Housing market sentiment and policy effectiveness: Evidence from Shanghai

Zhengyi Zhou[1]

This version: Apr, 2016

Abstract With a micro-level data, this paper constructs the first sentiment index for the Shanghai housing market. This is done by creatively applying the sentiment index construction approach in the stock market to the housing market. The look-ahead-bias-free monthly index is highly correlated with official indexes about consumer and investor confidence. In-sample tests show that positive sentiment helps explain the lower returns in the future, although sentiment does not have out-of-sample forecasting ability. Furthermore, we find that loosening policies in the housing market tend to increase sentiment, while there is no evidence about sentiment drops related to tightening policies. And if a tightening policy meets with high sentiment, the housing prices will rebound after an initial drop. The rebounding is more obvious in sub-markets where housing prices tend to grow whenever sentiment increases but rarely drop when sentiment decreases. Finally, we use a simple model to illustrate why sentiment affects policy outcomes.

Keywords Housing market, sentiment, Shanghai, policy
JEL P25 L52 G14

[1] Postal address: Shanghai Advanced Institute of Finance, Shanghai Jiao Tong University. 211 West Huaihai Road, Shanghai, China, 200030.
Email: zyzhou.12@saif.sjtu.edu.cn

I. Introduction

Owner-occupied housing constitutes the largest single source of wealth in the U.S. (DiPasquale and Wheaton, 1994), and also plays a crucial in the Chinese economy. The housing bubble (Kostovetsky, 2015) preceding the 2008 financial crisis further highlights the importance of the housing market. Unlike the stock market, the housing market features high percentage of individual traders, segmentation of market, asymmetry of information, and lack of short-sale mechanism, all of which makes it highly susceptible to sentiment-induced mispricing (Clayton, Ling, Naranjo, 2009; Hui, Wang, 2014). However, perhaps because of the limited access to micro transaction data (Zheng, Chau, Eddie, 2015), the empirical study of behavior bias in the housing market is quite limited relative to the huge size of the behavior finance literature. With a micro-level data, we try to fill this gap by focusing on the sentiment in the housing market. More specifically, we construct a sentiment index for the housing market in an important Chinese city, and study how sentiment status affects the outcomes of housing market policies. Our findings help people understand the housing market dynamics, and generate meaningful implications for policy makers who want to monitor the housing market or exert interventions.

Our research goal is achieved by four steps. Firstly, we create a method to construct a sentiment index for the housing market. The information about sentiment is obtained from several proxies mimicking the ones used to the construct the stock market sentiment index in Baker and Wurgler (2006). For example, we assume that newly-opened construction area in the housing market is analogous to the number of IPOs in the stock market. From regressions of these proxies on fundamental economic indicators, we get the residuals. Then we use the partial least squares (PLS) approach in Huang, Jiang, Tu and Zhou (2014) to generate a single sentiment index from these residuals. So the index is supposed to contain information about sentiment that cannot be explained by fundamentals. It turns out that our index has a high correlation with the housing purchase confidence index, the consumer confidence index, and the investor confidence index disclosed by official statistic bureaus. Compared to these official indexes, our index has higher frequency and a specific focus on the housing market.

Secondly, we explore whether market sentiment explains and predicts future housing market returns. The housing market returns are obtained from the housing price index, which is constructed by Zhou (2016) through the repeat sales approach (Case, Shiller, 1987). It turns out that positive sentiment has some in-sample explaining power for the lower returns in the subsequent three months. However, our-of-sample tests show that sentiment rarely forecasts future returns. This is not surprising because the Chinese housing market features frequent government interventions. Their timing is unpredictable but the influence on the housing prices may be non-trivial.

Considering the significant role of government interventions in the Chinese housing market, we are stimulated to conduct the third step. We examine whether policy shocks affect market sentiment, and investigate how sentiment status affects the policy outcomes. Some evidence shows that loosening policies are followed by a sentiment increase, while the sentiment changes around tightening months are insignificant. Here tightening/loosening policies refer to the government interventions towards the housing market rather than the monetary policies. We further find that if a tightening policy meets with a high-sentiment environment, the housing price will rebound after an initial drop. Moreover, the rebounding is more obvious in places where the housing prices tend to rise whenever sentiment increases but rarely drop when sentiment decreases. The cross-sectional finding that a submarket's rebounding speed depends on its responsiveness to sentiment confirms the significant role of sentiment on policy outcomes. Kurov (2010) find that investor sentiment plays a significant role in the effect of monetary policy on the stock market, that the effect of monetary news on sentiment depends on market conditions, and that monetary policy actions in bear market periods have a larger effect on stocks that are more sensitive to changes in investor sentiment and credit market conditions. Here we look at a different market and different policies, but we are in line with Kurov (2010) when emphasizing that sentiment affects policy outcomes, that the market sentiment is more sensitive to loosening policies which often comes in bear markets than to tightening policies which often comes in bull markets, and that the submarkets have different sensitiveness to sentiment changes and thus different response to policy actions. Bernanke and Kuttner (2005) mention that the large movements in excess returns associated with monetary policy changes may reflect overreaction of stock prices to policy actions. Our findings support their conjecture that policy outcomes may be influenced by market inefficiency driven by investor behavior, although from a real-estate angle.

Fourthly, we write a model based on Burnside, Eichenbaum, and Rebelo (2015), or "BER" afterwards, so as to theoretically illustrate why sentiment affects policy outcomes. Like in BER, in our model, the agents also have heterogeneous beliefs about the future, and those with firmer views are more likely to convert others to their beliefs. Our innovation lies in two model features: (1) A new policy makes the social dynamics return to the "initial state", where most agents hold neutral but uncertain opinions about the future, a small group of agents hold non-neutral but firm opinions, and another small group of agents are firmly neutral. (2) The non-neutral opinion is persistent over time. Under this setting, if a tightening policy coincides with high sentiment, the optimistic opinion in the society is still optimistic after the announcement day. Then a big number of agents with uncertain neutral opinions will be affected by these optimistic people and become willing to pay higher price, so the housing price rebounds.

The housing market of Shanghai in China is chosen as our research target. Apart from the data availability reason, Shanghai provides a meaningful and interesting research setting. It is one of the most important cities in China, a country with the second largest GDP in the world. Despite that Shanghai's residents occupy less than 2% of China's total population, its housing market accounts for 20% of the country's total residential property value (Chen, Hao, Stephens, 2010). And at the end of 2013, the loan balance of the real estate industry in Shanghai is 12.16% of that in the whole country. So if an international investor want to understand the Chinese housing market, it is crucial to have knowledge about the Shanghai housing market. Furthermore, as a huge metropolis with about 25 million permanent residents, Shanghai shares many common features with other big cities in the world. Hence, our findings have worldwide implications. Last but not least, the 2008 crisis highlights the importance of monitoring the housing market. Now that the central government and the local government have adopted various direct and indirect tools to intervene the Shanghai housing market, the rich experience here is worth referring to if a policy maker is considering about interventions.

The contribution of this paper is varied. Firstly, by creatively applying the traditional way of sentiment index construction in the stock market, we build a sentiment index for a housing market. This index is based on market patterns and straightforward to understand. In terms of methodology, our work is complementary to the textual-analysis approach of Soo (2013) and the approach of Hui and Wang (2014) which is based on the inter-arrival time of transactions. In terms of application, we provide the first sentiment index for the housing market in a big Chinese city, which can be conveniently used in other studies about the Chinese economy. The monthly index is also valuable for policy makers in their monitoring work, and for potential house buyers in their decision-making process. Secondly, with a measure of the sentiment, we effectively test how government interventions affect market sentiment, and evaluate how the outcomes of these interventions depend on current sentiment status. This knowledge helps policy makers avoid unexpected policy outcomes. Thirdly, our cross-sectional analysis indicates that the sensitiveness to changes in sentiment varies spatially, which leads to different market responses towards government interventions. Few papers show this point empirically. Fourthly, by showing that a modified BER model can explain our empirical findings, we provide BER with some empirical evidence.

II. Literature review

The housing market features a lack of short-sell mechanism and a high percentage of individual traders, which makes it quite likely that high sentiment creates bubbles. The model of Piazzesi and Schneider (2009) shows that even a small number of optimistic investors can have a large effect on prices. By extending that model, Burnside, Eichenbaum, and Rebelo (2015) theoretically explain why some housing market booms finally burst while others don't. They introduce heterogeneous belief and communication into the model, and show that a small number of optimistic agents can generate persistent boom as long as the uncertainty in the economy is not solved and the optimistic agents hold firmer views than skeptical agents. The Chinese housing market has witnessed fast price growth during the past two decades (Lau, Li, 2006; Fang, Gu, Xiong, Zhou, 2015), and the rapidly-developing rather than steady-state economy contains much uncertainty. So the model of Burnside, Eichenbaum, and Rebelo (2015) indicates that people's optimistic beliefs are quite relevant for the long-lasting boom of the Chinese housing market.

However, relative to the rich empirical works that focus on stock market sentiment, the evidence about sentiment's impact on the housing market is quite limited. Some evidence of sentiment's price impact in the housing market includes Jin, Soydemir, and Tidwell (2014), Clayton (1997), Gallimore and Gray (2002), Clayton, Ling, and Naranjo (2009), Kaplanski and Levy (2012), etc. But none of these works use data as micro-level as ours.

In terms of sentiment index construction, we are related to Soo (2013). Based on local newspaper articles, Soo (2013) uses a text-analysis approach to build housing sentiment indexes for U.S. cities. Hui and Wang (2014) build a housing market sentiment index for Hong Kong. Their approach concerns the waiting time between every two transactions, and then use complex econometric tools to estimate the expected duration for modeling the probability of sentiment-based trading. However, regarding the duration between two transactions as the single most important information about sentiment is inappropriate in Shanghai, where the government interventions often temporarily constrain people's trading behavior but not necessarily change their sentiment. So instead, we creatively apply the traditional method of sentiment index construction in the stock-market to the housing market. We refer to Baker and Wurgler (2006) when selecting proxies that are likely to contain sentiment information. Ling, Naranjo, and Scheick (2014) also follow the framework of Baker and Wurgler (2006) to construct a sentiment index for the commercial real estate market. Their index is based on eight underlying proxies related to REITs and the principal component analysis method. Our approach is different from theirs in that we use the partial least squares (PLS) method of Huang, Jiang, Tu, and Zhou (2014) to summarize the underlying proxies into a single sentiment index so as to decrease noise. Moreover, since there were no REITs in China before 2014, our underlying proxies are different from theirs.

To the extent that we address the policy issues in the housing market, we are also related to the papers discussing housing or land policies in China. For example, Lai (1998), Huang and Clark (2002), and Sato (2006) review the development of national housing policy. Wang (2012), Cao and Keivani (2014), Zou (2014), and Deng, Shen, and Wang (2011) analyze the effectiveness of a certain reform or program related to the housing or land market. Du, Ma, and An (2011) test the impact of the Chinese land policy evolution on the dynamic relation between housing and land prices. Du and Zhang (2015) evaluate the effects of the trial property taxes on Shanghai housing prices through a counterfactual analysis. We are different from them in that we emphasize on the market sentiment and how it interacts with government interventions to generate market dynamics. Kurov (2010) studies the role of sentiment in the effect of monetary policy, but the author focuses on the stock market rather than the housing market.

III. Institutional background and data description

3.1 Government intervention in the Shanghai housing market

The Shanghai housing market is subject to interventions from the central government and the local government. Central government interventions apply to the whole country, and are often followed by a local version of detailed regulation. Sometimes the local government also gives its own interventions.

Government interventions to the housing market are usually counter-cyclical. When the market is too hot, the government will discourage housing purchase by raising the down-payment ratio, posting restrictions on mortgage access or interest rates, raising the tax rates related to sales of high-turnover houses, or forbidding non-local people from purchasing houses in Shanghai. By doing so, they hope to reduce the speculation behavior in the housing market, thus increasing housing affordability and decreasing the risk of housing bubbles. On the other hand, the government will encourage house purchase when the market is too cold. In that case, they cut the down-payment ratio, give banks more freedom in deciding the mortgage availability and interest rates, or decrease tax rates related to house sales. Their incentive for doing this stems from the government's dependence on incomes from land sales, and from the wish to avoid housing market crash.

In March, 2005, the National 8 Rules signaled the first time that housing market interventions became political actions. **Table 1** documents the policy changes since then. The last two columns display the tools used in each tightening or loosening policy. Here the tools are expressed in abbreviation; the full content can be found in **Table A1** in the **Appendix**. For example, at the end of 2008, the government decreased the down-payment requirement for the second house of a household so as to encourage house purchase. In Oct, 2010, the government restricted the number of houses that a household can buy so as to prevent the market from overheating. Totally there are 14 (5) tightening (loosening) policies with concrete tools. The months witnessing a tightening (loosening) policy are marked as tightening (loosening) months; if the policy came at the second half of a month, we mark the next month instead. A month without any policy changes are called a non-policy month. The dummies signaling tightening or loosening months will be used in our regressions later.

3.2 Data description

3.2.1 Transaction records and the repeat sales price index

We make use of the secondary market transaction data that covers from Dec, 2006 to May, 2015. Only the data of houses with at least two transaction records are accessible to us. This introduces the selection bias that tends to overstate the investment incentives in the market; we will discuss about it in section 7.4. Houses involved in the secondary market generally can be classified into two types. The first type consists of the houses built by the government in the 1990s or earlier; about 84% of the houses in our sample belong to this type. The second type consists of developer-built houses that have experienced at least two resales. These houses were built after 2004.

The variables in our data include the house ID, transaction date, and transaction prices. There are also information about house characteristics, such as house size, address, house layout, house type (apartment or villa), the buyer's identity (citizen of Shanghai, citizen of other Chinese cities, or foreigners), and the buyer's type (nature person or legal person).

The sample includes transactions in 16 plates from 7 districts. The border of plate is decided by the government in the spirit that the neighborhoods in a plate should have similar infrastructure and demography character, so it is reasonable to believe that “plate” well characterizes submarkets. These 16 plates occupy 18.19% of the total secondary market transactions in Shanghai during 2010.1-2015.5. Our sample contains totally 97,132 transactions related to 45,104 houses. All the three plates in the Huangpu District are covered. Our data of the Pudong District also have a nice coverage; the 8 plates which we have access to contain more than 80% of the total secondary market transactions in Pudong during 2010.1-2015.5. Considering that Huangpu is the best representative of the “Old Shanghai” and Pudong is the best representative of the “New Shanghai”, our data is qualified for regional comparison.

Actually, this sample is the same as the secondary market data in Zhou (2016). With this sample, Zhou (2016) constructs a repeat sales housing price index, and confirms its reliability through comparison with alternative indexes. The author explains why the secondary market data is more suitable than the primary market one for the index construction purpose and the policy analysis purpose. In our analysis, we take advantage of this index. This is why we also focus on the secondary market transactions rather than the primary market ones.

In **Figure 1**, the line marked by “X” shows our repeat-sales housing price index of Shanghai. From Dec, 2006 to May, 2015, the index rose by 242.15%, or an average monthly return of 1.23%, which is higher than the 0.88% average monthly growth of urban per-capita disposable income in Shanghai during 2006-2014. The standard deviation of the monthly index return is 4.66%, thus generating a Sharp-ratio of 0.26. The vertical lines marked the dates when a new policy came out. As can be seen from the figure, loosening policies usually appeared after a period of slow price growth, such as at end of 2008, and at the end of Sep, 2014. These loosening policies are followed by obvious housing price growth. In contrast, tightening policies failed to stop housing price growth. For example, the year 2010 witnessed frequent tightening policies, but the housing price kept growing, although with more volatility. We are also consistent with Du and Zhang (2015) that the trial property taxes introduced in 2011 rarely affect the housing price in Shanghai.

3.2.2 Macroeconomic variables

The macroeconomic variables used to form the underlying proxies for the sentiment index are from the Wind database developed by the Wind Information Co., Ltd. This database is one of the most authoritative economic data source in China, and is widely used by the academic researchers as well as the finance industry. The official confidence indexes used to evaluate the effectiveness of our sentiment index are from the National Bureau of Statistics of China and the Bureau of Statistics of Shanghai.

IV. The housing market sentiment index

4.1 Sentiment proxies

We use five proxies to form the sentiment index. The first one is *NewhouseconR*, which is the newly-opened housing construction area each month in Shanghai divided by the six-month average of newly-released residential land supply in Shanghai before that month. Baker and Wurgler (2006) mention that the IPO market is often viewed as sensitive to sentiment, and use the number of IPOs as a proxy reflecting stock market sentiment. Similarly, it is likely that developers actively initiate new projects when the housing market sentiment is high. *NewhouseconR* is scaled on the land supply in previous months because we want to highlight developer behavior rather than urban land constraints. And we focus on new-house construction rather than land purchase because a firm may buy a piece of land and wait for a good time to resale it. In this case, land purchase does not necessarily reflect the firm's confidence in the housing market. In contrast, when a developer decide to build houses on the land, it is likely that the developer is optimistic about the housing market. Considering that the speed of housing construction in China can be very fast, the optimism captured by this approach should be more related to the short-term horizon than the long-term horizon.

The second proxy is *HouseinvR*, which is defined as the housing investment in Shanghai divided by the total real estate investment in Shanghai. It is also a monthly variable. In the stock market, the share of equity issues in total equity and debt issues is a measure of financing activity that may capture sentiment (Baker, Wurgler, 2006). In terms of the housing market, it is also possible that high *HouseinvR* predicts low housing market returns. A high *HouseinvR* indicates that the real-estate developers are more willing to invest in housing than in other types of real estate such as office building, thus signaling high sentiment in the housing market relative to other kinds of real-estate markets. The sentiment of the housing market may diverge from the sentiment of the general real estate market because the demand side of the former consists of both consumers and investors, while the demand side of commercial real estate mainly consists of pure investors.

The third proxy is *MedianIntv*. For each transaction in a month, we calculate the interval between the current transaction date and the last transaction date of this house. Then the nature log of the median interval is the *MedianIntv* of that month. Burnside, Eichenbaum, and Rebelo (2015) show that as the number of optimistic buyers grows, the sellers will enjoy higher probability of selling. Fisher, Gatzlaff, Geltner, and Haurin (2003) also emphasize that the complete change in the housing market's condition needs to be tracked by not only price, but also time on the market. But just like Zheng, Chau, and Eddie (2015), the information regarding time-on-market is also unavailable in our dataset. So instead, we construct *MedianIntv*. Roughly speaking, *MedianIntv* depends on the median time length that a house-owner maintains satisfied about her current house, plus the median time it takes for a seller to find a buyer. Supposing that the former part is relatively stable, then *MedianIntv* should be a proxy for time-on-market. Hence, we expect *MedianIntv* to be negatively correlated to sentiment level.

The fourth proxy is *SaleProb*, which is the area of transacted houses in Shanghai divided by the area of houses that is available for sale in Shanghai. The area of transacted houses is a monthly variable in the Wind database. The area of available-for-sale houses is a daily variable in the Wind database; we collapse it to the monthly level by taking average. Additional to *MedianIntv*, *SaleProb* is also a proxy for time-on-market. When the sentiment is low, we should observe that only a small proportion of the available-for-sale houses are actually transacted, so *SaleProb* should be positively correlated with sentiment level. If we think in the framework of Baker and Wurgler (2006), *SaleProb* actually reveals liquidity, which can serve as a sentiment proxy according to Baker and Stein (2004). Ling, Naranjo, and Scheick (2014) use the percentage of properties sold from the NPI each quarter as a proxy for aggregate liquidity in the private commercial real estate market, and our adoption of *SaleProb* embodies the same logic.

The fifth proxy is *SMB*. At the end of each year, we calculate the quintile breakpoints of house size based on transactions in that year. Then the houses transacted in the following year are classified into five groups by comparing their size with the latest breakpoints. Small (big) houses are those belonging to the first (fifth) group. Then we calculate the repeat sales index for small and big houses, respectively. The *SMB* is defined as the return of the small-house index minus the return of the big-house index. In the corporate finance literature, it has been found that overconfident CEOs are more likely to make an acquisition (Malmendier, Tate, 2008), and that overconfident managers overinvest when they have abundant internal funds but curtail investment when they require external financing (Malmendier, Tate, 2005). If individual home buyers suffer from overconfidence as CEOs do, they may also make aggressive purchase decisions when being very optimistic. And since the government policies often associate the second house of a household with high down-payment requirement (which decrease the "internal funds" of the buyer) and high mortgage interest rate (which increase the cost of "external funds" for the buyer), an aggressive purchase decision is more likely to be realized through choosing a bigger house rather than buying more houses. Therefore, we expect that *SMB* contains sentiment information, and is negatively associated with sentiment.

Table 2 shows the average value of each proxy since 2009, the earliest year when all the variables used to calculate the proxies are available. The *NewhouseconR* was very high in 2009, coinciding with the Four-Trillion-Plan of the central government to stimulate the economy after the crisis. Also, the *SalesProb* was high in 2009-2010, indicating that the loosening policies at the end of 2008 effectively encouraged house purchases. But we should be cautious about its 2009 average value; seasonality may bias it upward because *SaleProb* is unavailable until Apr, 2009. *SMB* is negative in 2009-2010, indicating that the stimulus package during 2009-2010 effectively enhanced the market sentiment. But the lowest *SMB* value appeared in 2012, which coincides with the start of a new leadership in the central government. Perhaps many potential buyers were expecting looser housing market policies and better economic conditions at that time.

4.2 The sentiment index

We standardize each of the five proxies so that their mean is zero and standard deviation is one. To remove business cycle variation from the proxies, we regress each of the standardized proxies on eight variables indicating economic fundamentals.

These variables include the Purchasing Managers' Index (*PMI*), the average profit margin of the real estate industry in the last year (*ReProf*), *CPI*, the growth of M2 (*M2G*), and the gross industrial output value above designated size in Shanghai (*SHBigIndProd*). We also include *ReLoan*, which is defined as the domestic loans of the real estate industry in Shanghai divided by the total investment of the real estate industry in Shanghai; this variable is supposed to reflect whether it is easy for developers to get loans. The *Defaultr* is included as well, which is the yield spread between AA corporate bonds and AAA corporate bonds. The last fundamental variable is *Term*, which is the interest rate difference between the benchmark interest rate of above-five-year loans and six-month loans. Except *ReProf*, all the other seven variables are updated monthly. Then the residuals from these regressions are "clean" proxies for sentiment that are orthogonal to economic fundamentals. To iron out idiosyncratic jumps, we smooth the proxies with three-month moving average values. The smoothed "clean" proxies are marked with a superscript *c*.

To the extent that some proxies may lead others, we make a choice between the current value and the lagged value as Baker and Wurgler (2006) do. More specifically, we include the current value of $NewhouseconR^c$, $HouseinvR^c$, $MedianIntv^c$, $SaleProb^c$, SMB^c as well as their one-month lagged values in a principal component analysis. Then we calculate the correlation between the first principal component (*Prin1*) and the current and lagged values of each of the proxies. The current or lagged variable is selected as the final proxy, whichever is more correlated with the first principal component.

Table 3 shows the correlations of current and lagged proxies with the first principal. Except for $HouseinvR^c$, the current values are more correlated with the first principal component. In addition, although we don't use the principal component analysis to generate the sentiment index as Baker and Wurgler (2006) do, the *Prin1* is still supposed to catch much information about sentiment level. As expected, $MedianIntv^c$ and SMB^c are negatively correlated with *Prin1*, while the others are positively correlated with it. In other words, when the housing market sentiment is high, houses are traded more frequently, big houses are more popular, more new projects are opened, a bigger percentage of real estate investment are related to housing rather than other types of real estate such as office building, and more houses on the market can find a buyer.

Following Huang, Jiang, Tu, and Zhou (2014), we apply the partial least squares (PLS) approach to construct a look-ahead-bias-free sentiment index. This is done through two steps. In the first step, for each proxy x_i at each month *t*, regression (1) is run, where ret_s is the housing market return at time *s*, which no later than *t*. Then the series of the loading $\pi_{i,1,t}$ captures the time-varying sensitivity of each proxy x_i to the market sentiment instrumented by future housing market return, with the issue of short term reversal taken into account. Here we assume that the sentiment proxies are related to the expected housing market returns and uncorrelated with the unpredictable return shocks. One difference with the first-stage regression in Huang, Jiang, Tu, and Zhou (2014) is that we include ret_{s-1} on the right side. This is because the housing market shows negative first-order autocorrelation, which is similar to the feature of short term reversals in the stock market (Da, Liu, and Schaumburg, 2014; Huang, Liu, Rhee, Zhang, 2010). According to Zhou (2016), the AR(1)-GARCH(1,1) regressions of the monthly housing market returns show that the AR(1) term has a coefficient of -0.4483 (p<0.0001). The feature of short term reversal is in contrast to the western housing markets that feature positive autocorrelation, but is closer to the Hong Kong housing market which has negative autocorrelation (Quan, Titman, 1999). The future return is not a clean instrument for the current sentiment unless the short-term reversals are taken into account through controlling for the current returns.

$$x_{i,s-1} = \pi_{i,0,t} + \pi_{i,1,t} ret_s + \pi_{i,2,t} ret_{s-1} + u_{i,s-1}\,, \quad s \le t \tag{1}$$

In the second step, for each month *t*, the cross-sectional regression (2) is run. The independent variable is the loading obtained in the first step. Then the slope *S* is the estimated market sentiment.

$$x_{i,t} = c_t + S_t \widehat{\pi_{i,1,t}} + v_{i,t} \tag{2}$$

In **Figure 1**, the line marked with dots shows the sentiment index. Since we require that the first-step regression has at least 10 observations, the sentiment index covers from Mar, 2010 to May, 2015. The mean (median) is 0.09 (0.03), the maximum (minimum) is 0.92 (-0.60), and the standard deviation is 0.31. The change in sentiment, or the first difference of the sentiment index, is positively correlated with the contemporaneous housing market return; the correlation is 0.26 with a p-value of 0.0425.

To confirm the reliability of our monthly sentiment index, we compare it is official confidence indexes. Since some of the official indexes are quarterly, we convert them into monthly by assigning the quarterly value to all months in the quarter. Considering that the housing market is supported by both consumption incentives and investment incentives (Han, 2013; Miller and Pandher, 2008), we include both consumer confidence indexes and investor confidence index in the analysis. As **Table 4A** shows, our sentiment index is significantly and positively correlated with the consumer confidence index of Shanghai, the investor confidence index of China, and the investor confidence index about domestic economic policies. Its correlations with the consumer confidence index of China and with the investor confidence index about domestic economic fundamentals are also positive. More importantly, its correlation with the housing purchase confidence index of Shanghai is as high as 0.45 (p=0.0007). Beginning in Mar, 2011, this quarterly confidence index is similar to the GTTB (i.e. Good time to buy) index mentioned in Croce and Haurin (2009), which they suggest be added into the list of leading indicators of the housing market. And when the housing purchase confidence index of Shanghai is negatively correlated with 4 out of the 5 other confidence index in 2011, our sentiment index still has a correlation of 0.22 with the housing purchase confidence index, as **Table 4B** shows. The above evidence shows that our index is strongly capable at capturing the housing market sentiment, even when the housing market sentiment diverges from the general confidence of consumers or investors.

4.3 Comparison with alternative methods

We build three alternative sentiment indexes. The first one is based on principal component analysis as in Baker and Wurgler (2006). More specifically, the first principal component of $NewhouseconR^c$, one-month lag of $HouseinvR^c$, $MedianIntv^c$, $SaleProb^c$, and SMB^c is regarded as a measure of sentiment, which we denote as S^{PCA}.

The second one is a look-ahead-bias-free index based on principal component analysis, which we denote as S^{BFPCA}. To calculate its value at time t, we use information up to time t only, so each period corresponds to a process of principal component analysis.

The third one is a full-sample index based on the PLS approach. Instead of calculating time-varying loadings in the first step, we use the full sample to obtain five loadings. So for each cross-sectional regression in the second step, the independent variable takes the same value, and the slope is a measure for the sentiment of that period. We denote this measure as S^{FSPLS}.

Our sentiment index S is advantageous over S^{PCA} and S^{BFPCA} in that it can get rid of the error component in the five sentiment proxies that is irrelative to returns. From **Figure 2**, we can see that S^{PCA} and S^{BFPCA} are more volatile than S, indicating the noise in them. Also, S is advantageous over S^{FSPLS} in that it is look-ahead-bias-free, thus being more practical. **Table 5A** gives the pairwise correlation among these four sentiment measures. The two measures using full-sample, S^{PCA} and S^{FSPLS}, are highly correlated. The two look-bias-free measures, S^{BFPCA} and our S, are also significantly and positively correlated. However, S^{PCA} and S^{BFPCA} have low correlation; S^{FSPLS} and S are also only modestly correlated. This is a comfort finding because it means the big difference among the measures is driven by the sample period from which we absorb sentiment information, rather than by the econometrical method.

Table 5B displays the correlation of each sentiment measure with the official confidence indexes. S^{PCA} has a significant correlation only with the consumer confidence index of Shanghai, but with the wrong sign. S^{BFPCA} has a positive correlation with the consumer confidence index of Shanghai, which is significant at the 10% level. But it has no significant correlation with other confidence indexes, including the housing purchase confidence index. S^{FSPLS} is highly correlated with the investor confidence indexes, especially the investor confidence index about domestic economic fundamentals. But this is not enough, considering that houses are important durable consumption goods in addition to investment tools. Even worse, S^{FSPLS} is negatively correlated with the housing purchase confidence index. Generally speaking, in terms of simultaneously capturing sentiment in the consumption incentive and the investment incentive, none of the three alternative sentiment measures performs as well as our S measure.

V. Predicting future returns

So far, there is still no agreement in the literature on whether sentiment can forecast asset returns, at least across various sentiment proxies and over various horizon. For the stock market, some papers find such evidence (Feldman, 2010; Brown, Cliff, 2005), while some don't (Brown, Cliff, 2004; Gupta, Hammoudeh, Modisen, Nguyen, 2014). Baker and Wurgler (2007) argue that stock market crashes tend to occur in high sentiment periods, but the timing of the crashes within these periods is hard to predict. The same thing is true for the housing market; there has been substantial debate about whether consumers' attitudes and expectations should or could have an independent effect on predictive accuracy in the housing market (Croce, Haurin, 2009). The Shanghai housing market hasn't witnessed a crash event till now, so it may be especially hard to find the predictive power of sentiment here. Nevertheless, we still examine whether sentiment can explain future housing market returns through in-sample tests and whether it has predictive power through out-of-sample tests. After all, we assume that future housing market returns is a proxy for current sentiment when making the sentiment index.

5.1 In-sample explanatory power of sentiment

We regress future housing market returns on the current sentiment level, as formula (3a) shows; the key independent variable is our sentiment index *S*. The dependent variable is the cumulative return from the month *t*+*a* to the month *t*+*b*. And unlike Zheng, Chau, and Eddie (2015), we additionally control for housing market seasonality. The *Spring* (*Autumn*) on the right side is a dummy that equals to one if the month is January or February (September or October). The cold weather and the Spring Festival in January or February may delay the housing purchase behavior. On the other hand, the Chinese have an old saying about the housing market: Golden September and silver October, which means that the housing market is very hot during these two months. Finally, we control for returns in month *t* because the housing market shows negative first-order autocorrelation.

$$R_{[t+a,t+b]} = \alpha + \beta_1 S_t + \beta_2 Spring_t + \beta_3 Autumn_t + \beta_4 ret_t + \varepsilon_t \tag{3a}$$

Table 6A shows the regression results. No matter we look at horizon of one-month ahead, three-month ahead, six-month ahead, nine-month ahead, or one-year ahead, the sentiment cannot explain future returns. In contrast, the ret_t term always has a significantly negative coefficient, except for the nine-month horizon. For example, an increase of 1% in the current month's return will predict a drop of 0.40% in the next month.

Considering that optimistic and pessimistic sentiment may have different explaining power, we discompose *S* into two parts as shown in (3b). More specifically, *PosiS* (*NegaS*) equals to *S* when *S* is positive (non-positive), and zero otherwise. **Table 6B** tells that for the one-month and three-month horizon, the positive sentiment *PosiS* has some explaining power for future returns. The negative coefficient indicates that high sentiment is likely to be followed by lower future returns. This provides some justification for using future housing market returns as a proxy for current sentiment. On the other hand, *NegaS* always has an insignificant coefficient, indicating that low sentiment rarely have any explaining power for future returns. Considering the lack of short-sale mechanism in the housing market, our result is consistent with Stambaugh, Yu, and Yuan (2012) that anomaly is stronger following high levels of sentiment in the stock market due to short-sale impediments.

$$R_{[t+a,t+b]} = \alpha + \beta_1 PosiS_t + \beta_2 NegaS_t + \beta_3 Spring_t + \beta_4 Autumn_t + \beta_5 ret_t + \varepsilon_t \quad \text{(3b)}$$

5.2 Out-of-sample predictive power of sentiment

In this part, we examine whether a prediction for future returns based on current sentiment outperforms the historical average. We use the R_{OS}^2 statistics in Campbell and Thompson (2008) to conduct out-of-sample analysis. Its calculation is shown in (4), where $\hat{r_t}$ is the return predicted by sentiment, and $\overline{r_t}$is the historical average return before time *t*. To get the predicted return of time *t*+1, we firstly run regression (5a) using information available up to time *t*. Then the predicted return of time *t*+1 is calculated as (5b) shows.

$$R_{OS}^2 = 1 - \frac{\sum_{t=1}^{T}(r_t - \hat{r_t})^2}{\sum_{t=1}^{T}(r_t - \overline{r_t})^2} \quad \text{(4)}$$

$$r_s = c_t + \beta_t S_{s-1} + \varepsilon_s, \quad s \le t \quad \text{(5a)}$$

$$\widehat{r_{t+1}} = \widehat{c_t} + \widehat{\beta_t} S_t \quad \text{(5b)}$$

Table 7A displays the results. As the last row shows, the R_{OS}^2 is always negative. The correlation between the predicted return and realized return is also negative. So sentiment has little out-of-sample predictive power for future returns, and hardly beats the historical average. Even if we use *PosiS* and *NegaS* instead of *S* as we did in section 5.1, the out-of-sample prediction performance is still not improved.

Now that the housing market features short-term reversal, it is interesting to examine if current returns help predict future returns. For each time *t*, we run regression (6a), and calculate the predicted return as (6b) shows. The prediction performance is shown in **Table 7B**. For the one-month-ahead horizon, the addition of current returns successfully makes our prediction outperform the historical average; the R_{OS}^2 is as high as 0.1181. The correlation between the predicted returns and the realized returns is as high as 0.29 (p =0.04). In terms of the three-month-ahead horizon, although R_{OS}^2 is negative, the predicted returns have a significantly positive correlation with the realized return, while that of the historical average is negative. Hence, the feature of short-term reversals improves the predictability of the housing market returns.

$$ret_s = c_t + \beta_{1,t} S_{s-1} + \beta_{2,t} ret_{s-1} + \varepsilon_s, \quad s \le t \quad \text{(6a)}$$

$$\widehat{ret_{t+1}} = \widehat{c_t} + \widehat{\beta_{1,t}} S_t + \widehat{\beta_{2,t}} ret_t \quad \text{(6b)}$$

VI. Policy and sentiment

Zhou (2016) finds that in the tightening months, the prices and trading volume in the housing market both drop. In the subsequent months, prices and trading volume rebound, and volume rebound quicker than prices. In this section, we examine if the rebounding is associated with the sentiment status in the policy months.

In regression (7), on the right side we include the sentiment index *S* and its interaction terms with the *Tight* and *Loose* dummies; *Tight* (*Loose*) is 1 if the month is a tightening month, and 0 otherwise. The coefficients of the *Tight* (*Loose*) dummies will tell whether the subsequent returns are higher or lower in the post-announcement-periods, and those of the interaction terms will tell whether the subsequent returns depend on the sentiment environment at the announcement month. Again, we take into account seasonality and short term reversals in the regression.

$$R_{[t+a,t+b]} = c + \beta_1 S_t + \beta_2 Loose_t + \beta_3 S_t \times Loose_t + \beta_4 Tight_t + \beta_5 S_t \times Tight_t \qquad (7)$$
$$+\beta_6 Spring_t + \beta_7 Autumn_t + \beta_8 ret_t + \varepsilon_t$$

Table 8A shows the results corresponding to horizons ranging from one-month ahead to one-year ahead. Since the sentiment index covers 2010/03-2015/05, no loosening months will be involved if we look at nine-month or one-year horizon, so no variables involving *Loose* are included in the last two regressions. And if we look at the three-month or six-month horizon, there is only one loosening month during the testing period, so we have to abandon $S\times Loose$ so as to keep *Loose* in the regression. Similar to the results in **Table 6A**, *S* still does not show explaining power for future returns. However, for the six-month horizon, $S\times Tight$ has a significantly positive coefficient, and for the 12-month horizon, the coefficient is marginal significant (p=0.11). This means if the government enacts a tightening policy when the sentiment is high, the returns will rebound in the next 6 months. In other words, given that the sentiment is high, the market returns after a tightening month will be higher than in normal times.

We look at separate plates to figure out whether or not the above patterns are driven by one or two plates. To begin with, we make sub-indexes for each plate; it is required that the degree of freedom is at least 1000 in the regression of the index-making process so as to ensure the index quality. Then there are 12 plates for which we can make sub-indexes. We plot the returns of the sub-indexes around the tightening months. As **Figure 3a** shows, six out of the twelve plates experienced negative returns during the tightening months, but the returns became positive again in the next month. In other words, the prices first dropped, and then began growing again. **Figure 3b** gives the pattern about the other 6 plates. Four of them didn't experienced negative returns during the tightening months. Only in the Lujiazui Plate and the Chuansha Plate, the price drop in the tightening months was followed by another drop in the next month. The dynamics of the 12 submarkets' indexes shows that the "drop-rebound" phenomenon is not driven by one or two plates.

Now we repeat regression (7) for each plate and focus on the six-month horizon; this is the horizon where $S\times Tight$ has a significant coefficient in **Table 8A**. Since now only one loosening month is included in our sample period, we drop $S\times Loose$ so as to keep *Loose* in the regression. As **Table 8B** shows, $S\times Tight$ has a positive coefficient in 8 out of the 12 regressions, and 3 of them are significant: Lujiazui, Waigaoqiao, and Yangjing. So although the "drop-rebound" pattern is common, the role of sentiment in the policy outcomes is associated with certain spatial features. One plausible aspect is the submarkets' sensitiveness to sentiment. For example, the places attracting wide attention may be more susceptible to sentiment.

To pin down the price sensitiveness to sentiment changes, we run regression (8) for each plate. Here *posichg* (*negachg*) is the change in sentiment from last month if the change is positive (*non-positive*), and zero otherwise. The coefficient β_1 and β_2 tells the responsiveness of the housing price to positive and negative sentiment changes, respectively. **Table 9** shows the estimated β_1 and β_2 for each plate. In the table, the plates are ranked first by β_1 in descending order and then by β_2 in ascending order. Among the 12 plates, the plate "Lujiazui" and the plate "Waigaoqiao" have very high β_1 and very low β_2. Lujiazui is the famous financial center of Shanghai. Waigaoqiao is famous for its tariff-free zone and free trade zone. Recall that Lujiazui and Waigaoqiao are among the plates where the sentiment plays a very significant role in the policy outcomes. So it is confirmed that sentiment matters for policy outcomes. Otherwise, we should not find such clear relation between the rebounding after tightening months and the responsiveness to sentiment in normal times.

$$ret_{i,t} = c_i + \beta_{i,1} posichg_t + \beta_{i,2} negachg_t + \beta_{i,3} ret_{i,t-1} + \beta_{i,4} Spring_t + \beta_{i,5} Autumn_t + \varepsilon_{i,t} \quad (8)$$

But why should high sentiment weaken the effectiveness of tightening policies? One possible explanation is the persistency of optimism. Optimism doesn't easily disappear; optimistic investors will always find a way for themselves. For example, when the policy limited the number of houses that a household could buy, some couples even divorced so as to buy an additional house, promising each other to marry again after the purchase.

To confirm the persistency of sentiment, we examine the sentiment changes around policy months. So the regression (9) is run. The dependent variable is the average sentiment in the post-announcement periods minus the average sentiment in the pre-announcement periods. In terms of horizon, *m* ranges from 1 to 6 so that at least a loosening month is involved.

$$S_{[t+1,t+m]} - S_{[t-m,t-1]} = c + \beta_1 Loose_t + \beta_2 Tight_t + \beta_3 Spring_t + \beta_4 Autumn_t + \varepsilon_t \quad (9)$$

From **Table 10**, we can see that *Loose* always has a positive coefficient. It is significant when *m*=3, 4, 5, and marginally significant when *m*=6. Although there is only two (one) loosening months when *m*=1 (*m*>1), this result still provides some evidence that loosening policies tend to lift the market sentiment. On the other hand, *Tight* has an insignificant coefficient for all those horizons. This confirms our conjecture about the persistency of optimism. The bottom line is, a tightening policy can hardly cool down the market sentiment within six months. Considering the counter-cyclical nature of the interventions, the above result is consistent with Kurov (2010) in that policy shocks have a stronger impact on investor sentiment in bear rather than bull market periods. Another finding is that the *Spring* dummy has a significantly positive coefficient when *m*≥2, which means the housing market sentiment usually rises after January and February.

VII. Robustness checks

7.1 Alternative horizon choice in the PLS process

In the first stage of our index construction process, we absorb information about the sentiment of month *t* from the housing market return of the month *t*+1. Then there comes the question whether using longer-horizon returns to estimate current sentiment can improve the sentiment index. So here we consider alternative horizons, and use the cumulative return from month *t*+1 to month *t*+*m* to instrument the sentiment in month *t*.

Table 11 shows how the sentiment indexes estimated by alternative horizons are correlated with the official confidence indexes. If we use the cumulative return in the next two months to instrument current sentiment, then the sentiment index is only significantly correlated with the consumer confidence index of Shanghai. If we use the next three months instead, then the sentiment index is significantly correlated with the consumer confidence index of Shanghai and of China, but it is negatively correlated with the housing purchase confidence index. Hence, using future returns in longer horizons does not improve the performance of the sentiment index.

7.2 Controlling for market returns in the cross-sectional analysis

In section VI, we argue that the rebounding after tightening policies is more obvious in plates where the housing prices easily goes up whenever sentiment increases but rarely drop when sentiment decreases. In the regression that tests each plate's sensitiveness to sentiment changes, we only include sentiment changes, lagged return, and seasonality dummies. Stemming from the spirit of the CAPM model and following the practice of Zheng, Chau, and Eddie (2015), here we include the market return as a risk factor, and run regression (10). The *MKT* variable is the return of the overall housing price index which is calculated from transactions in all plates.

$$ret_{i,t} = c_i + \beta_{i,1} posichg_t + \beta_{i,2} negachg_t + \beta_{i,3} ret_{t-1} + \beta_{i,4} Spring_t + \beta_{i,5} Autumn_t + \beta_{i,6} MKT_t + \varepsilon_{i,t} \quad (10)$$

Table 12 shows the coefficients of *posichg*, *negachg*, and *MKT*. The *rank_p* variable gives the descending rank of each plate in terms of β_1 value, the *rank_n* variable gives the ascending rank of each plate in terms of β_2 value, and *score* is the average of these two ranks. A small *score* means the housing prices easily go up whenever sentiment increases but rarely drop when sentiment decrease. It turns out that the 'Lujiazui" plate and the "Waigaoqiao" plate have the lowest *score*, thus confirming our previous findings.

7.3 The distinct role of sentiment

When making the sentiment index, we project the sentiment proxies on fundamental variables and use the residuals to form the final index. To confirm that the sentiment index represents information distinct from economic fundamentals, here we test if the final index can be explained by several groups of fundamental variables.

The first group includes the eight variables that we include as independent variables when calculating the "clean" sentiment proxies from the raw proxies: *PMI*, *Reprof*, *CPI*, *M2G*, *SHBigIndProd*, *ReLoan*, *Defaultr*, *Term*. They are defined at the beginning of section 4.2.

The second group includes 12 dummies that represent the 12 months in a year. In **Table 10,** *Spring* has a significantly positive coefficient when *m*>1**,** indicating that the sentiment generally goes up after January and February. Here we consider the seasonality issues more closely. Our index will be more interesting if it is not driven by something related to calendar.

The third group includes income growth and demographic dynamics. *IncomeG* is defined as the growth of the per capita disposable income of urban residents in Shanghai. For demographics, we consider the growth of total population in Shanghai (*TotpopG*), the growth of the number of households in Shanghai (*HousehdG*), the growth of urban population in Shanghai (*UrbanpopG*), the growth of water-users in Shanghai (*WateruserG*), the growth of population with Shanghai "Hukou" (*HujipopG*), and the growth of male-to-female ratio among the people with Shanghai "Hukou" (*GenderG*). Among the demographic variables provided by the National Bureau of Statistics and the Ministry of Public Security, these variables have a continuous record since 2005. Since the income and demographic variables are all yearly while our sentiment index is monthly, we assign the yearly value of these variables to each month in that year.

Then we run the following regression. *PopuG* stands for *TotpopG*, *HousehdG*, *UrbanpopG*, *WateruserG*, or *HujipopG*; since these five demographic variables are highly correlated, we put them into the regression one by one. The omitted month dummy is *Dec*; we use December as the benchmark.

$$\begin{aligned}
S_t &= c + \beta_1 IncomeG_t + \beta_2 PopuG_t + \beta_3 GenderG_t \\
&+ \beta_4 PMI_t + \beta_5 Reprof_t + \beta_6 CPI_t + \beta_7 M2G_t + \beta_8 SHBigIndProd_t + \beta_9 ReLoan_t + \beta_{10} Defaultr_t + \beta_{11} Term_t \\
&+ \beta_{12} Jan_t + \beta_{13} Feb_t + \beta_{14} Mar_t + \beta_{15} Apr_t + \beta_{16} May_t + \beta_{17} Jun_t + \beta_{18} Jul_t + \beta_{19} Aug_t + \beta_{20} Sep_t + \beta_{21} Oct_t + \beta_{22} Nov_t + \varepsilon_t
\end{aligned} \tag{11}$$

As **Table 13** shows, the eight fundamental variables in the first group rarely have coefficients that are significant at 5% level. This confirms that the sentiment index is not another face of fundamentals. The only variable with very significant coefficient is *SHBigIndProd*, suggesting that higher output level is associated with higher sentiment in the housing market.

In terms of the month dummies, *Apr* has a significant coefficient. This positive coefficient means the sentiment in April is especially high. But other month dummies have insignificant coefficients, confirming that our sentiment index reflects more than calendar effects.

Now let us look into income and demographics. The coefficient of *IncomeG* is insignificant in specification [1], [3], and [5], and significantly negative in [2] and [4]. Thus, it is unlikely that our sentiment index simply reflects shocks in the disposable income of urban families. *PopuG* has a significantly positive coefficient in [1], [2], [3], and [5]. So generally speaking, high sentiment coincides with population growth. People tend to become confident about further housing price growth when there are many people entering into the market as potential buyers. However, *PopuG* has a significantly negative coefficient in [4], where *PopuG* takes the value of *HujipopG*. People with Shanghai "Hukou" are native residents rather than new-comers. Hence, we should distinguish between two sources of population growth: new-comers from other cities and new-born babies in native families. While the former is associated with higher housing market sentiment, the latter is just the opposite. One explanation may be that when the sentiment is high, young native couples expect further housing price growth. The great pressure related to house purchase makes them delay the coming of their babies so as to make or save money. In terms of *GenderG*, it has a significantly positive coefficient except in [4]. Considering the findings in finance that men are more overconfident than women when making investment decisions (Barber, Odean, 2001), it is no surprising that higher male-to-female ratio is associated with higher sentiment.

Now that we find the sentiment index correlated with several fundamental variables like population growth, growth of male-to-female ratio, output level, and the month April, we want to examine whether sentiment has a distinct role in house pricing. So we run the regression (12). Here ΔS_t is S_t minus S_{t-1}. The results are shown in **Table 14**. Even after controlling for the above three groups of fundamental variables, the coefficient of ΔS_t is still positive and significant at 5% level. Therefore, sentiment changes and market dynamics are correlated in a way that can hardly be explained by economic fundamentals, highlighting the distinct role of sentiment for house pricing.

$$\begin{aligned} ret_t &= c + \alpha \Delta S_t + \beta_1 IncomeG_t + \beta_2 PopuG_t + \beta_3 GenderG_t \\ &+ \beta_4 PMI_t + \beta_5 Reprof_t + \beta_6 CPI_t + \beta_7 M2G_t + \beta_8 SHBigIndProd_t + \beta_9 ReLoan_t + \beta_{10} Defaultr_t + \beta_{11} Term_t \\ &+ \beta_{12} Jan_t + \beta_{13} Feb_t + \beta_{14} Mar_t + \beta_{15} Apr_t + \beta_{16} May_t + \beta_{17} Jun_t + \beta_{18} Jul_t + \beta_{19} Aug_t + \beta_{20} Sep_t + \beta_{21} Oct_t + \beta_{22} Nov_t + \varepsilon_t \end{aligned} \tag{12}$$

7.4 The selection bias

Since our transaction data only covers houses with at least two resales after Dec 2006, we need to take into account the selection bias problem. Considering that housing investors are likely to have shorter horizons than housing consumers, the repeat sales index may overstate investors' behavior in the market. And among the five raw sentiment proxies, *MedianIntv* and *SMB* are based on the transaction data, so the sentiment index may also overstate the sentiment of investors.

Zhou (2016) finds that housing consumption incentives are associated with more overreaction to policy changes than investment incentives are. This indicates that consumers are less rational than investors, and thus the sentiment issue should be more relevant for the former than the latter. Given that the transaction patterns and sentiment patterns documented in this paper tend to overstate investment incentives, it should become harder to find the relation between sentiment and market dynamics. Hence, the selection bias actually enhances our major conclusions.

VIII. A model

8.1 The Settings

By modifying the model of Burnside, Eichenbaum, and Rebelo (2015), we illustrate why sentiment is relevant for housing market dynamics and policy outcomes.

The economy is populated by a continuum of agents with measure one. All agents have linear utility and discount utility with rate β. Agents are either home owners or renters; each agent can only own one house and there is no short-selling. Assume that there is a fixed stock of houses, k<1, in the economy. There is a rental market with 1-k houses, which are produced by competitive firms at a cost of w per period, so the rental rate is always w. The momentum utilities associated with owning and renting a house are ε^h and ε^r, respectively. We assume that $\varepsilon=\varepsilon^h>\varepsilon^r-w$ so that home prices are positive. All these settings are the same as in Burnside, Eichenbaum, and Rebelo (2015), which we call BER.

As BER points out, there exist infrequent changes in the value of housing fundamentals, such as changes in regulation. The flow utility of owning a home under the current policy and the unknown future policy is denoted by ε *and* ε^f, respectively. At time zero, all agents agree that the possibility of policy change is $\emptyset$. Agents fall into three categories depending on their priors about ε^f. Following BER, we refer to these agents as "optimistic", "skeptical", and "vulnerable". Correspondingly, agent types are indexed by $j=o$, s, v and are assumed to be publicly observable. An agent of type j attaches the probability distribution function $f^j(e^f)$ to the distribution of ε^f. The expectation of optimistic agents about ε^f follows the process (13). If a policy change happens at time t, the expectation follows (14), where $E_{t-1}^o(\varepsilon_{old}^f)$ denotes optimists' expected future utility before the policy change. The time varying feature of the expectation is the point that distinguishes our model from BER's.

$$E_t^o(\varepsilon^f) = E_{t-1}^o(\varepsilon^f) + u_t\,. \tag{13}$$

$$E_t^o(\varepsilon^f) = E_{t-1}^o(\varepsilon_{old}^f) + u_t \tag{14}$$

Here u_t is a shock under the normal distribution N(0,σ). When a loosening (tightening) policy change takes place, u_t is σ (-σ). To the extent that future policy (at least its timing) is unpredictable, this assumption does not hurt the normal distribution of u_t. Basically, (13) and (14) says that optimists' belief is persistent, even if a policy change takes place.

Skeptical and vulnerable agents have neural beliefs. They expect that the flow utility of home-ownership will stay the same under the future policy:

$$E^s(\varepsilon^f) = E^v(\varepsilon^f) = \varepsilon\,.$$

Then the market sentiment S_t is defined as follows:

$$S_t = E_t^o(\varepsilon^f) - E_t^s(\varepsilon^f) = E_t^o(\varepsilon^f) - \varepsilon$$

Our definition of sentiment describes the deviation of the optimistic agents' view from that of the skeptical ones. It does not depend on the percentage of agents who are optimistic, a feature that is justified by Piazzesi and Schneider (2009) that even a small number of optimistic investors can have a large effect on prices.

Like BER, we define the fundamental value of a house for a given agent before the next policy arrives, assuming that this agent is the marginal buyer before the arrival of the next policy. The value for agent j at time t, P_t^j, is given by:

$$P_t^j = \beta\left\{\phi[E_t^j(\varepsilon^f) + \beta\frac{E_t^j(\varepsilon^f)}{1-\beta}] + (1-\phi)(\varepsilon + P_{t+1}^j)\right\} \tag{15}$$

When the optimistic agent is always the marginal home buyer, $P_t^o = P_{t+1}^o = P^o$. So we obtain:

$$P_t^o = \beta \frac{\phi E_t^o(\varepsilon^f)/(1-\beta) + (1-\phi)\varepsilon}{1-\beta(1-\phi)}$$

$$P^s = P^v = \beta \frac{\varepsilon}{1-\beta}$$

Following BER, we use the entropy of $f^j(e^f)$ to measure the uncertainty of an agent's views. A high value of e^j means agent j has high uncertainty about e^f.

$$e^j = -\sum_{i=1}^{n} f^j(\varepsilon_i^f)\ln[f^j(\varepsilon^f)]$$

Agents meet randomly at the beginning of the period. When agent l meets agent j, the agent j adopts the priors of agent l with probability γ^{lj}:

$$\gamma^{lj} = \max(1 - e^l / e^j, 0)$$

The above equation ensures that the entire population will converge to the view of the agent with the lowest entropy. We further assume that $e^s<e^v$ and $e^o< e^v$. Then we discuss two economic conditions:

Case I: $e^s<e^o$.

The law of social dynamics, which is public information, is shown by (16a) to (16c). Here o_t, s_t, and v_t stands for the percentage that each agent type occupies in the population.

$$o_{t+1} = o_t + \gamma^{ov} o_t v_t - \gamma^{so} o_t s_t \tag{16a}$$

$$s_{t+1} = s_t + \gamma^{sv} s_t v_t + \gamma^{so} o_t s_t \tag{16b}$$

$$v_{t+1} = v_t - \gamma^{ov} o_t v_t - \gamma^{sv} s_t v_t \tag{16c}$$

When a policy change takes place, the value of o, s, and v return to the initial value o_0, s_0, and v_0, and it is assumed that $o_0=s_0<< v_0$. The intuition of this assumption is that after a policy change, only a small group of people have firm views about possible future policies and the corresponding impacts, while most people are not quite sure. This is reasonable because the policy change brings a new environment, and people need time to learn about the policy impact, based on which they guess the government's next action. Then as time goes by, the number of optimistic agents will first go up, because many vulnerable agents convert to optimistic ones, while the number of skeptical ones is so small at the beginning that most optimistic agents don't have a chance to meet them. After that, the number of optimistic agents gradually decreases to zero because the conversion of vulnerable agents into the skeptical type increases the chance that an optimistic agent meets a skeptical one.

Similar to the logic in BER, before the coming of the new policy, the equilibrium price path is given by (17). Here t_1 marks the time point when the optimistic agents become the marginal buyers according to the law of social dynamic, i.e. $o_{t_1}>1-k$; t_2 marks the time point when the skeptical agents become the marginal buyers again.

$$P_t = \begin{cases} P^s + [\beta(1-\phi)]^{t_1-t}[E_t^s(P_{t_1}) - P^s], t < t_1 \\ P_t^o - [\beta(1-\phi)]^{t_2+1-t}[P_t^o - P^s], t_1 \le t \le t_2 \\ P^s, t > t_2 \end{cases} \tag{17}$$

The intuition is that, before t_1, skeptical and vulnerable agents are the marginal buyers. The housing price is their fundamental value plus the expected discounted capital gain, which will be realized if the new policy's arrival is later than t_1. When $t_1 \le t \le t_2$, the marginal buyers are the optimistic agents, and the price is their fundamental value minus the discounted expected capital loss, which will be realized if the new policy's arrival is later than t_2. After t_2, the marginal buyers are always skeptical agents, and the housing price stays constant unless the new policy arrives. Hence, before t_1, the housing price has an upward trend; when $t_1 \le t \le t_2$, the trend is downward; after t_2, the price is flat.

Case II: $e^o < e^s$.

In this case, the law of social dynamics is shown in (18a) to (18c). As time passes, the number of skeptical agents first goes up and then decreases to zero, while the optimistic type gradually occupies the whole population. After optimistic agents become the marginal buyers at t_1, the price will stay at their fundamental value until the new policy arrives.

$$o_{t+1} = o_t + \gamma^{ov} o_t v_t + \gamma^{os} o_t s_t \tag{18a}$$

$$s_{t+1} = s_t + \gamma^{sv} s_t v_t - \gamma^{os} o_t s_t \tag{18b}$$

$$v_{t+1} = v_t - \gamma^{ov} o_t v_t - \gamma^{sv} s_t v_t \tag{18c}$$

Before the coming of the new policy, the equilibrium price path is given by:

$$P_t = \begin{cases} P^s + [\beta(1-\phi)]^{t_1-t}[E_t^s(P_{t_1}) - P^s], t < t_1 \\ P_t^o, t \ge t_1 \end{cases}$$

8.2 Explain the dependence of policy outcomes on sentiment

During Mar, 2010 to May, 2015 (i.e. the period that our sentiment index covers), there were 6 tightening months but only 2 loosening months. So the conclusions about tightening months are likely to be more universal than those about loosening policies. And according to our empirical results, the outcome of tightening policies depends on sentiment, while that of loosening policies don't. Therefore, the theoretical analysis below emphasize on the announcement and post-announcement effects related to tightening policies.

Let us consider what will happen if a tightening policy arrives at time zero. Then $u_0 = -\sigma$. We denote the last time point before time zero by T_0. And the flow of home-ownership utility before and after the policy change is denoted by by ε^{old} and ε, respectively. We further assume that $-\sigma < \varepsilon - \varepsilon^{old} < 0$. That is, the tightening policy reduces the utility of owning a home, but the magnitude of the reduction is less than the reduction of optimistic agents' expectation for ε^f. Furthermore, considering the counter-cyclical nature of interventions, it is likely that the housing price was rising before a tightening policy arrives. Therefore, we assume that $T_0<t_1^{old}$, where t_1^{old} is the time when the positive price growth was supposed to end were it not for the policy change.

Suppose the economic condition fits into Case I before the policy change, and the entropy of skeptical agents is still lower than the optimistic ones after the change. Then we have following equations, where P_{old}^s is the fundamental value of skeptical agents before the policy change:

$$P_{T_0} = P_{old}^s + [\beta(1-\phi)]^{t_1^{old}-T_0}[E_{T_0}^s(P_{t_1^{old}}) - P_{old}^s]$$

$$P_0 = P^s + [\beta(1-\phi)]^{t_1}[E_0^s(P_{t_1}) - P^s]$$

$$P_0 - P_{T_0} = P^s - P_{old}^s + [\beta(1-\phi)]^{t_1}[E_0^s(P_{t_1}) - P^s] - [\beta(1-\phi)]^{t_1^{old}-T_0}[E_{T_0}^s(P_{t_1^{old}}) - P_{old}^s]$$

Because $\varepsilon<\varepsilon^{old}$, we have $P^s<P_{old}^s$. And it is easy to understand that $t_1>t_1^{old}-T_0$, because the social dynamic has to restart from the initial values before reaching the point when optimistic agents become marginal buyers. Then it can be proved that $P_0 < P_{T_0}$, the details of which is in **Appendix B**. This means when the tightening policy is announced, there is a drop in housing prices.

The interesting part of our model is that it indicates if a tightening policy is announced in a high-sentiment environment, there will be a price rebounding after the announcement. To see this, first recall formula (13) and that the social dynamic path is public information, which leads to:

$$E_s^j[E_t^i(\varepsilon^f)] = E_s^i(\varepsilon^f)\text{, } s\leq t \tag{19}$$

And then it follows that

$$E_t^s(P_{t_1}) - P^s = (P_t^o - P^s)\left\{1-[\beta(1-\phi)]^{t_2+1-t_1}\right\} = \frac{\phi\beta\left\{1-[\beta(1-\phi)]^{t_2+1-t_1}\right\}}{(1-\beta)[1-\beta(1-\phi)]}S_t$$

So according to the first line in (17), the difference between post-announcement price P_t and P_0 is:

$$P_t - P_0 = [\beta(1-\phi)]^{t_1-t}\frac{\phi\beta\left\{1-[\beta(1-\phi)]^{t_2+1-t_1}\right\}}{(1-\beta)[1-\beta(1-\phi)]}\left\{S_t - [\beta(1-\phi)]^t S_0\right\}\text{, } t<t_1. \tag{20}$$

Combining (20) with (13), **Appendix C** proves that $P_t>P_0$ as long as

$$E_{T_0}^o(\varepsilon^f) > \varepsilon - \frac{\sum_{\tau=1}^{t} u_\tau}{1-[\beta(1-\phi)]^t} - u_0 \tag{21}$$

In addition, for any $0<n<m<t_1$, the price change P_m-P_n is positive as long as:

$$E_{T_0}^o(\varepsilon^f) > \varepsilon - \frac{\sum_{\tau=n+1}^{m} u_\tau}{(1-[\beta(1-\phi)]^{m-n})} - \sum_{\tau=0}^{n} u_\tau \tag{22}$$

Therefore, when the sentiment at T_0 is high enough, the housing prices in the post-announcement periods will be higher than in the announcement period, and will keep growing until t_1. This leads to a typical housing price rebounding after the announcement of a tightening policy, which finds support in our empirical work. The key of the mechanism is that after a policy change, the social dynamics return to the initial values, and the number of optimistic agents will be increasing for a while.

Now, suppose the economic condition fits into Case II before the policy change, and the entropy of skeptical agents is still higher than the optimistic ones after the change. Then based on the assumption that $-\sigma<\varepsilon-\varepsilon_{old}<0$ and $t_1^{old}-T_0<t_1$, it can be easily proved that $P_0-P_{T_0}<0$. In addition, the sentiment threshold at T_0 that incurs rebounding in the post-announcement period is exactly the same as (21) and (22) shows.

If the negative policy change converts the economy condition from Case II to Case I, we still have the conclusion that housing prices will first drop and then rebound if a tightening policy meets with a high-sentiment environment. If the tightening policy converts the economy condition from Case I to Case II, the conclusion about price rebounding in the post-announcement period maintains, although the announcement effect depends on the model parameters. Therefore, our model always predicts a price rebounding after a tightening policy's arrival. And as long as it doesn't happen that the optimistic agents become more certain than skeptical ones after the arrival, the announcement effect will be negative.

IX. Conclusion

We create a market-pattern-based sentiment index for the Shanghai housing market. The monthly index is look-ahead-bias-free, and is highly correlated with official confidence indexes, investor confidence indexes, and the housing purchase confidence index. So our index is good at capturing the sentiment in the consumption incentive as well as the investment incentive of the housing market.

Then with this sentiment index, we study the role of sentiment in the Shanghai housing market, which is one of the biggest housing markets in China. It is found that positive sentiment helps explain the low housing market returns in the future, while negative sentiment does not explain high returns. We also investigate how government intervention affects sentiment, and how intervention's outcomes depend on sentiment. While some evidence shows that a loosening policy can lift the market sentiment, no evidence shows a tightening policy can quickly cool down the sentiment. Furthermore, if a tightening policy is announced in a high-sentiment environment, there will be a significant price rebounding in the subsequent months. The rebounding is more obvious in the places where the housing prices tend to rise whenever sentiment increases but rarely drop when sentiment decreases. The above findings reveal that even though a tightening policy can temporarily constrain the market participants' behavior, it is quite hard to eliminate the optimistic views, and a spatially varied price rebounding will follow these tightening policies.

Finally, based on the model of Burnside, Eichenbaum, and Rebelo (2015), we theoretically illustrates why sentiment affects policy outcomes. The key of the mechanism lies in that: (1) Most agents become highly uncertain about future policies after a new policy arrives, because people need time to learn the impact of a new policy, based on which they predict the government's next action. (2) There always exists a group of people with optimistic, firm, and persistent views. So if a tightening policy meets with a high-sentiment environment, then after the announcement day, the number of optimistic people will grow because a lot of people with uncertain opinions will become optimistic after they meet with the firmly optimistic people. As a result, the housing price rebound.

There are many directions worth exploring based on this study. For example, our method to make housing market sentiment indexes can be applied to other cities and countries. Also, it is interesting to investigate the relation of sentiment with time-on-market, return volatility, developer behavior, etc. It is also meaningful to include housing market sentiment into a macroeconomic model. As Hui and Wang (2014) point out, real estate market has contagious effects on other industries. So it is meaningful to studies how housing market sentiment interacts with the other parts in the economy and generates farther-reaching influence.

Reference

[1] Baker M, Stein J C. Market liquidity as a sentiment indicator. Journal of Financial Markets, 2004, 7(3): 271-299.

[2] Baker M, Wurgler J. Investor sentiment and the cross - section of stock returns. The Journal of Finance, 2006, 61(4): 1645-1680.

[3] Baker M, Wurgler J. Investor Sentiment in the Stock Market. Journal of Economic Perspectives, 2007, 21(2):129-152.

[4] Barber B M, Odean T. Boys will be boys: Gender, overconfidence, and common stock investment. Quarterly journal of Economics, 2001: 261-292.

[5] Bernanke B S, Kuttner K N. What explains the stock market's reaction to Federal Reserve policy? The Journal of Finance, 2005, 60(3): 1221-1257.

[6] Brown G W, Cliff M T. Investor sentiment and the near-term stock market. Journal of Empirical Finance, 2004, 11(1): 1-27.

[7] Brown G W, Cliff M T. Investor Sentiment and Asset Valuation. The Journal of Business, 2005, 78(2): 405-440.

[8] Burnside C, Martin Eichenbaum, and Sergio Rebelo. Forthcoming. Booms and Busts in Housing Markets. Journal of Political Economy.

[9] Campbell J Y, Thompson S B. Predicting excess stock returns out of sample: Can anything beat the historical average?. Review of Financial Studies, 2008, 21(4): 1509-1531.

[10] Cao A, Keivani R. The Limits and Potentials of the Housing Market Enabling Paradigm: An Evaluation of China's Housing Policies from 1998 to 2011. Housing Studies, 2014, 29(1): 44-68.

[11] Case K E, Shiller R J. Prices of single family homes since 1970: New indexes for four cities. New England Economics Review, 1987:45-56.

[12] Chen J, Hao Q, Stephens M. Assessing housing affordability in post-reform

China: a case study of Shanghai. Housing Studies, 2010, 25(6): 877-901.
[13]Clayton J, Ling D C, Naranjo A. Commercial real estate valuation: fundamentals versus investor sentiment. The Journal of Real Estate Finance and Economics, 2009, 38(1): 5-37.
[14]Clayton J. Are housing price cycles driven by irrational expectations?. The Journal of Real Estate Finance and Economics, 1997, 14(3): 341-363.
[15]Croce R M, Haurin D R. Predicting turning points in the housing market. Journal of Housing Economics, 2009, 18(4): 281-293.
[16]Da Z, Liu Q, Schaumburg E. A closer look at the short-term return reversal. Management Science, 2014, 60(3): 658-674.
[17]Deng L, Shen Q, Wang L. The emerging housing policy framework in China. Journal of Planning Literature, 2011, 26(2): 168-183.
[18]DiPasquale D, Wheaton W C. Housing market dynamics and the future of housing prices. Journal of urban economics, 1994, 35(1): 1-27.
[19]Du H, Ma Y, An Y. The impact of land policy on the relation between housing and land prices: Evidence from China. The Quarterly Review of Economics and Finance, 2011, 51(1): 19-27.
[20]Du Z, Zhang L. Home-purchase restriction, property tax and housing price in China: A counterfactual analysis. Journal of Econometrics, 2015.
[21]Fang H, Q Gu, W Xiong, L Zhou. Demystifying the Chinese Housing Boom, 2015, NBER Working Paper No. 21112.
[22]Feldman T. A more predictive index of market sentiment. Journal of Behavioral Finance, 2010, 11(4): 211-223.
[23]Fisher J, Gatzlaff D, Geltner D, Haurin D. Controlling for the impact of variable liquidity in commercial real estate price indices. Real Estate Economics, 2003, 31(2): 269-303.
[24]Gallimore P, Gray A. The role of investor sentiment in property investment decisions. Journal of Property Research, 2002, 19(2): 111-120.
[25]Gupta R, Hammoudeh S, Modise M P, Nguyen D K. Can economic uncertainty, financial stress and consumer sentiments predict US equity premium?. Journal of International Financial Markets, Institutions and Money, 2014, 33: 367-378.
[26]Han L. Understanding the puzzling risk-return relationship for housing. Review of Financial Studies, 2013, 26(4): 877-928.
[27]Huang D, Jiang F, Tu J, Zhou G. Investor sentiment aligned: a powerful predictor of stock returns. Review of Financial Studies, 2014: hhu080.
[28]Huang W, Liu Q, Rhee S G, et al. Return Reversals, Idiosyncratic Risk, and Expected Returns[J]. Review of Financial Studies, 2010, 23(1):147-168.
[29]Huang Y, Clark W A V. Housing tenure choice in transitional urban China: a multilevel analysis. Urban Studies, 2002, 39(1): 7-32.
[30]Hui E C, Wang Z. Market sentiment in private housing market. Habitat International, 2014, 44: 375-385.
[31]Jin C, Soydemir G, Tidwell A. The US housing market and the pricing of risk: Fundamental analysis and market sentiment. Journal of Real Estate Research, 2014, 36(2): 187-219.

[32] Kaplanski G, Levy H. Real estate prices: An international study of seasonality's sentiment effect. Journal of Empirical Finance, 2012, 19(1): 123-146.
[33] Kostovetsky L. Political capital and moral hazard. Journal of Financial Economics, 2015, 116(1): 144-159.
[34] Kurov A. Investor sentiment and the stock market’s reaction to monetary policy. Journal of Banking & Finance, 2010, 34(1): 139-149.
[35] Lai O K. Governance and the housing question in a transitional economy, the political economy of housing policy in China reconsidered. Habitat International, 1998, 22(3): 231-243.
[36] Lau K M, Li S M. Commercial housing affordability in Beijing, 1992–2002. Habitat International, 2006, 30(3): 614-627.
[37] Ling D C, Naranjo A, Scheick B. Investor Sentiment, Limits to Arbitrage and Private Market Returns. Real Estate Economics, 2014, 42(3): 531-577.
[38] Malmendier U, Tate G. CEO overconfidence and corporate investment. The Journal of Finance, 2005, 60(6): 2661-2700.
[39] Malmendier U, Tate G. Who makes acquisitions? CEO overconfidence and the market's reaction. Journal of financial Economics, 2008, 89(1): 20-43.
[40] Miller N, Pandher G. Idiosyncratic volatility and the housing market. Journal of Housing Research, 2008, 17(1): 13-32.
[41] Piazzesi M, Schneider M. Momentum traders in the housing market: survey evidence and a search model[R]. National Bureau of Economic Research, 2009.
[42] Quan D C, Titman S. Do real estate prices and stock prices move together? An international analysis. Real Estate Economics, 1999, 27(2):183-207.
[43] Sato H. Housing inequality and housing poverty in urban China in the late 1990s. China Economic Review, 2006, 17(1): 37-50.
[44] Soo C K. Quantifying animal spirits: news media and sentiment in the housing market. The Stephen M. Ross School of Business at the University of Michigan Research Paper Series, 2013.
[45] Stambaugh R F, Yu J, Yuan Y. The short of it: Investor sentiment and anomalies. Journal of Financial Economics, 2012, 104(2): 288-302.
[46] Wang S Y. Credit constraints, job mobility, and entrepreneurship: evidence from a property reform in china. Review of Economics and Statistics, 2012, 94(2): 532-551.
[47] Zheng X, Chau K W, Eddie C M. Liquidity risk and cross-sectional return in the housing market. Habitat International, 2015, 49: 426-434.
[48] Zhou Z. Overreaction to policy changes in the housing market: Evidence from Shanghai. Regional Science and Urban Economics, 2016, forthcoming. DOI:10.1016/j.regsciurbeco.2016.02.004.
[49] Zou Y. Contradictions in China's affordable housing policy: Goals vs. structure. Habitat International, 2014, 41: 8-16.

Figure 1 Repeat-sales index, policy changes, and sentiment index

The line marked with "X" is the repeat sales housing price index calculated by Zhou (2016). The line marked with dots is our sentiment index. The vertical lines mark the dates when government interventions took place, and the detailed information about these interventions are shown in **Table 1**.

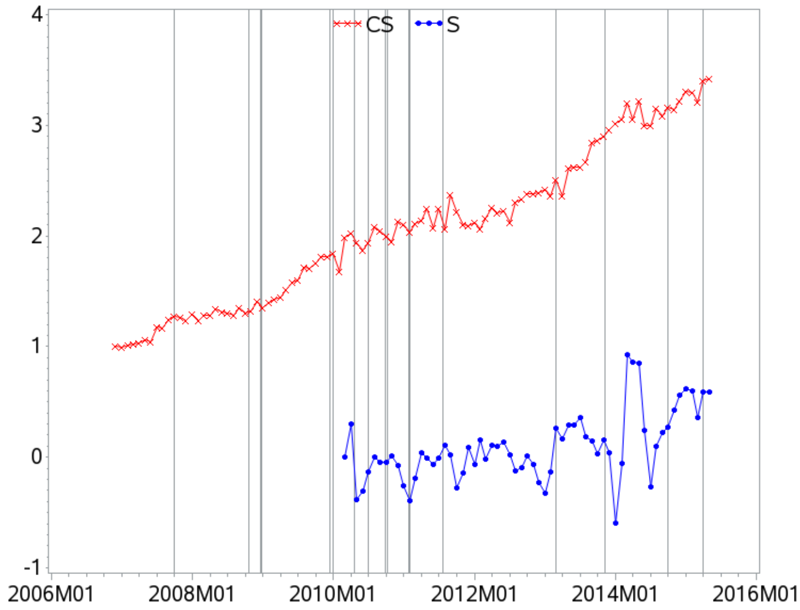

Figure 2 Four measures of housing market sentiment

The line marked with "X" is the sentiment index based on principal analysis method that is adopted in Baker and Wurgler (2006); the full sample is used when computing the first principal comonent. The line marked with dots is the sentiment index based on the PLS method adopted in Huang, Jiang, Tu, Zhou (2014); the first-stage regression uses the full sample. The line marked with circles is the sentiment index based on principal analysis; it is look-ahead-bias free in that only the information available up to time *t* is used to compute the index value of *t*. The line marked with "+" is our look-bias-free sentiment index based on the PLS method.

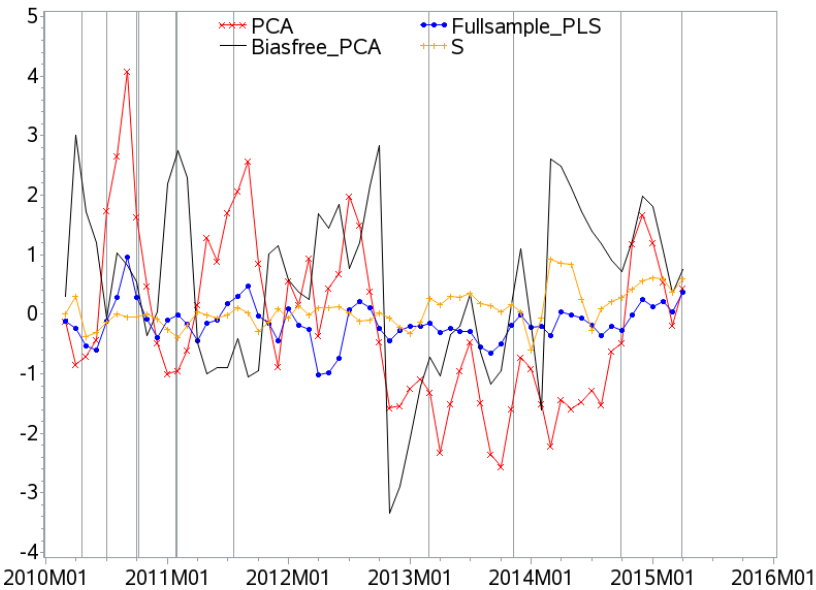

Figure 3a Plates with the drop-rebounding pattern related to policy changes

The figure shows the returns of repeat-sales housing price indexes around a tightening month. The Huamu, Jinqiao, Expo, Yangjing, Jiangqiao, and Waigaoqiao plates are included here. In these plates, the returns were negative in the tightening months, but then became positive in the next month. In other words, the prices first dropped, and then quickly began growing again.

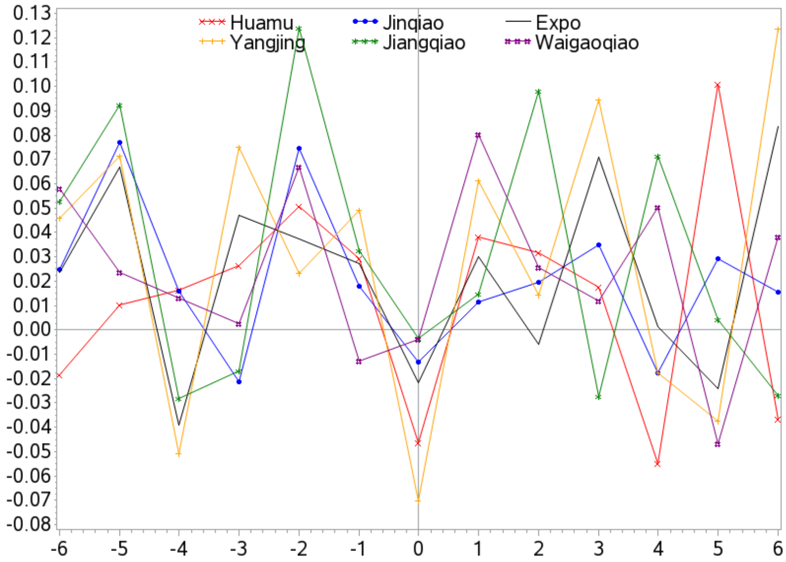

Figure 3b Other plates

The figure shows the returns of repeat-sales housing price indexes around a tightening month. The Tianlin, Sanlin, Zhongshan, Laoximen, Lujiazui, and Chuansha plates are included here. In these plates, the housing prices either didn't drop in the tightening months, or didn't grow in the next month.

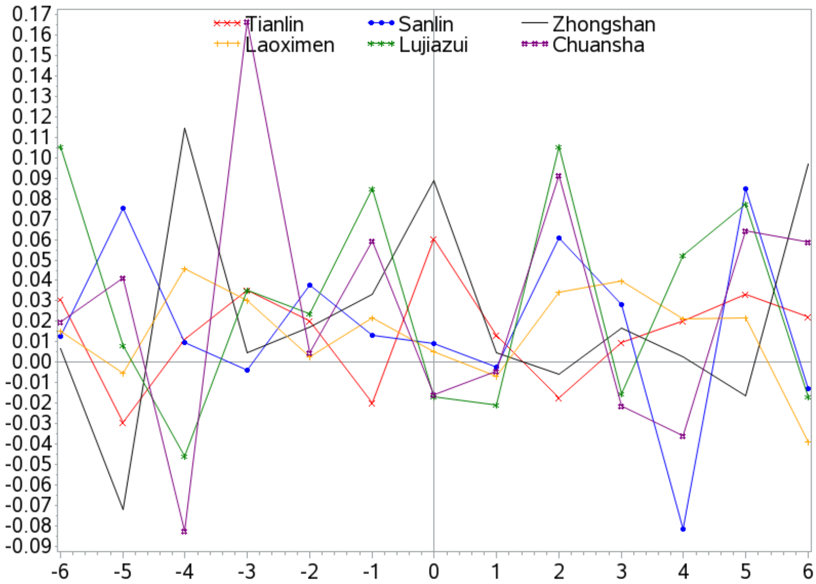

Table 1 Policy changes

The *Tighten* (*Loosen*) column lists the tools used in the tightening (loosening) policies.

Date	Policy name	Level	Tighten	Loosen
2005.3.17	317 Rules	Central	MR	
2005.3.26	National 8 Rules	Central		
2006.5.24	National 6 Rules, National 15 Rules	Central	T-RS, DP-1	
2007.9.27	927 New Policy	Central	DP-2, MR-2	
2008.10.22	1022 Rules	Central		DP, MR
2008.12.20	National 13 Rules	Central		T-RS,DP-2,MR-2, T-O
2008.12.27	Shanghai 8 Rules	Local		DP-2, MR-2, DP-PFL-2, Tot-PFL-2
2009.12.14	National 4 Rules	Central		
2009.12.29	Shanghai 15 Rules	Local	T-RS, T-O	
2010.1.7	National 11 Rules	Central	DP-2	
2010.4.17	New National 4 Rules	Central	DP-1, DP-2, MR-2, MA-NSH	
2010.9.29	929 Rules	Central	DP-1, MA-3	
2010.10.7	Shanghai 12 Rules	Local	DP-PFL-1,Tot-PFL-1, DP-PFL-2,Tot-PFL-2, PFLA-3，RP-SH, RP-NSH	RR
2011.1.26	New National 8 Rules	Central	T-RS, DP-2, RP-SH, RP-NSH	
2011.1.28	Property tax pilot	Central	T-O	
2011.1.31	Shanghai 9 Rules	Central	DP-PFL-2,MR-PFL-2	
2011.7.20	720 Rules	Local	NHP	
2012.2.27	227 Rules	Local	Restate old tools	
2012.7.26	Shanghai 6 Rules	Local	Restate old tools	
2013.2.26	National 5 Rules	Central	T-O	
2013.3.30	New Shanghai 6 Rules	Central	Restate old tools	
2013.11.8	Shanghai 7 Rules	Local	DP-2, RP-NSH	
2014.9.29	930 New Policy	Central		MR-1,DP-2,MR-2
2015.3.30	330 New Policy	Central		DP-2, DP-PFL-1, DP-PFL-2

Table 2 Descriptive statistics of sentiment proxies: Average value by year

NewhouseconR is the newly opened residential construction area in Shanghai divided by the average newly-released residential land supply in the past 6 months. *HouseinvR* is the ratio between residential housing investment and total real estate investment in Shanghai. *MedianIntv* is the log of the median days between the last sale and the current sale of a same house. *Saleprob* is the ratio between the transacted area of residential houses and the area of residential houses that are available for sale. *SMB* is the difference between the housing price returns of small and big houses, where small (big) houses are those in the bottom (top) size quintile. Since *SaleProb* is not available until Apr, 2009, its average value of 2009 is the average from April to December.

Year	NewhouseconR	HouseinvR	MedianIntv	SaleProb	SMB
2009	3.8739	0.6210	6.3869	0.4369	-0.0020
2010	2.4459	0.6211	6.1316	0.2320	-0.0015
2011	1.5880	0.6448	6.3337	0.1432	0.0239
2012	2.1560	0.6128	6.8520	0.1202	-0.0238
2013	1.4247	0.5709	7.2133	0.1585	-0.0036
2014	1.6945	0.5435	7.4030	0.1190	0.0036
2015	1.2084	0.5382	7.5169	0.1218	0.0201

Table 3 Correlation of the first principal component with current and lagged sentiment proxies

The superscript "*c*" means the proxy is orthogonal to the variables about economic fundamentals. It is the residual in the regression of the raw proxy on the economic fundamental variables, and is smoothed with the three-month moving average. *Lag* refers to the one-month lag of the corresponding variable. Significant level of 1%, 5%, 10% are marked by *, **, and ***.

	Correlation with *Prin1*	Significant level	Obs		Correlation with *Prin1*	Significant level	Obs
NewhouseconR^{c}	0.7532***	<.0001	71	**Lag_NewhouseconR^{c}**	0.7386***	<.0001	71
HouseinvR^{c}	0.6152***	<.0001	71	**Lag_HouseinvR^{c}**	0.7400***	<.0001	71
MedianIntv^{c}	-0.7801***	<.0001	71	**Lag_MedianIntv^{c}**	-0.7635***	<.0001	71
SaleProb^{c}	0.0972	0.4200	71	**Lag_SaleProb^{c}**	0.0775	0.5206	71
SMB^{c}	-0.3273***	0.0053	71	**Lag_SMB^{c}**	-0.2410**	0.0429	71

Table 4A Correlation of the sentiment index with official confident indexes

CC_SH is the consumer confidence index of Shanghai, *HP_SH* is the housing purchase confidence index of Shanghai, *CC_CN* is the consumer confidence index of China, *IC_CN* is the investor confidence index of China, *IC_CN_F* is the investor confidence index about domestic economic fundamentals, and *IC_CN_P* is the investor confidence index about domestic economic policies. Significant level of 1%, 5%, 10% are marked by *, **, and ***.

	CC_SH	**HP_SH**	**CC_CN**	**IC_CN**	**IC_CN_F**	**IN_CN_P**
Correlation	0.4214***	0.4501***	0.1769	0.2290*	0.1419	0.3010**
p-value	0.0006	0.0007	0.1654	0.071	0.2673	0.0165
Obs	63	53	63	63	63	63

Table 4B Correlation between the sentiment index and confident indexes: 2011

	S	**CC_SH**	**HP_SH**	**CC_CN**	**IC_CN**	**IC_CN_F**	**IN_CN_P**
S	1	-0.5578*	0.2190	0.4483	-0.4852	-0.5305*	-0.3313
		0.0595	*0.4940*	*0.1439*	*0.1099*	*0.0760*	*0.2928*
CC_SH	-0.5578*	1	-0.8655***	0.1398	0.5447*	0.5683*	0.0678
	0.0595		*0.0003*	*0.6648*	*0.0671*	*0.0539*	*0.8341*
HP_SH	0.2190	-0.8655***	1	-0.5122*	-0.4644	-0.4150	0.2083
	0.4940	*0.0003*		*0.0887*	*0.1283*	*0.1798*	*0.5159*
CC_CN	0.4483	0.1398	-0.5122*	1	0.0865	-0.0095	-0.2659
	0.1439	*0.6648*	*0.0887*		*0.7893*	*0.9767*	*0.4035*
IC_CN	-0.4852	0.5447*	-0.4644	0.0865	1	0.9708***	0.6902**
	0.1099	*0.0671*	*0.1283*	*0.7893*		*<.0001*	*0.0130*
IC_CN_F	-0.5305*	0.5683*	-0.4150	-0.0095	0.9708***	1	0.7208***
	0.0760	*0.0539*	*0.1798*	*0.9767*	*<.0001*		*0.0082*
IN_CN_P	-0.3313	0.0678	0.2083	-0.2659	0.6902**	0.7208***	1
	0.2928	*0.8341*	*0.5159*	*0.4035*	*0.0130*	*0.0082*	

Table 5A Comparison among sentiment indexes: Pairwise correlation

S^{PCA} is the sentiment index based on principal component analysis and full sample. S^{FSPLS} is the sentiment index based on the PLS approach and full sample. S^{BFPCA} is the look-ahead-bias-free sentiment index based on the principal component analysis. S is our look-ahead-bias-free sentiment index based on the PLS approach. Significant level of 1%, 5%, 10% are marked by *, **, and ***.

	S^{PCA}	S^{FSPLS}	S^{BFPCA}	S
S^{PCA}	1	0.60795***	0.0790	-0.0867
		<.0001	0.5415	0.5030
S^{FSPLS}	0.60795***	1	0.05712	0.11112
	<.0001		0.6592	0.3899
S^{BFPCA}	0.0790	0.05712	1	0.2973**
	0.5415	0.6592		0.0189
S	-0.0867	0.11112	0.2973**	1
	0.5030	0.3899	0.0189	

Table 5B Sentiment index comparison: Correlation with official confidence indexes

Corr is the correlation between a sentiment index and a confidence index; p-value is the significant level of the correlation. *Obs* is the number of observations.

		CC_SH	HP_SH	CC_CN	IC_CN	IC_CN_F	IN_CN_P
S^{PCA}	**Corr**	-0.2701**	-0.0372	0.1527	0.0345	-0.0541	-0.0588
	p-value	0.0218	0.7915	0.2003	0.7737	0.6518	0.6239
	Obs	72	53	72	72	72	72
S^{BFPCA}	**Corr**	0.2449*	0.0598	0.0773	0.0870	-0.0241	0.0132
	p-value	0.0531	0.6705	0.5472	0.4979	0.8514	0.918
	Obs	63	53	63	63	63	63
S^{FSPLS}	**Corr**	-0.0472	-0.0244	0.0702	0.2695**	0.3604***	0.2119*
	p-value	0.6982	0.8638	0.5634	0.024	0.0022	0.0783
	Obs	70	52	70	70	70	70

Table 6A Regression of future housing market returns on current sentiment level

$$R_{[t+a,t+b]} = \alpha + \beta_1 S_t + \beta_2 Spring_t + \beta_3 Autumn_t + \beta_4 ret_t + \varepsilon_t$$

The parameter *a* and *b* are displayed in the form of [*a*,*b*] in the first row. Numbers in italic are p-values. Significant level of 1%, 5%, 10% are marked by *, **, and ***.

	[1,1]	[1,3]	[1,6]	[1,9]	[1,12]
Intercept	0.0224***	0.0363***	0.0691***	0.0957***	0.1264***
	0.0023	0.0000	0.0000	0.0000	0.0000
S	-0.0110	-0.0162	-0.0039	0.0095	-0.0016
	0.5445	0.4241	0.9004	0.8044	0.9718
Spring	-0.0140	0.0091	-0.0166	0.0113	-0.0013
	0.3570	0.5780	0.5201	0.7124	0.9713
Autumn	-0.0269	-0.0117	-0.0221	-0.0460	-0.0005
	0.0742	0.4681	0.3149	0.1124	0.9875
ret	-0.3973***	-0.5440***	-0.3894**	-0.1882	-0.4956**
	0.0004	0.0000	0.0155	0.3150	0.0339
Obs	62	60	57	54	51
R^2	24.19%	32.43%	12.41%	8.41%	9.69%

Table 6B The explaining power of positive sentiment and negative sentiment for future returns

$$R_{[t+a,t+b]} = \alpha + \beta_1 PosiS_t + \beta_2 NegaS_t + \beta_3 Spring_t + \beta_4 Autumn_t + \beta_5 ret_t + \varepsilon_t$$

PosiS equals to *S* when *S* is positive, and zero otherwise. *NegaS* equals to *S* when *S* is non-positive, and zero otherwise.

	[1,1]	[1,3]	[1,6]	[1,9]	[1,12]
Intercept	0.0330***	0.0465***	0.0784***	0.1045***	0.1347***
	0.0004	*0.0000*	*0.0000*	*0.0000*	*0.0000*
PosiS	-0.0437*	-0.0490*	-0.0349	-0.0209	-0.0296
	0.0816	*0.0821*	*0.4069*	*0.6831*	*0.6283*
NegaS	0.0788	0.0705	0.0743	0.0844	0.0717
	0.1231	*0.2036*	*0.3362*	*0.3580*	*0.5259*
ret_t	-0.4293***	-0.5775***	-0.4205***	-0.2175	-0.5241**
	0.0001	*0.0000*	*0.0102*	*0.2542*	*0.0283*
Spring	-0.0039	0.0191	-0.0076	0.0199	0.0075
	0.8063	*0.2685*	*0.7771*	*0.5362*	*0.8484*
Autumn	-0.0306**	-0.0151	-0.0251	-0.0492*	-0.0034
	0.0405	*0.3449*	*0.2567*	*0.0930*	*0.9223*
Obs	62	60	57	54	51
R^2	28.81%	35.85%	14.51%	9.96%	10.72%

Table 7A Out-of-sample predicting power of sentiment

Pred refers to the return predicted by past sentiment. *Hisavg* refers to the historical average return. *Corr* is the correlation with the realized return, p-value is the significant level of the correlation, and *Obs* is the number of observations. For each month *t*, the forecasting target is the cumulative return from month $t+a$ to month $t+b$, with a and b displayed in the form of $[a,b]$ in the first row. Significant level of 1%, 5%, 10% are marked by *, **, and ***.

		[1,1]	[1,3]	[1,6]	[1,9]	[1,12]
Pred	**Corr**	-0.4583***	-0.2061	-0.3622**	-0.3502**	-0.58358***
	p-value	0.0006	0.1599	0.0184	0.0363	0.0007
	Obs	52	48	42	36	30
Hisavg	**Corr**	-0.2560*	-0.1131	-0.3234**	-0.4644***	-0.5946***
	p-value	0.0464	0.4023	0.0206	0.0013	<.0001
	Obs	61	57	51	45	39
$\mathbf{R_{OS}^2}$		-0.0661	-0.2852	-0.3759	-0.1773	-0.0557

Table 7B Out-of-sample predicting power of sentiment and lagged returns

Pred refers to the return predicted by past sentiment and past returns.

		[1,1]	[1,3]	[1,6]	[1,9]	[1,12]
Pred	**Corr**	0.2901**	0.2234	-0.1886	-0.3080*	-0.4450**
	p-value	0.0370	0.1270	0.2318	0.0677	0.0137
	Obs	52	48	42	36	30
Hisavg	**Corr**	-0.2560*	-0.1131	-0.3234**	-0.4644***	-0.5946***
	p-value	0.0464	0.4023	0.0206	0.0013	<.0001
	Obs	61	57	51	45	39
$\mathbf{R_{OS}^2}$		0.1181	-0.0940	-0.3846	-0.2117	-0.0330

Table 8A Sentiment and policy outcomes

$$R_{[t+a,t+b]} = c + \beta_1 S_t + \beta_2 Loose_t + \beta_3 S_t \times Loose_t + \beta_4 Tight_t + \beta_5 S_t \times Tight_t + \beta_6 Spring_t + \beta_7 Autumn_t + \beta_8 ret_t + \varepsilon_t$$

The parameter *a* and *b* are shown in the form of [*a*,*b*] in the first row. The dependent variable is the cumulative returns in the [*t*+*a*, *t*+*b*] window after month *t*. *S* is our sentiment index. *Tight* (*Loose*) is a dummy that equals to 1 if a tightening policy took place in that month; if the policy came in the second half of a month, then the next month rather than the month itself is marked. *Spring* (*Autumn*) is a dummy that equals to 1 if the month is January or February (September of October). And *ret* is the return in month *t*. Significant level of 1%, 5%, 10% are marked by *, **, and ***. Numbers in italic are p-values.

	[1,1]	**[1,3]**	**[1,6]**	**[1,9]**	**[1,12]**
Intercept	0.0239***	0.0346***	0.0722***	0.0963***	0.1286***
	0.0028	*0.0001*	*0.0000*	*0.0000*	*0.0000*
S	-0.0185	-0.0159	-0.0239	-0.0050	-0.0219
	0.3536	*0.4583*	*0.4603*	*0.9002*	*0.6482*
Loose	0.0084	0.0461	0.0477		
	0.9276	*0.3467*	*0.4582*		
S×Loose	0.0170				
	0.9315				
Tight	-0.0081	0.0147	-0.0099	0.0050	-0.0021
	0.6738	*0.4753*	*0.7123*	*0.8781*	*0.9571*
S×Tight	0.0656	-0.0093	0.2009*	0.1823	0.2352
	0.3693	*0.9053*	*0.0545*	*0.1430*	*0.1108*
Spring	-0.0125	0.0091	-0.0132	0.0157	0.0037
	0.4262	*0.5904*	*0.6058*	*0.6089*	*0.9197*
Autumn	-0.0281*	-0.0164	-0.0268	-0.0461	-0.0009
	0.0836	*0.3411*	*0.2371*	*0.1107*	*0.9790*
ret	-0.4235***	-0.5270***	-0.4433***	-0.2180***	-0.5458***
	0.0004	*0.0001*	*0.0068*	*0.2566*	*0.0229*
R^2	26.00%	34.25%	20.15%	12.54%	14.94%
Obs	62	60	57	54	51

Table 8B Sentiment and policy outcomes: By plate

$$R_{[t+1,t+6],i} = c + \beta_1 S_t + \beta_2 Loose_t + \beta_3 Tight_t + \beta_4 S_t \times Tight_t + \beta_5 Spring_t + \beta_6 Autumn_t + \beta_7 ret_{t,i} + \varepsilon_{t,i}$$

We run the above regression separately for each plate *i*. Numbers in italic are p-values. Significant level of 1%, 5%, 10% are marked by *, **, and ***.

	Lujiazui	**Waigaoqiao**	**Jiangqiao**	**Expo**	**Laoximen**	**Sanlin**	**Yangjing**	**Tianlin**	**Chuansha**	**Jinqiao**	**Zhongshan**	**Huamu**
Intercept	0.0948***	0.0828***	0.1442***	0.0772***	0.0563***	0.0719***	0.0927***	0.0715***	0.0722**	0.0693***	0.0514**	0.0586***
	0.0004	*0.0001*	*0.0015*	*0.0007*	*0.0005*	*0.0010*	*0.0000*	*0.0000*	*0.0207*	*0.0005*	*0.0372*	*0.0005*
S	-0.0669	-0.0417	-0.1615	-0.0031	0.0121	-0.0005	-0.0506	0.0060	0.0255	-0.0235	-0.0004	0.0119
	0.3817	*0.4900*	*0.2102*	*0.9624*	*0.7901*	*0.9936*	*0.4047*	*0.8860*	*0.7795*	*0.6676*	*0.9953*	*0.7975*
Loose	-0.0219	-0.0183	0.2454	0.0599	0.1428	0.0466	0.1198	0.0074	-0.0534	-0.0198	-0.1299	0.0293
	0.8851	*0.8793*	*0.3356*	*0.6485*	*0.1178*	*0.7075*	*0.3216*	*0.9287*	*0.7691*	*0.8543*	*0.3746*	*0.7455*
Tight	-0.0427	0.0355	-0.1135	0.0290	-0.0237	0.0137	0.0416	0.0237	0.0304	0.0171	0.0381	-0.0060
	0.4965	*0.4727*	*0.2815*	*0.5858*	*0.5235*	*0.7905*	*0.4127*	*0.4945*	*0.6854*	*0.7049*	*0.5295*	*0.8821*
S×Tight	0.5613**	0.7019***	0.2513	-0.0554	-0.0596	0.0663	0.5956***	0.1018	-0.2007	-0.2142	0.0173	0.0785
	0.0237	*0.0005*	*0.5358*	*0.7932*	*0.6811*	*0.7381*	*0.0029*	*0.4454*	*0.4894*	*0.2173*	*0.9413*	*0.5868*
Spring	-0.0691	-0.0908*	0.0007	0.0231	-0.0048	-0.0533	0.0087	-0.0539	0.0967	-0.0116	-0.0194	-0.0126
	0.2535	*0.0593*	*0.9949*	*0.6514*	*0.8931*	*0.2847*	*0.8557*	*0.1065*	*0.1821*	*0.7880*	*0.7407*	*0.7255*
Autumn	-0.0013	0.0045	-0.1570*	-0.0357	0.0138	0.0119	-0.0322	-0.0202	-0.0349	0.0016	0.0507	0.0088
	0.9810	*0.9182*	*0.0843*	*0.4315*	*0.6625*	*0.7860*	*0.4458*	*0.4906*	*0.5829*	*0.9678*	*0.3247*	*0.7788*
ret_t	-0.4363***	-0.3856***	-0.4222**	-0.6337***	-0.5155***	-0.4270**	-0.5959***	-0.5309***	-0.2668	-0.4176**	-0.4057***	-0.3003*
	0.0020	*0.0013*	*0.0375*	*0.0001*	*0.0043*	*0.0154*	*0.0000*	*0.0004*	*0.2435*	*0.0125*	*0.0037*	*0.0805*
Obs	57	57	57	57	57	57	57	57	57	57	57	57
R^2	29.63%	42.47%	17.14%	30.21%	21.56%	13.37%	48.35%	29.70%	11.53%	18.56%	20.89%	8.61%

Table 9 Sensitiveness to sentiment changes: By plate

$$ret_{i,t} = c_i + \beta_{i,1} posichg_t + \beta_{i,2} negachg_t + \beta_{i,3} ret_{i,t-1} + \beta_{i,4} Spring_t + \beta_{i,5} Autumn_t + \varepsilon_{i,t}$$

The variable *posichg* is change in sentiment from last period if the change is positive, and zero otherwise; *negachg* is change in sentiment from last period if the change is non-positive, and zero otherwise. Numbers in italic are p-values. Significant level of 1%, 5%, 10% are marked by *, **, and ***.

No.	Dependent variable	β_1	p-value	β_2	p-value
1	$ret_{Lujiazui}$	0.0944	0.3433	-0.1391	0.1758
2	$ret_{Waigaoqiao}$	0.0823	0.4128	-0.0254	0.8085
3	$ret_{Jiangqiao}$	0.0802	0.5168	0.0169	0.8947
4	ret_{Expo}	0.0669	0.4160	0.1048	0.2191
5	$ret_{Laoximen}$	0.0447	0.3903	0.0254	0.6361
6	ret_{Sanlin}	0.0405	0.5591	0.0494	0.4915
7	$ret_{Yangjing}$	0.0351	0.7180	0.0836	0.4124
8	$ret_{Tianlin}$	0.0184	0.7357	0.0123	0.8277
9	$ret_{Chuansha}$	0.0173	0.8440	0.0115	0.8992
10	$ret_{Jinqiao}$	0.0113	0.8532	0.1763***	0.0061
11	$ret_{Zhongshan}$	0.0022	0.9828	0.1719	0.1040
12	ret_{Huamu}	-0.0364	0.5623	0.0164	0.8004

Table 10 Policy and sentiment changes

The dependent variable is the change in sentiment, or the average sentiment in the [1,*m*] window minus the average sentiment in the [-*m*,1] window. Numbers in italic are p-values. Significant level of 1%, 5%, 10% are marked by *, **, and ***.

$$S_{[t+1,t+m]} - S_{[t-m,t-1]} = c + \beta_1 Loose_t + \beta_2 Tight_t + \beta_3 Spring_t + \beta_4 Autumn_t + \varepsilon_t$$

	m=1	m=2	m=3	m=4	m=5	m=6
Intercept	-0.0204	-0.0035	0.0117	0.0338	0.0543	0.0599**
	0.7366	0.9532	0.8161	0.4155	0.1191	0.0316
Loose	0.2429	0.3391	0.5750*	0.5773**	0.3755*	0.2548
	0.3745	0.3662	0.0730	0.0272	0.0768	0.1161
Tight	-0.0342	-0.0800	-0.0347	-0.0009	0.0245	0.0492
	0.8315	0.6009	0.8046	0.9939	0.7925	0.4929
Spring	0.1991	0.2152*	0.2339**	0.2373**	0.1821**	0.1325**
	0.1341	0.0908	0.0320	0.0108	0.0224	0.0308
Autumn	-0.0141	-0.0025	-0.0693	-0.1342	-0.1448*	-0.1214**
	0.9154	0.9846	0.5340	0.1407	0.0549	0.0376
R^2	5.28%	7.11%	14.27%	22.34%	21.44%	22.40%
Obs	61	59	57	55	53	51

Table 11 Performance of sentiment indexes estimated by future returns in longer horizons

We use the cumulative returns from month t+1 to month t+m to instrument the sentiment in month t. *Corr* is the correlation with the official confidence indexes. The p-value shows the significance level of the correlation. *Obs* is the number of observations. Significant level of 1%, 5%, 10% are marked by *, **, and ***.

		CC_SH	HP_SH	CC_CN	IC_CN	IC_CN_F	IC_CN_P
m=2	Corr	0.3826***	0.1838	0.1185	0.1562	0.0177	0.1364
	p-value	0.0021	0.1878	0.3592	0.2254	0.8917	0.2905
	Obs	62	53	62	62	62	62
m=3	Corr	0.2703**	-0.0090	0.2645**	0.0028	0.0161	0.0331
	p-value	0.0351	0.9491	0.0394	0.9829	0.9022	0.8004
	Obs	61	53	61	61	61	61

Table 12 Sensitiveness to optimistic and pessimistic sentiment: Control for market return

$$ret_{i,t} = c_i + \beta_{i,1} posichg_t + \beta_{i,2} negachg_t + \beta_{i,3} ret_{t-1} + \beta_{i,4} Spring_t + \beta_{i,5} Autumn_t + \beta_{i,6} MKT_t + \varepsilon_{i,t}$$

The *MKT* is the overall housing market return, which is based on transactions in all plates. The *rank_p* variable gives the descending rank of each plate in terms of β_1 value, the *rank_n* variable gives the ascending rank of each plate in terms of β_2 value, and *score* is the average of these two ranks. Significant level of 1%, 5%, 10% are marked by *, **, and ***.

	MKT	p-value	posichg	p-value	negachg	p-value	rank_p	rank_n	score
$ret_{Lujiazui}$	1.1699***	0.0019	0.0717	0.4364	-0.2112**	0.0326	1	1	1
$ret_{Waigaoqiao}$	1.0704***	0.0028	0.0457	0.6266	-0.0634	0.5199	3	3	3
$ret_{Jiangqiao}$	2.0346***	0.0000	0.0037	0.9695	-0.0890	0.3790	8	2	5
ret_{Sanlin}	0.7859***	0.0008	0.0150	0.8136	0.0003	0.9966	5	6	5.5
$ret_{Laoximen}$	0.0099	0.9554	0.0443	0.4033	0.0249	0.6505	4	8	6
ret_{Expo}	0.4124	0.1375	0.0517	0.5277	0.0817	0.3399	2	10	6
$ret_{Tianlin}$	0.2172	0.2402	0.0101	0.8537	0.0003	0.9951	7	7	7
$ret_{Yangjing}$	0.4647	0.1690	0.0134	0.8908	0.0649	0.5243	6	9	7.5
ret_{Huamu}	0.4933**	0.0178	-0.0553	0.3640	-0.0107	0.8656	12	4	8
$ret_{Chuansha}$	0.3747	0.2141	0.0012	0.9889	-0.0074	0.9356	11	5	8
$ret_{Jinqiao}$	0.2579	0.2057	0.0018	0.9760	0.1621**	0.0121	10	11	10.5
$ret_{Zhongshan}$	-0.0017	0.9959	0.0022	0.9826	0.1720	0.1120	9	12	10.5

Table 13 Income, demographic, and sentiment

IncomeG is the growth of per capita disposable income of urban residents in Shanghai. *PopuG* is population growth. In specification [1], it stands for the growth of urban population in Shanghai (*UrbanpopG*); in specification [2], it stands for the growth of total population in Shanghai (*TotpopG*); in specification [3], it stands for the growth of households in Shanghai (*HousehdG*); in specification [4], it stands for the growth of population with Shanghai "Hukou" (*HujipopG*); in specification [5], it stands for the growth of water-users in Shanghai (*WateruserG*). *GenderG* is the growth of male-to-female ratio among people with Shanghai "Hukou". *Jan* is a dummy that equals to 1 for January, and 0 otherwise. Other month dummies are defined similarly. *PMI*, *ReProf*, *ReLoan*, *SHBigIndProd*, *CPI*, *M2G*, *Defaultr*, and *Term* are defined at the beginning of section 4.2.

$$\begin{aligned}S_t &= c + \beta_1 IncomeG_t + \beta_2 PopuG_t + \beta_3 GenderG_t \\ &+ \beta_4 PMI_t + \beta_5 Reprof_t + \beta_6 CPI_t + \beta_7 M2G_t + \beta_8 SHBigIndProd_t + \beta_9 ReLoan_t + \beta_{10} Defaultr_t + \beta_{11} Term_t \\ &+ \beta_{12} Jan_t + \beta_{13} Feb_t + \beta_{14} Mar_t + \beta_{15} Apr_t + \beta_{16} May_t + \beta_{17} Jun_t + \beta_{18} Jul_t + \beta_{19} Aug_t + \beta_{20} Sep_t + \beta_{21} Oct_t + \beta_{22} Nov_t + \varepsilon_t\end{aligned}$$

	[1]		[2]		[3]		[4]		[5]	
	Coeff.	p	Coeff.	p	Coeff.	p	Coeff.	p	Coeff.	p
Intercept	-2.78	0.43	-14.31*	0.09	1.47	0.54	3.78*	0.09	-2.92	0.42
IncomeG	9.64	0.18	-8.04***	0.01	2.39	0.56	-11.77***	0.00	10.92	0.15
UrbanpopG	12.29**	0.03								
TotpopG			236.36**	0.03						
TothousehdG					65.77**	0.03				
HujipopG							-288.98**	0.03		
WateruserG									14.49**	0.03
GenderG	784.44**	0.03	2389.94**	0.03	258.46**	0.03	-357.66*	0.05	883.39**	0.03
PMI	-0.04	0.31	-0.04	0.31	-0.04	0.31	-0.04	0.31	-0.04	0.31
ReProf	0.55*	0.08	1.87**	0.04	0.10	0.52	0.12	0.46	0.56*	0.07
ReLoan	-2.45	0.41	-2.45	0.41	-2.45	0.41	-2.45	0.41	-2.45	0.41
SHBigIndProd	2E-4***	0.01	2E-4***	0.01	2E-4***	0.01	2E-4**	0.01	2E-4***	0.01
CPI	-13.37	0.36	-13.37	0.36	-13.37	0.36	-13.37	0.36	-13.37	0.36
M2G	0.82	0.87	0.82	0.87	0.82	0.87	0.82	0.87	0.82	0.87
Defaultr	-6.25	0.59	-6.25	0.59	-6.25	0.59	-6.25	0.59	-6.25	0.59
Term	-60.60	0.22	-60.60	0.22	-60.60	0.22	-60.60	0.22	-60.60	0.22
Jan	0.26	0.36	0.26	0.36	0.26	0.36	0.26	0.36	0.26	0.36
Feb	0.16	0.49	0.16	0.49	0.16	0.49	0.16	0.49	0.16	0.49
Mar	0.32	0.13	0.32	0.13	0.32	0.13	0.32	0.13	0.32	0.13
Apr	0.45**	0.04	0.45**	0.04	0.45**	0.04	0.45**	0.04	0.45**	0.04
May	0.26	0.22	0.26	0.22	0.26	0.22	0.26	0.22	0.26	0.22
Jun	0.08	0.68	0.08	0.68	0.08	0.68	0.08	0.68	0.08	0.68
Jul	0.05	0.85	0.05	0.85	0.05	0.85	0.05	0.85	0.05	0.85
Aug	0.07	0.71	0.07	0.71	0.07	0.71	0.07	0.71	0.07	0.71
Sep	0.10	0.55	0.10	0.55	0.10	0.55	0.10	0.55	0.10	0.55
Oct	0.01	0.95	0.01	0.95	0.01	0.95	0.01	0.95	0.01	0.95
Nov	0.00	1.00	0.00	1.00	0.00	1.00	0.00	1.00	0.00	1.00
R^2	66.64%		66.64%		66.64%		66.64%		66.64%	
Obs	57		57		57		57		57	

Table 14 Housing market return and sentiment changes

$$ret_t = c + \alpha \Delta S_t + \beta_1 IncomeG_t + \beta_2 PopuG_t + \beta_3 GenderG_t$$
$$+\beta_4 PMI_t + \beta_5 Reprof_t + \beta_6 CPI_t + \beta_7 M2G_t + \beta_8 SHBigIndProd_t + \beta_9 ReLoan_t + \beta_{10} Defaultr_t + \beta_{11} Term_t$$
$$+\beta_{12} Jan_t + \beta_{13} Feb_t + \beta_{14} Mar_t + \beta_{15} Apr_t + \beta_{16} May_t + \beta_{17} Jun_t + \beta_{18} Jul_t + \beta_{19} Aug_t + \beta_{20} Sep_t + \beta_{21} Oct_t + \beta_{22} Nov_t + \varepsilon_t$$

The dependent variable *ret* is the monthly return of the repeat-sale index. ΔS_t is S_t minus S_{t-1}. Other variables are the same as in **Table 13**.

	[1]		**[2]**	
	Coeff.	p	Coeff.	p
Intercept	-0.5753	0.4175	-0.2488	0.5890
ΔS_t	0.0730**	0.0244	0.0730**	0.0244
IncomeG	0.1722	0.9020	-0.8929	0.2428
UrbanpopG	0.6112	0.5753		
HujipopG			-14.3763	0.57533
GenderG	48.7669	0.4780	-8.0504	0.8184
PMI	0.0116	0.1560	0.0116	0.1560
ReProf	0.0256	0.6730	0.0042	0.8981
ReLoan	-0.6313	0.2943	-0.6313	0.2943
SHBigIndProd	0.0000	0.8707	0.0000	0.8707
CPI	2.5257	0.4186	2.5257	0.4186
M2G	0.1417	0.9020	0.1417	0.9020
Defaultr	-0.5506	0.8117	-0.5506	0.8117
Term	-4.2206	0.6626	-4.2206	0.6626
Jan	0.0426	0.4654	0.0426	0.4654
Feb	-0.0203	0.6666	-0.0203	0.6666
Mar	0.0165	0.7022	0.0165	0.7022
Apr	0.0147	0.7588	0.0147	0.7588
May	0.0457	0.2928	0.0457	0.2928
Jun	0.0040	0.9206	0.0040	0.9206
Jul	0.0206	0.6967	0.0206	0.6967
Aug	0.0355	0.3759	0.0355	0.3759
Sep	0.0416	0.2343	0.0416	0.2343
Oct	0.0080	0.8583	0.0080	0.8583
Nov	-0.0305	0.4128	-0.0305	0.4128
Lagret	-0.4344***	0.0065	-0.4344***	0.0065
R^2	54.92%		54.92%	
Obs	57		57	

Appendix A

Table A1 Abbreviation of intervention tools

"PFL" refers to the Housing Provident Fund Loans, which is part of China's social welfare system.

Tool	Abbreviation
Down payment: All houses	DP
Mortgage rate: All houses	MR
Down payment: First house	DP-1
Mortgage rate: First house	MR-1
Down payment: Second house	DP-2
Mortgage rate: Second house	MR-2
Mortgage Availability: Third house	MA-3
Ceiling of PFL-total value ratio: First house	DP-PFL-1
Ceiling of PFL: First house	Tot-PFL-1
Ceiling of PFL-total value ratio: Second house	DP-PFL-2
Ceiling of PFL: Second house	Tot-PFL-2
PFL Mortgage rate: Second house	MR-PFL-2
PFL Availability: Third house	PFLA-3
Restriction of purchase: with "Hukou"	RP-SH
Restriction of purchase: without "Hukou"	RP-NSH
Tax on resale	T-RS
Other taxes	T-O
Resale of relocation houses	RR
Mortgage availability: without "Hukou"	MA-NSH
Ask price of first-hand houses	NHP

Appendix B: Proof of $P_0 < P_{T_0}$

Recall formula (13). Since the social dynamic path is public information, we have:

$$E_s^j[E_t^o(\varepsilon^f)] = E_s^o(\varepsilon^f)\text{ , } s<t$$

Therefore,

$$E_0^s(P_{t_1}) - P^s = P_0^o - [\beta(1-\phi)]^{t_2+1-t_1}(P_0^o - P^s) - P^s = (P_0^o - P^s)\left\{1-[\beta(1-\phi)]^{t_2+1-t_1}\right\}$$

$$E_{T_0}^s(P_{t_1^{old}}) - P_{old}^s = P_{T_0}^o - [\beta(1-\phi)]^{t_2^{old}+1-t_1^{old}}(P_{T_0}^o - P_{old}^s) - P_{old}^s = (P_{T_0}^o - P_{old}^s)\left\{1-[\beta(1-\phi)]^{t_2^{old}+1-t_1^{old}}\right\}$$

Since the social dynamic is deterministic, we have

$$t_2 + 1 - t_1 = t_2^{old} + 1 - t_1^{old}$$

Notice that,

$$P_0^o - P^s = \beta\frac{\phi E_0^o(\varepsilon^f)/(1-\beta) + (1-\phi)\varepsilon}{1-\beta(1-\phi)} - \beta\frac{\varepsilon}{1-\beta} = \frac{\phi\beta[E_0^o(\varepsilon^f) - \varepsilon]}{(1-\beta)[1-\beta(1-\phi)]} = \frac{\phi\beta S_0}{(1-\beta)[1-\beta(1-\phi)]}$$

$$P_{T_0}^o - P_{old}^s = \frac{\phi\beta S_{T_0}}{(1-\beta)[1-\beta(1-\phi)]}$$

$$S_0 - S_{T_0} = E_{T_0}^o(\varepsilon^f) + u_0 - \varepsilon - [E_{T_0}^o(\varepsilon^f) - \varepsilon_{old}] = \varepsilon_{old} - \varepsilon + u_0$$

Because of the tightening policy that arrives at time zero, u_0=-σ.
By assumption, $-\sigma<\varepsilon-\varepsilon^{old}<0$. Hence,

$$E_0^s(P_{t_1}) - P^s < E_{T_0}^s(P_{t_1^{old}}) - P_{old}^s\ .$$

Recall that

$$P_0 - P_{T_0} = P^s - P_{old}^s + [\beta(1-\phi)]^{t_1}[E_0^s(P_{t_1}) - P^s] - [\beta(1-\phi)]^{t_1^{old}-T_0}[E_{T_0}^s(P_{t_1^{old}}) - P_{old}^s]\ ,$$

where $P^s < P_{old}^s$ and $t_1 > t_1^{old} - T_0$.

So, $P_0 - P_{T_0} < 0$.

Appendix C: Proof of the condition for $P_t>P_0$

$$P_t - P_0 = [\beta(1-\phi)]^{t_1-t}[E_t^s(P_{t_1}) - P^s] - [\beta(1-\phi)]^{t_1}[E_0^s(P_{t_1}) - P^s]$$

$$= [\beta(1-\phi)]^{t_1-t}\frac{\phi\beta\left\{1-[\beta(1-\phi)]^{t_2+1-t_1}\right\}}{(1-\beta)[1-\beta(1-\phi)]}\left\{S_t - [\beta(1-\phi)]^t S_0\right\}$$

$$S_t - [\beta(1-\phi)]^t S_0 = E_t^o(\varepsilon^f) - \varepsilon - [\beta(1-\phi)]^t[E_0^o(\varepsilon^f) - \varepsilon]$$

$$= E_t^o(\varepsilon^f) - [\beta(1-\phi)]^t E_0^o(\varepsilon^f) - (1-[\beta(1-\phi)]^t)\varepsilon$$

$$= E_0^o(\varepsilon^f) + \sum_{\tau=1}^{t} u_\tau - [\beta(1-\phi)]^t E_0^o(\varepsilon^f) - (1-[\beta(1-\phi)]^t)\varepsilon$$

$$= E_{T_0}^o(\varepsilon^f) + \sum_{\tau=0}^{t} u_\tau - [\beta(1-\phi)]^t E_{T_0}^o(\varepsilon^f) - u_0[\beta(1-\phi)]^t - (1-[\beta(1-\phi)]^t)\varepsilon$$

So, when $E^o_{T_0}(\varepsilon^f) > \varepsilon - \dfrac{\sum_{\tau=1}^{t} u_\tau}{1-[\beta(1-\phi)]^t} - u_0$,

we have $S_t - [\beta(1-\phi)]^t S_0 > 0$, and thus $P_t > P_0$

Similarly, for any 0< $n<m<t_1$, we have

$$P_m - P_n = [\beta(1-\phi)]^{t_1-m} \frac{\phi\beta\left\{1-[\beta(1-\phi)]^{t_2+1-t_1}\right\}}{(1-\beta)[1-\beta(1-\phi)]}\left\{S_m - [\beta(1-\phi)]^{m-n} S_n\right\}$$

$$\begin{aligned}
&S_m - [\beta(1-\phi)]^{m-n} S_n \\
&= E^o_m(\varepsilon^f) - \varepsilon - [\beta(1-\phi)]^{m-n}[E^o_n(\varepsilon^f) - \varepsilon] \\
&= E^o_n(\varepsilon^f) + \sum_{n+1}^{m} u_\tau - [\beta(1-\phi)]^{m-n} E^o_n(\varepsilon^f) - (1-[\beta(1-\phi)]^{m-n})\varepsilon \\
&= (1-[\beta(1-\phi)]^{m-n}) E^o_n(\varepsilon^f) + \sum_{n+1}^{m} u_\tau - (1-[\beta(1-\phi)]^{m-n})\varepsilon
\end{aligned}$$

Therefore, as long as:

$$E^o_{T_0}(\varepsilon^f) > \varepsilon - \frac{\sum_{n+1}^{m} u_\tau}{(1-[\beta(1-\phi)]^{m-n})} - \sum_{\tau=0}^{n} u_\tau$$

we have $P_m > P_n$.