Loss Aversion and Residential Property Development Decisions in China: A Semi-Parametric Estimation

Helen X. H. Bao
Department of Land Economy, University of Cambridge, CB3 9EP, UK
Email: hxb20@cam.ac.uk
Tel: +44 1223 337 116
Fax: +44 1223 337 130

and

Charlotte Chunming Meng
Department of Land Economy, University of Cambridge, CB3 9EP, UK
Email: cm794@cam.ac.uk
Tel: +44 1223 337 116
Fax: +44 1223 337 130
Abstract

Loss aversion is a core concept in prospect theory that refers to people’s asymmetric attitudes with respect to gains and losses. More specifically, losses loom larger than gains. With the capability of loss aversion to explain economic phenomena, some of which are puzzling under expected utility theory, this concept has received significant attention. This study develops a behavioral model of loss aversion to explain the development decisions by residential property developers in the People’s Republic of China. Under the leasehold property right system, real estate development has two stages—first to lease land from the government, and then to develop the property according to the lease terms. This presents a unique opportunity to test the presence and effect of loss aversion in real estate development decisions. More specifically, this study determines when the land premium paid by a developer is substantially higher than the market value, whether and how this “paper loss” will affect the pricing of the housing products and development time of the project in future development. We use a sample of land and house transaction records from Beijing to test the hypothesis. This is the first study to use a semi-parametric model in estimating developers’ loss aversion. Results show that developers are most prone to loss aversion bias around the reference point or when facing large losses. The results also suggest that loss aversion contributes to the cyclical trading pattern in housing markets.
1. INTRODUCTION

Loss aversion is one of the core concepts in prospect theory. The theory is often suggested as an alternative paradigm to classical expected utility theory and has been proven to be more useful in explaining choice behavior under uncertainty. Under prospect theory, people’s utility does not come from the absolute level of wealth but from the loss and gain that are derived from a comparison with a reference point. When the reference point is determined, the utility function that is defined based on the losses and gains is S-shaped, which determines the other two crucial features of prospect theory other than reference-dependence. One feature is the decrease in the marginal change of utility when gain or loss levels increase (diminishing sensitivity). The other feature is losses looming larger than gains; this phenomenon is usually referred to as loss aversion. In other words, a loss-averse decision maker would experience more utility drops when facing a loss than the utility increase from a gain of the same magnitude. This feature is the focus of this study. Solid evidence has confirmed that loss aversion plays an important role in decision-making processes under uncertainty (Bleichrodt, Pinto, and Wakker 2001; Berkelaar, Kouwenberg, and Post 2004; Abdellaoui, Bleichrodt, and Paraschiv 2007; Pope and Schweitzer 2011). The concept also helps to explain a variety of economic phenomena, some of which are puzzling circumstances under expected utility theory.¹

We explored whether real estate developers’ pricing behavior are prone to the behavioral bias of loss aversion. Previous research has revealed one puzzling feature of the housing market: house price and trading volume are positively correlated (Norman and Michael 1986). In an up market, the trading volume would increase substantially despite the high selling prices. However, in a down market, most of the houses would sit on the market with a very high price and would be eventually pulled out without a sale. This cyclical trading pattern was demonstrated by Genesove and Mayer (2001) as the disposition effect on the housing market. Such effect implies that when prices are high in an up market, most sellers are in a gain domain and would obtain a deal quickly, whereas when prices are high in a down market, sellers are reluctant to adapt their selling price and would hold on to the property and spend more time waiting.

In this paper, we offer a behavioral explanation of this puzzling phenomenon by including three innovative elements: a new focus on property developers rather than household sellers and buyers, an expectation-based reference point rather than using initial purchase prices, and a semi-parametric estimation rather than an ordinary least squares estimation. A developer-focused perspective is very important, yet is missing in existing literature. Developers are extremely significant in determining market conditions, especially for a market dominated by new homes. A semi-parametric estimation model is used so as to allow the degree of loss aversion to vary with the magnitude of losses and gains.

One of the challenges faced by loss aversion studies on the real estate market, as well as other markets, is determining the reference point, the benchmark upon which each outcome is coded as a gain or a loss. The coding is crucial because it determines two features of the decision maker: (i) risk attitudes and (ii) sensitivity. The coding is

¹ Such economic puzzles include, among others, the equity premium puzzle (see, for example, Benartzi and Thaler 1995; Gneezy and Potters 1997; Gneezy, Kapteyn, and Potters 2003), the endowment effect (see, for example, Kahneman, Knetsch, and Thaler 1990; van Dijk and van Knippenberg 1996, 1998), and the disposition effect (see, for example, Odean 1998, Barberis and Xiong 2009, Da Costa et al. 2013).
important to risk attitudes because utility in the gain domain is concave, whereas it is convex in the loss domain, which implies risk aversion for gains and risk seeking for losses. Moreover, the coding is important to sensitivity because people are more sensitive to losses, which implies different slopes in the two domains and a kink at the reference point. However, no agreement has been reached on the location of the reference point in the existing literature. Potential candidates include the status quo, the certain equivalent, the previous purchase price, and the recent expectation. However, these candidates have not been compared. Real estate studies, such as the influential study of Genesove and Mayer (2001) on the Boston condominium market and the follow-up studies by Bokhari and Geltner (2011), Leung and Tsang (2013), and Anenberg (2011), showed that the previous purchase price is the most used candidate. Undoubtedly, being the initial cost, previous purchase price could perform naturally as a benchmark for monetary gains and losses. More importantly, it is observable. Nonetheless, initial purchase prices do not incorporate new market information that could lead to adaptation of the reference point. Therefore, Koszegi and Rabin (2006) proposed an expectation-based reference point and argued that it is superior to previous purchase price. This study adopted the proposal by Koszegi and Rabin (2006) and used the expected selling price as the reference point. Specifically, we chose average land purchase price within 3 miles as the reference land price. The expectation-based selling price as the reference point is an alternative to the previous purchase price used by previous loss aversion studies; hence our finding helps to test whether the previous findings are robust, which is one of the contributions of this study.

The attempt to fill the research gap of loss aversion studies in the People’s Republic of China (PRC) real estate market is another contribution of this study. The PRC has a booming and influential housing market with great potential. Therefore, understanding the behavior of the participants in this market is highly crucial for researchers, practitioners, and policy makers. However, many differences exist between developing and developed markets in terms of the market participants’ knowledge and experience, market regulation, and market efficiency. These differences highlight the necessity to explore whether the findings derived from developed markets are applicable to the PRC. Two of the few behavioral studies on the PRC are those of Leung and Tsang (2013) and He and Asami (2014). The study conducted by Leung and Tsang (2013) empirically confirmed the existence of loss aversion. However, it is based on the data from the Hong Kong, China housing market, which differs from most cities in the PRC in the housing transaction system and trading preferences. Therefore, to determine whether the finding is applicable to PRC cities, further exploration is needed. The study of He and Asami (2014) focuses on the endowment effect and employs survey data instead of transaction data. Under the leasehold property right system, real estate development in the PRC has two stages—leasing land from the government and developing the property according to the lease terms. This two-stage nature presents a unique opportunity to test the presence and effect of loss aversion in real estate development decisions. Thus, this paper constructed a model based on this nature and contributes to the limited literature on the behavior of an important housing market participant in the PRC. Our research could assist developers, investors, and policy makers in judging the market better and in making more effective decisions.

We analyzed the September 2003–June 2014 market data of Beijing, which were derived from property development projects (land purchase, house construction, and sales). Semi-parametric estimation of the value function in prospect theory shows that the loss aversion level changes with the magnitude of the losses and gains. Specifically, developers are most loss averse around their reference point and when they face large losses or gains. When facing medium-sized losses or gains, the loss aversion effect disappears.
2. LITERATURE REVIEW

In this section, we first present an overview of prospect theory and an introduction to loss aversion. Second, we provide a discussion of two important issues in current loss aversion studies: determining the reference point and further methods of identifying or measuring loss aversion. Finally, we present an empirical implementation issue about the data collection method used in the current loss aversion studies.

2.1 Prospect Theory

Although expected utility theory has been regarded as a dominant descriptive and normative model that characterizes the rational choice behavior of agents (Friedman and Savage 1948), researchers often claim that the choices of people deviate systematically from the optimal outcome of expected utility theory (Allais 1953, Samuelson and Zeckhauser 1988). To account for the discrepancy, Kahneman and Tversky (1979) proposed prospect theory as an alternative. Both field and experiment data have shown that prospect theory is probably the most descriptively valid model to measure behavior under risk and uncertainty (Starmer 2000).

Prospect theory distinguishes two phases in the decision-making process: an early phase of editing and a subsequent phase of evaluation. In the editing phase, people reorganize and reformulate the given prospects to obtain a simpler presentation of these prospects. When the final presentation is obtained, the decision-making process proceeds to the second phase, in which people evaluate each of the edited prospects and then choose the prospect with the highest value.

The value of each prospect is a weighted average of the values of the outcomes. However, unlike expected utility theory, which uses given probabilities as weights and the levels of outcome as values to be weighted, prospect theory defines a new weighting function \( \pi \) and value function \( v \).

Let \((x, p; y, q)\) denote a simple prospect or gamble with at most two non-zero outcomes.\(^2\) The decision maker receives either \(x\) with a possibility of \(p\) or \(y\) with a possibility of \(q\). If \(V\) represents the overall utility that one gains from this prospect or gamble, then the evaluation process described in the previous section can be formulated as Equation (1).

\[
V(x, p; y, q) = \pi(p)v(x) + \pi(q)v(y)
\]  

The weighting function \(\pi(\cdot)\) assigns each possibility \(p\) with a decision weight \(\pi(p)\), which measures the level of people’s subjective probability distortion. The function is monotonic and increasing with \(p\), with \(\pi(0) = 0\) and \(\pi(1) = 1\). A salient property of \(\pi\) is that it is not equal to the pure likelihood of the event (i.e., \(\pi(p) \neq p\)) in most cases. Specifically, the weighting function overestimates low possibilities and underestimates moderate and high possibilities. A detailed discussion can be found in Kahneman and Tversky (1979), but it is omitted here because the weighting function is not related to loss aversion.

\(^2\) An extended model with more than two outcomes is easily obtainable. For a model with a large number of outcomes, see discussions about cumulative prospect theory by Tversky and Kahneman (1992).
The value function $v(\cdot)$ assigns a value $v(x)$ to each outcome $x$, which reflects the subjective attitudes of the decision maker. The value function is characterized by the following properties:

(i) The value function is defined by the changes of wealth relative to a reference point, that is, $x$ equals the gain or loss, instead of the absolute wealth level such as that in expected utility theory. If $r$ represents the reference point and $x_0$ equals the wealth level after obtaining the outcome, then $x = x_0 - r$. A detailed discussion on the different choices of the reference point is in section 2.2.

(ii) The value function exhibits diminishing sensitivity toward changes in $x$ (either positive or negative) as magnitudes increase. For example, people experience more happiness or sadness when gains or losses increase from 100 to 200 than when the increase is from 1,000 to 1,100. This feature implies that the value function is generally concave for gains (i.e., $v'(x) \leq 0, x > 0$) and commonly convex for losses (i.e., $v'(x) \geq 0, x < 0$).

(iii) "Losses loom larger than gains." Therefore, with the same amount of change in $x$, the value change in the loss domain becomes larger than the value change in the gain domain. The value function is therefore steeper for losses than for gains, thus creating an imperfection at the reference point. This feature is commonly referred to as loss aversion.

There is no agreement on the functional form of the value function in the existing literature till now. A very widely accepted value function example that features the three abovementioned elements is suggested by Tversky and Kahneman (1992):

$$v(x) = \begin{cases} x^\alpha, & x \geq 0 \\ \lambda(-x)^\beta, & x < 0 \end{cases}$$

(2)

where $x$ is the gain or loss that is compared with the reference point, $\alpha$ and $\beta$ are positive values between 0 and 1, and $\lambda$ is the coefficient that measures the degree of loss aversion. A $\lambda$ that is greater than 1 captures the effect of loss aversion. The implementation of the equation and the measurement of loss aversion are crucial in verifying the existence and degree of loss aversion. Other alternatives of measurement are discussed in section 2.3.

The concept of loss aversion is widely employed in (i) the financial market (equity premium puzzle by Benartzi and Thaler [1995]; disposition effect by Odean [1998], Kyle et al. [2006], and Henderson [2012]); (ii) consumption choice problems (people’s asymmetric response to price changes by Putler [1992] and Ray et al. [2015]; asymmetric response to promotions of different brands by Hardie, Johnson, and Fader [1993] and Bronnenberg and Wathieu [1996]; and endowment effect by Thaler [1980] and Kahneman, Knetsch, and Thaler [1990]); and (iii) effort provision problems (downward-sloping labor supply by Camerer et al. [1997], Kőszegi and Rabin [2006], Farber [2008], and Crawford and Meng [2011]). Other promising applications include the optimal contract form and compensation schemes for loss-averse chief executive officers (de Meza and Webb 2007, Dittmann et al. 2010, and Herweg et al. 2010); consumption insensitivity to bad news of future income (Bowman, Minehart, and Rabin 1999); optimal ordering when managers or news vendors are loss averse (Ho et al. 2010, Wang and Webster 2007, and Wang 2010); and so on.
Evidence shows that loss aversion bias exhibits different levels in different contexts. For example, the level depends on the nature of the outcome: in the context of monetary outcomes, the loss aversion level is estimated to be 2.25 by Tversky and Kahneman (1992); in the context of health decisions, the level is estimated to be 3.06 by Bleichrodt, Pinto, and Wakker (2001). Regional differences are also identified. Abdellaoui et al. (2013) estimate loss aversion in dimensions of risk and time and find that loss aversion levels are consistently higher in Rotterdam than in Paris. A number of other factors cause loss aversion to vary among individuals; examples include experience (Haigh and List 2005, List 2003), education (Booij and van de Kuilen 2009), gender (Booij and van de Kuilen 2009, Brooks and Zank 2005), framing (Keysar et al. 2012), and so on. The results highlight that loss aversion levels vary for different decisions and that they exhibit heterogeneity among individuals.

### 2.2 Role of Reference Points

As the starting point for prospective losses and gains, a reference point is essential to loss aversion studies. It is the benchmark that each outcome or wealth level is compared against prior to coding and evaluating as a gain or loss (Kahneman 1992). The coding determines two important features of the decision maker: (i) risk attitudes and (ii) sensitivity. The coding is significant to risk attitudes because utility in the gain domain is concave, whereas that in the loss domain is convex, which implies risk aversion for gains and risk seeking for losses. The coding is important to sensitivity because people are more sensitive to losses, which implies different slopes in the two domains and a kink at the reference point. Kahneman and Tversky (1979, 1992) mentioned a number of potential factors such as the status quo, the formulation of the offered prospects, and the expectations of the decision maker. However, they did not provide clear guidance regarding the reference point. Despite the increasing number of studies on people’s irrational reaction toward the departures from the reference point, no agreement has been reached on the true nature of the reference point and the way it is formed. Precise determination of the reference point is difficult because the point is an intermediate variable and cannot be observed directly (Paraschiv and Chenavaz 2011). Such variables are unavailable from existing data sources and are difficult to measure accurately. Different people may use different reference points, and even the reference points of the same people may vary over time (Winer 1986; Hardie, Johnson, and Fader 1993); these situations increase the complexity of the problem. Evidence shows that reference points shift on different occasions. Most of the existing studies assumed that a reference point is set without discussion about the validity of the reference point and that the model and analysis are based on this assumption. Nonetheless, without the validity of the reference point, the whole analysis could be completely inaccurate and unreliable.

Several potential candidates for the reference point include the status quo, the initial cost for the same goods, the anticipated price, or the combination of the three. The most widely discussed and used reference point is the status quo (i.e., the current state of the decision maker) because it is believed that people generally intend to maintain the current state. Such belief parallels one of the behavioral anomalies called status quo bias, which refers to people’s reluctance to move from the current status (Samuelson and Zeckhauser 1988). The status quo reference point has been used in various studies such as those of Ert and Erev (2013); Booij, van Praag, and van de Kuilen (2010); and Barberis et al. (2001).
Koszegi and Rabin (2006) proposed the decision maker’s recent expectation of the outcome as an alternative for the reference point. Such a proposal was aimed at reconciling some seemingly contradictory predictions on risk attitudes and other behavioral biases. They argue that people’s expectation, which is determined endogenously by the current economic environment, makes predictions that are better than those of the status quo. The existing studies that incorporate expectation as the reference point include those of Winer (1986); Paul and Koszegi (2008); Loomes and Sugden (1986); and Ang, Bekaert, and Liu (2005).³ This type of reference point is advantageous because for fleeting activities where no ownership is involved, such as shopping, entertainment, travel, and surgery, status quo is absent. In these cases, expectation is the reasonable reference point (Koszegi and Rabin 2006). For example, when a person is told to undergo a dental surgery, an expectation is formed. Therefore, when a subsequent checkup confirms that the surgery is not necessary, a gain is experienced, which is different from the expectation. This example is beyond the reach of the status quo reference point.

In the housing market, measuring the reference point is even more difficult because of the high level of heterogeneity of housing products, infrequent transactions, and the subsequent lack of transaction data. In addition, intermediary evaluations of the house value as well as the future price expectation may also affect the reference formation. At present, two types of reference points have been proposed in the housing market: previous purchase price and anticipated future house price.

As the prior cost, the initial purchase price of the property is regarded as a natural reference because it marks clearly whether money is gained or lost through a transaction. When one obtains a higher price than the price one paid to acquire the property, a prospective gain exists; otherwise, a loss is expected. The initial purchase price has already been regarded as a reference in some highly influential studies on the housing market (Genesove and Mayer 2001, Leung and Tsang 2013, Anenberg 2011) and on the commercial real estate market (Bokhari and Geltner 2011).

However, the housing market is not constant, and therefore the seller’s reference point should be regularly updated, especially when the time between two sales is relatively long. This situation resulted in the proposal of Koszegi and Rabin (2006) to use the expected house price as the reference point rather than the previous price or the adapted current price. Buyers would adapt the reference point because they would certainly gain information or stimuli about the new market condition. The problem has been pointed out by Genesove and Mayer (2001), who proposed that the fixed previous purchase price reference may cause the bizarre coefficients for housing price hedonic regressions. In this research, the practice of Koszegi and Rabin (2006) was followed; thus, we assumed that the developers use their anticipated price as reference point. The underlying reason for this assumption is that decision makers are fully aware of the market condition; thus, the expected price is determined endogenously by the market condition. This assumption is reasonable because the agents in our study are real estate developers who are very experienced and professional participants.

### 2.3 Identification of Loss Aversion

Numerous studies have explored the existence of loss aversion either in experimental settings or in real market conditions. Their identification methods are labeled by the

---

³ Loomes and Sugden (1986) and Ang, Bekaert, and Liu (2005) used expectation that is endogenously determined by a certain equivalent as the reference point. We also considered it as expectation-based because, in some cases, people form their expectations based on the certain equivalent value.
authors as a direct measure or an indirect measure depending on whether these methods estimated the value function and loss aversion parameter itself. This section explores the different direct loss aversion measurements in the present literature as well as some indirect methods that have been used to identify loss aversion.

Table 1 lists some definitions of loss aversion for direct measurements from the existing literature. The first definition is introduced with prospect theory by Kahneman and Tversky (1979) and denotes that utility drop is greater than utility increase when the losses and gains have the same magnitude. This definition by Tversky and Kahneman (1992) is similar to a special case where \( x = 1 \). Wakker and Tversky (1993) were the first to use derivatives in loss aversion measurements. Their definition implies that, for the same level of losses and gains, the marginal increase caused by a small loss decrease, that is, the slope, should be at least the same as the marginal increase caused by the small gain increase. Subsequently, Bowman, Minehart, and Rabin (1999) proposed a significantly stronger measurement: the ratio of the lower bound of the slope of the value function in the loss domain and the upper bound of the slope of the value function in the gain domain. In this definition, loss aversion means that regardless of the losses and gains, a marginal change in the gain domain is lower than a marginal change in the loss domain. Köbberling and Wakker (2005) provided the first and only non-global definition, which states that loss aversion is the ratio between the left derivative and the right derivative of \( v(\cdot) \) at the reference point.

Table 1: Definitions of Loss Aversion in Direct Measurements

<table>
<thead>
<tr>
<th>Authors</th>
<th>Definition</th>
<th>Literature</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kahneman and Tversky (1979)</td>
<td>( -\frac{v(-x)}{v(x)} )</td>
<td>Bleichrodt, Pinto, and Wakker (2001)</td>
</tr>
<tr>
<td>Tversky and Kahneman (1992)</td>
<td>( -\frac{v(-1)}{v(1)} )</td>
<td>Booij, van Praag, and van de Kuilen (2010)</td>
</tr>
<tr>
<td>Wakker and Tversky (1993)</td>
<td>( \frac{v'(x)}{v'(x)} )</td>
<td>Schmidt and Traub (2002)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Pennings and Smidts (2003)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Gurevich, Kliger, and Levy (2009)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>von Gaudecker, van Soest, and Wengström (2011)</td>
</tr>
<tr>
<td>Bowman, Minehart, and Rabin (1999)</td>
<td>( \inf \frac{v'(-x)}{v'(y)} )</td>
<td>Abdellaoui et al. (2007)</td>
</tr>
<tr>
<td>Köbberling and Wakker (2005)</td>
<td>( \frac{v'_l(0)}{v'_r(0)} )</td>
<td>Booij and van de Kuilen (2009)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Abdellaoui et al. (2013)</td>
</tr>
</tbody>
</table>

Notes: \( v(x) \) is the value function defined in 2.1. All \( x, y > 0 \). \( \inf \frac{v'(-x)}{v'(y)} \) is the lower bound of the slope of the value function in the loss domain, whereas \( \sup v'(y) \) is the upper bound of the slope of the value function in the gain domain. \( v'_l(0) \) denotes the left derivative of \( v(\cdot) \) at the reference point, and \( v'_r(0) \) denotes the right derivative.

Two types of direct definitions are included: global and local. There is a clear difference between the two types. To attain the global index, as proposed by Kahneman and Tversky (1979); Tversky and Kahneman (1992); Wakker and Tversky (1993); and Bowman, Minehart, and Rabin (1999), the entire data range should be scanned, which leads to ambiguous results, depending on the different points chosen. However, as proposed by Köbberling and Wakker (2005), only one specific number is returned for the local reference point. Another difference lies in the incorporation of the curvature of the value function by the global measurements and the possibility of distinguishing loss aversion from the curvature by the local definitions. Furthermore, identification of loss aversion based on global definitions can be too strict for empirical purposes since it includes the whole domain of definition whereas the local definitions use only several
points in the domain. Therefore, global definitions are suitable for theoretical purposes, whereas local definitions could remain empirical.

Another method of identifying loss aversion is the indirect way. Unlike the previous loss aversion measurement method that is based on estimation or hypothesis of the value function, the indirect way involves no value function but focuses on the direct effect of losses and gains on the final decision, such as price, time, or effort. This method is widely employed when transactional data instead of experimental data are used. For example, when housing studies probe whether sellers on the real estate market exhibit loss aversion, prices are assumed as a function of the value of gain and loss. The estimated coefficient of loss and gain captures the effect of loss aversion. The model can be summarized as follows:

\[ p = f(\mu, \text{Loss}, \text{Gain}) \]  

(3)

where \( \text{Loss} \) and \( \text{Gain} \) are defined as the truncated deviation from the reference point. In a loss, the variable \( \text{Loss} \) equals the deviation and \( \text{Gain} \) equals zero. In a gain, the variable \( \text{Gain} \) equals the deviation and \( \text{Loss} \) equals zero. When the magnitude of the coefficient of \( \text{Loss} \) is greater than that of \( \text{Gain} \), sellers are prone to loss aversion.

The choice of using direct or indirect measurements in the empirical work depends largely on the type of data used. On the one hand, when experimental data are available, direct measurement of loss aversion is appropriate and reliable because under the specially designed experiments, controlling for a variety of disturbing factors and eliciting the utility function are significantly easier. In fact, all the literature listed in Table 1 is based on experiments. On the other hand, indirect measurement is more suitable for studies using transactional data because eliciting the utility function can be very difficult given the complexity of real-life decisions. In this sense, we followed the indirect method because our research is based on transactional data.

2.4 Data Collection Method

Apart from the theoretical issues discussed above, another crucial issue in the empirical implementation in loss aversion studies is related to data. Laboratory data are widely used in loss aversion studies because loss aversion is not directly observed and is difficult to distinguish from other related factors in real market settings. However, in a laboratory setting, researchers have the luxury of controlling for other factors and eliciting information to calculate the net effect of loss aversion. This approach enhances the conceptual validity of the study. Nonetheless, it is problematic and not applicable to housing studies. Nearly 80% of laboratory-based loss aversion studies recruited students as experimental subjects. Student participation is achieved by designing the experiment as an academic requirement or by giving a small monetary reward. This approach is easy to implement and cost-efficient for academic researchers. However, measurement errors may arise because of the lack of experience of the students. The problem worsens in the case of housing studies because most students do not have experience or involvement in buying houses. Another source of measurement error is the absence of a market mechanism (Berkeley, Kouwenberg, and Post 2004). Therefore, using experimental data in housing studies may lead studies further away from the truth.

Due to the problem related to laboratory data, the efforts of researchers in using market data to empirically test loss aversion in real market settings have been growing. Genesove and Mayer (2001) are credited as the first and most influential researchers to have identified loss aversion in the residential housing market. They used
weekly data from the Boston condominium market between 1990 and 1997. Losses and gains were calculated based on the previous purchase price. Therefore, only properties with repeated sales within the time period were included. Evidence proved that condominium sellers in Boston are subject to loss aversion. When exposed to a prospective loss, they set the asking price 25%–35% higher than the difference between the original purchase price and current expected selling price. Following their method, Bokhari and Geltner (2011) applied the same test to the United States commercial real estate market. Anenberg (2011) conducted similar tests on the San Francisco secondhand housing market and Leung and Tsang (2013) on the Hong Kong, China secondhand housing market. The results of those studies also confirmed the existence of loss aversion in the seller’s behavior.

In the behavioral analysis of the real estate market, transaction data are more reliable than experiment data in various aspects. First, housing transactions are very different from the traditional goods market and the financial market in such aspects as the large stakes involved, the low frequency of transactions at the individual and market levels, and the combination of high heterogeneity of products and high information asymmetry. These aspects make prior knowledge and market experience extremely important to the reliability of the experiment data. In most cases, however, the subjects lack experience and provide answers based only on their imagination, as discussed earlier in this section. Therefore, the validity of such data is debatable. Second, even if all participants have been or are involved in similar decisions, their answers may still be random due to the lack of real incentive. By contrast, transaction data are decisions that have actually been made in market settings. Therefore, the issues stated above are not true for transaction data. The many advantages of transaction data over experiment data motivated us to use transaction data in our research.

3. CONCEPTUAL FRAMEWORK

In this study, we built a conceptual framework to test the hypothesis on whether loss aversion will affect house prices. Developers derive gain–loss utility from land purchase. They hold an expected land price as the reference price level. If the actual transaction land price is higher than the reference, they experience negative utility from the loss; otherwise, it is positive utility from the gains. In our model, for property \( i \), a reference point \( (\text{ref}_i) \) is chosen as the log average land price of land purchase within a distance of 3 miles. Losses and gains are the difference between the reference point and the actual log land purchase price, that is, \( \text{ref} - L_i \). If it is a positive value, it is a gain; otherwise, it is a loss.

We then divide the differences into two variables, namely, gain and loss, which represent perceived gains and losses, respectively, as depicted in Equation (4).

\[
\begin{align*}
gain_i &= \begin{cases} 
0, & \text{ref} < L_i \\
\text{ref} - L_i, & \text{ref} \geq L_i
\end{cases} \\
loss_i &= \begin{cases} 
L_i - \text{ref}, & \text{ref} < L_i \\
0, & \text{ref} \geq L_i
\end{cases}
\end{align*}
\] (4)

We assumed that the log housing price, \( P \), is a linear function of the observable attributes, the indicator of the quarter of selling the house, and an indicator of loss or gain:

\[
P_{lt} = \alpha_0 + X_t\beta + \delta_t + \alpha_{11}\text{gain}_i + \alpha_{12}\text{loss}_i + \epsilon_{it}
\] (5)
where $P_{it}$ is the log house selling price; $X_i = (x_1, x_2, \ldots)'$ is a matrix of house hedonic attributes; and $\delta_t$ are yearly dummies, which equals to 1 only in the year when the houses are sold and 0 otherwise. $\varepsilon_{it}$ is the error term. If $\alpha_{11} > 0$, then a gain would increase the housing price, and vice versa. If $\alpha_{12} > 0$, then a loss would increase the housing price, and vice versa. If a loss aversion effect exists, $\alpha_{12}$ should be larger in magnitude than $\alpha_{11}$.

We estimated Equation (5) using two methods: one using the entire dataset and the other using piecewise regression. For the piecewise regression, we divide the sample into the four groups that represent low losses/gains (Group 1), medium-low losses/gains (Group 2), medium-high losses/gains (Group 3), and high losses/gains (Group 4). Cutoff points for the division are chosen to give subsamples a similar sample size while retaining a decent size of each subsample (no less than 25). More specifically, Group 1 consists of the properties with losses and gains smaller than 0.3 in magnitude; Group 2 consists of the properties with losses and gains greater than 0.3 and smaller than 0.6 in magnitude; Group 3 consists of the properties with losses and gains greater than 0.6 and smaller than 0.9 in magnitude; and Group 4 consists of the properties with losses and gains greater than 0.9 in magnitude.

Given that the functional form of the loss aversion effect is unclear, we also employ a semi-parametric model to achieve a more accurate estimation of the loss–gain effect. We determine that the semi-parametric model performs better than the non-parametric model because the sample size required to yield reliable results becomes extremely large with the growing number of non-parametric regressors (the curse of dimensionality). Given that the Beijing dataset contains only 130 observations, the semi-parametric model is more reliable than non-parametric estimation.

The model for estimation is given in Equation (6). The actual log house prices are regressed on a vector of housing characteristics ($X_i$) and a vector of gain and loss variables ($gain_i, loss_i$). This model is a partial linear model in which $g(gain_i, loss_i)$ is the non-parametric part, whereas all other attributes are linearly incorporated in the model as previously done in Equation (3).

$$P_{it} = \alpha_0 + X_i \beta + \delta_t + g(gain_i, loss_i) + \varepsilon_{it} \tag{6}$$

We used thin plate smoothing spline, one of the penalized least squares estimation methods, to estimate the model. A smooth parameter $\lambda$ is in the penalty to control the balance between the goodness of fit and the smoothness of the approximation.

4. DATA SOURCE AND SUMMARY

The data came from Beijing, the capital of the PRC. Beijing has experienced dramatic population growth and sprawled considerably in the last few years. The number of permanent residents jumped from 2.03 million in 1949 to 21.15 million in 2013, and the number of foreign residents increased from 0.06 million in 1949 to 8 million in 2013.\(^5\)

The growing population density pushed the demand for houses and gave rise to residential property developments. Consequently, land prices and house prices rapidly

---

\(^4\) Ordinary least squares regression is actually a special case of partial linear regression when we impose a linear functional form for $g(gain_i, loss_i)$.

increased, especially after 2005. Thus, the residential property market has drawn much attention from both the government and the research field.

Figure 1 illustrates some important statistics for residential property development in Beijing. The figure shows the increasing housing prices in the past decade and the steady growth of the Beijing residential property market both on the supply side (investment and land purchase) and on the demand side (sold area). Included in the figure are the annual land area purchased by real estate developers, annual sold area of residential property, average price of residential property, and annual property investment. Residential property development investment grew steadily throughout the period, increasing sevenfold from 1999 (CNY23.66 billion) to 2013 (CNY172.46 billion). Land area purchased was highest in 2002 (20.93 million square meters [m²]) and did not exhibit the sharp increase shown by sales and prices. However, the absolute area value remained very high (approximately 8 million m² every year). House prices tripled from CNY4,847/m² in 1999 to CNY17,854/m² in 2013. The increasing trend began in 2006 and peaked in 2013. Despite the extremely high house prices, the sold area remained stable at 10 million–15 million m² per year after peaking in 2005 (28 million m²). Notably, a slight drop in the four statistics was observed during 2011 because of the house purchase restrictions issued in 2010. The restrictions were placed to control the excess demand and the extremely sharp house price increase in Beijing. However, the drop was only minor and the trend resumed its increase after 2 years. In conclusion, a massive increase in development projects and housing transactions is taking place in Beijing. This implies the great potential of the booming residential property market. Due to its significance and representativeness, we focused on the Beijing residential property market.

Figure 1: Annual Residential Property Development in Beijing (1999–2013)

CNY = yuan, m² = square meter.
Note: Annual residential property investment is on the right axis. The other three are on the left axis.
Our sample consists of properties that were built between 2003 and 2010 and sold between 2006 and 2014. The variables include land and house characteristics, land purchase details (date and price), house transaction details (date and price), and developers’ characteristics. The data were obtained from the Hang Lung Center for Real Estate of Tsinghua University (http://www.cre.tsinghua.edu.cn), a leading center for real estate research and education in the PRC. Official statistics, which supplemented the data, were obtained from the Beijing Municipal Bureau of Statistics (http://www.bjstats.gov.cn).

Table 2: Variable Definitions and Descriptive Statistics

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Standard Deviation</th>
<th>Definitions</th>
</tr>
</thead>
<tbody>
<tr>
<td>green</td>
<td>34.52</td>
<td>8.44</td>
<td>green area ratio = gross green area/gross land area</td>
</tr>
<tr>
<td>parking</td>
<td>0.96</td>
<td>0.35</td>
<td>no. of parking units/no. of flats</td>
</tr>
<tr>
<td>FAR</td>
<td>19.19</td>
<td>63.24</td>
<td>floor area ratio = gross floor area/gross land area</td>
</tr>
<tr>
<td>fee</td>
<td>120.57</td>
<td>388.97</td>
<td>property management fee per square meter per month</td>
</tr>
<tr>
<td>no.of houses</td>
<td>924.60</td>
<td>743.41</td>
<td>no. of flats in this development project</td>
</tr>
<tr>
<td>dist_underground</td>
<td>2,560.93</td>
<td>2,898.72</td>
<td>distance to the nearest underground station (m)</td>
</tr>
<tr>
<td>dist_center</td>
<td>21,749.76</td>
<td>10,219.48</td>
<td>distance to city center (m)</td>
</tr>
<tr>
<td>dist_hospital</td>
<td>12,653.20</td>
<td>7,984.79</td>
<td>distance to the nearest hospital (m)</td>
</tr>
<tr>
<td>dist_park</td>
<td>7,960.42</td>
<td>6,096.56</td>
<td>distance to the nearest park (m)</td>
</tr>
<tr>
<td>dist_school</td>
<td>12,440.65</td>
<td>8,511.00</td>
<td>distance to the nearest primary school (m)</td>
</tr>
<tr>
<td>land price</td>
<td>12,608.52</td>
<td>10,549.20</td>
<td>actual land price (CNY/m²)</td>
</tr>
<tr>
<td>reference land</td>
<td>11,872.56</td>
<td>6,646.37</td>
<td>reference land price (CNY/m²)</td>
</tr>
<tr>
<td>price</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>house price</td>
<td>20,820.73</td>
<td>12,953.77</td>
<td>house sales price (CNY/m²)</td>
</tr>
</tbody>
</table>

CNY = yuan, m = meter, m² = square meter.

We use house prices, the monthly average price (CNY/m²) for each development project, as the dependent variables. The explanatory variables used in the estimation and the descriptive statistics are summarized in Table 2. Total house number (housenum) in a development project is a proxy for development size. Floor area ratio (FAR), green area ratio (green), parking space (parking), and property management fee (fee) are features of a project that affect the house prices. Spatial characteristics are represented by the natural logarithm of the distance to the nearest underground station (dist_underground); the city center, which is Tiananmen Square (dist_center); the hospital (dist_hospital); the park (dist_park); and the primary school (dist_school). The dataset also includes developer’s features, such as the ownership of the developer (central state-owned enterprise, noncentral state-owned enterprise, or private enterprise) and whether the developer is a listed company. Although these data are not shown in the table, they are included in the regression, as are dummy variables. After dropping observations with missing values, we retained 130 property projects.
5. EMPIRICAL RESULTS

This section presents the empirical results of the models described in section 3. The first step is to calculate the expectation-based reference point. The loss and gain variables are firstly obtained after the reference point calculation. Subsequently, we tested the hypothesis about loss aversion and pricing decisions based on the linear regression of Equation (5) and semi-parametric regression of Equation (6).

A histogram to show the distribution of the gains and losses is given in Figure 2. They are the difference between the reference point and the actual log land purchase price, that is, \( \text{ref} - L_i \). If it is a positive value, it is a gain; otherwise, it is a loss. The dotted line is the normal distribution curve fitted for the data.

<table>
<thead>
<tr>
<th>Table 3: Ordinary Least Squares Regression Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Entire Dataset</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>gain coefficient</td>
</tr>
<tr>
<td>loss coefficient</td>
</tr>
<tr>
<td>No. of observations</td>
</tr>
<tr>
<td>R-square (%)</td>
</tr>
</tbody>
</table>

We use both the entire dataset regression and piecewise regression to see if loss aversion levels change with the magnitude of the losses and gains. Coefficients for the observable attributes \( X_i \) are reasonably consistent with those of the literature and are omitted from the table. The entire dataset regression does not appear to support the loss aversion effect. The coefficient on loss (0.488) is smaller in magnitude than the
coefficient on gain (−0.627). One possibility is that the linear relationship between loss/gain and house price may not be globally the same. Therefore, the piecewise regression is employed to effectively capture the nonlinear relationship. The results show that small losses seem to loom larger than small gains (0.91 vs. 0.13); the same is true for large losses (0.57 vs. 0.31). However, when the magnitude falls between 0.3 and 0.9, the pattern seems to be reversed, that is, gains have a stronger effect than losses. However, the pattern we obtained is not rigorously justified if only piecewise regression is used. Either the cutoff point selection or the linearity imposed in each regression could bias the estimation results.

To capture the effect of losses and gains on house prices more accurately, a semi-parametric model based on Equation (6) is estimated using the TPSPLINE method. Figure 3 is an illustration of the results. It shows the utility from the land purchase implied in the house prices, with the assumption that other attributes are the same for all the projects. The negative values in the x-axis represent losses, whereas the positive values represent gains. In other words, it is \( \text{ref} - L_i \) in the x-axis. This transformation aims to make the results easily comparable with the value functions in prospect theory, as the dotted line shown in Figure 3, and hence the loss aversion effect directly observable. Utility function is obtained through some basic assumptions: (i) at the reference point, \( v = 0 \); and (ii) one unit of house price increase (holding all variables except loss/gain constant) implies one unit of negative utility, while one unit of house price decrease implies one unit of positive utility.

**Figure 3: Value Function Estimated from the Semi-Parametric Estimation**

Figure 3 demonstrates that developers would set a price higher than the price they would otherwise set in response to a paper loss increase; however, they would set a price lower than the price they would otherwise set in response to a paper gain increase. As for the slopes, the biggest difference is around the reference point. Two dotted lines are added to show the slopes for the two domains. For gains, the marginal utility is 0.32, that is, a one-unit increase of gain from land purchase raises a developer’s utility level by 0.32 unit. For losses, the slope is 0.6, that is, when a one-unit loss is experienced close to the reference point, a developer’s utility drops by 0.6 unit. From this estimation, loss aversion level is nearly 2 around the reference point.
When the magnitudes of losses and gains grow, the effect of loss aversion disappears. This is from the decreasing slopes in the loss domain and increasing slopes in the gain domain around the magnitude of 0.5. When the magnitude is within the range of 0.5–1, loss effects and gain effects are almost the same. When developers face losses that are greater than 1, losses loom significantly larger than gains again. To conclude, developers are most loss averse under two circumstances: (i) around the reference point and (ii) when facing very high losses.

The findings show that transactions in the land market affect transactions in the house market. High land transaction prices would lead to a house price disposition effect. The impact takes the form of the loss aversion bias, and the level of the impact varies with the magnitude of the losses and gains. The loss aversion impact is the strongest around the market expectation price. As the magnitude of a loss or a gain grows, their impact is attenuated. But when land transaction prices deviate too much, the impact of loss aversion strengthens again.

6. CONCLUSION AND FUTURE RESEARCH DIRECTIONS

This study is an extension of previous loss aversion studies on the real estate market (e.g., Genesove and Mayer 2001, Bokhari and Geltner 2011, Leung and Tsang 2013, Anenberg 2011). Using the September 2003–June 2014 land purchase and house sales data from Beijing, this study has shown that loss aversion affects pricing decisions. Specifically, losses in the land purchase phase would lead to disposition behavior in the later stage of house sales. The effect is most distinct around the reference point as well as when losses or gains are high in magnitude.

We improve the previous studies in the following ways: employing a new expectation-based reference point, focusing on the property developers, and allowing for variant loss aversion levels using semi-parametric analysis. Moreover, our research adds validity to behavioral findings based on transaction data rather than on experiments. Despite the increasing number of studies on loss aversion since the concept's introduction, most of them have continued to use data from student experiments. The validity of the findings from these experiments in real market conditions has become debatable because (i) students lack the knowledge and experience; and (ii) most of these experiments were only concerned with small-stakes gambles, which are considerably simpler than real-life decisions. In housing studies, the problems related to experiment data have worsened; thus, such studies cannot produce ecologically valid findings, given the complexity of both the decision-making process and the nature of the product. Real market data, such as property transaction prices and dates, have become the preferred sources. In this sense, the use of Beijing property development and transaction data highlights the validity and significance of this research.

The results of this research have broad implications for our understanding of the PRC’s real estate markets. First, housing prices are determined not only by house characteristics but also by the behavioral biases of developers and sellers as well. This implication indicates that the market is far from perfect and it is more complicated than the market predicted by classical economic theory. Second, the positive correlation between housing price and volume, which has been identified in previous research along with the strong stickiness of housing prices in a down market, cannot be explained by perfect asset models. The behavioral bias of the developer, namely loss aversion, plays a significant role in this cyclical trading pattern.
In this study, however, we failed to directly observe the reference points, and unknown measurement errors occurred. These are important areas to address in the future. Given the nature of the available data and existing literature, this study, at best, could only provide an indication of the underlying developer behavior. Future research should pay attention to the reference formation process and incorporate more elements that may affect the reference point. For instance, the historical peak of the changing outcome plays an important role in the reference adapting process (Gneezy 2005). In addition, the adaptation of reference point is a huge issue. As Chen and Rao (2002) highlighted in their experiments, although people’s reference points shift immediately after a stimulus occurs, such a shift is incomplete. The magnitude of the shift depends on the time difference between two stimuli. Reference points are important to the validity of the whole research. Therefore, they should be given special attention.

After identifying the overall effect of loss aversion, another important future research topic is whether any moderating factors—for example, developer ownership and developer transparency—interact or influence the level of loss aversion. Such research will be crucial in providing instructive insights and practical guidelines for developers, policy makers, and home buyers.
REFERENCES


