



Research article

Comparing ask and transaction prices in the Swiss housing market

Diego Ardila^{1,*}, Ahmed Ahmed^{2,†} and Didier Sornette^{1,3}

¹ Department of Management, ETH Zurich, Technology, and Economics, Zurich, Switzerland

² Department of Economics, University of Chicago, Chicago, USA

³ Swiss Finance Institute, University of Geneva, Switzerland

* **Correspondence:** Email: dardila@ethz.ch; Tel: +41446329956.

Abstract: We analyze the relationship between ask and transaction prices in the Swiss residential real estate market over the 2005–2015 period. First, we present strong evidence that ask and transaction prices are co-integrated across different market segments, but they do not Granger-cause one another. Second, we analyze the cross-sectional distributions of ask and transaction prices/per living space and conclude that they do not follow the same distribution, with the distribution of transaction prices close to a log normal distribution and the distribution of ask prices exhibiting slightly fatter tails. Finally, we show significant evidence that transaction prices tend to exceed ask prices during protracted booms and bubble regimes. We discuss these empirical patterns in light of theoretical housing search models, and provide support for the hypothesis that the 2005–2015 Swiss market has been dominated by an auction-like dynamics. Hence, although ask prices constitute a suitable proxy to follow the development of the Switzerland’s real estate market, especially given the sparsity of available transaction data, they might be prone to underestimate the extent of price increases when the market is booming, and the magnitude of the correction when the market enters the bust phase of the housing cycle.

Keywords: ask prices; transaction prices; log normal distribution; co-integration; granger causality; bubbles; LPPLS; search models

JEL Codes: C48, C58, G01, R31

1. Introduction

The development of residential property prices in Switzerland has raised concerns about the possibility of a bubble (Ardila et al., 2013). Classic macroeconomic indicators, such as price-to-rent and price-to-income ratios, have steadily increased since 2008, while Swiss household debt remains among

[†] The author contributed to this work during his time at ETH Zurich while being funded through the SNF grant 160305.

the highest of OECD economies. The situation prompted the Swiss National Bank to issue several macro-prudential policies in order to manage the risk that the housing boom may pose to the economy (Basten and Koch, 2015).

The effectiveness of these measures and the overall evolution of the market has been difficult to assess by the general public. The difficulties can be mostly attributed to two factors. First, the Swiss housing market is highly illiquid, with a volume of transactions consistently low. Second, contrary to countries such as the UK and the US, there is no single database encompassing the totality of the transactions; in fact, the consolidation of an open and transparent database has been historically discouraged by an institutional environment marked by banking secrecy and a tendency to favor a lean regulatory framework (Vogler, 2006). In this context, the use of ask prices, and in particular Internet data, has surged as a plausible alternative to monitor the development of the market. Nevertheless, the complicated relationship between ask and transaction prices remains poorly understood.

The subject has received some attention from both theoretical and empirical perspectives, in the general literature on real estate in the World, and especially in the US. For example, due to limited human processing capabilities, Diaz et al. (1999) argue that potential home buyers often use the ask price in the housing market as a prediction for the transaction final price. Therefore, establishing an ask price that is expected to result in the highest final transaction price may be an important strategy. Knight et al. (1994) and Horowitz (1992) argue that ask prices can play a vital role in the transaction process through the strategic influence of anchoring heuristic and signaling. Based on the former, listing a housing unit for a higher ask price relative to the market value has been found to be positively correlated with securing a higher transaction price. Based on the later, extraordinary ask prices (either high or low) may convey information to potential buyers about the features of the property and the motivations of the seller. That is, although sellers may strategically choose a high ask price assuming that it may lead to a higher final transaction one, this may also decrease the rate at which buyers arrive and consider buying the unit. Using data from Tokyo between 2005 and 2009, Shimizu et al. (2016) compared the initial asking price, final asking price, contract price, and transaction price, and showed that there are significant differences in the distribution of these prices and in the distribution of the attributes of properties. Anenberg and Laufer (2017) showed that an asking price-based index can accurately forecast the Case–Shiller index for the nine largest US cities in the years 2008–2012.

We are not aware of any previous research on the statistical relationship between the ask and transaction prices in the context of the Swiss housing market. This paper aims to help fill this gap. We examine the relationship between ask and transaction prices in Switzerland at different levels of aggregation. Our analysis combines state-of-the-art panel co-integration techniques, Granger-causality tests for dynamic panel data models, density and tail tests to compare the cross-sectional distributions of ask and transaction prices, and bubble tests to explore the relationship between the two price types during different market regimes.

Our results provide strong evidence that ask and transactions prices are co-integrated across different market segments, while they do not constitute lagging indicators of one another, i.e. there is no Granger-causality in either direction. This the co-integration relation might be explained by the search housing process as well as by other exogenous factors, possibly connected to the micro-structure of the market (Timmermans et al., 2017).

The analysis of the cross-sectional distributions of prices/per living space suggests that they are different, though they both seem to display thin tails. We interpret this result as evidence that ask and

transaction prices do not only tend to move together in the long term, but also do not deviate significantly from each other. Therefore, the phenomenon of extreme bids (Levin and Pryce, 2007) is not, at least at the aggregate level, statistically extreme. Nonetheless, we observe a tendency of transaction prices to be higher than ask prices. Indeed, time series regressions suggest that this behavior is connected to expected and unexpected rate of house price increase, as well as to the presence of a housing bubble in the market.

Our results are thus inconsistent with the assumptions made by standard search models (Wheaton, 1990; DiPasquale and Wheaton, 1996). According to these models, ask prices often constitute an upper bound for transaction prices. Sellers determine ask prices, wait for offers, and sell the housing unit for transaction prices that are less than or equal the ask prices, provided the offer is above their reservation price; otherwise, they withdraw the units from the market. On the contrary, our results are consistent with search models that suggest that during a booming market, transaction prices will tend to exceed ask prices, as the seller's strategy changes to an auction based strategy (Días and Jerez, 2013; Albrecht et al., 2016).

Overall, we argue that ask prices are a suitable alternative to follow the development of the market. As ask and transaction prices are closely co-integrated, and there is no obvious causality from one another, agents cannot use ask prices to anticipate price rises, which could lead to further price rises and to a destabilizing market. Alternatively, transaction prices do not reveal all the information contained in the ask prices, which would make the use of the later redundant. Nevertheless, since we find that transaction prices tend to be higher than ask prices during bubble regimes, the use of the later to monitor the market may admittedly underestimate observed price increases when the market is booming, and the extent of the correction when the market has entered the bust phase of the housing cycle.

The rest of the paper is organized as follows. Section 2 describes the data sets and its sources. Section 3 briefly describes our empirical strategy. Section 4 presents our results. Finally, section 5 concludes the paper.

2. Description of the Swiss housing market and data

Two databases were used in this study, one containing the ask prices and the other containing the transaction prices. The ask prices data were based on residential ads and it was collected incrementally by comparis.ch between January 2005 and July 2015. The property market division of comparis.ch gathers data from the 17 largest property portals in Switzerland, creating a rich view on the market, but also introducing a large and un-estimated number of duplicate ads (by 2015-Q2, 6.2 million records are present in the raw data). These duplicates advertise the same property, during the same period, and sometimes, with conflicting information. Within the scope of this study, the identification of the duplicates was crucial, as they could potentially affect the price indices described in section 3.1, on which we conducted the analysis. We implemented a procedure based on the Support Vector Machine (SVM) algorithm (Scholkopf and Smola, 2001) and string distance measures (Cohen et al., 2003) in order to identify the duplicate ads. The procedure determined, in a given zip code and a given quarter, the ads that represented the same residential property by analyzing the similarity between their different attributes (e.g. their title, description, and number of rooms). In this study, we have only included ads with positive price and living space, which amount to about 80% of the data, as this information was essential to develop the price indices on which we base the analysis. In addition, ads with different prices were considered different since this study did not intend to track the price changes of the properties on sale.

Although the ads database contains a timely and rich view of the Swiss real estate market, a valid concern is whether it appropriately reflects the developments of market prices. To examine this issue, we compared the *comparis.ch* database against the Swiss Real Estate Datapool (SRED) database for apartments in the period between 2005-Q1 and 2015-Q2 (the overlapping period for which we had access to both databases). SRED is an association that aims to promote market efficiency and transparency in the Swiss housing market. Its database covers approximately 40% of all residential transactions in Switzerland, and it is arguably the highest quality data source available for the most liquid part of the market. To simplify the comparison, we limited the analysis to ads and transactions in the 32 districts in which at least five transactions of apartments per quarter were observed. Table 1 gives summary statistics for this subset of the database. As it can be observed, the two databases differ substantially in terms of volume. There is roughly a 5:1 ratio between their respective total number of observations (4.6 at district level and 4.6 at national level). The corresponding price developments, on the other hand, seem to behave similarly. The ratio of average growth rate of prices per district per quarter is close to unity at the national, cantonal, and district level.

Table 1. Volume (#Obs) and price change aggregates of apartments for the ask (*comparis.ch*) and transaction (SRED) databases. Statistics correspond to the 2005-Q1/2015-Q2 period. **Tx** (resp. **Ask**) stands for transaction (ask).

| Level | Source | #Obs. | Average #Obs. per quarter | Min. average #Obs. per quarter | Max average #Obs. per quarter | Average quarterly price growth rate | Std. deviation price growth rate |
|-----------------|---------------|--------|------------------------------------|--|---|---|--|
| National | Tx | 77000 | 1300 | 728 | 2065 | 1.1 | 2.2 |
| | Ask | 351000 | 8400 | 2430 | 15900 | 1.2 | 2.6 |
| | Ask/Tx | 4.6 | 6.4 | 3.3 | 7.7 | 1.1 | 1.2 |
| Cantonal | Tx | 8500 | 144 | 67 | 243 | 1.1 | 6.4 |
| | Ask | 39000 | 930 | 270 | 1874 | 1 | 3.9 |
| | Ask/Tx | 4.6 | 6.4 | 4.1 | 7.7 | 0.9 | 0.6 |
| District | Tx | 2394 | 40.6 | 16.2 | 78.6 | 1.1 | 9.7 |
| | Ask | 11000 | 261.5 | 71.3 | 547.4 | 1.0 | 5.3 |
| | Ask/Tx | 4.6 | 6.4 | 4.4 | 7.0 | 1.0 | 0.5 |

One simple reason for these patterns is the fact that not every advertised property might lead to a transaction. In addition, there are idiosyncratic reasons that deserve some comments. The SRED database covers only about 40 percent of the market, and its transactions are concentrated on the most populated and liquid cantons of Switzerland (i.e. Zurich and Geneva). Indeed, only 32 districts out of the 166 districts in this database contain more than 5 transactions per district per quarter. On the other hand, the ads' database aggregates the information from several online sources. As a consequence, it may also be prone to contain duplicate or invalid records, which might have escaped the de-duplication procedure that we conducted.

The top panel of Figure 1 shows changes in the mean of the cross-sectional logarithmic price per square meter for the housing ask and transaction price distributions over the period from 2005 to 2015,

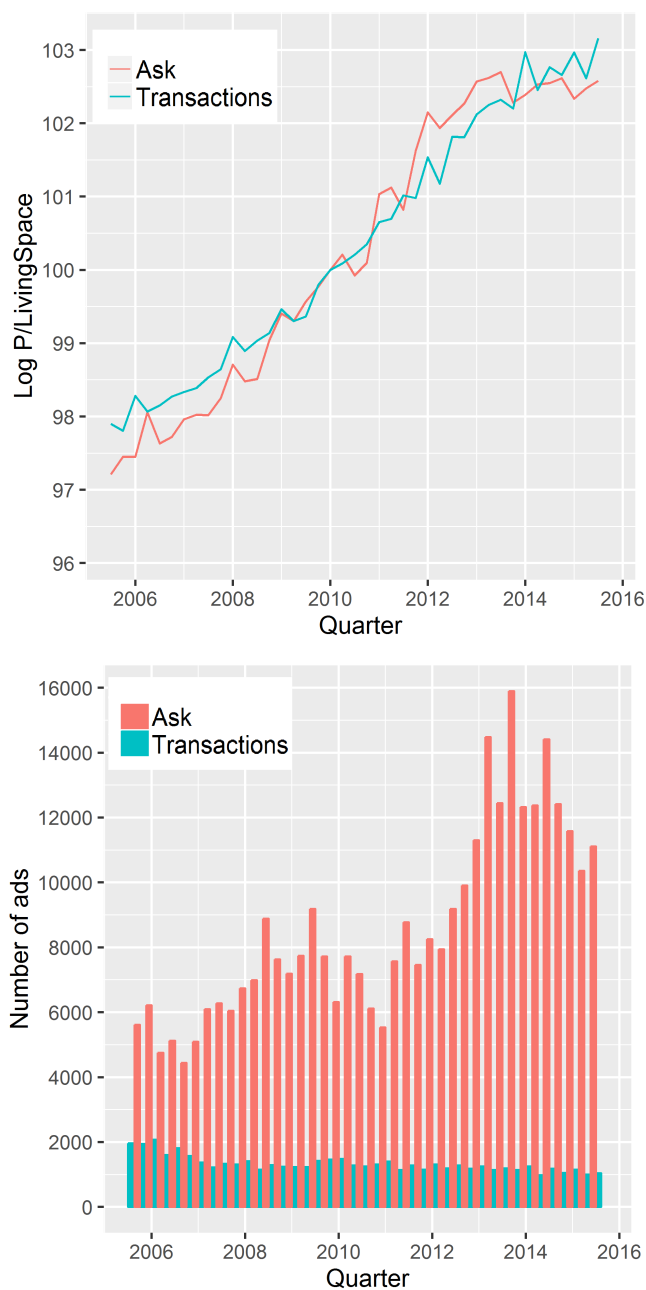


Figure 1. Top panel: mean of the cross-sectional logarithmic price per square meter for the apartment ask (comparis.ch) and transaction (SRED) price distributions over the period from 2005 to 2015. The logarithmic scale of the price means that going from 98 to 103 corresponds to a 48% relative increase of the price ($\exp(105)/\exp(100) = 1 + 0.48$), which matches the perceived relatively strong price increase over the period of the sample. Bottom panel: Ask (comparis.ch) and transaction (SRED) volume over the same period.

while the bottom panel of Figure 1 shows the ask and transaction volume over the same period. This figure shows that the mean logarithmic price per square meter exhibits a slow growth over the period of the analysis for both ask and transaction prices. Regarding the volume panel, there are large scale fluctuations for the asking volume over the period of the analysis, while the transaction volume has been decreasing since 2005 at a very slow pace.

3. Empirical strategy

In order to thoroughly study the relationship between ask and transaction prices, we follow a fourfold approach. First, we build quantile regressions and test for panel co-integration across different quantiles. Second, we estimate a dynamic panel data model to test for Granger causality between changes in ask and transaction prices. Third, we compare the time evolution of the cross-sectional distribution of ask and transaction prices. Finally, we use price-based bubble tests to study the relationship between ask and transaction prices during different price regimes. In the next section, we elaborate on the statistical tools that we employ to conduct this analysis before presenting the results in section 4.

3.1. Quantile regressions

We start the analysis with the study of the time series properties of both databases across different quantiles. To do so, we used quantile regression to compute district-level quarterly price indices corresponding to the τ -conditional quantile, allowing size and time effects to vary across quantiles.

A quantile regression estimates a conditional quantile function, in which a quantile of the conditional distribution of the response variable is expressed as a function of the covariates. It allows its estimates to vary with the corresponding quantile. This is useful when quantile effects might exist, as a result of non-Gaussian structures of the residuals and/or coexistence of several sub-populations. The conditional quantile enables us to explore differences in the development of prices across different segments of the market, as housing characteristics might be valued differently at different points of the distribution. In addition, they also control for outliers, as by construction, the quantile loss function is robust to their presence. Each district-level index for district i has the form,

$$Q_{\tau}(\log p_{t,i}|size, \mathbf{T}) = \mathbf{X}\beta_i^{\tau'} = [1, Size, \mathbf{T}] [\alpha_i^{\tau}, \beta_{size,i}^{\tau}, \beta_{\mathbf{T},i}^{\tau}]' \quad (1)$$

where $Q_{\tau}(\bullet)$ denotes the conditional quantile function for the τ -quantile (which we want to estimate), \mathbf{X} is the vector of co-variates, and $\beta_i^{\tau'}$ is the vector of corresponding coefficients, including an intercept α_i^{τ} . \mathbf{X} contains the size of the property $Size$, and a vector of time dummy variables for each quarter, denoted as \mathbf{T} . β_i^{τ} is obtained by solving

$$\beta_i^{\tau} = \arg \min_{\alpha_i^{\tau}} \rho_{\tau} \left(\frac{1}{n} \sum_{j=1}^n \log p_{t,i,j} - \left(\alpha_{0,i} + \alpha_{i,size} Size_{i,j} + \sum_{i=1}^T \alpha_{i,T_i} T_{i,j} \right) \right) \quad (2)$$

n being the total number of observations, and ρ_{τ} the check function weighting the residual μ_j ,

$$\rho_{\tau}(\mu_j) = \begin{cases} \tau\mu_j & \text{if } \mu_j \geq 0 \\ (1 - \tau)\mu_j & \text{otherwise} \end{cases} \quad (3)$$

which is asymmetric when $\tau \neq 1/2$. For $\tau = 1/2$, this recovers the conditional median function, i.e. a calibration in the sense of the medians of the residuals (or so-called L_1 norm). The resulting minimization problem is formulated as a linear function of parameters, and can be solved very efficiently by linear programming methods.

3.2. Co-integration

The use of co-integration techniques to test for the presence of long term relationship among integrated variables has enjoyed growing popularity. Given the low power of these techniques when applied to short time series, a natural extension has consisted of expanding them to panel data, while allowing as much as possible heterogeneity of the individual time series.

In order to formally test for co-integration between ask and transaction prices, we employed the sets of statistics proposed by Westerlund (2005) and Pedroni (2004). They are designed to test the null hypothesis of no co-integration between time series x and y , containing N cross-sectional units, by inferring whether the residuals of a regression of y on x contain a unit root or not. Our motivation to employ multiple statistics was to examine the robustness of our findings. Specifically, consider the least square regression,

$$y_{i,t} = d_i \hat{\delta}_i + x_{i,t} \hat{\beta}_i + \hat{e}_{i,t} \quad (4)$$

where $i = 1 \dots N$ denotes the cross-sectional units, and d_i is a vector of deterministic components with coefficients $\hat{\delta}_i$. The residuals of equation 4 are stationary when x and y are co-integrated. Thus, testing the null hypothesis of no co-integration is equivalent to testing the regression residuals for a unit root. Equation 4 is able to accommodate individual specific short-run dynamics, individual specific fixed effects, as well as deterministic trends, and it does not constrain the slope coefficients to be the same across cross-sectional units.

Westerlund (2005)'s variance ratio tests might be regarded as panel data generalizations of (Breitung, 2002) and are based on the value taken by the autoregressive parameter ρ_i in Equation 5:

$$\hat{e}_{it} = \rho_i \hat{e}_{it} + u_{it} \quad (5)$$

Consider $\hat{E}_{it} = \sum_{j=1}^t \hat{e}_{ij}$ and $\hat{R}_i = \sum_{t=1}^T \hat{e}_{it}^2$. Then, the first statistic

$$VR_P = \left(\sum_{i=1}^N \sum_{t=1}^T \hat{E}_{it}^2 \right) \left(\sum_{j=1}^N \hat{R}_j^{-1} \right) \quad (6)$$

is constructed under the maintained assumption that the autoregressive parameter ρ_i is the same for all the units. That is, the null and alternative hypotheses are formulated as $H_0 : \rho_i = 1$ for all i versus $H_1 : \rho_i = \rho$ and $\rho < 1$ for all i . Hence, rejection of the null hypothesis should be taken as evidence of co-integration for the entire panel.

The second statistic,

$$VR_G = \sum_{i=1}^N \sum_{t=1}^T \hat{E}_{it}^2 \hat{R}_i^{-1} \quad (7)$$

is constructed under the maintained assumption that the autoregressive parameter may vary across units. Thus, the null and alternative hypotheses are formulated as $H_0 : \rho_i = 1$ for all i versus $H_1 : \rho_i < 1$ for $i = 1, \dots, N_1$ and $\rho_i = 1$ for $N_1 + 1, \dots, N$, where we require $N_1/N = \xi \in (0, 1]$ as N goes to infinity.

Hence, rejection of the null hypothesis should be taken as evidence of co-integration for a non-vanishing fraction of the panel. The asymptotic distributions of VR_P (6) and VR_G (7) are respectively

$$T^{-2}N^{-1/2}VR_P - N^{1/2}\Theta_{w,3} \implies N(0, \Sigma_{w,33}) \quad (8)$$

$$T^{-2}N^{-1/2}VR_P - N^{1/2}\Theta_{w,1}\Theta_{w,2}^{-1} \implies N(0, \phi_w\tilde{\Sigma}_w\phi) \quad (9)$$

where $\tilde{\Sigma}$ denote the upper left 2×2 sub-matrix of Σ_w and $\phi_w = (\Theta_{w,2}^{-1}, -\Theta_{w,1}\Theta_{w,2}^{-2})$. $\Theta_{w,1}$, $\Theta_{w,2}$ and Σ_w are moments of a vector Brownian motion functional, which Westerlund (2005) computes via MonteCarlo simulations. These values are constant and do not depend on the data. They only depend on whether equation 4 includes a trend or not.

Pedroni (2004) also proposes a set of statistics that supports panel and group alternative hypotheses. Let $\tilde{e}_{it} = (\Delta\hat{e}_{it}, \hat{e}_{it-1})'$ and $A_i = \sum_{t=1}^T \tilde{e}_{it}\tilde{e}_{it}'$. Then he defines the following test statistics for the null of no co-integration in heterogeneous panels,

$$Z_{\hat{\rho}_{NT-1}} = \left(\sum_{i=1}^N A_{22i} \right)^{-1} \sum_{j=1}^N (A_{21j} - T\hat{\lambda}_j) \quad (10)$$

$$Z_{\hat{i}_{NT}} = \left(\tilde{\sigma}^2 \sum_{i=1}^N A_{22i} \right)^{-1/2} \sum_{j=1}^N (A_{21j} - T\hat{\lambda}_j) \quad (11)$$

$$\tilde{Z}_{\hat{\rho}_{NT-1}} = \sum_{i=1}^N (A_{22i})^{-1} (A_{21i} - T\hat{\lambda}_i) \quad (12)$$

$$\tilde{Z}_{\hat{i}_{NT}} = \sum_{i=1}^N (\tilde{\sigma}_i^2 A_{22i})^{-1/2} (A_{21i} - T\hat{\lambda}_i) \quad (13)$$

where $\hat{\mu}_{it} = \hat{e}_{it} - \hat{\rho}_i\hat{e}_{i,t-1}$, $\hat{\lambda}_i = T^{-1} \sum_{s=1}^K w_{sK} \sum_{t=s+1}^T \hat{\mu}_{it}\hat{\mu}_{i,t-s}$ for some choice of lag window $w_{sK} = 1 - 1/(1 - K)$, $\hat{s}_i^2 = T^{-1} \sum_{t=2}^T \hat{\mu}_{it}^2$, $\hat{\sigma}_i^2 = \hat{s}_i^2 + 2\hat{\lambda}_i^2$, and $\tilde{\sigma}_{NT}^2 = N^{-1} \sum_{i=1}^N \hat{\sigma}_i^2$.

Similar to Westerlund (2005)'s statistics, rejection of the null hypothesis using $Z_{\hat{\rho}_{NT-1}}$ and $Z_{\hat{i}_{NT}}$ should be interpreted as evidence of co-integration for the whole panel, while rejection of the null using $\tilde{Z}_{\hat{\rho}_{NT-1}}$ and $\tilde{Z}_{\hat{i}_{NT}}$ should be interpreted as evidence of co-integration for a non-vanishing fraction of the panel. The asymptotic distributions of Pedroni (2004)'s statistics as $(T, N \rightarrow \inf)_{seq}$ are

$$T\sqrt{N}Z_{\hat{\rho}_{NT-1}} - \Theta_2\Theta_1^{-1}\sqrt{N} \implies N(0, \phi'_{(2)}\psi(2)'\phi_{(2)}) \quad (14)$$

$$Z_{\hat{i}_{NT}} - \Theta_2(\Theta_1(1 + \Theta_3))^{-1/2}\sqrt{N} \implies N(0, \phi'_{(3)}\psi(3)'\phi_{(3)}) \quad (15)$$

$$TN^{-1/2}\tilde{Z}_{\hat{\rho}_{NT-1}} - \tilde{\Theta}_1\sqrt{N} \implies N(0, \tilde{\psi}_{11}) \quad (16)$$

$$N^{-1/2}\tilde{Z}_{\hat{i}_{NT}} - \tilde{\Theta}_2\sqrt{N} \implies N(0, \tilde{\psi}_{22}) \quad (17)$$

where the values for ϕ_j are given by

$$\phi'_{(1)} = -\Theta_1^{-2}, \quad \phi'_{(2)} = (-\Theta_2\Theta_1^{-2}, \Theta_1^{-1}), \quad (18)$$

and

$$\phi'_{(3)} = \left(-\frac{1}{2}\Theta_2\Theta_1^{-3/2}(1 + \Theta_3)^{-1/2}, \Theta_1^{-1/2}(1 + \Theta_3)^{-1/2}, -\frac{1}{2}\Theta_2\Theta_1^{-1/2}(1 + \Theta_3)^{-3/2} \right). \quad (19)$$

As in (Westerlund, 2005), Θ , $\tilde{\Theta}$, Ψ , and $\tilde{\Psi}$ are moments of functionals that do not depend on the data, but on whether the data generating process contains a trend. We do not include a trend for the calculation of any of the statistics, but time demean the indices to take into account that cross-sectional independence, an assumption of the statistics, is arguably violated across districts of the residential Swiss Market.

3.3. Causation

We explored possible causal relationship between changes in ask and transaction prices. We considered regressions of the form

$$\Delta \log p_{i,t}^{dv} = \alpha_{i,0} + \sum_{i,k}^N \alpha_k \Delta \log p_{i,t-k}^{dv} + \sum_{i,k}^M \delta_k \Delta \log p_{i,t-k}^{iv} + \epsilon_{i,t} \quad (20)$$

where dv and iv denote the dependent and independent variables. They correspond either to ask and transactions prices or transactions and ask prices respectively, depending on the direction of the causality that we study. The test of whether iv does not cause dv is simply a test of the joint hypothesis $\delta_k = 0, \forall k = 1..M$. This can be done using standard F-tests.

The estimation of equation 20 requires more care. It is common practice to take the first-difference of the model in order to deal with the inconsistency introduced by the individual specific effects $\alpha_{i,0}$,

$$\tilde{\Delta} \log p_{i,t}^{dv} = \sum_{i,k}^N \alpha_k \tilde{\Delta} \log p_{i,t-k}^{dv} + \sum_{i,k}^M \delta_k \tilde{\Delta} \log p_{i,t-k}^{iv} + \epsilon_{i,t} - \epsilon_{i,t-1} \quad (21)$$

where $\tilde{\Delta}$ denotes the difference operation conducted to eliminate $\alpha_{i,0}$. However, OLS estimation of equation 21 is also inconsistent because the lagged dependent variables introduce correlation with the error term $\epsilon_{i,t} - \epsilon_{i,t-1}$. Therefore, we use the difference Generalized Method of Moments (GMM) estimator for dynamic models with panel data (Arellano, 2003). The GMM panel data uses lagged dependent variables as valid instruments in equation 21. For example, with $N = 1$, $\Delta \log p_{i,t-2}^{dv}$ becomes available as an instrument for $\tilde{\Delta} \log p_{i,t-1}^{dv} = \Delta \log p_{i,t-1}^{dv} - \Delta \log p_{i,t-2}^{dv}$, since $\Delta \log p_{i,t-2}^{dv}$ is not correlated with $\epsilon_{i,t} - \epsilon_{i,t-1}$.

A more efficient estimator can be obtained by using additional lags of the dependent variable. For example, both $\Delta \log p_{i,t-2}^{dv}$ and $\Delta \log p_{i,t-3}^{dv}$ might be used as instruments for $\tilde{\Delta} \log p_{i,t-1}^{dv}$. Furthermore, the number of instruments available is highest for the dependent variable observed at time t closest to the final period: in period 3, there is only one available instrument, in period 4 there are two, and so on. Arellano and Bond (1991) proposes panel GMM estimation using these wider unbalanced instrument sets, which is known as the Arellano-Bond estimator. To simplify matters, we do not employ this more efficient estimator, but explore the use of different number of instruments to check for robustness in our results.

3.4. Distributional and tail tests

In order to quantify the difference between the cross-sectional distributions of ask and transaction prices, we use a two-sample Anderson-Darling test. The test is a modification of the Kolmogorov-Smirnov

(KS) test that gives more weight to the tail of the distribution and is thus a much better choice when there is a special interest in testing the tail. We test the null hypothesis of equal yearly cross-sectional distributions of *Price/LivingSpace* in every district and use a rolling one-year window, with a step of one quarter. As we apply the test multiple times for every district and in multiple quarters, we adjusted the p-values using the Benjamini, Hochberg, and Yekutieli (BHY) procedure that controls the false discovery rate (FDR), while allowing for positive dependence among the statistics (Benjamini and Yekutieli, 2001). The FDR is the expected proportion of false discoveries among the rejected hypotheses. Tests that control the FDR are more powerful than those that control the more stringent family-wise error rate.

In addition, we analyze the tail of the distributions using the framework described by Clauset et al. (2009) for discerning and quantifying power-law behavior in empirical data. In every district and for each yearly rolling cross-sectional distributions of ask and transaction *Price/LivingSpace*, we test the null hypothesis that the data is generated from a power law distribution, against the alternative that the data is not generated from a power law distribution. Similarly, we use the Vuong's test statistic to evaluate the hypothesis that the data is generated by power law distributions against the log normal alternative (Malevergne et al., 2011). Specifically, we test the null hypothesis that both distributions are equally far from the true distribution, against the alternative that one of these distributions is closer to the true distribution, and also examine the sign of the statistic to determine whether the power law is a better alternative. As in the equal-distributional tests, we adjusted the p-values using the BHY method, since we are again in a multiple hypothesis setup.

3.5. Ask and transaction prices during bubbles

Finally, we investigate the possible biases that the use of ask prices as a proxy for transaction prices might introduce to an analysis. We employ the conditional quantile indices described in section 3.1, and, following Haurin et al. (2013), we estimate an OLS regression of the form:

$$\log p_{i,j,t}^{tx} - \log p_{i,j,t}^{ask} = \alpha + \beta_{XRATE}XRATE_{i,j,t} + \beta_{UXRATE}UXRATE_{i,j,t} + \beta_B BUBBLE_{i,t} + \beta_{BXRATE}(BUBBLE : XRATE)_{i,j,t} + \beta_{Dist} \mathbf{District}_j \quad (22)$$

as well as a logit regression of the form:

$$I(\log p^{tx} > \log p^{ask})_{i,j,t} = \Lambda(\alpha + \beta_{XRATE}XRATE_{i,j,t} + \beta_{UXRATE}UXRATE_{i,j,t} + \beta_B BUBBLE_{i,t} + \beta_{BXRATE}(BUBBLE : XRATE)_{i,j,t} + \beta_{Dist} \mathbf{District}_j) \quad (23)$$

where $i \in 0.1, 0.15, \dots, 0.85, 0.9$ denotes the i -th conditional quantile, and j and t indicate the j -th district and the t -th period, respectively. $XRATE_{i,j,t}$ and $UXRATE_{i,j,t}$ correspond to the expected and unexpected price increases, $BUBBLE_{j,t}$ is an indicator variable that denotes whether there is evidence of a bubble in district j at time t , and $BUBBLE : XRATE_{i,j,t}$ corresponds to an interaction term between the expected price increase and the bubble indicator. $\mathbf{District}_i$ is a vector of dummy variables that controls for fixed-effects at the district level, $I(\cdot)$ denotes the indicator function for the sign of $(\log p^{tx} > \log p^{ask})_{i,j,t}$, and $\Lambda(\cdot)$ is the logit function.

Equations 22 and 23 allow us to study the relationship between a bubble regime and the ask and transaction prices conditioned on previous returns; they provide insights regarding the search mechanism that characterizes the sellers' strategy in the market. As already mentioned, according to standard search models, transaction prices should rarely exceed ask prices, as this should only occur as

a result of unexpected changes in price due to exogenous demand shifts. The effect should be transient, since households expectations should eventually adjust. Hence, standard search models anticipate that the coefficient of $UXRATE_{i,j,t}$ is significantly positive, while that of $XRATE_{i,j,t}$ is not. In contrast, endogenous search models predict that during prolonged booming periods transaction prices will tend to exceed ask prices, as it is rational for households to switch to an auction-like mechanism. In this case, the ask price will tend to be used as a lower bound instead of as an upper one, and both $UXRATE_{i,j,t}$ and $XRATE_{i,j,t}$ are expected to exhibit significantly positive coefficients. Similarly, since bubbles can be seen as protracted nonlinear booms, we argue that, according to endogenous search models, ask prices should also tend to be lower than transaction prices, consistent with the idea of a very hot market. Consequently, β_B should be positive and significant.

As Haurin et al. (2013), we estimate the expected price increases as moving averages

$$XRATE_{i,j,t} = (1/K) \sum_{k=1}^K \Delta \log p_{i,j,t-k} = (1/K) \sum_{k=1}^K (\log p_{i,j,t-k} - \log p_{i,j,t-k-1})$$

and make

$$UXRATE_{i,j,t} = \Delta \log p_{i,j,t} - XRATE_{i,j,t}$$

We explore $K \in 1, 2, 3, 4$ to allow expectations to be based on different forms of moving averages and observe that this does not impact our conclusions. To determine whether there is a bubble in the market, we use the LPPLS (log-periodic power law singularity) model embedded in a JTest setup for non-nested model selection (Davidson and MacKinnon, 1981). We test the null hypothesis of no bubble in district j at time t , against the bubble alternative. This bubble detection test is based on the identification of a transient super-exponential trend in the dynamics of the log prices. It was first proposed on the basis of empirical observations in (Sornette et al., 1996; Feigenbaum and Freund, 1996), and later justified by Johansen et al. (2000) and Johansen et al. (1999) within the framework of rational expectation model of bubbles. From a theoretical view point, Johansen et al. argue that the no arbitrage condition, together with a hierarchical self-reinforcing organization of the market, and the need of investors to be compensated for the risk of the crash, generate a power law finite-time singular price dynamics as the bubble approaches its end. As a result of the positive feedbacks, such super-exponential dynamics is unsustainable as it ends in a finite time singularity, which signals a change of regime (the end of the bubble) (Sornette and Cauwels, 2015).

Take the log transaction price time series $\ln p_t$ with $t = 1, \dots, [\tau T], [\tau T] + 1, \dots, T$, $\tau \in (0, 1)$ and $[\tau T]$ denoting the greatest integer smaller than or equal to τT . $[\tau T]$ corresponds to the starting period in which the bubble is detected. If there is enough evidence to reject the null hypothesis of a non-explosive process in favor of the alternative super-exponential trend, the bubble hypothesis can be supported. Specifically, we compare a stationary $AR(1)$ process in the log returns against $\Delta \ln \hat{p}_{t_{sexp}}$, the log returns predicted by a fitted super-exponential trend in the subsample between $[\tau T]$ and T :

$$\Delta \ln p_t = \rho \Delta \ln p_{t-1} + \alpha_{sexp} \Delta \ln \hat{p}_{t_{sexp}} + \epsilon_t, \text{ for } t = [\tau T], [\tau T] + 1, \dots, T \quad (24)$$

where ϵ_t is a white noise process. The null hypothesis of no bubble after $[\tau T]$ period is rejected if the t-statistic $\hat{t}_{\alpha_{sexp}}$ for the estimate of α_{sexp} exceeds the corresponding critical value^{*†}. When the starting

* As a remark, $\rho \geq 1$ would also suggest super-exponential behavior, but we chose not to test this alternative to focus on the super-exponential trend.

† Critical values can be obtained via Monte Carlo simulations.

date of the super-exponential trend is not known, the statistic takes the following form:

$$\text{supSEXP}(\tau_0) = \sup_{\tau \in [0, \tau_0]} \hat{t}_{\alpha_{sexp}} \quad (25)$$

where τ_0 determines the interval in which the super exponential trend is tested. The super-exponential specification to obtain $\Delta \ln \hat{p}_{t_{sexp}}$ is given by the log periodic power law singularity (LPPLS) model (Filimonov and Sornette, 2013)

$$\ln p_{t_{sexp}} = A + (t_c - t)^m [B + C \cos(\omega \ln(t_c - t) - \phi)] \quad (26)$$

where $0 < m < 1$, $B < 0$, $3 < \omega < 15$, and $|C|(\omega^2 + m^2)^{1/2} \leq |B|m$. t_c corresponds to the non-random time of the termination of the bubble. These last two conditions ensure that the instantaneous expected return diverges at t_c . In practice, it does not of course, but the hypothesis is that the average price trajectory can be approximated over a time interval until close to its turning point by such a process with increasing returns. As calibration of equation 26 on quarterly data can be difficult due to the low frequency of the volatility of house prices and the relatively large number of parameters (7 in total, 3 nonlinear, 4 linear after the reformulation performed in (Filimonov and Sornette, 2013)).

4. Results

4.1. Co-integration

In this section, we formally test whether ask and transaction prices tend to move together. To do so, we create ask and transactions conditional quantile indices for 32 Swiss districts, and test for co-integration among them. We test each pair of indices individually, as well as the whole panel.

Figures 2 and 3 show the development of ask and transaction housing prices at selected districts and at the national level, respectively. The national index corresponds to an average of median logarithmic prices, comprising the 32 districts in which at least five transactions per quarter were observed. Visual inspection already suggests that transaction and ask prices are co-integrated till a discrepancy between the indices started in 2013 in which the transaction prices exceeded the ask prices at both selected districts (except for Bulach) and national levels as displayed in Figures 2 and 3. The price premium might be a consequence of the measures issued by the SNB to mitigate the bubble risk in the housing market. With prices expected to stop rising and demand remaining unassuaged, a gap between asking and transaction prices emerged in which transactions were (on average) conducted at higher prices than those originally advertised.

Table 2 reports the results for the individual district co-integration tests and for the complete panel data, when using the median property indices. The Intercept and Slope columns correspond to the estimated values of equation 4, when the deterministic component does not include a trend. For all the statistics, large negative values should be interpreted as rejection of the null hypothesis of no co-integration.

The individual tests yield mostly evidence of co-integration. On the one hand, the statistics ρ_G , ρ_P , t_{NPP} , t_{NPG} , VR_P have all negative values, and are significant at a 1% significance level. On the other hand, VR_G is mostly insignificant and we are unable to reject the null hypothesis for any of the districts. The tests applied on the whole panel suggest strong evidence of co-integration. All values are negative, and well below the critical values. This observation remains true even when demeaning the time series



Figure 2. Logarithm of the median ask and transaction prices on selected districts. The figures correspond to conditional median indices for ask and transaction log prices, computed using the quantile regression specifications described in section 3.1. As for Figure 1, the logarithmic scale of the price means that going from, say, 99 to 104 corresponds to a 48% relative increase of the price ($\exp(104)/\exp(99) = 1 + 0.48$), which matches the perceived relatively strong price increase over the period of the sample.

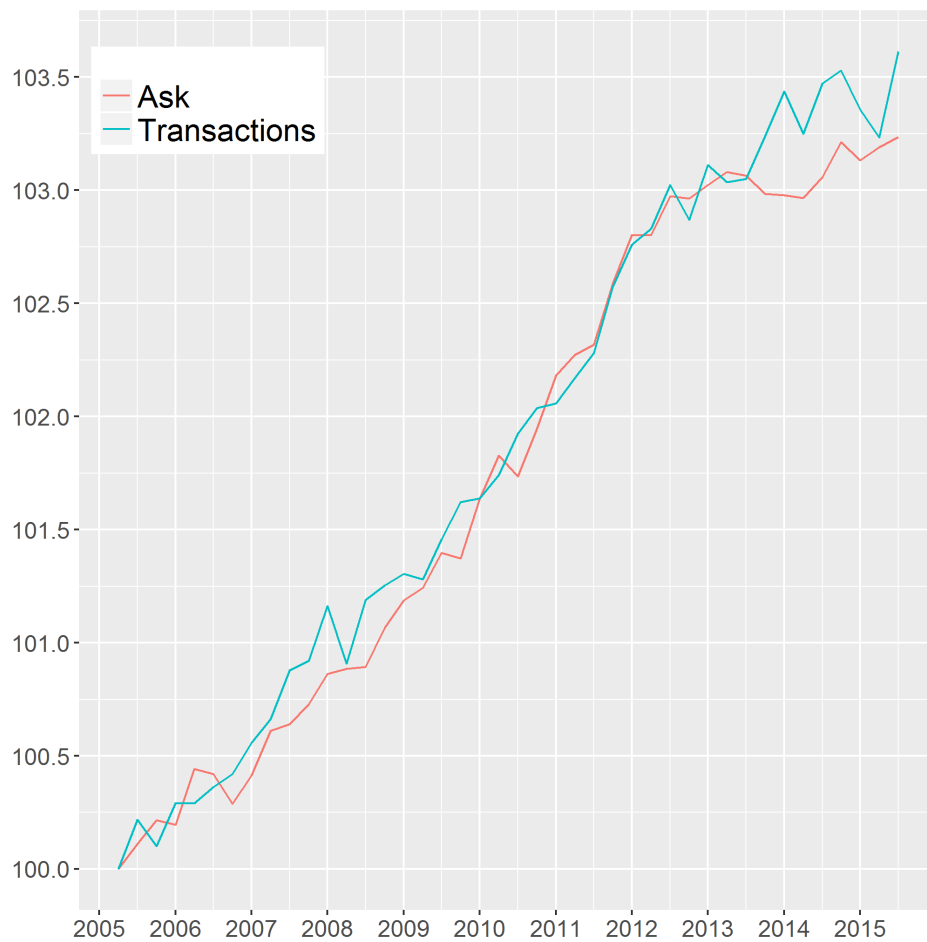


Figure 3. Logarithm of the median ask and transaction prices on the national level. The indices correspond to the respective average of the 32 conditional median indices, computed at the district level. As for Figure 1, the logarithmic scale of the price means that going from, say, 99 to 104 corresponds to a 48% relative increase of the price ($\exp(104)/\exp(99) = 1 + 0.48$), which matches the perceived relatively strong price increase over the period of the sample.

Table 2. Individual and panel co-integration test statistics between median ask and transaction property indices, for selected districts. The indices were built using median quantile regressions, as described in section 3.1. The standardized test statistics of Equations 6–7 and 10–13, also explained in section 3.2, are asymptotically normal. Critical values are thus 1.645, 1.96, 2.575 for, respectively, the 0.9, 0.95, 0.99 confidence levels.

| District | Intercept | Slope | VR_G | VR_P | ρ_G | ρ_P | t_{NPP} | t_{NPG} |
|-----------------------|-----------|-------|--------|--------|----------|----------|-----------|-----------|
| Affoltern | 3.30 | 0.76 | -1.19 | -1.01 | -3.67 | -4.48 | -3.15 | -3.37 |
| Bülach | -0.22 | 1.02 | -1.21 | -1.02 | -5.47 | -6.40 | -4.72 | -5.24 |
| Dielsdorf | 1.62 | 0.88 | -1.23 | -1.03 | -5.06 | -5.97 | -5.81 | -6.53 |
| Hinwil | 2.00 | 0.85 | -1.14 | -0.98 | -4.15 | -5.00 | -4.13 | -4.54 |
| Horgen | -0.38 | 1.02 | -1.24 | -1.04 | -4.63 | -5.51 | -5.03 | -5.61 |
| Meilen | 1.16 | 0.91 | -0.76 | -0.76 | -3.17 | -3.95 | -2.66 | -2.79 |
| Pfäffikon | 0.61 | 0.96 | -1.00 | -0.90 | -3.34 | -4.12 | -3.41 | -3.68 |
| Uster | 1.67 | 0.87 | -1.15 | -0.99 | -4.18 | -5.02 | -3.80 | -4.14 |
| Winterthur | 0.88 | 0.93 | -1.17 | -1.00 | -4.01 | -4.84 | -3.76 | -4.09 |
| Dietikon | 2.76 | 0.79 | -1.18 | -1.01 | -5.05 | -5.96 | -4.52 | -5.00 |
| Zurich | 1.83 | 0.86 | -0.93 | -0.86 | -5.54 | -6.49 | -4.03 | -4.41 |
| Zug | 1.33 | 0.91 | -1.12 | -0.97 | -6.31 | -7.31 | -4.96 | -5.51 |
| Basel-Stadt | 0.48 | 0.96 | -1.20 | -1.01 | -3.67 | -4.48 | -3.54 | -3.83 |
| Arlesheim | -0.04 | 1.00 | -1.27 | -1.05 | -5.82 | -6.79 | -7.38 | -8.40 |
| Albula | -0.70 | 1.06 | -1.03 | -0.92 | -5.73 | -6.69 | -4.73 | -5.24 |
| Prättigau-Davos | -0.46 | 1.03 | -1.03 | -0.92 | -3.90 | -4.73 | -4.45 | -4.91 |
| Surselva | 7.09 | 0.46 | -1.19 | -1.01 | -4.48 | -5.35 | -4.31 | -4.74 |
| Baden | 2.73 | 0.80 | -1.14 | -0.98 | -4.06 | -4.90 | -3.37 | -3.63 |
| Bremgarten | 2.68 | 0.80 | -0.97 | -0.89 | -2.14 | -2.84 | -2.91 | -3.09 |
| Locarno | -0.92 | 1.05 | -1.24 | -1.04 | -5.00 | -5.91 | -4.67 | -5.17 |
| Lugano | -5.25 | 1.37 | -1.26 | -1.05 | -4.44 | -5.31 | -5.68 | -6.37 |
| Aigle | 5.69 | 0.56 | -1.01 | -0.91 | -4.32 | -5.17 | -3.51 | -3.79 |
| Lausanne | -0.85 | 1.05 | -0.95 | -0.87 | -4.85 | -5.75 | -3.73 | -4.06 |
| Lavaux-Oron | 1.83 | 0.86 | -1.06 | -0.93 | -5.96 | -6.94 | -4.45 | -4.92 |
| Morges | -2.75 | 1.20 | -0.71 | -0.74 | -6.24 | -7.24 | -4.22 | -4.65 |
| Nyon | -0.98 | 1.07 | -1.00 | -0.90 | -2.36 | -3.07 | -2.21 | -2.26 |
| Riviera-Pays-d'Enhaut | 0.10 | 0.98 | -0.74 | -0.75 | -3.14 | -3.91 | -2.39 | -2.47 |
| Martigny | -2.73 | 1.20 | -1.25 | -1.04 | -5.72 | -6.68 | -5.26 | -5.87 |
| Monthey | 0.98 | 0.91 | -1.10 | -0.96 | -5.25 | -6.18 | -3.95 | -4.32 |
| Sierre | 4.29 | 0.66 | -1.21 | -1.02 | -3.71 | -4.52 | -3.97 | -4.35 |
| Sion | -1.80 | 1.13 | -1.19 | -1.01 | -5.23 | -6.16 | -5.28 | -5.90 |
| Geneva | -1.42 | 1.08 | -0.67 | -0.71 | -4.30 | -5.15 | -3.35 | -3.61 |
| No time demeaned | | | -6.11 | -5.31 | -25.62 | -31.67 | -23.61 | -25.89 |
| Time demeaned | | | -3.74 | -4.24 | -24.58 | -30.65 | -22.99 | -23.97 |

to control for dependence among districts. In this case, the absolute values of the statistics decrease slightly, but they remain strongly significant. We thus conclude that the median ask and transaction prices are co-integrated.

Table 3. Panel co-integration test statistics between ask and transaction property indices. The panel contains the conditional τ -quantile indices for the districts listed in Table 2. The indices were built using τ -quantile regressions, as described in section 3.1. The standardized test statistics of equations 6–7 and 10–13, also explained in section 3.2, are asymptotically normal. Critical values are thus 1.645, 1.96, 2.575 for, respectively, the 0.9, 0.95, 0.99 confidence levels.

| Type | Conditional Quantile τ | VR_G | VR_P | ρ_G | ρ_P | t_{NPP} | t_{NPG} |
|------------------|--------------------------------|--------|--------|----------|----------|-----------|-----------|
| No time demeaned | 0.1 | -5.82 | -5.06 | -27.22 | -35.33 | -26.26 | -27.84 |
| | 0.25 | -6.12 | -5.25 | -24.06 | -29.14 | -22.52 | -25.29 |
| | 0.5 | -6.10 | -5.31 | -25.62 | -31.66 | -23.61 | -25.89 |
| | 0.75 | -5.95 | -5.30 | -27.24 | -32.27 | -23.84 | -26.28 |
| | 0.9 | -5.39 | -4.97 | -26.49 | -32.10 | -23.22 | -25.02 |
| Time demeaned | 0.1 | -5.10 | -5.10 | -28.70 | -36.31 | -27.61 | -27.69 |
| | 0.25 | -4.12 | -4.43 | -24.49 | -29.87 | -22.08 | -23.93 |
| | 0.5 | -3.73 | -4.23 | -24.58 | -30.64 | -22.99 | -23.96 |
| | 0.75 | -3.50 | -3.78 | -24.98 | -30.40 | -22.15 | -24.47 |
| | 0.9 | -5.38 | -5.00 | -27.28 | -32.30 | -24.59 | -26.98 |

In Table 3, we explore whether the co-integration conclusion extends to additional conditional quantiles. The examined quantiles cover the 0.1–0.9 range, which correspond to the most representative segments of the market. The null hypothesis of no co-integration is rejected across all quantiles and by all tests. These results also extend to the demeaned time series, which control for the possible violation of the cross-sectional independence. Hence, there is strong evidence that ask prices reflect the dynamics of the transactions of the Swiss apartment market, at least in the studied districts. Unfortunately, we are unable to expand this analysis to the other districts, as there are simply not enough transactions. Nevertheless, in our understanding, nothing suggests that less liquid districts could behave differently.

4.2. Causality

We now examine causality between ask and transaction prices. Table 4 presents the results. Estimates are based on the GMM estimator, described in section 3.3, which uses lag variables as instruments. Reported critical values correspond to bootstrapped estimates in order to control for small sample effects. With any number of lags and in any direction, there is no evidence of a causal relationship. The null hypothesis of no Granger-causality cannot be rejected in either of the cases. The co-integration between ask and transaction prices does not appear to originate from Granger causality; exogenous factors are therefore more likely to explain the co-movement among these two variables.

As discussed in section 1 and based on these empirical results, the absence of a positive causal relationship from ask to transaction prices suggests that there is no arbitrage opportunity and the agents could not use the public information of ask prices to anticipate price rises, which in turn could lead

to further price rises. The absence of a causal relationship from transaction to ask prices justifies the validity of the bubble diagnosis based on the later. In this sense, this paper represents a strong support for the use of ask prices in real estate and bubble analysis.

4.3. Cross sectional analysis

Figures 4 and 5 present the probability density functions (PDFs) and the cumulative distribution functions (CDFs) of the cross-sectional *Price/LivingSpace* distribution, respectively, on a biannual basis from 2005 to 2015 for the ask and transaction *Price/LivingSpace*. The blue lines in Figure 5 represent the CDF of the standard log-normal distribution. The PDFs are slightly skewed to the right for both ask and transaction prices, and they shift slowly to the right over time as a consequence of the trend in the data. This might be explained by the nature of the two databases. Very high ask prices tend to be publicly unreported, as the interested buyer is typically invited to request for more information. Similarly, very high transaction prices are not reported either, as these records are removed from the transaction database to protect the identity of the buyer.

Visual inspection suggest that ask and transaction prices do not appear to follow the same distribution. To formally test for equality of distribution, we employed the Anderson-Darling tests on a yearly rolling window basis. The results, presented in Figure 6, show that ask and transaction prices have historically deviated from one another in at least 13 districts for at least 3 consecutive years. In total, we reject the null hypothesis of equal distributions 44% of the times, though towards the end of sample, it is only rejected in 5 out of the 32 districts that we analyze. Hence, although evidence from the previous section led us to conclude that ask and transaction prices are co-integrated, they seem to move away from each other over prolonged periods of time.

However, notwithstanding that in Figure 5 ask prices seem to exhibit fatter tails than the log-normal distribution, the analysis of the tail suggests that the deviations are not very strong, at least at the aggregate district level. Both ask and transaction prices present very weak evidence of heavy-tail behavior. Although the null hypothesis of a power law distribution is seldom rejected (Figures 7a and 7b), the Vuong's test statistic to compare distributions is positive and above the critical value in only 8.49% of the cases for ask prices and never for transaction prices. Thus, a log normal distribution tends to describe well or better the data, despite rejection rates of the power law hypothesis of 8.43% and 0.3%, for ask and transaction prices respectively. In addition, in Figure 7c, we observe that the evidence in favor of heavy-tailedness in ask prices is mostly restricted to four districts (Geneva, Morges, Albula and Uster). This indicates that the heavy-tails, if present, are confined to a very few regions.

4.4. Ask and transaction prices during bubbles

Figure 8 shows the quarterly results of the application of the price-based bubble test described in section 3.5 on the 32 qualified districts' transaction prices data. A blue square indicates that a bubble was detected at the given district-quarter combination. At the last quarter tested, 2015-Q2, there are 18 districts with evidence of bubbles, while over the last two years of the analysis, there are 13 districts that exhibit bubble signals consistently over every quarter. Out of the bubble districts, Geneva and Nyon stand out, as they present bubble signals in all quarters (except one quarter in the case of Nyon). On the other hand, there are 8 districts that did not show any bubble signals over the sample period.

In light of the discussion of the previous section, these results are by themselves interesting as they suggest that there is evidence of real estate bubbles despite the fact that the cross-sectional price

Table 4. Granger causality tests. The table reports GMM estimations and Granger causality tests using $\Delta \log p_{i,t}^{dv} = \alpha_{i,0} + \sum_{i,k}^N \alpha_k \Delta \log p_{i,t-k}^{dv} + \sum_{i,k}^M \delta_k \Delta \log p_{i,t-k}^{iv} + \epsilon_{i,t}$, with $M = N = 4, 6, 8$. iv denotes the independent variable. dv denotes the dependent variable. The F – statistics tests whether $\delta_k = 0, \forall k = 1..M$. Estimations employ the median conditional indices described in section 3.1.

| DV = Median ask prices | | | | | | |
|--------------------------------|-------|-------|-------|-------|-------|-------|
| | 4 | | 6 | | 8 | |
| | DV | IV | DV | IV | DV | IV |
| 1 | -0.67 | 0.03 | -0.84 | 0.02 | -0.88 | -0.04 |
| 2 | -0.45 | 0.02 | -0.76 | 0.03 | -0.83 | -0.09 |
| 3 | -0.22 | 0.03 | -0.64 | 0.04 | -0.72 | -0.14 |
| 4 | -0.13 | 0.01 | -0.58 | 0.07 | -0.70 | -0.13 |
| 5 | | | -0.45 | 0.07 | -0.62 | -0.13 |
| 6 | | | -0.37 | 0.17 | -0.65 | -0.04 |
| 7 | | | | | -0.29 | -0.20 |
| F-statistic | 0.36 | | 0.30 | | 0.25 | |
| Critical Value | 2.06 | | 1.57 | | 1.15 | |
| DV = Median transaction prices | | | | | | |
| | DV | IV | DV | IV | DV | IV |
| 1 | -0.92 | -0.15 | -1.15 | -0.09 | -1.22 | 0.04 |
| 2 | -0.66 | -0.19 | -1.16 | -0.15 | -1.26 | 0.09 |
| 3 | -0.34 | -0.22 | -0.99 | -0.22 | -1.13 | 0.06 |
| 4 | -0.15 | -0.04 | -0.82 | -0.11 | -0.98 | -0.01 |
| 5 | | | -0.54 | -0.21 | -0.73 | 0.01 |
| 6 | | | -0.17 | -0.26 | -0.42 | -0.08 |
| 7 | | | | | -0.26 | 0.23 |
| 8 | | | | | -0.18 | 0.21 |
| F-tatistic | 1.17 | | 1.24 | | 1.06 | |
| Critical value | 6.94 | | 5.55 | | 4.28 | |

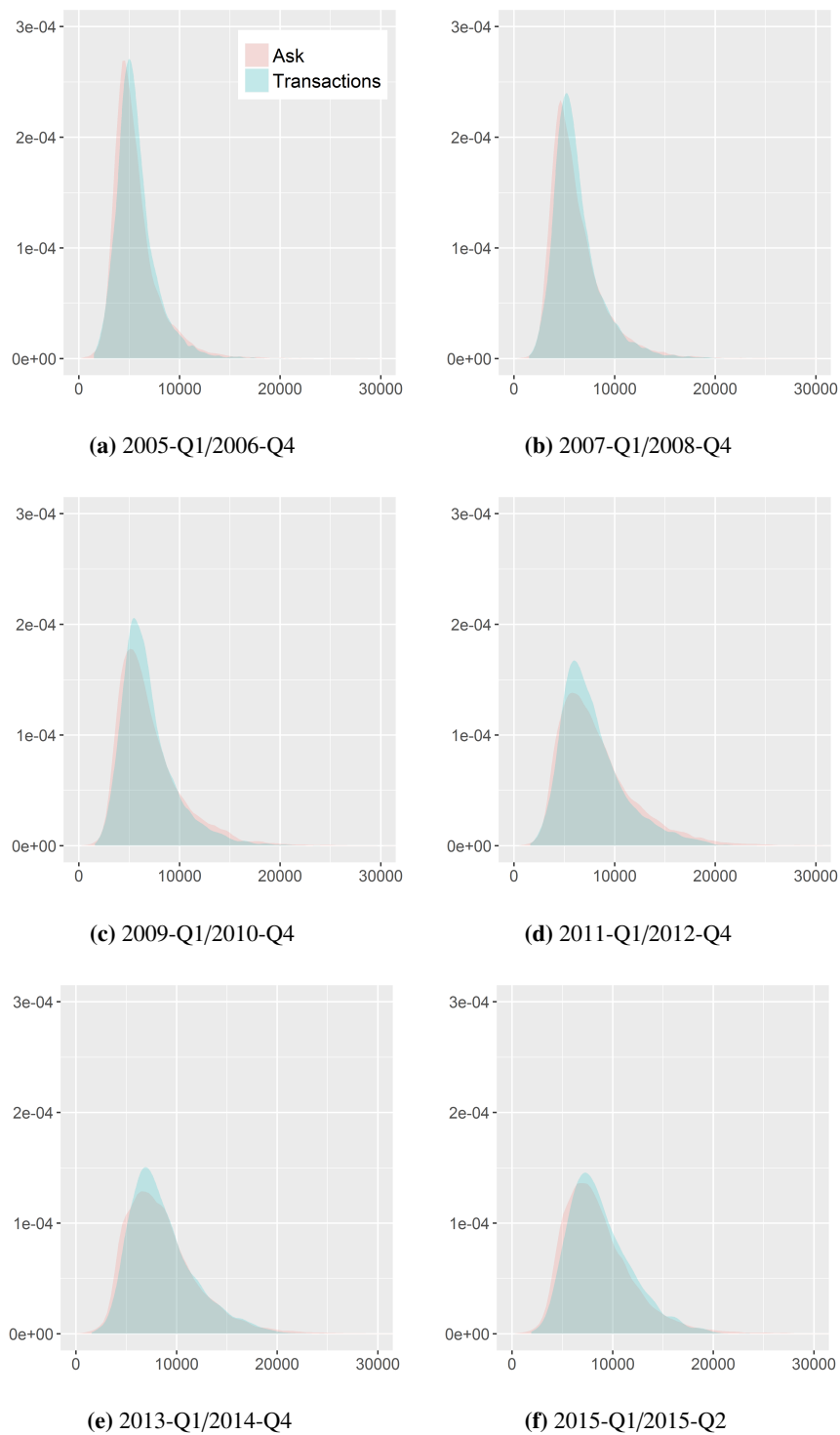


Figure 4. Probability density functions (PDFs) of the biannual cross-sectional ask and transaction $Price/LivingSpace$ of apartments, aggregated over the selected 32 districts.

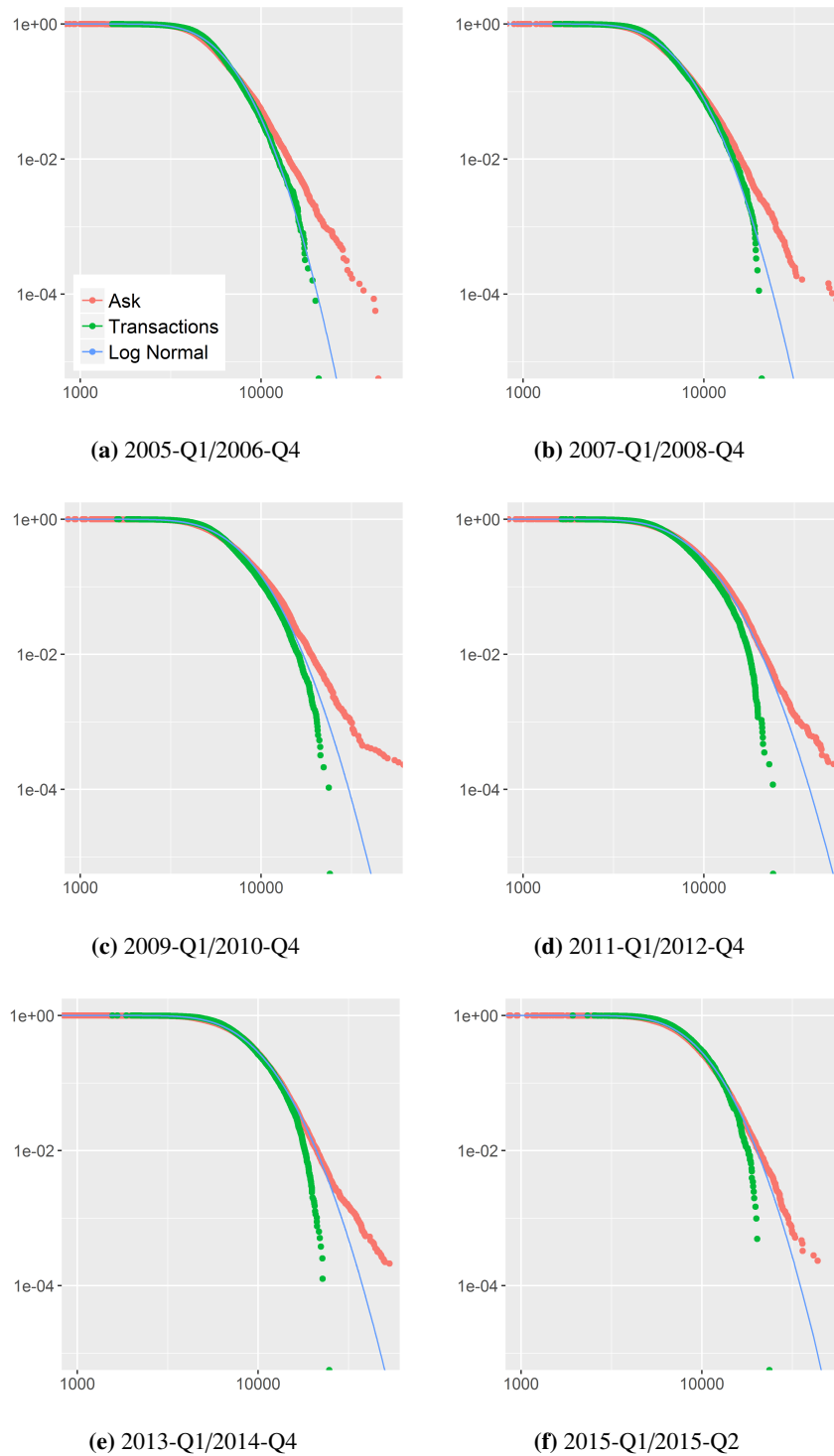


Figure 5. Cumulative distribution functions (CDFs) of the biannual cross-sectional ask and transaction $Price/LivingSpace$ of apartments, aggregated over the selected 32 districts.

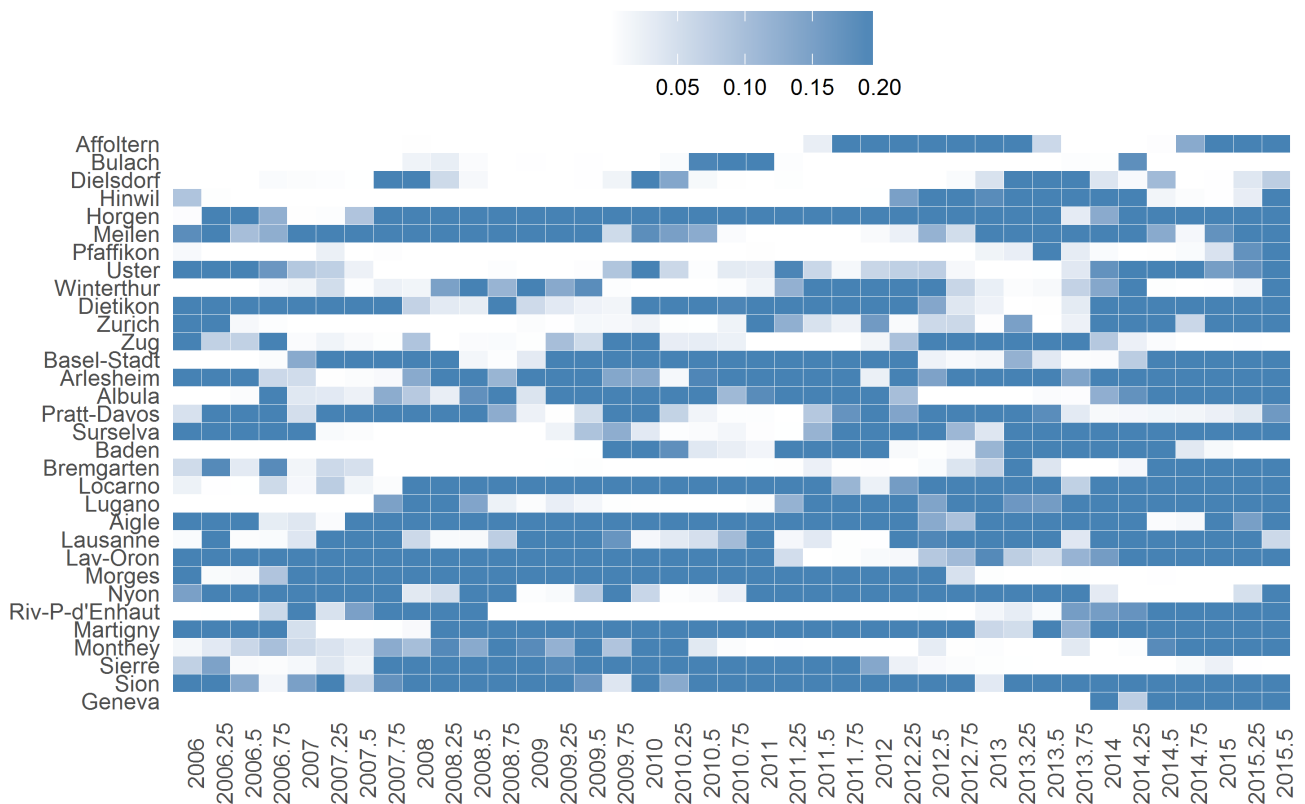


Figure 6. Anderson Darling adjusted p-values for yearly cross-sectional distribution functions of ask and transaction *Price/LivingSpace*. p-values were adjusted according to the Benjamini, Hochberg, and Yekutieli method, which controls the false discovery rate, i.e. the expected proportion of false discoveries amongst the rejected hypotheses.

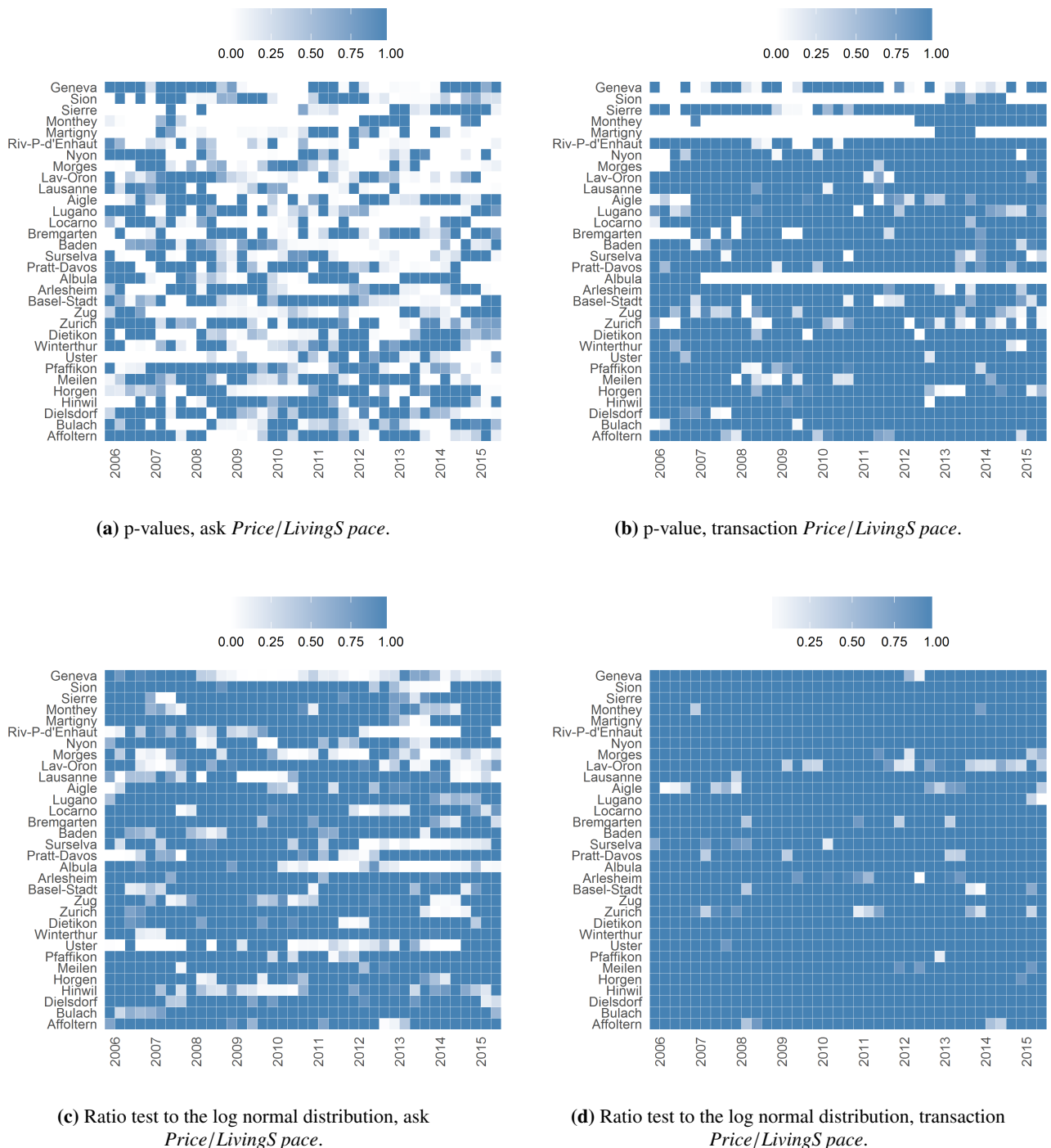


Figure 7. Adjusted p-values for power law and ratio tests, as described in section 3.4. p-values were adjusted according to the Benjamini, Hochberg, and Yekutieli method, which controls the false discovery rate, i.e. the expected proportion of false discoveries amongst the rejected hypotheses.

assumes that the transaction price rarely exceed the ask price (Horowitz, 1992). Rather, as explained by Haurin et al. (2013), they are consistent with the endogenous mechanism model in which households shift from a standard search model to an auction-like model during protracted booms.

Table 5. Summary statistics for OLS and logit regressions 22 and 23. Reported are the estimated coefficients (Est.), the standard errors (SD), the t-statistics, p-values, and bootstrapped p-values (columns $Pr(> t)^*$ and $Pr(> z)^*$).

| | $DV : \log p^{tx} - \log p^{ask}$ | | | | | $DV : I(\log p^{tx} > \log p^{ask})$ | | | | |
|------------------|-----------------------------------|--------|-----------|-------------|-------------|--------------------------------------|-------|-----------|-------------|-------------|
| | Est. | SD | \hat{t} | $Pr(> t)$ | $Pr(> t)^*$ | Est. | SD | \hat{z} | $Pr(> z)$ | $Pr(> z)^*$ |
| XRATE | 1.56 | 0.02 | 62.73 | 0.00 | 0 | 3.67 | 0.10 | 35.68 | 0.00 | 0 |
| BUBBLE | 0.19 | 0.02 | 12.35 | 0.00 | 0.01 | 0.39 | 0.05 | 7.90 | 0.00 | 0.03 |
| UXRATE | 0.63 | 0.01 | 97.86 | 0.00 | 0 | 1.43 | 0.03 | 48.46 | 0.00 | 0 |
| XRATE:BUBBLE | -0.19 | 0.04 | -4.64 | 0.00 | 0.98 | -0.72 | 0.14 | -5.03 | 0.00 | 0.58 |
| (Intercept) | -0.65 | 0.03 | -21.79 | 0.00 | | -1.64 | 0.10 | -16.59 | 0.00 | |
| Bulach | -0.04 | 0.04 | -1.07 | 0.28 | | -0.36 | 0.14 | -2.54 | 0.01 | |
| Dielsdorf | 0.66 | 0.04 | 15.85 | 0.00 | | 1.76 | 0.13 | 13.56 | 0.00 | |
| Hinwil | 0.42 | 0.04 | 10.25 | 0.00 | | 1.19 | 0.13 | 9.45 | 0.00 | |
| Horgen | 1.47 | 0.04 | 35.47 | 0.00 | | 3.69 | 0.17 | 21.85 | 0.00 | |
| Meilen | 0.24 | 0.04 | 5.79 | 0.00 | | 0.86 | 0.13 | 6.71 | 0.00 | |
| Pfaffikon | 1.15 | 0.04 | 27.57 | 0.00 | | 3.31 | 0.15 | 21.69 | 0.00 | |
| Uster | 0.67 | 0.04 | 16.24 | 0.00 | | 2.16 | 0.13 | 16.47 | 0.00 | |
| Winterthur | 0.22 | 0.04 | 5.26 | 0.00 | | 0.34 | 0.13 | 2.55 | 0.01 | |
| Dietikon | 0.56 | 0.04 | 13.64 | 0.00 | | 1.45 | 0.13 | 11.33 | 0.00 | |
| Zurich | 0.70 | 0.04 | 16.98 | 0.00 | | 1.85 | 0.13 | 14.21 | 0.00 | |
| Zug | 0.95 | 0.04 | 22.86 | 0.00 | | 2.57 | 0.14 | 18.58 | 0.00 | |
| Basel-Stadt | 1.47 | 0.04 | 35.38 | 0.00 | | 3.91 | 0.17 | 23.10 | 0.00 | |
| Arlesheim | 0.76 | 0.04 | 18.22 | 0.00 | | 1.88 | 0.13 | 14.24 | 0.00 | |
| Albula -0.49 | 0.04 | -11.63 | 0.00 | | -0.78 | 0.16 | -4.86 | 0.00 | | |
| Pratt-Davos 1.41 | 0.04 | 33.63 | 0.00 | | 3.28 | 0.15 | 21.72 | 0.00 | | |
| Surselva | -0.49 | 0.04 | -11.77 | 0.00 | | -0.37 | 0.15 | -2.44 | 0.01 | |
| Baden | 0.53 | 0.04 | 12.88 | 0.00 | | 1.35 | 0.13 | 10.68 | 0.00 | |
| Bremgarten | 1.25 | 0.04 | 29.91 | 0.00 | | 3.16 | 0.15 | 21.40 | 0.00 | |
| Locarno | 0.88 | 0.04 | 21.15 | 0.00 | | 2.20 | 0.14 | 16.11 | 0.00 | |
| Lugano | 1.84 | 0.04 | 44.32 | 0.00 | | 5.40 | 0.30 | 18.30 | 0.00 | |
| Aigle | -0.60 | 0.04 | -14.41 | 0.00 | | -0.36 | 0.15 | -2.40 | 0.02 | |
| Lausanne | 0.17 | 0.04 | 4.17 | 0.00 | | 0.47 | 0.13 | 3.61 | 0.00 | |
| Lav-Oron | 0.61 | 0.04 | 14.87 | 0.00 | | 1.52 | 0.13 | 11.48 | 0.00 | |
| Morges | 0.87 | 0.04 | 21.12 | 0.00 | | 2.02 | 0.13 | 15.19 | 0.00 | |
| Nyon | 0.60 | 0.04 | 14.32 | 0.00 | | 1.48 | 0.13 | 11.31 | 0.00 | |
| Riv-P-d'Enhaut | 0.13 | 0.04 | 3.11 | 0.00 | | 0.49 | 0.13 | 3.78 | 0.00 | |
| Martigny | 0.10 | 0.04 | 2.48 | 0.01 | | 0.55 | 0.14 | 4.03 | 0.00 | |
| Monthey | 0.77 | 0.04 | 18.46 | 0.00 | | 1.99 | 0.14 | 14.50 | 0.00 | |
| Sierre | -0.24 | 0.04 | -5.74 | 0.00 | | -0.11 | 0.14 | -0.73 | 0.47 | |
| Sion | -0.39 | 0.04 | -9.36 | 0.00 | | -0.45 | 0.15 | -3.05 | 0.00 | |
| Geneva | 0.68 | 0.04 | 16.23 | 0.00 | | 1.69 | 0.13 | 12.72 | 0.00 | |

5. Conclusion

Using two data sets consisting of ask and transaction prices in the Swiss housing market from 2005 to 2015, we studied the relationship between ask and transaction prices. We found that ask prices are co-integrated with transaction prices, so that they tend to move together in the long term. The co-integration is most likely originated from exogenous factors, as Granger-causality cannot explain it. In addition, our analysis

shows that the cross-sectional distributions of the logarithmic transaction prices adjusted for the size of the property are in general close to a log normal distribution, in most districts and during most of the sample period. The distribution of logarithmic ask price tend to exhibit fatter tails.

Based on the lack of (linear) predictability of returns, the co-movement of prices, and the almost no evidence of heavy-tailedness, we argue that ask prices are indeed informative and a sound alternative to monitor the market, especially in light of the scarcity and sparsity of transactions taking place in Switzerland. However, as we have also found evidence that the Swiss housing market appears to follow an auction-like dynamics, in which transaction prices regularly exceed ask prices, conclusions drawn from ask prices might underestimate the extent of price increases while the market is booming, and the magnitude of the correction when the market enters the bust phase of the housing cycle.

Acknowledgments

This research was supported in part by an SNF grant to the Chair of Entrepreneurial Risks at the ETH Zurich (160305).

Conflict of interest

The authors declare no conflict of interest.

References

- Albrecht J, Gautier P, Vroman S (2016) Directed search in the housing market. *Rev Econ Dyn* 19: 218–231.
- Anenberg E, Laufer S (2017) A more timely house price index. *Rev Econ Stat* 99: 722–734.
- Ardila D, Cauwels P, Sanadgol D, et al. (2013) Is there a real estate bubble in Switzerland? Diagnostic of Q4/2012. *Swiss Real Estate J* 6: 38–47.
- Arellano M (2003) *Panel data econometrics*, Oxford University Press.
- Arellano M, Bond S (1991) Some tests of specification for panel data: Monte carlo evidence and an application to employment equations. *Rev Econ Stud* 58: 277–297.
- Basten CC, Koch C (2015) Higher bank capital requirements and mortgage pricing: evidence from the Countercyclical Capital Buffer (CCB), BIS Working Paper.
- Benjamini Y, Yekutieli D (2001) The control of the false discovery rate in multiple testing under dependency. *Ann Stat*, 1165–1188.
- Breitung J (2002) Nonparametric tests for unit roots and cointegration. *J Econometrics* 108: 343–363.
- Clauset A, Shalizi CR, Newman ME (2009) Power-law distributions in empirical data. *SIAM Rev* 51: 661–703.
- Cohen W, Ravikumar P, Fienberg S (2003) A comparison of string metrics for matching names and records, In *Kdd workshop on data cleaning and object consolidation*, 3: 73–78.

- Davidson R, MacKinnon JG (1981) Several tests for model specification in the presence of alternative hypotheses. *Econometrica J Econometric Soc*, 781–793.
- Días A, Jerez B (2013) House prices, sales, and time on the market: a search-theoretical framework. *Int Econ Rev* 54: 837–872.
- Diaz J, Zhao R, Black R (1999) Does contingent reward reduce negotiation anchoring? *J Prop Investment Financ* 17: 374–379.
- DiPasquale D, Wheaton W (1996) *Urban Economics and Real Estate Markets*, Prentice-Hall, Englewood Cliffs, NJ.
- Feigenbaum JA, Freund P (1996) Discrete scale invariance in stock markets before crashes. *Int J Modern Phys B* 10: 3737–3745.
- Filimonov V, Sornette D (2013) A stable and robust calibration scheme of the log-periodic power law model. *Phys A* 392: 3698–3707.
- Haurin D, McGreal S, Adair A, et al. (2013) List price and sales prices of residential properties during booms and busts. *J Housing Econ* 22: 1–10.
- Horowitz JL (1992) The role of the list price in housing markets: theory and an econometric model. *J Appl Econometrics* 7: 115–129.
- Johansen A, Ledoit O, Sornette D (2000) Crashes as critical points. *Int J Theor Appl Financ* 3: 219–255.
- Johansen A, Sornette D, Ledoit O (1999) Predicting financial crashes using discrete scale invariance. *J Risk* 1: 5–32.
- Knight JR, Sirmans C, Turnbull GK (1994) List price signaling and buyer behavior in the housing market. *J Real Estate Financ Econ* 9: 177–192.
- Levin EJ, Pryce GB (2007) A statistical explanation for extreme bids in the house market. *Urban Stud* 44: 2339–2355.
- Malevergne Y, Pisarenko V, Sornette D (2011) Testing the Pareto against the lognormal distributions with the uniformly most powerful unbiased test applied to the distribution of cities. *Phys Rev E* 83: 036111.
- Pedroni P (2004) Panel cointegration: asymptotic and finite sample properties of pooled time series tests with an application to the ppp hypothesis. *Econometric Theory* 20: 597–625.
- Scholkopf B, Smola AJ (2001) *Learning with kernels: support vector machines, regularization, optimization, and beyond*, MIT press.
- Shimizu C, Nishimura KG, Watanabe T (2016) House prices at different stages of the buying/selling process. *Reg Sci Urban Econ* 59: 37–53.
- Sornette D, Cauwels P (2015) Financial bubbles: mechanisms and diagnostics. *Rev Behav Econ* 2: 279–305.

Sornette D, Johansen A, Bouchaud JP (1996) Stock market crashes, precursors and replicas. *J de Physique I* 6: 167–175.

Takaaki O, Takayuki M, Chihiro S (2011) The evolution of house price distribution.

Timmermans M, Heijmans R, Daniels H (2017) Cyclical patterns in risk indicators based on financial market infrastructure transaction data.

Vogler RU (2006) *Swiss banking secrecy: origins, significance, myth*, Association for Financial History (Switzerland and Principality of Liechtenstein).

Westerlund J (2005) New simple tests for panel cointegration. *Econometric Rev* 24: 297–316.

Wheaton W (1990) Vacancy, search, and prices in a housing market matching model. *J Polit Economy* 98: 1270–1292.



AIMS Press

©2021 the Author(s), licensee AIMS Press. This is an open access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>)