschews the constraints and trade-offs faced by AVM developers.

In this paper, we build on our experience of statistical model building and implementation of an AVM. We discuss the stages that are necessary to implement an AVM and present and compare particular solutions to the challenges and trade-offs appearing at each stage. Throughout the paper, we maintain the perspective demanded by an AVM's aim: providing good appraisals for average properties based on a manageable, but sufficiently sophisticated, statistical model.

The results of our analysis can be summarized as follows. First, outliers can cause large appraisal errors. It is therefore imperative to remove outliers during the model development step. We use the robust version of the Mahalanobis distance for this purpose. The distance identifies properties with characteristics that are too far away from the center of the data.

Second, the specification of the statistical model requires a compromise between bias-reducing flexibility and variance-reducing parsimony. Flexibility would be ensured by using only very recent observations and by employing a highly data-adaptive mod-

elling approach. This risks, however, that the developed model could perform poorly out-of-sample. We approach the challenge by using cross validation during the model development step. This ensures that the specification is already based on out-of-sample performance. We use a flexible parametric regression model that is linear in the coefficients and allows a finite set of nonlinear Box-Cox-Type transformations for continuous regressors. The model can be estimated easily and can be implemented straightforwardly for the provision of predictions.

Third, regarding the proper dependent variable, AVM developers have to trade-off the statistical advantages of log-prices (which strongly dampen the heteroscedasticity present in untransformed prices) with the convenience of working directly with untransformed prices (thus avoiding the retransformation problem). In our Monte Carlo simulation, we find that models with the price as dependent variable can lead to very high mean squared prediction errors. Log prices appear, therefore, to be the better choice—provided that adjustment is made for the bias introduced by undoing the log transformation to obtain appraisals. We find that log predictions corrected with the adjustment suggested by Garderen (2001) produce the best out-of-sample performance.

Acknowledgements

We have benefited from comments from Wolfgang Härdle, Verity Watson, Dong Zhi, and participants at presentations at the AREUEA International Meeting 2013, the University of Aberdeen, the Humboldt-Universität zu Berlin, and the University of Reading. We are grateful for information provided by Eberhard Hackel and Rebecca Holter from Fitch Ratings and Jonathan Livingstone and Deniz Yegenaga from Moody’s Investors Service Ltd. The usual disclaimer applies. Financial support from the Deutsche Forschungsgemeinschaft, CRC 649 *Economic Risk*, is gratefully acknowledged.

A Appendix

A.1 Transformation function and cross-validation

The transformation function is

$$T_\lambda(x) = \begin{cases} \lambda^{-1} \left[\{\sigma_x^{-1}(x + a_\lambda)\}^\lambda - 1 \right] & \text{for } \lambda \in \Lambda, \\ \ln\{\sigma_x^{-1}(x + a_0)\} & \text{for } \lambda = 0 \end{cases} \quad (\text{A1})$$

with $\Lambda = \{-2, -1, -0.5, 0.5, 1, 2\}$. x is a continuous explanatory variable, σ_x is the standard deviation of x in the sample and λ is the parameter that determines the transformation. A particular value of λ implies a specific constant a_λ . These constants are computed according to the suggestions made in Bunke et al. (1999) and aim to make, for any given λ , the transformation as nonlinear as possible.

We choose λ_i for each of the three continuous variables simultaneously using the cross-validation criterion

$$\boldsymbol{\lambda}^* = \arg \min_{\boldsymbol{\lambda}} \sum_{n=1}^N \{y_n - \hat{y}_{-n}(\boldsymbol{\lambda})\}^2, \quad (\text{A2})$$

where $\boldsymbol{\lambda}$ is the vector comprised of all λ_i s. Here, N is the number of observations in the estimation sample, $\hat{y}_{-n}(\boldsymbol{\lambda})$ denotes the predicted value of y_n from an OLS regression of y on all explanatory variables, including the continuous variables, which have been transformed according to the particular value of λ under consideration. The subscript $-n$ indicates that observation n has been omitted from the regression used to predict the observed price of house n . By leaving out the observation used for evaluating the model fit the cross validated choice of $\boldsymbol{\lambda}^*$ is optimal in the sense of minimizing an estimate of the expected squared prediction error (Bunke et al., 1999). The R^2 -type, standardized cross-valuation criterion

$$\text{CVS} = 1 - \frac{\sum_{n=1}^N \{y_n - \hat{y}_{-n}(\boldsymbol{\lambda}^*)\}^2}{\sum_{n=1}^N (y_n - \bar{y})^2} \quad (\text{A3})$$

is the according measure of fit.

A.2 Market value predictors

For the MVPs based on the log regressions, we use two different estimators for the retransformation adjustment. For comparison, we also compute the market value predictor without adjustment

$$\text{MVP1} = \exp\{\hat{w}(\mathbf{x})\}. \quad (\text{A4})$$

This predictor is biased and inconsistent for the market value. It is, however, a consistent predictor of the median transaction price if log prices are symmetrically distributed. This follows because any order statistic of a monotonic function equals the function of the order statistic. Predictions of the median price may be an useful and easy to

understand alternative to market value predictors for (aspiring) property owners who want to get a feel for the market.

To understand the first retransformation adjustment, observe that we would obtain the market value if we multiplied MVP1 with $\exp\{-E[\ln U]\}$, see Eq. 4. Duan (1983) and Brown and Mariano (1984) suggested an estimator for $\exp\{-E[\ln U]\}$ that uses $E[U] = 1$ and exploits

$$\begin{aligned}\exp\{-E[\ln U]\} &= E[U] \exp\{-E[\ln U]\} \\ &= E[\exp\{\ln U - E[\ln U]\}] \\ &= E[\exp\{u\}].\end{aligned}$$

Replacing the unobserved u with the residuals from the hedonic regression gives the first adjusted predictor

$$\text{MVP2} = \exp\{\widehat{w}(\mathbf{x})\} \left(\frac{1}{N} \sum_{n=1}^N \exp\{\widehat{u}_n\} \right). \quad (\text{A5})$$

Observe that the retransformation adjustment is identical for all predictions.

The second adjusted predictor assumes that u follows a normal distribution. In addition to adjusting for the retransformation, the predictor also considers that the coefficients in the hedonic regression are estimated. Goldberger (1968) derived an unbiased estimator for the predicted variable. This estimator requires numerical integration. Garderen (2001) proposed the computationally simpler

$$\text{MVP3} = \exp\{\widehat{w}(\mathbf{x})\} \sum_{i=0}^{\infty} \frac{z^i}{i!m}, \quad (\text{A6})$$

where the second term is the hypergeometric function with $m = 0.5(N - K)$ and

$$z = 0.5m(1 - \widehat{\sigma}_{\widehat{w}(\mathbf{x})}^2)\widehat{\sigma}_u^2. \quad (\text{A7})$$

Observe that z depends on the estimated standard error of the log price prediction for a house with characteristics \mathbf{x} . The retransformation adjustment for the market value prediction is, therefore, house-specific. Under the assumption of log-normal house prices, MVP3 is unbiased and mean square error efficient.

The fourth predictor comes from the estimated regression with the price as dependent variable, see Eq. 5. This predictor is

$$\text{MVP4} = \widehat{V}(\mathbf{x}). \quad (\text{A8})$$

References

- Anglin, P. M. and Gencay, R.: 1996, Semiparametric estimation of a hedonic price function, *Journal of Applied Econometrics* **11**, 633–648.
- Bin, O.: 2004, A prediction comparison of housing sales prices by parametric versus semi-parametric regressions, *Journal of Housing Economics* **13**, 68 – 84.
- Brown, B. W. and Mariano, R. S.: 1984, Residual-based stochastic prediction and estimation in a nonlinear simultaneous system, *Econometrica* **52**, 321–343.
- Bunke, O., Droge, B. and Polzehl, J.: 1999, Model selection, transformations and variance estimation in nonlinear regression, *Statistics* **33**, 197–240.
- Cameron, A. C. and Trivedi, P. K.: 2005, *Microeconometrics. Methods and Applications*, Cambridge University Press, New York.
- Case, B., Clapp, J., Dubin, R. and Rodriguez, M.: 2004, Modeling spatial and temporal house price patterns: A comparison of four models, *Journal of Real Estate Finance and Economics* **29**, 167–191.
- Case, B. and Quigley, J. M.: 1991, The dynamics of real estate prices, *Review of Economics and Statistics* **73**, 50–58.
- Clapp, J. M.: 2004, A semiparametric method for estimating local house price indices, *Real Estate Economics* **32**, 127–160.
- Downie, M. L. and Robson, G.: 2007, Automated Valuation Models: an international perspective, *CML research*, Council of Mortgage Lenders.
- Duan, N.: 1983, Smearing estimate: A nonparametric retransformation method, *Journal of the American Statistical Association* **78**, 605–610.
- Fitch Ratings Structured Finance: 2012, Criteria for automated valuation models in EMEA RMBS, *Residential Mortgage/EMEA August*, Fitch, New York.
- Garderen, K. J. V.: 2001, Optimal prediction in loglinear models, *Journal of Econometrics* **104**, 119–140.

- Gençay, R. and Yang, X.: 1996, A forecast comparison of residential housing prices by parametric versus semiparametric conditional mean estimators, *Economic Letters* **52**, 129–135.
- Goetzmann, W. N.: 1992, The accuracy of real estate indices: Repeat sale estimators, *Journal of Real Estate Finance and Economics* **5**, 5–53.
- Goetzmann, W. N. and Peng, L.: 2002, The bias of the RSR estimator and the accuracy of some alternatives, *Real Estate Economics* **30**, 13–39.
- Goh, Y. M., Costello, G. and Schwann, G.: 2012, Accuracy and robustness of house price index methods, *Housing Studies* **27**, 643–666.
- Goldberger, A. S.: 1968, The interpretation and estimation of Cobb-Douglas functions, *Econometrica* **36**, 464–472.
- Haupt, H., Schurbus, J. and Tschernig, R.: 2010, On nonparametric estimation of a hedonic price function, *Journal of Applied Econometrics* **25**, 894–901.
- Hill, R. C., Knight, J. and Sirmans, C.: 1997, Estimating capital asset price indexes, *Review of Economics and Statistics* **79**, 226–233.
- IAAO: 2003, Standard on automated valuation models (AVMs), *Assessment standards*, International Association of Assessing Officers, Chicago.
- LeSage, J. and Pace, R. K.: 2009, *Introduction to Spatial Econometrics*, Statistics: Textbooks and Monographs, Chapman & Hall, Boca Raton.
- Myers, R. H.: 1989, *Classical and Modern Regression with Applications*, PWS-KENT.
- Moody's Investors Service: 2008, Moody's approach to automated valuation models in rating UK RMBS, *International Structured Finance August*, Moody's, New York.
- Moody's Investors Service: 2009, What drives UK mortgage loans to default, *International Structured Finance August*, Moody's, New York.
- Moody's Investors Service: 2012, Moody's approach to rating RMBS in Europe, Middle East, and Africa, *Rating Methodology June*, Moody's, New York.

- Nattagh, N.: 2007, Current issues in validating automated valuation models in the U.S. mortgage industry, Paper presented to the European Real Estate Society conference. Geostat Advisory LLC.
- Palmquist, R. B.: 1980, Alternative techniques for developing real estate price indexes, *Review of Economics and Statistics* **62**, 442–448.
- Palmquist, R. B.: 1991, Hedonic methods, in J. B. Braden and C. D. Kolstad (eds), *Measuring the demand for environmental quality*, Contributions to Economic Analysis, North Holland, Amsterdam, chapter 4, pp. 77–120.
- Rousseeuw, P. J.: 1985, Multivariate estimation with high breakdown point, in W. Grossmann, G. Pflug, I. Vincze and W. Wertz (eds), *Mathematical Statistics and Applications*, Vol. B, Reidel, Dordrecht, pp. 283–297.
- Shiller, R. J. and Weiss, A. N.: 1999, Evaluating real estate valuation systems, *Journal of Real Estate Finance and Economics* **18**, 147–161.
- Thibodeau, T. G.: 2003, Marking single-family property values to market, *Real Estate Economics* **31**, 1–22.
- Yatchew, A.: 2003, *Semiparametric Regression for the Applied Econometrician*, Themes in Modern Econometrics, Cambridge University Press, Cambridge.

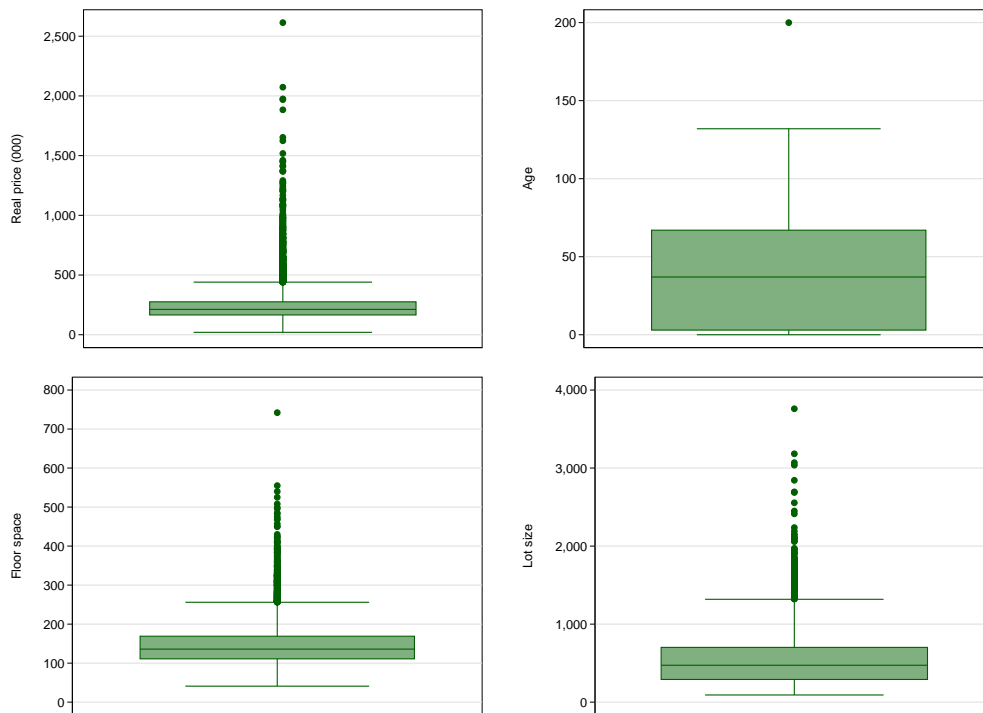


Figure 1: Box plots of transaction price and continuous house characteristics before automated outlier detection. Box plots use 8,851 observations from 2000-2005. Real price is transaction price deflated with Berlin's CPI (base year 2005), in thousand EUR. Floor and lot size are in meters squared.

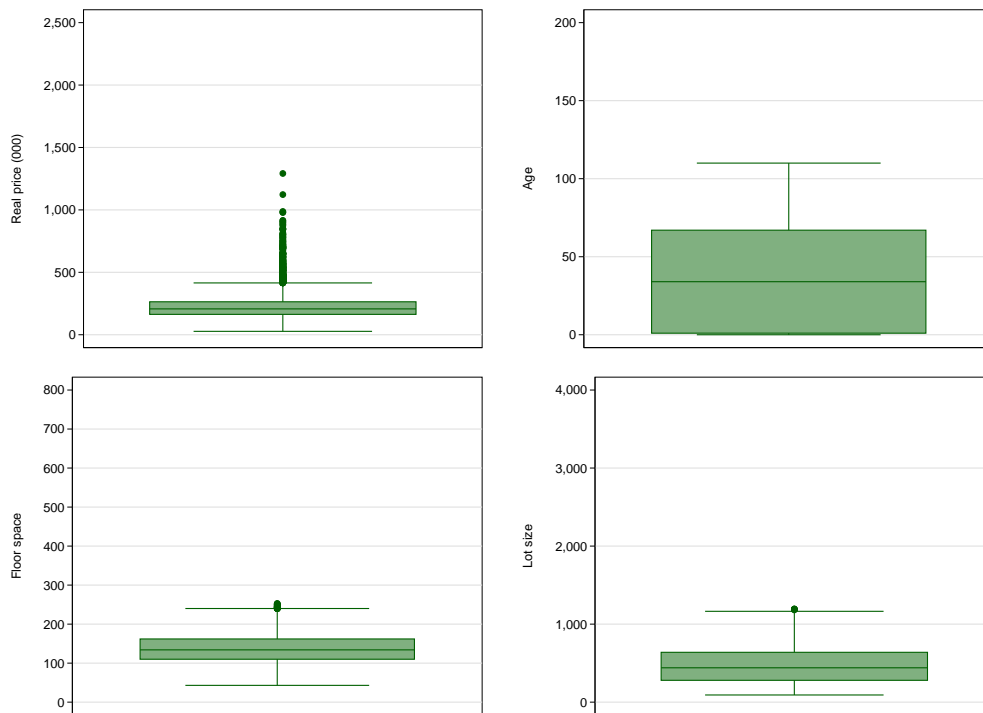


Figure 2: Box plots of transaction price and continuous house characteristics after automated outlier detection. Box plots use 8,429 observations from 2000-2005. Real price is transaction price deflated with Berlin's CPI (base year 2005), in thousand EUR. Floor and lot size are in meters squared.

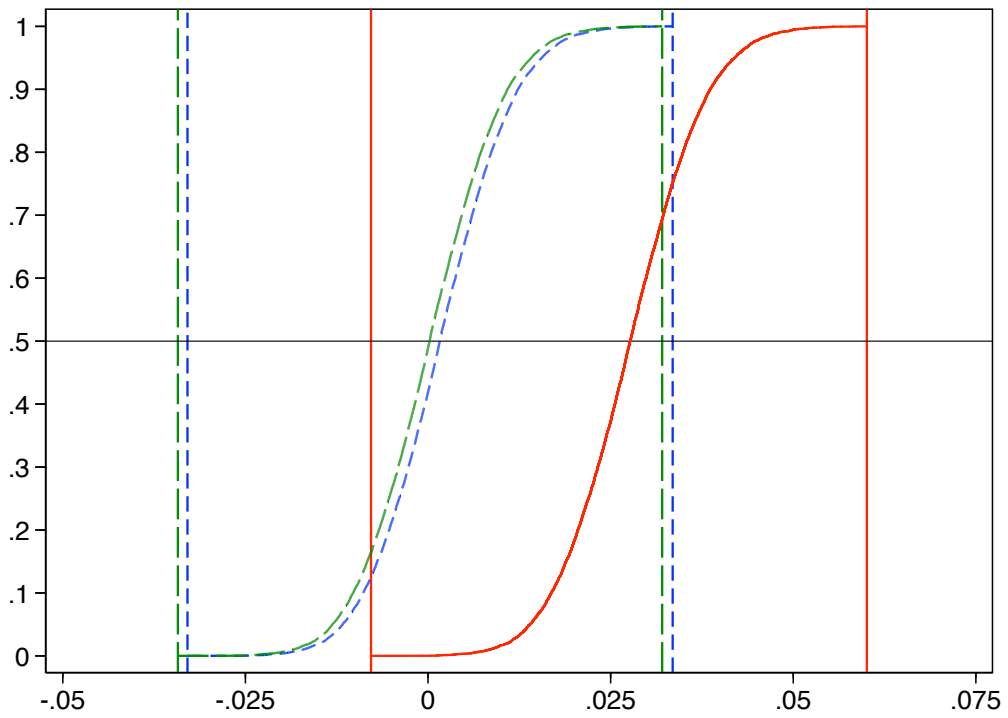


Figure 3: Empirical distribution of Mean Prediction Errors in 10,000 Monte Carlo replications. Solid red line is for MVP1. Blue dashed line is for MVP2. Green long dashed line is for MVP3. Vertical lines indicate the smallest and largest Mean Prediction Error obtained for the respective market value predictor.

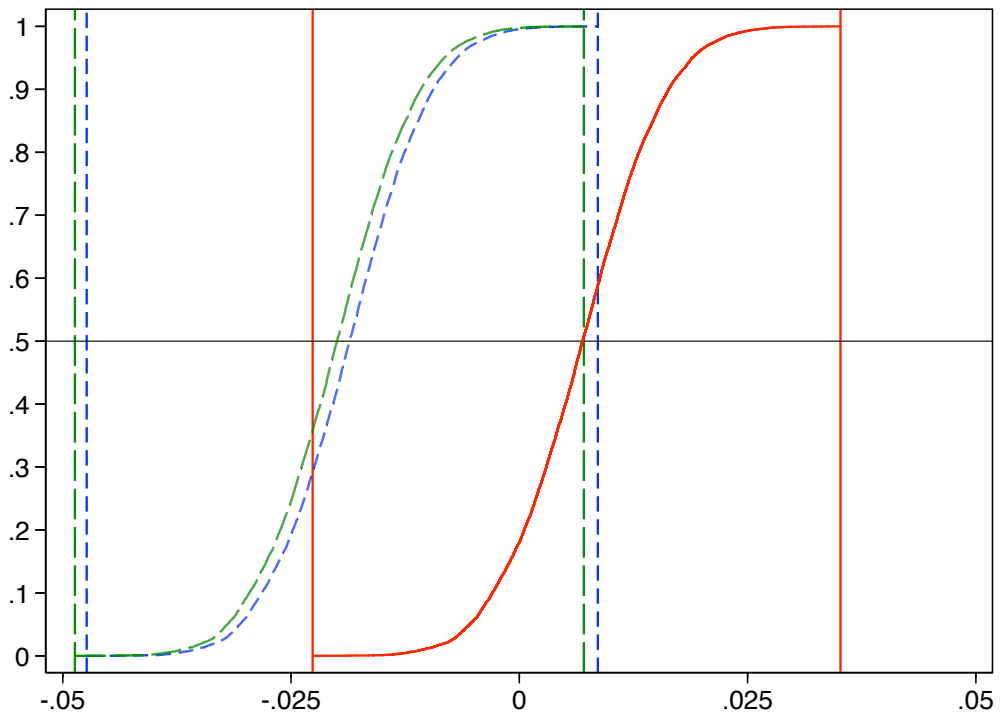


Figure 4: Empirical distribution of Median Prediction Errors in 10,000 Monte Carlo replications. Solid red line is for MVP1. Blue dashed line is for MVP2. Green long dashed line is for MVP3. Vertical lines indicate the smallest and largest Median Prediction Error obtained for the respective market value predictor.

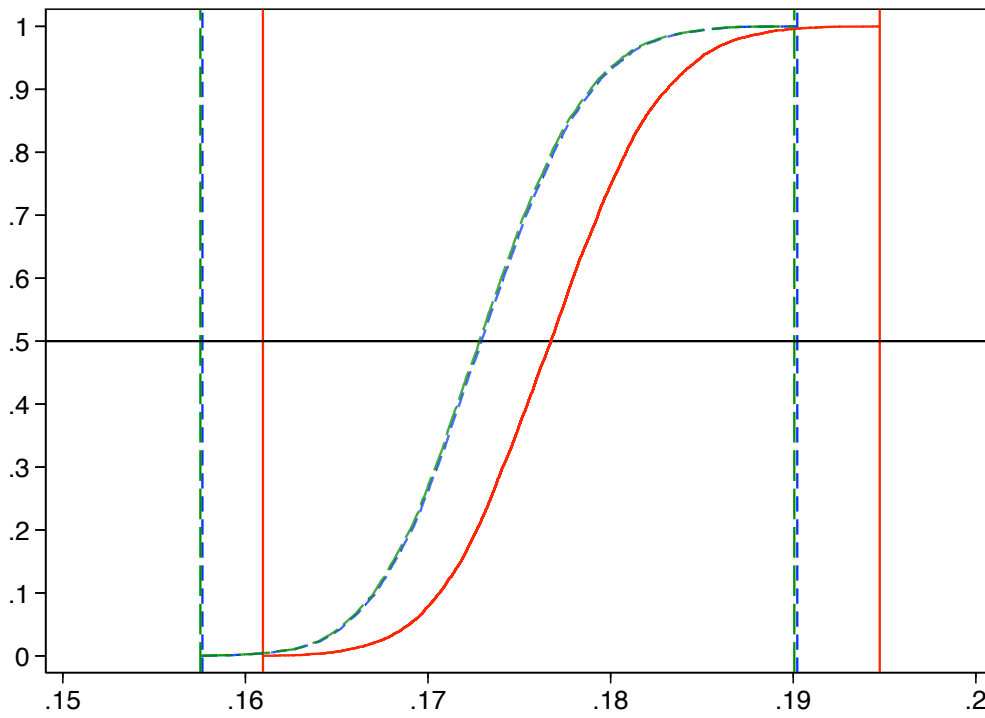


Figure 5: Empirical distribution of Mean Absolute Prediction Errors in 10,000 Monte Carlo replications. Solid red line is for MVP1. Blue dashed line is for MVP2. Green long dashed line is for MVP3. Vertical lines indicate the smallest and largest Mean Absolute Prediction Error obtained for the respective market value predictor.

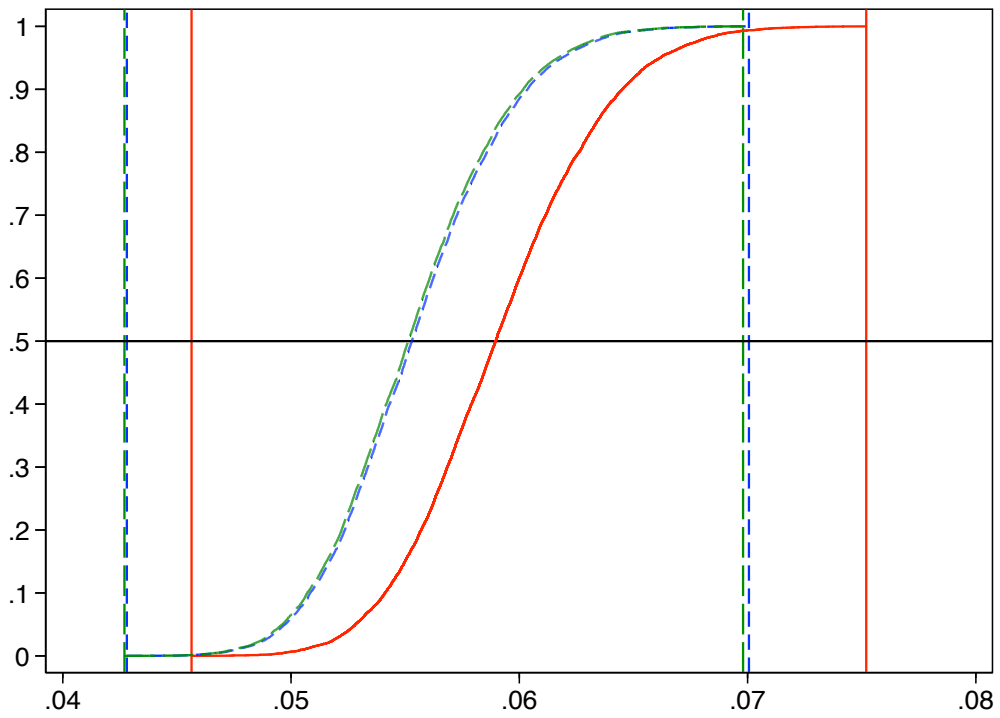


Figure 6: Empirical distribution of Mean Squared Prediction Errors in 10,000 Monte Carlo replications. Solid red line is for MVP1. Blue dashed line is for MVP2. Green long dashed line is for MVP3. Vertical lines indicate the smallest and largest Mean Squared Prediction Error obtained for the respective market value predictor.

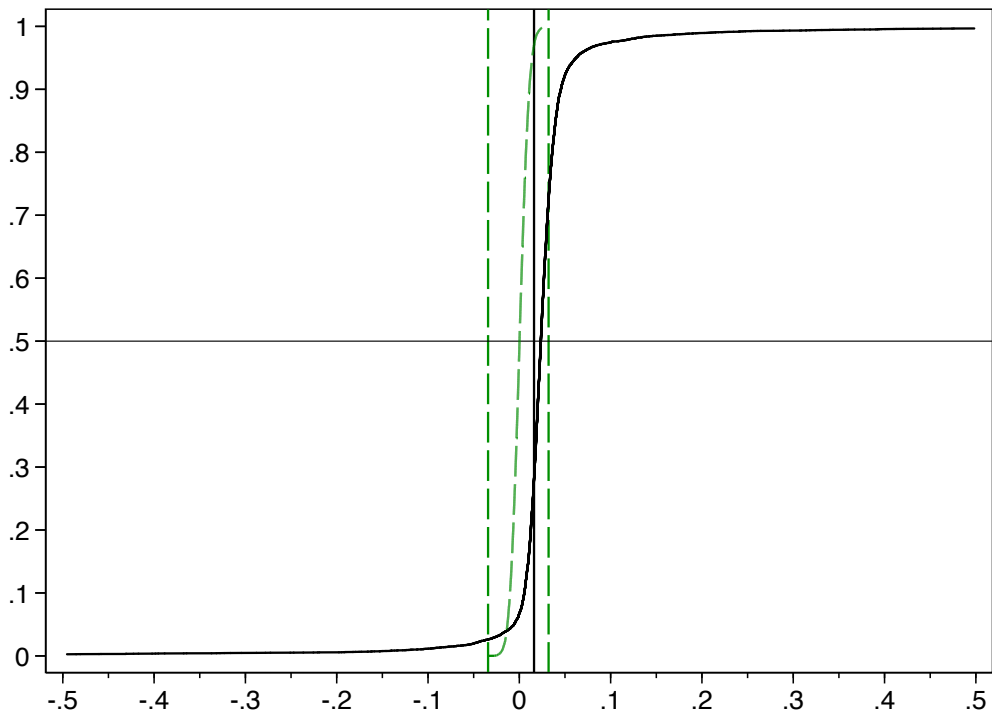


Figure 7: Empirical distribution of Mean Prediction Errors in 10,000 Monte Carlo replications. Solid black line is for MVP4. Green long dashed line is for MVP3. Vertical lines indicate the smallest and largest Mean Prediction Error obtained for the respective market value predictor.

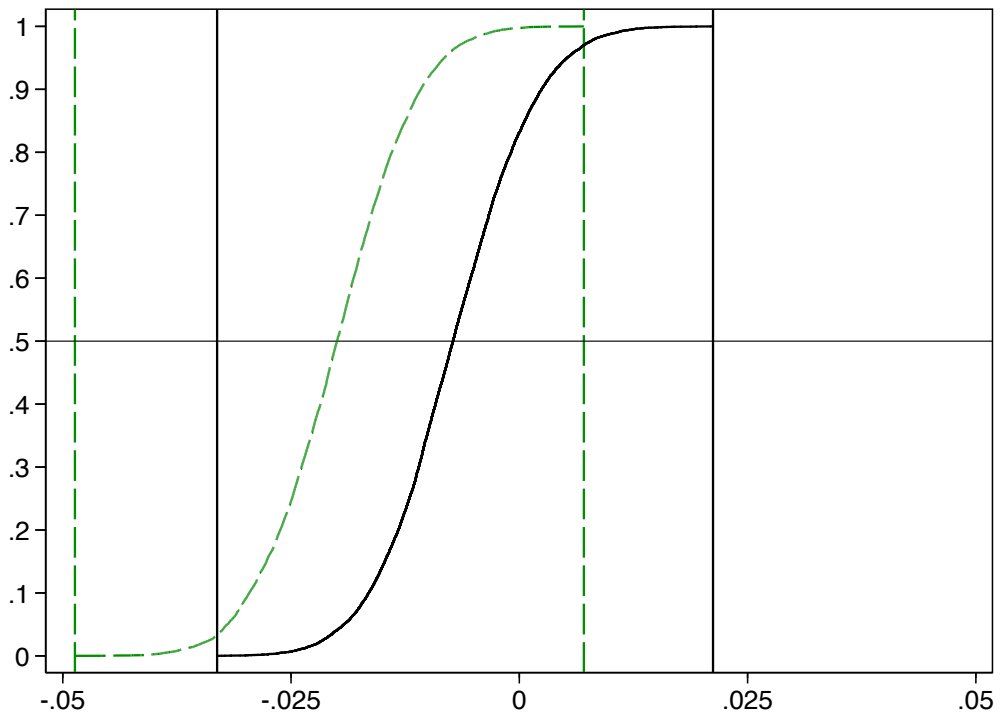


Figure 8: Empirical distribution of Median Prediction Errors in 10,000 Monte Carlo replications. Solid black line is for MVP4. Green long dashed line is for MVP3. Vertical lines indicate the smallest and largest Median Prediction Error obtained for the respective market value predictor.

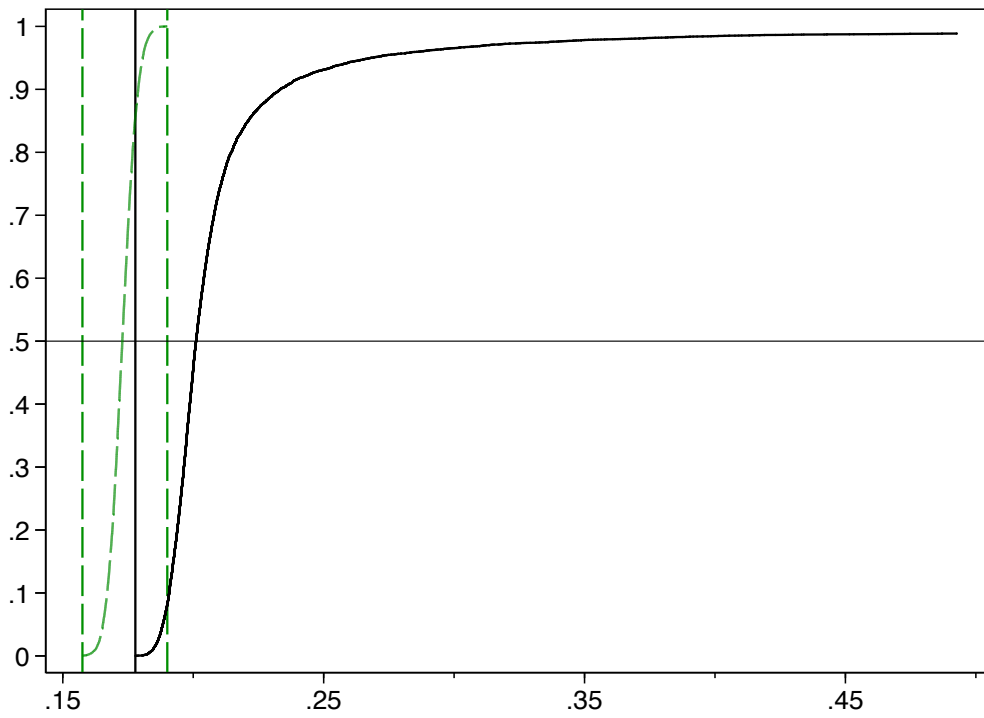


Figure 9: Empirical distribution of Mean Absolute Prediction Errors in 10,000 Monte Carlo replications. Solid black line is for MVP4. Green long dashed line is for MVP3. Vertical lines indicate the smallest and largest Mean Absolute Prediction Error obtained for the respective market value predictor.

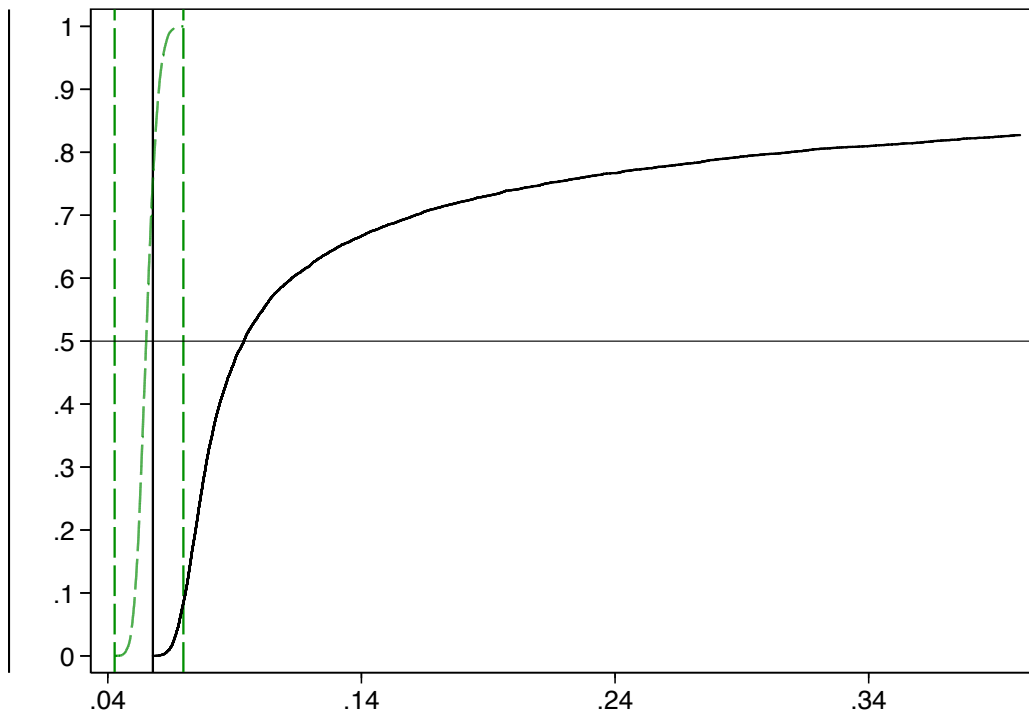


Figure 10: Empirical distribution of Mean Squared Prediction Errors in 10,000 Monte Carlo replications. Solid black line is for MVP4. Green long dashed line is for MVP3. Vertical lines indicate the smallest and largest Mean Squared Prediction Error obtained for the respective market value predictor.

Table 1: Summary statistics for transacted single-family houses. Reports mean of variables. Standard deviations are in brackets. Real price is transaction price deflated with Berlin’s CPI (base year 2005), in thousand EUR. Floor and lot size are in meters squared. Attic means that the attic is upgraded for living. Expert-based location rating is an ordinal ranking of house’s neighborhood.

	Development sample		Validation sample	
<i>Price and continuous characteristics</i>				
Real price	231.83	[113.35]	219.32	[111.27]
Age of building	36.89	[29.67]	37.89	[31.58]
Floor size	139.57	[40.38]	141.31	[42.50]
Lot size	497.61	[257.48]	501.98	[255.88]
<i>Discrete house characteristics</i>				
Detached house	0.46		0.51	
Semi-detached house	0.34		0.32	
Row house	0.20		0.17	
One storey	0.52		0.51	
Two storeys	0.45		0.44	
Three storeys	0.03		0.05	
Attic	0.56		0.55	
Flat roof	0.11		0.13	
Full cellar	0.73		0.68	
Part cellar	0.11		0.10	
No cellar	0.16		0.22	
Stove heating	0.06		0.03	
Poor state of repair	0.37		0.42	
Medium state of repair	0.55		0.54	
Good state of repair	0.08		0.04	
Direct lake access	0.00		0.01	
Lake view	0.00		0.01	
<i>Unusual features of the house</i>				
Legal	0.23		0.23	
Physical	0.03		0.02	
<i>Expert based location rating</i>				
Low	0.37		0.30	
Medium	0.43		0.51	
High	0.18		0.18	
Excellent	0.02		0.02	
Number of obs.	8,429		10,015	

Table 2: Model performance in development sample 2000-2005. Reports standardized cross-validation criteria CVS and transformation parameters λ^* for best model given dependent variable and length of estimation window. All models are fitted for the 8,429 observations of the development sample.

Dependent variable	Length of estimation window			
	1 Year	2 Years	3 Years	6 Years
ln Price	0.723 (1, 0.5, -2)	0.729 (1, 0.5, -2)	0.733 (0.5, 0.5, -2)	0.732 (0.5, 0.5, -2)
Price	0.675 (0.5, 2, -1)	0.691 (0.5, 2, -1)	0.688 (0.5, 0.5, 1)	0.713 (0.5, 1, -2)

Table 3: Hedonic regression results. Reports OLS estimates of Eq. 6 for the second three-year estimation window in the development sample. Dependent variable is the log price and $\lambda^* = (0.5, 0.5, -2)$. Coefficients for time and district dummies are not reported. Heteroscedasticity robust standard errors are reported in brackets. *** significant at 1%-level ** significant at 5%-level * significant at 10%-level.

$T(\text{Age})$	-0.323	[0.358]
$T(\text{Age})^2$	-2.316 ***	[0.568]
$T(\text{Floor-size})$	-0.105 ***	[0.022]
$T(\text{Floor-size})^2$	0.038 ***	[0.008]
$T(\text{Lot-size})$	0.247 ***	[0.027]
$T(\text{Lot-size})^2$	0.010	[0.008]
$T(\text{Age}) \times T(\text{Floor-size})$	0.547 ***	[0.064]
$T(\text{Age}) \times T(\text{Lot-size})$	-0.188 ***	[0.063]
$T(\text{Floor-size}) \times T(\text{Lot-size})$	-0.024 ***	[0.012]
Semi-detached	-0.045 ***	[0.012]
Row house	-0.065 ***	[0.017]
2 Storeys	0.030 **	[0.012]
3 Storeys	0.123 ***	[0.023]
Attic	0.043 ***	[0.010]
Flat roof	-0.032 **	[0.012]
Stove heating	-0.114 ***	[0.020]
No cellar	-0.130 ***	[0.012]
Part cellar	-0.058 ***	[0.014]
Good state of repair	0.132 ***	[0.014]
Poor state of repair	-0.231 ***	[0.018]
Direct lake access	0.331 ***	[0.078]
Lake view	0.167 ***	[0.052]
Low location quality	-0.035 ***	[0.011]
High location quality	0.100 ***	[0.015]
Excellent location quality	0.162 ***	[0.047]
Unusual legal	-0.035 ***	[0.010]
Unusual physical	-0.104 ***	[0.104]
Number of obs.	4,502	
R^2	0.723	
CVS	0.706	

Table 4: Summary statistics for performance measures in Monte Carlo simulation. Reports mean of respective performance measure. Standard deviations are in brackets. Number of replications is 10,000.

	MPE	MDPE	MAPE	MSPE
MVP1	0.028 [0.009]	0.007 [0.007]	0.177 [0.005]	0.059 [0.004]
MVP2	0.002 [0.008]	-0.019 [0.007]	0.173 [0.005]	0.055 [0.004]
MVP3	0.000 [0.008]	-0.020 [0.007]	0.173 [0.005]	0.055 [0.004]
MVP4	0.022 [0.419]	-0.007 [0.007]	0.224 [0.418]	1.811 [1.905]

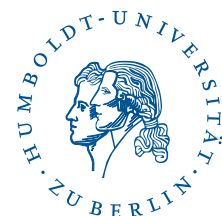
SFB 649 Discussion Paper Series 2013

For a complete list of Discussion Papers published by the SFB 649, please visit <http://sfb649.wiwi.hu-berlin.de>.

- 001 "Functional Data Analysis of Generalized Quantile Regressions" by Mengmeng Guo, Lhan Zhou, Jianhua Z. Huang and Wolfgang Karl Härdle, January 2013.
- 002 "Statistical properties and stability of ratings in a subset of US firms" by Alexander B. Matthies, January 2013.
- 003 "Empirical Research on Corporate Credit-Ratings: A Literature Review" by Alexander B. Matthies, January 2013.
- 004 "Preference for Randomization: Empirical and Experimental Evidence" by Nadja Dwenger, Dorothea Kübler and Georg Weizsäcker, January 2013.
- 005 "Pricing Rainfall Derivatives at the CME" by Brenda López Cabrera, Martin Odening and Matthias Ritter, January 2013.
- 006 "Inference for Multi-Dimensional High-Frequency Data: Equivalence of Methods, Central Limit Theorems, and an Application to Conditional Independence Testing" by Markus Bibinger and Per A. Mykland, January 2013.
- 007 "Crossing Network versus Dealer Market: Unique Equilibrium in the Allocation of Order Flow" by Jutta Dönges, Frank Heinemann and Tijmen R. Daniëls, January 2013.
- 008 "Forecasting systemic impact in financial networks" by Nikolaus Hautsch, Julia Schaumburg and Melanie Schienle, January 2013.
- 009 "'I'll do it by myself as I knew it all along': On the failure of hindsight-biased principals to delegate optimally" by David Danz, Frank Hüber, Dorothea Kübler, Lydia Mechtenberg and Julia Schmid, January 2013.
- 010 "Composite Quantile Regression for the Single-Index Model" by Yan Fan, Wolfgang Karl Härdle, Weining Wang and Lixing Zhu, February 2013.
- 011 "The Real Consequences of Financial Stress" by Stefan Mittnik and Willi Semmler, February 2013.
- 012 "Are There Bubbles in the Sterling-dollar Exchange Rate? New Evidence from Sequential ADF Tests" by Timo Bettendorf and Wenjuan Chen, February 2013.
- 013 "A Transfer Mechanism for a Monetary Union" by Philipp Engler and Simon Voigts, March 2013.
- 014 "Do High-Frequency Data Improve High-Dimensional Portfolio Allocations?" by Nikolaus Hautsch, Lada M. Kyj and Peter Malec, March 2013.
- 015 "Cyclical Variation in Labor Hours and Productivity Using the ATUS" by Michael C. Burda, Daniel S. Hamermesh and Jay Stewart, March 2013.
- 016 "Quantitative forward guidance and the predictability of monetary policy – A wavelet based jump detection approach –" by Lars Winkelmann, April 2013.
- 017 "Estimating the Quadratic Covariation Matrix from Noisy Observations: Local Method of Moments and Efficiency" by Markus Bibinger, Nikolaus Hautsch, Peter Malec and Markus Reiss, April 2013.
- 018 "Fair re-valuation of wine as an investment" by Fabian Y.R.P. Bocart and Christian M. Hafner, April 2013.
- 019 "The European Debt Crisis: How did we get into this mess? How can we get out of it?" by Michael C. Burda, April 2013.

SFB 649, Spandauer Straße 1, D-10178 Berlin
<http://sfb649.wiwi.hu-berlin.de>

This research was supported by the Deutsche
Forschungsgemeinschaft through the SFB 649 "Economic Risk".



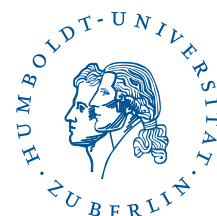
SFB 649 Discussion Paper Series 2013

For a complete list of Discussion Papers published by the SFB 649, please visit <http://sfb649.wiwi.hu-berlin.de>.

- 020 "Disaster Risk in a New Keynesian Model" by Maren Brede, April 2013.
- 021 "Econometrics of co-jumps in high-frequency data with noise" by Markus Bibinger and Lars Winkelmann, May 2013.
- 022 "Decomposing Risk in Dynamic Stochastic General Equilibrium" by Hong Lan and Alexander Meyer-Gohde, May 2013.
- 023 "Reference Dependent Preferences and the EPK Puzzle" by Maria Grith, Wolfgang Karl Härdle and Volker Krättschmer, May 2013.
- 024 "Pruning in Perturbation DSGE Models - Guidance from Nonlinear Moving Average Approximations" by Hong Lan and Alexander Meyer-Gohde, May 2013.
- 025 "The 'Celtic Crisis': Guarantees, transparency, and systemic liquidity risk" by Philipp König, Kartik Anand and Frank Heinemann, May 2013.
- 026 "State Price Densities implied from weather derivatives" by Wolfgang Karl Härdle, Brenda López-Cabrera and Huei-Wen Teng, May 2013.
- 027 "Bank Lending Relationships and the Use of Performance-Sensitive Debt" by Tim R. Adam and Daniel Streitz, May 2013.
- 028 "Analysis of Deviance in Generalized Partial Linear Models" by Wolfgang Karl Härdle and Li-Shan Huang, May 2013.
- 029 "Estimating the quadratic covariation of an asynchronously observed semimartingale with jumps" by Markus Bibinger and Mathias Vetter, May 2013.
- 030 "Can expert knowledge compensate for data scarcity in crop insurance pricing?" by Zhiwei Shen, Martin Odening and Ostap Okhrin, May 2013.
- 031 "Comparison of Methods for Constructing Joint Confidence Bands for Impulse Response Functions" by Helmut Lütkepohl, Anna Staszewska-Bystrova and Peter Winker, May 2013.
- 032 "CDO Surfaces Dynamics" by Barbara Choroś-Tomczyk, Wolfgang Karl Härdle and Ostap Okhrin, July 2013.
- 033 "Estimation and Inference for Varying-coefficient Models with Nonstationary Regressors using Penalized Splines" by Haiqiang Chen, Ying Fang and Yingxing Li, July 2013.
- 034 "Robust Estimation and Inference for Threshold Models with Integrated Regressors" by Haiqiang Chen, July 2013.
- 035 "A new perspective on the economic valuation of informal care: The well-being approach revisited" by Konstantin Kehl and Stephan Stahlschmidt, July 2013.
- 036 "Herding in financial markets: Bridging the gap between theory and evidence" by Christopher Boortz, Simon Jurkatis, Stephanie Kremer and Dieter Nautz, July 2013.
- 037 "Default Risk Calculation based on Predictor Selection for the Southeast Asian Industry" by Wolfgang Karl Härdle and Dedi Dwi Prastyo, August 2013.
- 038 "ECB monetary policy surprises: identification through cojumps in interest rates" by Lars Winkelmann, Markus Bibinger and Tobias Linzert, August 2013.
- 039 "Limited higher order beliefs and the welfare effects of public information" by Camille Cornand and Frank Heinemann, August 2013.
- 040 "Privacy Concerns, Voluntary Disclosure of Information, and Unraveling: An Experiment" by Volker Benndorf, Dorothea Kübler and Hans-Theo Normann, September 2013.

SFB 649, Spandauer Straße 1, D-10178 Berlin
<http://sfb649.wiwi.hu-berlin.de>

This research was supported by the Deutsche
Forschungsgemeinschaft through the SFB 649 "Economic Risk".



SFB 649 Discussion Paper Series 2013

For a complete list of Discussion Papers published by the SFB 649, please visit <http://sfb649.wiwi.hu-berlin.de>.

- 041 "Goodness-of-fit Test for Specification of Semiparametric Copula Dependence Models" by Shulin Zhang, Ostap Okhrin, Qian M. Zhou and Peter X.-K. Song, September 2013.
- 042 "Volatility linkages between energy and agricultural commodity prices" by Brenda López Cabrera and Franziska Schulz, September 2013.
- 043 "Testing the Preferred-Habitat Theory: The Role of Time-Varying Risk Aversion" by Till Strohsal, September 2013.
- 044 "Assortative matching through signals" by Friedrich Poeschel, September 2013.
- 045 "Intertemporal Consumption and Debt Aversion: An Experimental Study" by Thomas Meissner, September 2013.
- 046 "Automated Valuation Modelling: A Specification Exercise" by Rainer Schulz, Martin Wersing and Axel Werwatz, October 2013.

SFB 649, Spandauer Straße 1, D-10178 Berlin
<http://sfb649.wiwi.hu-berlin.de>

This research was supported by the Deutsche
Forschungsgemeinschaft through the SFB 649 "Economic Risk".

