

Article

# Dynamics of the US Housing Market: A Quantal Response Statistical Equilibrium Approach

Özlem Ömer 

Department of Economics, The New School for Social Research, 6 E 16th Street, New York, NY 10011, USA; omero679@newschool.edu; Tel.: +90-542-723-8050

Received: 11 August 2018; Accepted: 19 September 2018; Published: 30 October 2018



**Abstract:** In this article, we demonstrate that a quantal response statistical equilibrium approach to the US housing market with the help of the maximum entropy method of modeling is a powerful way of revealing different characteristics of the housing market behavior before, during and after the recent housing market crash in the US. In this line, a maximum entropy approach to quantal response statistical equilibrium model (QRSE) is employed in order to model housing market dynamics in different phases of the most recent housing market cycle using the S&P Case Shiller housing price index for 20 largest- Metropolitan Regions, and Freddie Mac housing price index (FMHPI) for 367 Metropolitan Cities for the US between 2000 and 2015. Estimated model parameters provide an alternative way to understand and explain the behaviors of economic agents, and market dynamics by questioning the traditional economic theory, which takes assumption for the behavior of rational utility maximizing representative agent with self-fulfilled expectations as given.

**Keywords:** housing market crash; statistical equilibrium; quantal response; informational entropy; the maximum entropy method

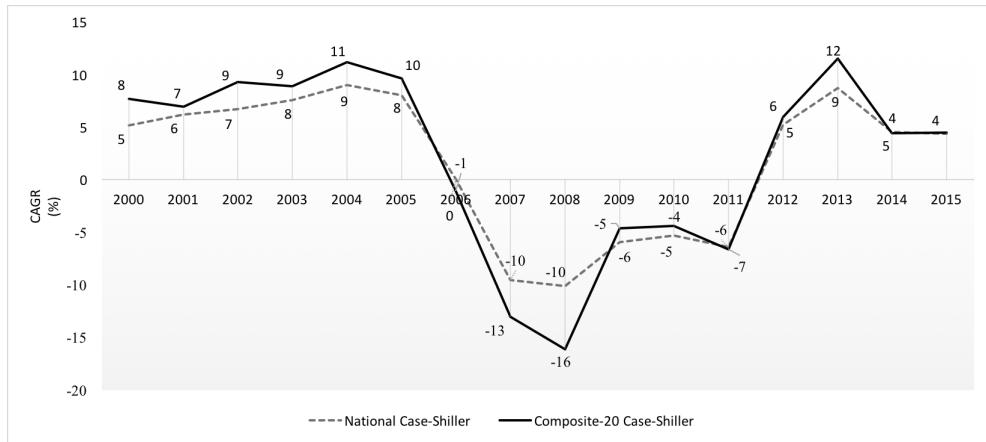
**JEL Classification:** C18; D89; D90; E30; G01; R39

---

## 1. Introduction

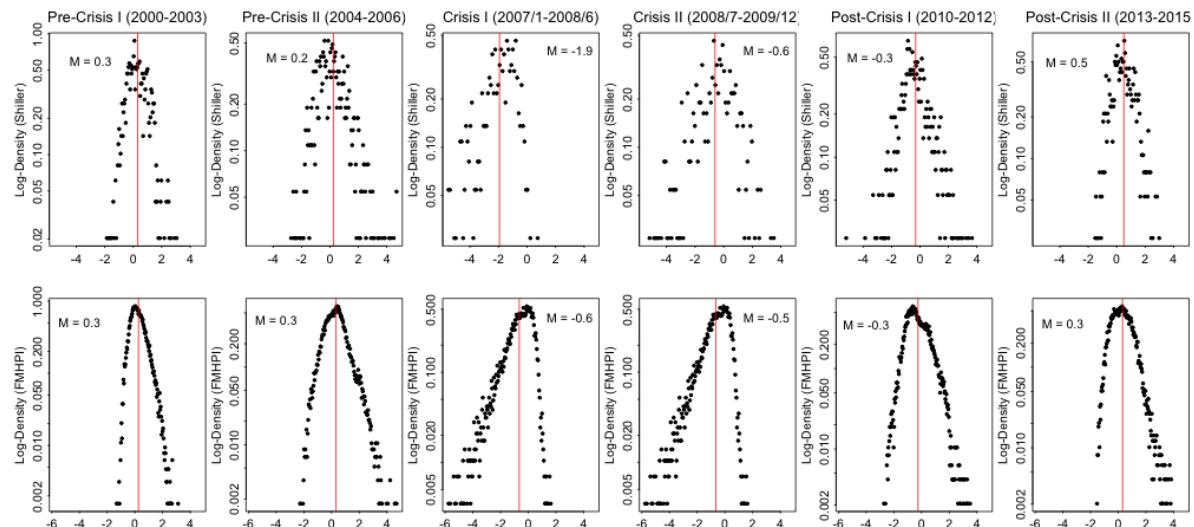
The US housing market experienced one of its key periods of boom-bust cycle between 2000 and 2015, urging researchers to understand and explain the movements of house price changes throughout the cycle. As stated by many authors [1,2] and evident from Figure 1, between 2000 and 2005, seems to enter a positive price feedback loop—in which any house price increase brings about a further rise in house price growth rates, signaling a boom period. Beginning in 2006, the rapid house price growth reverses with slowing appreciation, followed by a catastrophic collapse between 2007 and 2008, in which house price growth rates simultaneously drop nationwide. Decline in house price growth rates alleviates in 2009, but does not experience positive growth again until 2012. The period 2007–2009, therefore, displays a distinct divergence from the periods of 2000–2006 and 2010–2015.

To capture the behavior of the US housing market for several sub-periods of before, during and after the market collapse, Figure 2 represents the log density distributions of monthly house price growth rates for all available metropolitan areas represented in two different monthly house price indices, Case-Shiller Housing Price Index and Freddie Mac Housing Price Index (FMHPI). Distributions have a striking single peaked shape throughout the cycle indicating a central tendency of the system around an average monthly growth rate. Some distortions in the distribution appear to arise at the end of the boom (2000–2006) and during the crash (2007–2009) by pointing out a change in the characteristics of the market behavior.



**Figure 1.** Average (Compound) Annual House Price Growth Rate, CAGR is calculated based on monthly inflation adjusted Case-Shiller National and Composite-20 Indices as:

$$\text{CAGR} = \left( \frac{\text{Last Month House Price Index}}{\text{First Month House Price Index}} \right)^{\frac{1}{\text{Number of Years}}} - 1.$$



**Figure 2.** Log Density Plots of Monthly House Price Growth Rates: Calculated as the histogram plots on a log density scale of house price growth rates. The first row shows the histogram plots on a log density scale of monthly house price growth rates based on Case-Shiller Price Index growth rates (%) for each 20 largest metropolitan areas while the second row is based on FMHPI, including monthly price growth rates for the DC region and 367 metropolitan areas in the US. “M” represents the average growth rate for the panel data in each window. Each plot includes all the cross sectional data represented by the specified index, and for the specified time period. Pooled house price growth rate data is then used to calculate log density distributions. More detailed explanation about the data sources will be provided in Section 2.

Over the past couple decades, several modeling approaches have been put forward to explain large fluctuations in housing markets, such as the one represented in Figures 1 and 2. Search and Matching models of housing markets [3,4], for example, represent temporary house price movements away from the “competitive Walrasian equilibrium” as a result of “exogenous” shocks affecting the demand supply schedules in the market via market frictions, e.g., vacancy rate in the market in which one of the most important problems of modeling interactions among the individuals is reduced to the bargaining process between “homogeneous” sellers and buyers. Simple heterogeneous agent

models (HAMs) carry the analysis one step further by introducing two (or three) different types of individuals and their interactions with each other [5,6]. As a result, they are able to reproduce endogenous fluctuations caused by the interactions between heterogeneous individuals (e.g., Chartists vs fundamentalists); however, they use some other simplifications, such as stationary rules to represent the behaviors of individuals and their interactions, and the rules to represent the law of price motions which also raise questions about how realistic the model results are in explaining the real causes behind the market behaviors. Computational Agent Based Models (ABMs) try to solve the problems of HAMs by providing an evolutionary process for the determination of behaviors of heterogeneous agents [7–9]. They are able to model a larger scale of more complex interactions between the agents as the main cause of endogenous out-of-equilibrium dynamics of the markets. On the down side, their highly complex and ambitious agenda complicates the estimation process of the model which weakens the robustness of the estimation results. Additionally, since the ABMs are structurally complex, they blur the causal relationships behind the market behaviors.

The Quantal Response Statistical Equilibrium (QRSE) model, introduced by Scharfenaker and Foley [10] employed in this paper, provides a method to explain the observed regularities in highly complex housing markets by taking into account social interactions between a large number of heterogeneous buyers/sellers and their reactions to house price changes. The main objective is to analyze the market dynamics in a simple but realistic way without relying on ad-hoc normative assumptions (For example, models based on traditional economic theory assume complete and transitive set of preferences for “homogeneous” individuals so that they can reduce the problem of providing a micro-foundation for the macroeconomic system to a simple aggregation of all “representative agents’” decisions). The model is built upon three main concepts: Statistical equilibrium of targeted economic variable(s) representing the macroeconomic state of the system, the assumption of quantal response behavior of the market participants, and the maximum entropy principle as a method of inference. In this model, the concept of statistical equilibrium from statistical mechanics (pioneered by, Boltzmann [11], Gibbs [12], Maxwell [13] and Jaynes [14]) allows for the relaxation of traditional assumptions of equilibrium. Therefore, deviating from the traditional concept of Walrasian (General) Equilibrium, the QRSE approach focuses on information theoretic *statistical equilibrium*, in which equilibrium takes the form of a distribution representing all possible states of the system (QRSE models introduce a general framework for equilibrium where the conventional market clearing general equilibrium exists just as a special and highly unlikely state of the economic systems). The maximum entropy principle as a method of inference, is the least biased method among all, because it maximizes the uncertainty of the system and gets the least informative state with no additional assumptions other than the existing knowledge of the researcher and observed data. Therefore, it is a useful tool for dealing with the hardship of modeling complex systems, such as housing markets. Finally, by introducing an assumption on the behavior of “purposeful” individuals represented as quantal response logit functions, QRSE models deal with modeling heterogeneity of interacting agents, hence reject the assumption of rational representative agent and replace it with boundedly rational typical agent whose preferences do not have to be well-defined, complete, and transitive. As a result, QRSE models introduce a platform to model social interactions among heterogeneous individuals, uncertainties in individuals’ decision making processes, and endogeneity of the system dynamics without imposing any ad-hoc assumptions.

In this study, observations of the US housing market between 2000 and 2015 from Figures 1 and 2 are used to study the questions of “which forces behind the market dynamics give rise to the boom-bust cycles and the collapse, and how can one capture them?”. To achieve this goal, a Quantal Response Statistical Equilibrium Model is introduced, in which two systemic forces govern equilibrium/out-of equilibrium dynamics of the US housing market. To define these market forces, the model sets a behavioral assumption that typical agents, based on the house price changes in the market, try to maximize the expected returns from their transactions subject to a constraint on the entropy of their mixed strategy over transactions. Moreover, the actions of buying/selling in the housing market create

a negative feedback effect on the house price growth rates themselves. Therefore, both the central tendencies in the distribution of house price growth rates and the deviations around these tendencies are the natural results of this process governed by the relationship between the changes in house price growth rates and the actions taken by the individuals in the system.

Based on Figure 2, we argue that periods before and after the market collapse of 2007 indicate a strong statistical equilibrium while the distortion in the equilibrium state starts to rise around late boom period of 2004–2006, and crisis of 2007–2009 signals a strong divergence from the equilibrium. After the crash, the system seems to recover towards a new statistical equilibrium between 2010–2015.

This paper sets out to (i) analyze the observed regularities in the housing market during the most recent housing cycle based on the statistical equilibrium approach, (ii) provide a model to explain the systemic forces behind the behavior of the housing market for the consequent periods of the cycle, and (iii) combine the model results with a solid economic interpretation. Section 2 will present the data sources employed in our analysis, and discuss observed regularities and the divergence from these regularities in the US housing market based on house price growth rates for a large number of metropolitan areas between 2000 and 2015. In Section 3, a maximum entropy method to quantal response statistical equilibrium model by Scherfenaker & Foley (2017) will be applied to the housing market dynamics. Section 4 will focus on the results of the model fit estimations for each sub periods and robustness of parameter estimations. Finally, Section 5 will provide a brief discussion about the results.

## 2. Observed Regularities in the House Price Growth Rates

Sampling and modeling techniques of repeat sales house price indices have been largely criticized in the literature [15]. To set an argument on a sound basis and check the robustness of the results, two different monthly house price index data for several metropolitan areas were used. First one is Freddie Mac House Price monthly index, FMHPI, which is available for DC area and 367 Metropolitan statistical areas (MSAs) throughout the US, and the second one is monthly S&P Case-Shiller Index for 20 largest US metropolitan cities (See Appendix A for more details about each data sets). To capture the regularities in the data, both data sets are deflated using the consumer price index (CPI). Corresponding real house price indices are then, used to calculate monthly real house price growth rates for each region in both data sets. Inflation adjusted monthly house price growth rates, hence, are calculated as  $x_t = \frac{100(g_t - g_{t-1})}{g_{t-1}}$ , where  $g_t$  is the house price index for time  $t$ .

Growth rates of price indices rather than the levels, are used in this analysis because the rate of change of housing prices play a more significant role than the levels of prices during the decision making process of economic agents because it provides information about the possible returns on the investment, and since we are dealing with cross sectional data, which each region carries different characteristics across time (i.e., peak levels of house prices, and timing to reach the peak differ among different regions), it is more convenient to use house price growth rates as a metric. Also, the rate of change of prices in logarithmic scale is able to capture the exponential or power law types of distributions from the data. It is important to capture extreme values in the data especially during the market crash since house prices do not follow random walk. Instead, they seem to persist in the short run while there exists a mean reversion in the long run [16,17] (Random walk theory pre-supposes the efficient market hypothesis, in which profit maximizing agents are rational, and can reach all available information freely in the market. Therefore, on the average, competition will cause the full effects of new information on intrinsic values of assets to be reflected “instantaneously” in actual prices, see [18]).

### *Visualizing Statistical Equilibrium: Patterns of House Price Growth Rates (2000–2015)*

As argued by Jaynes [14], Farjoun and Machover [19], Foley [20], and Scharfenaker and Semieniuk [21], the statistical equilibrium framework (from statistical mechanics) to describe macroeconomic systems, is a highly suitable approach, because laws governing macroeconomic variables such as price changes

and profit rates carry a statistical character, so that one must employ probabilistic considerations in order to capture such laws. Statistical mechanics was pioneered by Boltzmann [11], Maxwell [13] and Gibbs [12], in order to explain the physical macroscopic properties of matter from the microscopic properties of atoms and molecules which comprise the system as a statistical large population. In developing the kinetic theory of gasses, Boltzmann [11] derived that pressure and temperature of an ideal gas are a function of the average kinetic energy of its particles. Since determination of each micro-state requires specification of the position and momentum of each gas molecule, which is a cumbersome task due to high degree of freedom for the entire system, the theory of statistical mechanics reduces the problem to analysis of macroscopically observable parameters via a statistical analysis of the molecules of which gas is made up. Probabilistic reasoning, hence allows us to reach informative results concerning the equilibrium behavior of the complex systems without relying on strong assumptions about the detailed microscopic behavior of the individual components of the system [14,19].

Based on the same logic, statistical equilibrium in social systems; such as housing markets, also result from the interaction of a large number of economic agents with a certain level of uncertainty in their actions which in turn, collectively have an impact on the outcome variable, such as house price growth rates. Following classical political economy tradition of [22,23], we argue that individual economic agents seeking for a higher return than the economy-wide average from their transactions in a market would generate a “tendential gravitation” of each sub-market’s house price growth rate around that average as an unintentional result of the interactions among many individual buyers and sellers. Therefore, their decisions in the process of competing for higher expected returns determine the change in house prices at any point in time creating an average rate of price change in the market as a whole. This process can be interpreted as the statistical equalization process of house price changes in the system. As a result, statistical regularities (single peaked Laplace-like log frequency distributions) occur due to buying and selling decisions that are conditional on the expected changes in the housing prices for maximizing their returns from their transactions. The process of deciding when and where to sell/buy works as a competitive process, and creates a negative feedback mechanism while payoff maximizing sellers are seeking for the regional price increase to be above their expected fundamental rate of price increase of their houses to sell, and buyers are looking for markets whose regional price increase is below their expected fundamental rate of price increase to buy. Unintentionally, actions of sellers cause a decline in housing price increase in sub-markets, while behavior of buyers causes an increase in housing price change. Concept of statistical equilibrium as a remarkable representation of economic-laws based on classical political economy tradition, hence, can help us to capture the governing forces behind the system’s central tendency and deviations from it. This way we can detect and interpret empirical regularities in the housing market.

As briefly mentioned in introduction, Figure 2 shows the log density distributions of monthly house price growth rates of the panel data for “pre-crisis periods (2000–2003) and (2004–2006)”, “Crisis periods (2007–June 2008) and (July 2008–December 2009)”, and finally “Post-Crisis periods (2010–2012) and (2013–2015)” for two different cross sectional house price index data mentioned above in order to determine statistical equilibrium dynamics in the US housing market. Each plot includes monthly real house price growth rates for each and every metropolitan area represented in the data for the specified periods. This way, one will be able to observe the evolution of the distribution of house price growth rates; regularities during the statistical equilibrium periods, and the divergence from the equilibrium during the crisis, if they exist.

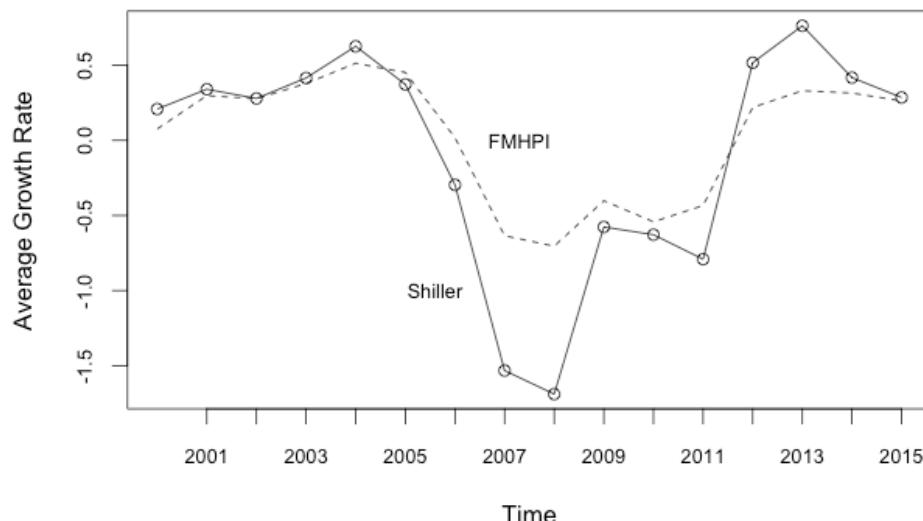
For both data sets, log density distributions of monthly real house price growth rates for the US housing market seem to follow a symmetric Laplace-like single peaked distribution between 2000–2003 and 2010–2015 suggesting a strong statistical equilibrium in the housing market before and after the collapse. Results from FMHPI data emphasizes changes in skewness among the sample periods, while results from the Case-Shiller index data emphasizes the loss of definition of the distribution in the late pre-crisis II and crisis periods. According to results based on FMHPI index, between 2004 and 2006,

symmetry in the system starts to deteriorate although the market forces still try to push the system back to the equilibrium. Symmetry in the distribution is not a necessary criterion for the statistical equilibrium. However, it can be interpreted as an indicator of rapid appreciation/depreciation of house prices in some sub markets before/after the crash. When the house market collapses, during 2007–2009 distribution becomes more dispersed for the Case-Shiller House price index signaling a deterioration of market forces whose roles are to push the system to the statistical equilibrium, while FMHPI distributions remains highly skewed to the left implying the increasing house price depreciation in some metropolitan areas (Note: Although each data sets are used for the same purpose in the analysis, some variation between the two is expected because S&P Case-Shiller house price index represents 20-largest metropolitan cities such as Chicago, Las Vegas and New York, while FMHPI consist of DC region and 367 metropolitan areas throughout the US representing a finer grained data. However, results follow similar patterns through time).

Another way of examining the central tendencies for house price growth rates is to look at the average price growth rate movements in the system as a whole. This is because when the system is at or near the statistical equilibrium, in the sense that macroscopic conditions and macroscopic parameters, such as average change in prices, appear to be constant or vary comparatively slowly [19]. Yet, this stable conditions in equilibrium is still dynamic and statistical, allowing microscopic parameters such as house price growth rate in any sub-market to change rapidly. The system as a whole evolves through time so do the macroscopic parameters. When the system is disrupted by a period of crisis, macroscopic parameters in hand can change drastically.

As Figure 3 points out, average monthly price growth rates calculated as the average price growth rate of the pooled cross sectional data for each year for each data sets are quite stable between 2000–2003 around 0.2%–0.34% with a further increase between (2003–2005), while the crisis period shows meaningful changes, in which stability in house price appreciation is replaced by a drastic decline in growth rates around  $-1.9\%$  for Shiller house price index. The decline in average price growth rate is more moderate in FMHPI, which can also be interpreted as the indicator of severity of crisis experienced by different metropolitan areas. This may be because Case-Shiller price index include 20 largest metropolitan cities which experienced the largest price appreciation during the boom, and the largest depreciation during the crisis, while FMHPI covers DC region and 367 US cities causing finer results on average. This might be signaling one of the stylized facts of US housing market, which is that different metropolitan areas experienced different price appreciation (depreciation) before (after) the crisis. This is why focusing on only the national data would leave out some of the most important characteristics of boom and bust cycles in the housing markets. One can argue that the stability of average monthly house price growth rates during the statistical equilibrium of the pre and post-crisis periods indicates some systemic forces in the housing market, which are dysfunctional during the crisis, and also the purpose of this paper to explain.

For the sake of understanding the forces that push the housing market to the statistical equilibrium, it is crucial to keep in mind that housing market dynamics involve social interactions of a large number of economic agents through socially determined structures in the market, from which the statistical distribution of housing price growth rates is determined. As a result of these interactions, based on individuals' quantal actions (buy and sell), and their engagement with the resulting outcomes of these actions in a highly "competitive" environment, system generates complex dynamics, which are hard to capture from the data directly. As Scharfenaker and Foley (2017) argues, the difficulty of capturing the dynamics of the economic systems from the data arises due to the dependency of observable variables on unobserved ones, which in turn, causes incomplete information about the problem in hand, raising the importance of the method of estimating the parameters in the model.



**Figure 3.** Average Monthly Housing Price Growth Rates (%). Calculated as the average price growth rate of the pooled cross sectional data for each year.

In the next section, the maximum entropy-quantal response model of statistical equilibrium, called “QRSE”, introduced by Scharfenaker and Foley [10], will be applied to housing market to overcome incomplete information problem in modeling, and to explain the dynamics/behaviors of the system during the most recent housing market cycle (2000–2015).

### 3. A Quantal Response Statistical Equilibrium Model of the Housing Market

The Quantal Response Statistical Equilibrium (QRSE) model predicts the equilibrium as the most probable state of the system in the form of the statistical equilibrium probability distribution of monthly house price growth rates,  $f[x]$ , which is derived by maximizing the entropy of the unknown distribution,  $H[f[x]] = -\sum f[x] \log[f[x]]$  subject to constraints including all information from the observations or theory related issues [14,24] (See Farjoun and Machover [19] Foley [25], dos Santos and Scharfenaker [26], Scharfenaker and Semieniuk [21] and Yang [27] for other applications of the maximum entropy statistical equilibrium approach). It includes two variables, an action variable,  $A$ , and an outcome variable  $x : X \rightarrow R$ . In this case, the action variable is a binary variable,  $A = \{a, \bar{a}\}$  where  $a$  and  $\bar{a}$  represent selling and buying actions respectively, and the outcome variable,  $x$  represents the monthly real house price growth rate (%) (In real applications, measurement based on  $x$  variable will be coarse-grained in finite number of bins). The heart of the model lies on the interaction between the action and outcome variable.

The model, which takes the interactions among buyers and sellers into account, is a joint distribution of actions  $A = \{\text{sell}, \text{buy}\}$  and house price growth rates,  $x$  which will be written as  $f[A, x]$  with the marginal and conditional frequencies  $f[x], f[A], f[A|x], f[x|A]$

$$f[x] = \sum f[A, x]; f[A] = \int f[A, x] dx$$

$$f[A|x] = \frac{f[A, x]}{f[x]} \quad \text{if } f[x] > 0, f[x|A] = \frac{f[A, x]}{f[A]} \quad \text{if } f[A] > 0.$$

In this set up, conditional distributions represent the causal relationships and hence, include the explanatory power for the model to overcome the problem of incomplete information. Quantal response of individuals’ actions assumes that individuals make their decisions based on the outcome variable, in this case, house price growth rates. Therefore, not only do individuals’ actions

(selling/buying) affect the house price changes ( $x$ ), but the change in house prices also affects individuals' selling/buying decisions.

The interdependence between selling/buying actions of individuals and the change in house prices are therefore, represented by the conditional action distributions:

$$f[A|x] = \{f[a|x] = f[sell|x], f[\bar{a}|x] = f[buy|x]\}. \quad (1)$$

The causal relation that links the buying-selling behavior of individuals to the house price growth rate,  $x$ , symmetrically determines the conditional outcome frequencies;

$$f[x|A] = \{f[x|a] = f[x|sell], f[x|\bar{a}] = f[x|buy]\}. \quad (2)$$

Overall, conditional frequency distributions Equations (1) and (2) allow us pursuing a theory of interaction through the inter-dependency of actions and the outcome variable in the case of housing market representation. Therefore, statistical equilibrium in the housing market model is a joint distribution, which determines two conditional distributions,  $f[A|x]$  and  $f[x|A]$ .

The problem of modeling the housing market occurs because we end up having house price changes as outcome variable, which is determined by individual actions in the housing markets. In general, if one knows one of the conditional distributions via observations or theoretical considerations, and also the corresponding marginal distributions from observations, such as  $f[A|x]$  and  $f[x]$ , then they would be enough to determine the joint distribution  $f[A,x] = f[A|x]f[x]$  from which one can derive the other marginal and conditional distributions. In this case, observed data would provide full information to estimate the statistical equilibrium. Yet, most of the time this is not the case. For example, in our case, we only have the marginal distribution of house price changes  $f[x]$ , and need to learn about the conditional distributions  $f[A|x]$  and  $f[x|A]$ . Buying-selling actions of individuals could be easily observable for some data sets, but for our data sets it is not the case. Individual actions of house buyers and sellers cannot be observed simultaneously apart from house price growth rate data. This means, when we are observing house price changes, we cannot measure the intensity of buying and selling actions of individual agents in the housing market from the same data directly. Here, we want to estimate the joint distribution in order to infer the causal links, but we only have a partial information about the joint distribution, such as observations on house price growth rates that determine the marginal distribution of housing price changes  $f[x]$ .

This type of a quantal response social interaction problem with only a partial information about the joint distribution of quantal actions and the outcome variable that is determined by these actions, such as an observed marginal distribution, necessitates using a theory that offers a prior over one of the conditional distribution using Shannon's information theory [28], and Jaynes' (1957) principle of the maximum entropy inference. This means model constraints in the entropy maximization problem play the key role.

### 3.1. Model Constraints

As emphasized by [24,29], it is important to choose the right constraints either using the observable data or theoretical frameworks, when employing the maximum entropy program to generate the most likely statistical equilibrium distribution of the system. Here, there are two important constraints, which are needed to represent the housing market dynamics. The first constraint represents the behavior of the individuals in the housing market based on the quantal response, and the second one is based on the impact of individuals' actions on the house price changes.

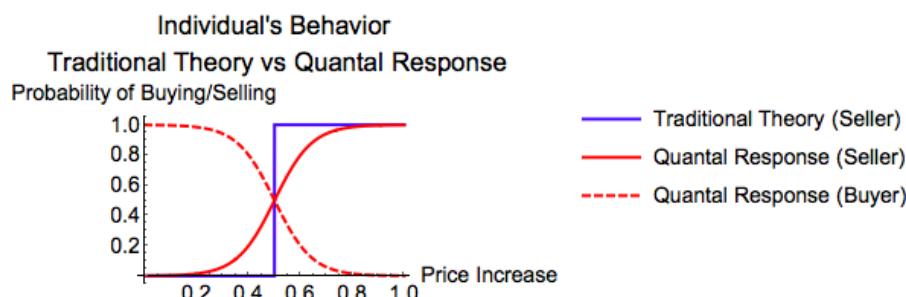
#### 3.1.1. Constraint 1: The Impact of House Price Changes on the Actions of Economic Agents (Quantal Response Behavior of Individuals)

Traditional economic theory assumes that decision makers have complete and transitive preferences over outcomes represented by a utility function [30–32]. According to this approach,

when utility maximizing rational individuals face with a decision to buy/sell a house in a particular market based on their payoff functions (i.e., the difference between the rate of increase in prices in the region and their estimate of the fundamental rate of price increase of the house), they would almost never sell (buy) the house when prices in the region rise slower (faster) than their expected fundamental rate of price increase of the house, and would wait until their expectations about house price increase coincide with actual price increase in the market (See Foley [33] for more methodological and theoretical details).

Yet, as it has been shown both experimentally and empirically, this approach does not represent the reality [34,35]. To provide a more realistic representation of individual behavior, the first constraint of our model assumes that individuals follow a binary quantal response function when they decide to sell/buy in the market. Therefore, we argue that when the regional price increase is well below sellers' estimate of fundamental rate of price increase of the house they would almost never sell. However, when the regional price increase gets closer to their expected rate of fundamental increase of the house, they sell more frequently depending on the regional price increase; even though their expectations about the price increase of the house do not perfectly coincide with the actual price increase in the region. Opposite is true for the buyers. When a buyer's estimate of fundamental rate of price increase for a house is well above the regional price increase (regional price appreciation rate is smaller than what she/he expected), they would almost always buy. However, when the actual regional increase gets closer to their estimates, they would buy less frequently.

Figure 4 demonstrates the difference between these two approaches in explaining the behavior of a seller/buyer as explained above. According to traditional theory, a rational individual would follow a step function while a typical individual in our model follows a logistic S-curve.



**Figure 4.** Individual's behavior: Traditional Theory vs. Quantal Response.

Modeling quantal response of market participants in the maximum entropy framework assumes that the agent may respond to a payoff function  $u[A, x]$ . In the case of housing market, it represents the payoff to the typical individuals who sell/buy houses in the market. If the agent chooses a mixed strategy  $f[A|x] : AxX \rightarrow (0, 1)$  over actions  $\{a, \bar{a}\} = \{\text{sell}, \text{buy}\} \in A$  to maximize the expected payoff subject to a constraint on a lower bound on the informational entropy of the mixed strategy,  $\sum_A f[A, x] u[A, x]$  as:

$$\begin{aligned} \text{Max}_x \quad & f[A, x] u[A, x] \\ \text{subject to} \quad & \sum_A f[A, x] = 1 \\ & H[f_{A,x}] = - \sum f[A, x] \log[f[A, x]] \geq H_{min} \end{aligned}$$

the resulting function turns out to be the quantal response (Gibbs) function  $f[A|x] \propto e^{u[A,x]/T}$ . When the parameter  $T > 0$ , recovered link functions will assign probabilities to different responses which are represented by S-shaped logit quantal functions for selling and buying actions.

Scharfenaker and Foley [10] presents the difference between the payoffs of selling and buying as a linear function by introducing a shift parameter  $\mu$ , defined as  $u(a|x) - u(\bar{a}|x) = x - \mu$ . Here,  $x$  is the regional house price change while  $\mu$  represents the expected fundamental rate of price increase for the house which the agent is willing to sell/buy. This relation tells us that higher house price growth rate increases the payoff of selling action while payoff of buying diminishes, thus the gap between the payoffs increases. Moreover, if the gap between the two is zero individuals become indifferent between buying and selling. In this case,  $\mu$  determines the indifference point of the house price growth rate in which the probability of selling is 50%. It provides useful information about the willingness of market participant's to sell at a certain house price growth rate, which will be discussed in more detail in following sections.

With this new specification, the impact of house price changes on the actions of selling and buying will be modeled as:

$$\begin{aligned} f[sell|x] &= \frac{1}{1 + e^{-\frac{(x-\mu)}{T}}}, \\ f[buy|x] &= \frac{e^{-\frac{(x-\mu)}{T}}}{1 + e^{-\frac{(x-\mu)}{T}}}. \end{aligned} \quad (3)$$

As a result, introducing an informational entropy constraint into the maximization problem allows one to introduce the bounded rationality framework explaining the individual's decision making process. Therefore, we are able to take into account individual's uncertainty of her/his actions which could be resulted from difficulty in reaching the perfect information, or in assessing market signals etc. during her/his decision making process (See Sims [36]). This way, we discard the assumption of "representative rational agent"—with complete and transitive preferences over price increases—which is the backbone of the traditional theory.

### 3.1.2. Constraint 2: The Impact of Actions on the House Price Growth Rates

The second constraint that will help us to model the impact of selling/buying actions on house price changes in the market is the conditional outcome constraint. As mentioned before, the act of selling houses itself tends to lower the house price growth rate in the sub-market relative to what it would have been if selling behavior had no downward pressure on the price change in the market. The effectiveness of buying and selling process can be represented in a maximum entropy formalism by a constraint on the difference between the weighted expected house price growth rates conditional on selling and buying actions of the individuals which can be written as:

$$\begin{aligned} f[sell]E[x|sell] - f[buy]E[x|buy] &\leq \delta \\ \int (f[sell, x] - f[buy, x])xdx &\leq \delta \end{aligned} \quad (4)$$

where  $\delta$  represents the "effectiveness" of the competition in the market. When  $\delta$  gets larger, the effect of selling and buying actions on the house price growth rate worsens indicating a deterioration in negative feedback mechanism from actions to house price change.

Based on these two constraints, it is argued that individuals make their decisions to sell or buy a house by comparing the house price increase in a particular region with their "expected fundamental rate of price (value) increase of their houses". To illustrate, in a housing market, sellers tend to sell more often if house price increase in the region is higher than her expected fundamental rate of price increase of the house expecting lower returns in the future. On the other hand, buyers tend to buy more often if the regional price increase is lower than their expected fundamental rate of price increase of a house expecting more appreciation in price increase for the house in the future. The collective result of the actions, in turn, is strong enough to impact the house price growth rate in the market (It is argued that each individual's action creates a "small" impact on regional price changes). This means, when sellers

sell their properties in a particular market, their actions set forces for a negative feedback mechanism on the house price change, and house price growth rate declines in the market. The opposite happens when the buyers buy houses so that upward price changes tend to increase. As a result of this statistical process, actual transactions among typical agents give rise to a statistical equilibrium distribution of house price growth rates around an average which the system tends to evolve towards. The regularities in the system, therefore is represented as a distribution of house price changes.

### 3.2. Solution to the Maximum Entropy Program of the QRSE Model for the Housing Market

Given the observed frequency data for monthly house price growth rates  $\bar{f}[x]$ , the theory of dependency of buying/selling actions on house price growth rates, and dependency of house price growth rates on actions, we aim to infer a quantitative model of housing market behavior in the form of a joint probability distribution  $f[A, x]$  over the action,  $A$ , and observed house price growth rates by using the maximum entropy method.

As the result of maximizing entropy of the joint distribution subject to three constraints—the constraint on the difference between the weighted conditional expected outcomes, Equation (4), a constraint on the mean value of the outcome,  $\int f[x]xdx = \xi$ , and the assumption that the conditional action is a logit function, Equation (3) with parameters  $\mu$  and  $T$ —the maximum entropy (recovered) marginal distribution  $\hat{f}[x]$  is inferred as the most probable distribution representing the statistical equilibrium in the housing market as (See Appendix B for the derivation of the maximum entropy program):

$$\hat{f}[x] = \frac{e^{H_{\mu,T}[x]} e^{-\gamma x} e^{\beta \operatorname{Tanh}[\frac{x-\mu}{2T}]x}}{\int e^{H_{\mu,T}[x]} e^{-\gamma x} e^{\beta \operatorname{Tanh}[\frac{x-\mu}{2T}]x} dx}. \quad (5)$$

$\hat{f}[x]$  is predicted as a Kernel to the maximum entropy program, and together with  $T, \mu, \beta$  and  $\gamma$  provides us with a multinomial model of the house price growth rates. The hyperbolic Tangent function, ( $\operatorname{Tanh}$ ), which also is a symmetric function, represents the kurtosis of the system, and determines how spread out the marginal distribution can get.

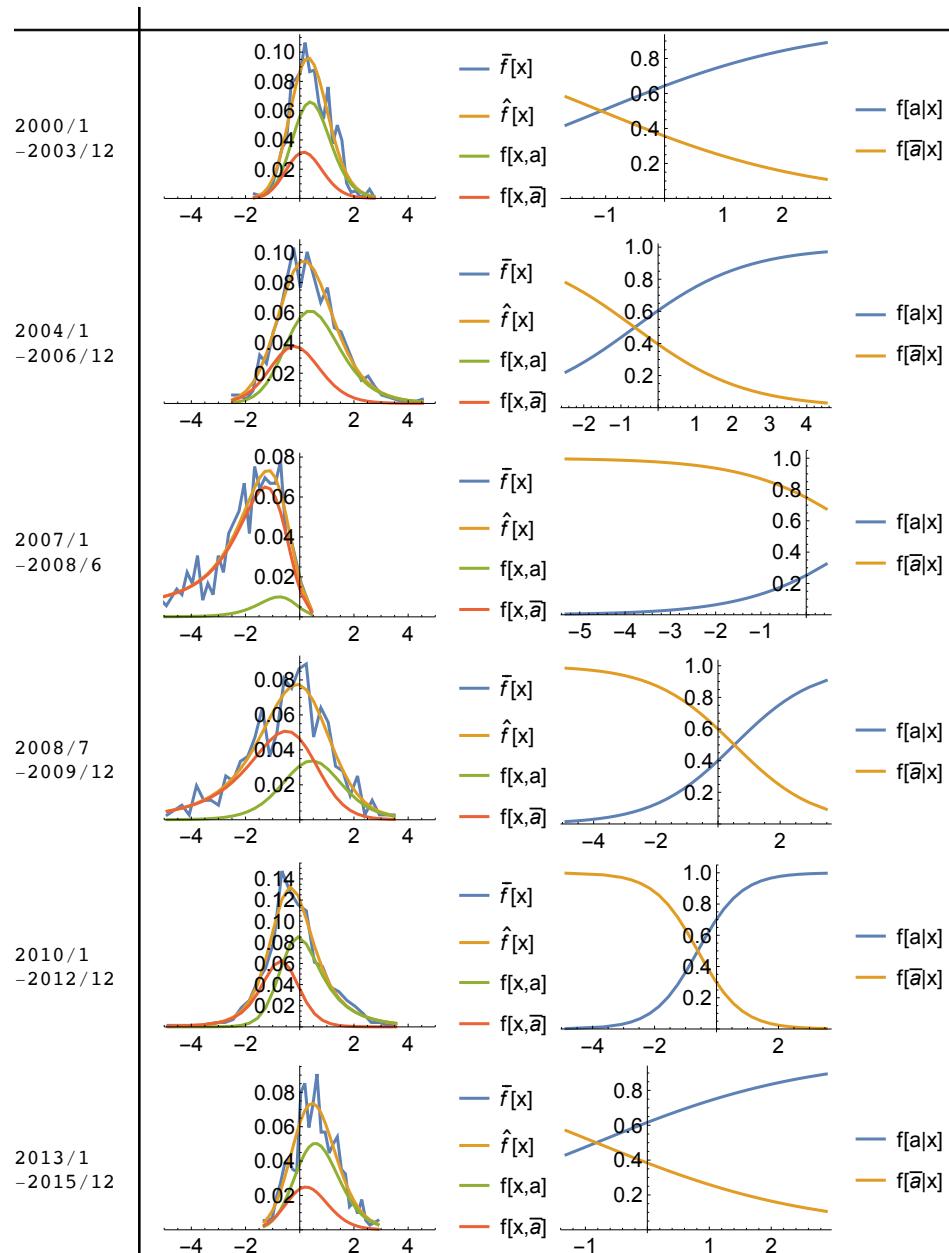
$T$  is called “responsiveness” parameter, and gives information about how the change in house prices affect the behavior of the individuals. Therefore, it is useful to understand how strong is the relationship between the outcome and the action, i.e., how much the change in house price growth rate affects the decisions of individuals’ action of selling houses. It provides useful information about the fluctuations and uncertainties in agents’ decision making process. The impact parameter  $\beta$ , which is the Lagrange multiplier for  $\delta$ , measures the impact of intensity of actions on the sub-regional outcomes i.e., how the actions of agents (selling/buying) affect the house price growth rate in a particular region. Another important parameter,  $\mu$  represents the expectations of individuals about the fundamental rate of price increase for a house, which can also be interpreted as perceptual behavior point in which the agents are active buyers and/or sellers in the sub-regions. Therefore, equilibrium behavior of the system could be measured as a combined result of parameters  $T, \mu$  and  $\beta$ . Parameter  $\gamma$  represents the shadow price for the mean constraint, which reveals information about the skewness of the marginal distribution. It also helps us to understand the relationship between expected fundamental rate of price increase for the house,  $\mu$ , and actual statistical average price increase in the data,  $\xi$ .  $\gamma$  becomes zero if  $\mu = \xi$ , while  $\gamma < 0$  represents positive skewness in the marginal distribution (See Appendix C, for the posterior probability estimation of the parameters, which provides the credible intervals for the estimated parameters).

## 4. Model Inference and Results

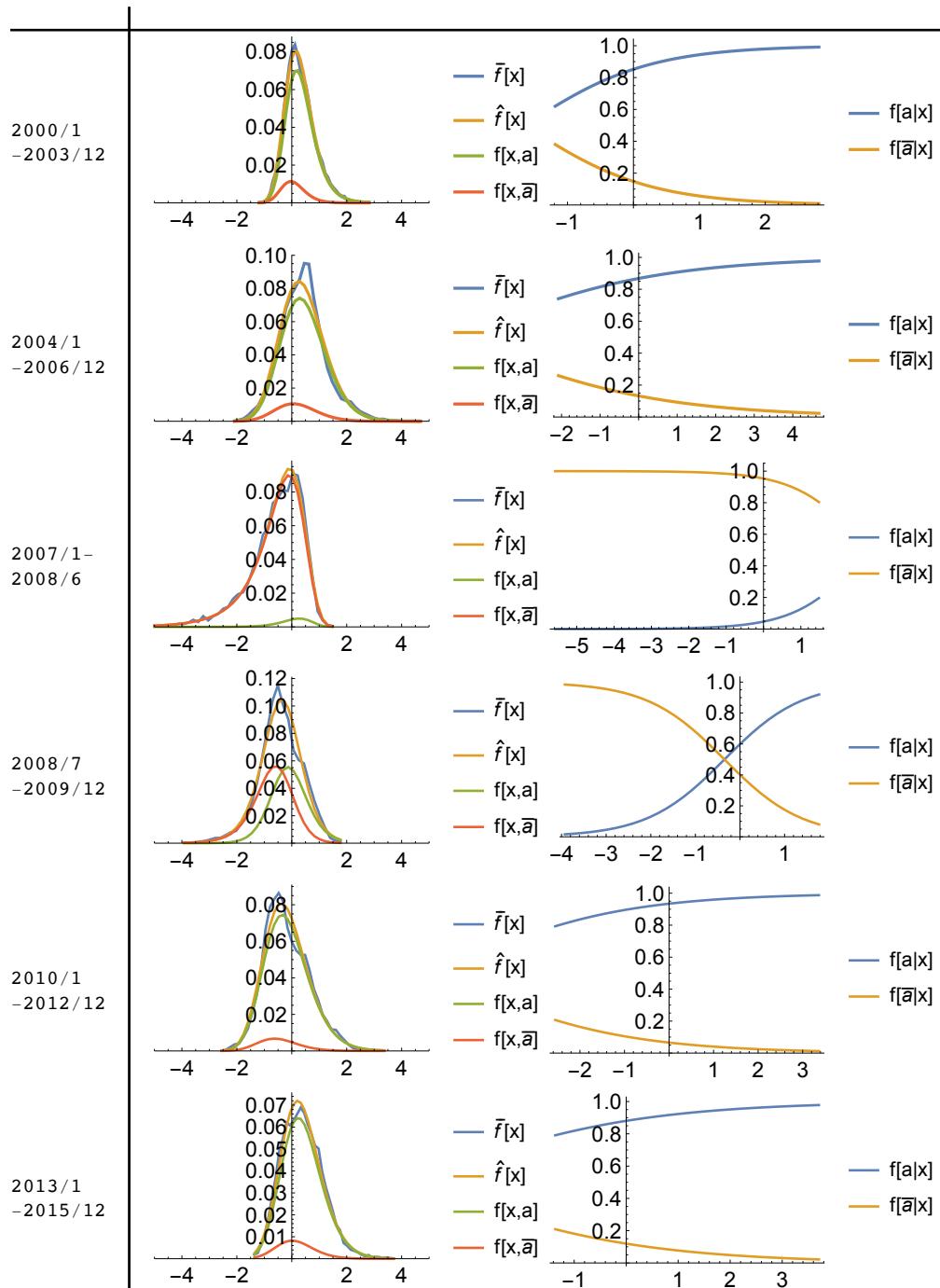
Figures 5 and 6 visualize the solution to the maximum entropy program subject to constraints from Appendix B. Parameters are estimated for subsequent time periods; 2000–2003, 2004–2006, January 2007–June 2008, July 2008–December 2009, 2010–2012 and 2013–2015 using the panel data of monthly house price growth rates for the metropolitan areas represented in the data sets. Sub-periods are chosen

based on distinctive differences in house price growth rate distributions which are represented in Figure 2 by taking information about historical house price movements in the US housing market into account. Then, for each sub-period, monthly house price growth rates of the entire MSAs represented by the data are pooled together for the estimation. Here, the objective is to capture the evolution of housing market behavior both qualitatively and quantitatively throughout the cycle.

Figures represent the observed marginal frequencies;  $\bar{f}[x]$ , the fitted (inferred) marginal frequencies;  $\hat{f}[x]$ , the maximum entropy conditional outcome frequencies;  $f[x, a], f[x, \bar{a}]$  (left), and the estimated conditional action frequencies;  $f[a|x], f[\bar{a}|x]$  (right), in which  $a$  represents selling action while  $\bar{a}$  represents buying action in the market.



**Figure 5.** Model Fit Plots (2000–2015) (Based on Inflation Adjusted S&P Case-Shiller Index Monthly House Price Growth Rate for each MSA represented in S&P Case-Shiller Index). Observed marginal frequencies;  $\bar{f}[x]$ , fitted (inferred) marginal frequencies;  $\hat{f}[x]$ , maximum entropy conditional outcome frequencies;  $f[x, a], f[x, \bar{a}]$  (left), estimated conditional action frequencies;  $f[a|x], f[\bar{a}|x]$  (right), in which  $a$  represents selling action while  $\bar{a}$  represents buying action in the market.



**Figure 6.** Model Fit Plots (2000–2015) (Based on Inflation Adjusted FMHPI Monthly House Price Growth Rate for each MSA presented in FMHPI index). Observed marginal frequencies;  $\bar{f}[x]$ , fitted (inferred) marginal frequencies;  $\hat{f}[x]$ , maximum entropy conditional outcome frequencies;  $f[x,a], f[x,\bar{a}]$  (**left**), estimated conditional action frequencies;  $f[a|x], f[\bar{a}|x]$  (**right**), in which  $a$  represents selling action while  $\bar{a}$  represents buying action in the market.

Table 1 presents the values of the estimated parameters and related variables obtained from the model fit results.

**Table 1.** Estimation results of parameters and related variables based on case-shiller index (vs. FMHPI).

	ID	KL Div.	$\mu$	$T$	$\beta$	$\delta$	$\gamma$	$\xi$	Marg. Action Prob.
									margin( $a$ ) margin( $\bar{a}$ )
2000-1 & 2003-12	0.96 (0.99)	0.04 (0.01)	-1.07 (-1.64)	1.82 (0.94)	4.41 (10.12)	0.26 (0.28)	-2.01 (-8.07)	0.39 (0.30)	0.68 (0.88) 0.32 (0.12)
2004-1 & 2006-12	0.98 (0.99)	0.02 (0.01)	-0.62 (-4.7)	1.47 (2.5)	1.49 (10.02)	0.44 (0.37)	-0.54 (-7.9)	0.35 (0.41)	0.64 (0.88) 0.36 (0.12)
2007-1 & 2008-6	0.97 (0.99)	0.03 (0.01)	1.36 (2.79)	1.25 (0.93)	6.16 (13.3)	1.62 (0.56)	6.07 (12.3)	-1.83 (-0.55)	0.09 (0.04) 0.91 (0.96)
2008-7 & 2009-12	0.96 (0.99)	0.04 (0.01)	0.52 (-0.35)	1.29 (0.86)	0.76 (1.68)	0.75 (0.31)	0.29 (0.36)	-0.47 (-0.42)	0.36 (0.49) 0.64 (0.51)
2010-1 & 2012-12	0.99 (0.99)	0.01 (0.01)	-0.6 (-5.17)	0.7 (1.95)	0.93 (10.71)	0.49 (-0.10)	-0.09 (-8.83)	-0.15 (-0.17)	0.60 (0.92) 0.40 (0.08)
2013-1 & 2015-12	0.97 (0.99)	0.03 (0.01)	-0.83 (-4.09)	1.74 (2.05)	3.15 (11.13)	0.36 (0.34)	-1.57 (-9.02)	0.57 (0.37)	0.68 (0.89) 0.32 (0.11)

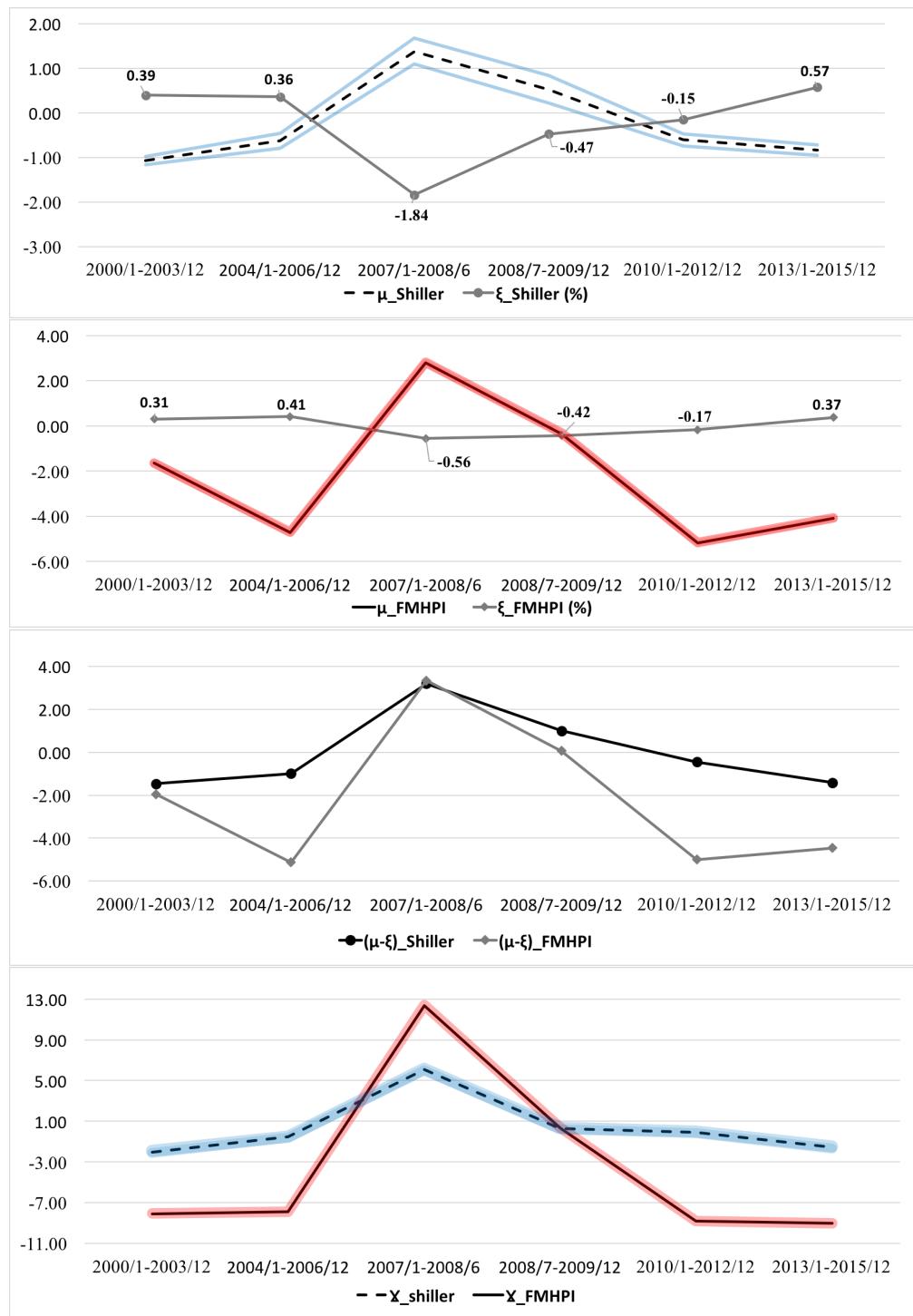
Note: ID represents the information distinguishability measure, which is a measure of closeness between the estimated and observed marginal distributions, Kullback Leibler Divergence is KL Div. (See Appendix C for more details on ID measure and KL Divergence), and the predicted marginal action probabilities of selling and buying are  $\text{marg}(a)$  and  $\text{marg}(\bar{a})$ .

Model fit results, show that the observed distributions of house price growth rates can be explained as the statistical equilibrium (or dis-equilibrium for some periods) produced by the interactions of buyers and sellers in the sub-housing markets according to a logit quantal response and a given effectiveness of buying/selling in affecting the house price growth rates in these regions.

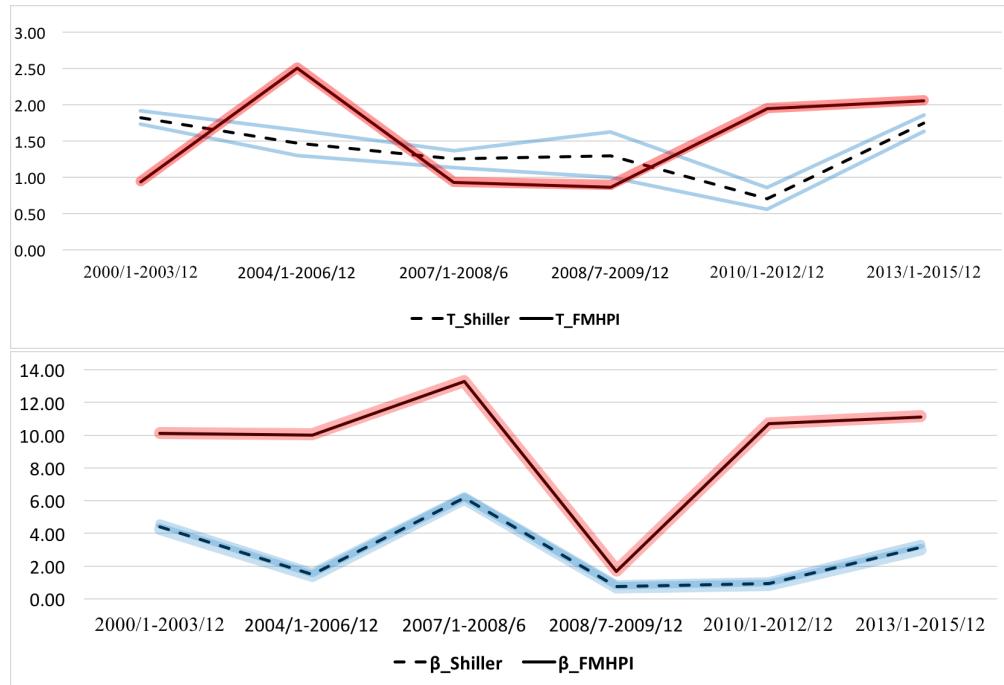
As can be followed from the Appendix C in more detail, model parameters  $T, \mu, \beta$  and  $\gamma$ , are estimated jointly by minimizing the Kullback-Leibler divergence between the observed marginal frequencies,  $\tilde{f}[x]$ , and inferred marginal frequencies  $\hat{f}[x]$ , where the inferred frequencies  $\hat{f}[x; T, \mu, \beta, \gamma]$  is obtained from Equation (5). Results from this process indicate that given the estimated parameters,  $T, \mu, \beta, \gamma$ , the QRSE model generates almost a perfect fit for the observed marginal frequencies in both data sets. Calculated information distinguishability (ID) measures, which show the closeness of the model fit to the observed frequencies, indicates that fitted distributions of marginal frequencies,  $\hat{f}[x]$  capture approximately 96 to 98% of the information content of the observed frequencies for each sub-period.

#### 4.1. Interpretation of Estimated Model Parameters

Using the model fit results, Figures 7 and 8 present the time series of estimated parameters,  $T, \mu, \beta, \gamma$ , with highest posterior density regions (HPD) (See Appendix C).



**Figure 7.** Estimated Parameters with 95% Credible Intervals (HPD):  $\mu$ , actual average house price growth rate  $\xi$ ,  $(\mu - \xi)$  and  $\gamma$ . Red and Blue lines represent the 95% credible intervals for estimated parameters.



**Figure 8.** Estimated Parameters with 95% Credible Intervals (HPD):  $T$  and  $\beta$ . Red and Blue lines represent the 95% credible intervals for estimated parameters.

#### 4.1.1. Interpreting $\mu$ , $\gamma$ and Their Relationship with Actual Average House Price Growth Rate $\xi$ through the Cycle

Parameter  $\mu$ —expected rate of fundamental price increase of a house—provides useful information about the expectations of economic agents. If  $\mu \neq \xi$ , it means that expectations of individuals about the fundamental house price increase do not coincide with the actual average price increase of the whole market signaling “unfulfilled” expectations of the individuals [33].

As Figure 7 shows, results from the model fit for both data sets indicate that individuals' expectations about the price increase stay below the actual average house price increase ( $\mu < \xi$ ) before and after the crisis for both data sets. Moreover, as can be followed from the Table 1, marginal probability of selling,  $\text{marg}(a)$  seems to be higher than the probability of buying,  $\text{marg}(\bar{a})$  in the market pointing out that the market participants were more willing to sell before and after the crash. This proves that higher actual (average) price increase than what market participants expected encouraged the sellers (buyers) to supply (demand) their houses more (less) frequently signaling the characteristics of a boom period. This is because the opportunity cost of keeping an empty house for a seller during the boom periods is higher than it is during the bust since there is a chance that house values might decline in the next period [37]; hence, the payoff of selling becomes higher during the boom, while it is the opposite for the bust and crash periods. During the crisis,  $\mu$  exceeds  $\xi$  signaling a significant deterioration in individuals' expectations and willingness to sell which shows itself as a low probability of selling action because the price changes dropped significantly. After the crisis, their perceptions about the market seem to return to pre-crisis periods where  $\mu < \xi$ . It is important to emphasize that expectations of economic agents never perfectly coincided with the actual average price increase— $\mu$  was never equal to  $\xi$  throughout the cycle. Another way of presenting the individuals' expectations and willingness of selling/buying can be followed by the measure,  $(\mu - \xi)$ , from Figure 8. Increasing gap between individuals' expectations and actual average price increase becomes more obvious during the collapse. By reaching its peak during the crash,  $(\mu - \xi)$  with a low probability of selling action renders how overcautious the market participants become when the system enters to crisis.

In addition to the relationship between  $\mu$  and  $\xi$ , another important piece of information that  $\mu$  provides us with is that housing market was more inclined to be “frothy” and speculative before and after the crash meaning that the market was more vulnerable to create a housing market bubble. This can be observed from the expectation parameter  $\mu$  being negative for pre and post-crisis periods for both data sets (except during the crash it becomes positive).

The estimation of  $\gamma$  parameter informs us about the skewness of the marginal distributions, which in turn, helps us to interpret the relationship between  $\mu$  and the actual mean  $\xi$ . When  $\mu$  is equal to  $\xi$ ,  $\gamma = 0$  indicates perfect symmetry in the marginal distribution,  $\bar{f}[x]$ . On the other hand,  $\mu < \xi$  implies positive skewness, meaning that some of the regions are experiencing house price increases well above the market average. This might mean that the market is actually in the boom cycle and building upon the housing bubble. As can be seen from Figure 8,  $\gamma$  was negative before and after the crash demonstrating a positive skewness in the marginal distributions, which also coincides with the boom periods in the housing markets while house price growth rates were increasing. During the crisis,  $\gamma$  shifts from negative to positive values indicating an increase in skewness to the left caused by negative price growth rates in most of the regions due to the market crash. Based on these results, one can argue that estimated  $\gamma$  parameter is a useful indicator for differentiating between boom and bust periods in the market, and the divergence between the expectations and the real outcome.

Overall, after the early pre-crisis sub-period (2000–2003), sellers start to become more cautious about their actions which can be observed from the relationship between  $\mu$ ,  $\xi$  and  $\gamma$ . Although marginal probability of selling was higher than the buying action during the periods before the crash, it starts to decline after 2003, and buying becomes more attractive than before, and the system becomes less positively skewed to the right before the crash because regional prices started to depreciate around late 2006. Periods of crisis, in which sharp declines in the house price growth rates occurred in most of the regions, could be represented as vanishing willingness to sell in the market (increasing  $\mu - \xi$ ), and increasing negative skewness of marginal distribution of the house price growth rates (positive values of  $\gamma$ ).

After the crash, both  $\mu$  and  $\gamma$  decline and get closer to their pre-crisis levels emphasizing the similarities in the behavior of the system between pre-crisis and post-crisis periods.

#### 4.1.2. Interpreting Negative Feedback Mechanism, $\beta$ , through the Cycle

As mentioned before,  $\beta$  is the Lagrangian multiplier for the impact constraint,  $\delta$  that limits the impact of selling on the change of house price growth rates in sub-regions. Therefore, it represents the impact of actions on the house price changes providing information about how effective is the market. The lower the  $\beta$  is, the lower the market effectiveness. This, in turn, means that the negative feedback mechanism from actions on house price changes, was deteriorated. For example, in a well functioning market, an increase in house price growth rate in a particular region would attract sellers to sell. Once the actions are taken, if the market has an highly effective negative feedback mechanism, it would cause a decline in house price growth rate of this particular region. On the other hand, if the mechanism does not work properly, house price growth rate in the region would keep increasing.

Time series of estimated  $\beta$  for both data sets from Figure 7 show that impact of selling/buying actions on forcing house price growth rates in the regions towards the average growth rate of the market was much weaker between 2004–2006 than 2000–2003 indicating a deterioration in the negative feedback mechanism right before the crash. This behavior of the market creates a plausible environment for unsustainable housing price bubble before the crash.

During the first half of the crisis,  $\beta$  increases to keep the system in equilibrium but since the willingness of selling action drops dramatically ( $\mu > \xi$ ) which also shows itself as a drastic drop in probability of selling (low supply of houses), even a strong negative feedback mechanism (higher  $\beta$ ) is not able to prevent market from collapsing. As a result, after the first half of the crisis,  $\beta$  reaches its lowest level indicating a dysfunctional negative feedback from the actions. To illustrate, if deterioration in the negative feedback mechanism did not occur between 2004–2006, strong willingness of selling

action of individuals (higher supply than demand), represented by the relationship between  $\mu$  and  $\xi$ , would cause house price growth rates to decline preventing prices from growing further. Since this is not the case between 2004 and 2006, price growth rates keep appreciating. When the appreciation reaches to unsustainable levels, market collapses in 2007 leading to further decline in the effectiveness of negative feedback mechanism between July 2008 and December 2009. After the collapse,  $\beta$  starts to rise again signaling the recovery of the market forces, which brings the system back to statistical equilibrium.

Overall, the trends in  $\beta$  and  $\mu$  provide a strong evidence that the cause of the bubble cannot be explained only by a simple supply and demand relation (e.g., the larger demand for the housing than the supply of the housing in the market) because the intrinsic mechanism behind the boom and the housing market bubble was the relationship between the actions and the outcome price changes. Traditional explanations for the housing market behavior such as a positive feedback mechanism on house price appreciation from the buyers, or speculations in the market [2,38] could only explain some part of housing bubble creation. If the negative feedback mechanism between the actions and house price changes worked effectively during the boom, supply of housing by sellers would be able to prevent the bubble from happening, and the system would not be able to diverge from the strong statistical equilibrium of the early 2000's. Wrong (unfulfilled) expectations seem to feed into both increasing prices and reckless actions (e.g., buyers ready to buy at a higher rate than normal times) without disrupting the bubble creation.

#### 4.1.3. Interpreting Responsiveness Parameter, $T$ through the Cycle

As briefly mentioned before, logit parameter from the quantal response behavior of individuals,  $T$ , determines the endogenous fluctuations of the house price growth rates as an indicator of the responsiveness of the economic agents' actions to the house price changes. The higher the  $T$  is the flatter the conditional action frequencies  $f[A|x]$ —flatter S-curves—become implying a weaker dependence of selling/buying actions on the house price changes. As a result, fluctuations and individuals' uncertainty of house price change become large. When  $T$  gets closer to zero, the dependence of actions on house price changes becomes stronger. This means even a small change in house price growth rate results in an instantaneous response of individuals on choosing their actions.

Responsiveness parameter,  $T$ , thus provides us with two quantitative measures to interpret the behavior of the individuals in a certain time period. First, it measures how responsive the action frequency of economic agents (buying/ selling) to changes in house price growth rates at the average house price growth rate, because derivative of the conditional action function at  $x = \mu$  is  $\frac{1}{4T}$ . To illustrate  $T = 1.82$  for the early pre-crisis period (January 2000–December 2003) based on Case-Shiller index, indicates that a 1% increase in the house price growth rate near the average growth rate would raise the frequency of selling action by 0.13%. Based on FMHPI index results,  $T = 0.94$  for the same period indicates that the same increase in house price growth rate near the average would raise the frequency of selling action by 0.26%. The difference between the two results tells us that any increase in house price growth rate would be more effective to attract sellers to take action in the market with a smaller  $T$ , who are more responsive to the changes in house price growth rates.

Second,  $T$  defines the region where the action is sensitive to house price changes: the value of  $x$  for which the conditional action frequency is  $p$ , with  $0 < p < 1$  is  $x[p, T] = T \log[p/(1-p)]$ . For example, based on Case-Shiller Index data,  $T = 1.82$  implies that the region of house price growth rate deviations where the frequency of selling lies between 0.05 and 0.95 is  $\{-5.36\%, 5.36\%\}$  while it is  $\{-2.76\%, 2.76\%\}$  based on FMHPI data for the period of 2000–December 2003. Accordingly, this tells us about how tight the quantal response link functions are (See Figures 5 and 6 for the quantal response link function plots, which represent conditional action distributions  $f[a|x]$  and  $f[\bar{a}|x]$ ).

The time series of  $T$ , estimated from the Case-Shiller Index data in Figure 7 demonstrates a stable pattern with a smooth decline between 2000 and 2009 from  $T = 1.82$  to  $T = 1.29$ . In the early post-crisis

period (2010–2012), it drops sharply to its minimum level,  $T = 0.7$ , and starts to increase dramatically after 2012 nearly reaching to its early pre-crisis period level,  $T = 1.74$  in the end of the cycle.

From the FMHPI index,  $T$  seems to follow a more volatile pattern as compared to results from the Case-Shiller Index. Starting from,  $T = 0.94$ , in early pre-crisis period, it reaches its maximum level of  $T = 2.5$ , in late pre-crisis period (2004–2006). A sharp decline in  $T$  occurs during the crisis (2007–2009) reaching approximately the same level as in the case of Case-Shiller Index data,  $T = 0.86$ . After the crash, it starts increasing steadily above its late pre-crisis period level of  $T = 2.05$  towards the end of the cycle between 2013 and 2015.

Patterns in the behavior of the temperature parameter,  $T$ , based on Case-Shiller Index data, which represent the largest metropolitan cities in the US, imply that typical sellers/buyers were less responsive to the changes in the house price growth rates at the beginning of the cycle, but they started to eliminate the uncertainties about the house price changes, that in turn, strengthened the dependency between actions and house price changes. In general, one would expect increasing uncertainties of individuals' actions of buying and selling to changes in house prices when the market gets closer to collapse. Yet, here, individuals seem to build stronger ties to the house price movements before the crisis, which might be due to increasing optimism about the future of the housing market, such that every increase in growth rate of house prices encouraged house sellers to be more responsive to house price increases in the market.

On the other hand, in case of FMHPI index data, which represents a larger number of metropolitan regions,  $T$  follows a sharp increase during the late pre-crisis period (2004–2006) signaling the worsening responsiveness of individuals' actions to house price changes. When the crash occurs responsiveness to the market signals seems to be regained but starts to deteriorate again right after the crash.

## 5. Discussion

The relationship between individuals' expected rate of fundamental price increase of the houses ( $\mu$ ) from the model estimations, which represents the expectations of individuals, and the actual average price increase in the market ( $\xi$ ) raises important questions about how models based on traditional theory might be misleading. To illustrate, taking the housing market case as an example, the traditional approach [30–32] would assume that sellers' expectations about the price increase of a house would be equal to actual average price increase in the market ( $\mu = \xi$ ). This is because individuals would have "correct (fulfilled)" expectations about the price changes. Moreover, if they had wrong expectations, the market would clear immediately since the individuals with wrong expectations would either be dropped out of the market, or would correct their expectations immediately as they process all available information related to prices. As a result, the differences in price changes in different sub-regions would cancel each other out, and the market would reach to the equilibrium instantaneously such that we would end up having a symmetric frequency distribution of house price changes for the whole market. Additionally, the process of market clearing would not impose any cost on market participants.

Even though, this scenario approximates the time periods where the market wanders around a statistical equilibrium (e.g., 2000–2003 and 2013–2015), it cannot explain the observed positive skewness in the distribution between 2004–2006, and negative skewness during the first half of the crisis (2007–June 2008) in the Figures 5 and 6. As opposed to statistical equilibrium periods, they indicate significant divergence from the equilibrium signaling "unfulfilled (wrong)" expectations of individuals.

In contrast with the traditional theory, this study shows that individuals' expectations about fundamental rate of price increases never coincided with the actual average price change of the market ( $\mu \neq \xi$ ). Instead, when the expectations about the fundamental rate of price increase was lower than the actual price increase ( $\mu < \xi$ ), sellers enjoyed higher price appreciation than expected while buyers might have had to buy at a "too high rate" believing that prices would have continued appreciating in the future. This market behavior seems to result itself in a boom period which in turn, explains

the positive skewness in house price changes before the crash. Since price increase kept appreciating during the boom, positive skewness could be interpreted as a “market punishment for the buyers” since they kept buying at an higher price increase than the normal times due to their optimistic expectations. During the crash, on the other hand, market dynamics took the opposite direction in which individuals’ expectations about the fundamental rate of price increase reached to a higher level than the actual average rate of price increase ( $\mu >> \xi$ ) so that sellers’ willingness to sell declined as they expected higher rate of price increase for their houses than the market average. As a result, a sharp, further decline occurred in the observed actual average price increase ( $\xi$ ) well below what individuals expected. Even though probability of buying (demand of housing) was much higher than the previous periods, market was frozen due to lack of housing supply (very low probability of selling). The period of crash with a declining house prices in most of the regions shows itself as a negatively skewed frequency distribution of house price changes in Figures 5 and 6. Since the values of houses kept depreciating with declining prices, the negative skewness could be interpreted as the “market punishment for the sellers [33]” during the crisis.

Based on these results, one can argue that housing market does not clear immediately, and the process of getting out of disequilibrium dynamics creates certain costs on market participants.

## 6. Conclusions

In this paper, it has been shown that employing the maximum entropy method with plausible theoretical frameworks as the constraints, we are able to overcome the incomplete information problem resulted from unavailability of data for individuals’ actions. As a result, we were able to fully determine the statistical model explained as an equilibrium joint distribution of observed house price growth rate data and unobservable buy/sell actions of individuals given the predicted model parameters,  $T, \mu, \beta$  and  $\gamma$ . The recovered joint distribution of house price growth rates,  $x$  and individual actions  $\{a, \bar{a}\}$ , which implies the marginal distribution of house price growth rates,  $\hat{f}[x]$ , was then compared to the observed distribution of  $x$ ,  $\tilde{f}[x]$  using the Kullback-Leibler (KL) divergence. Since KL Divergence is a good approximation to log posterior probability for the multinomial model for any parameters, we were able to estimate the conditional posterior distributions of each model parameter holding others at their maximum posterior estimates.

Estimations from the both data sets used in the analysis (S&P Case Shiller Index and FMHPI index) follow strikingly similar patterns throughout the cycle proving the strength of the QRSE model’s explanatory power of defining the behavior and dynamics of social systems.

Predicted parameters of the model  $T, \mu, \beta$  and  $\gamma$  reveal significant characteristics of the US housing market dynamics for pre-crisis, crisis and post-crisis periods during the most recent housing market cycle in the US. Model fit estimations show strong statistical equilibrium before the crisis. Deviations from the equilibrium start to arise between 2004 and 2006, and the first half of the crisis period (2007–June 2008) represents the disequilibrium in the housing market. It is worth mentioning that QRSE model is a statistical equilibrium model, and it might not be able to produce powerful results for the out-of-equilibrium cases such as housing market collapse because forces that keep the system in the statistical equilibrium might not work as well during the crash. In our case, self-organizing dynamics of the system seem to be weakened during the crisis; however, they were still strong enough to be measured in this study. The system seems to start recovering during the second half of the collapse (July 2008–December 2009). After the crisis, the market returns back to statistical equilibrium between 2010 and 2015. Positive average house price growth rates above early pre-crisis period between 2013 and 2015 could be signaling a new housing boom cycle which might necessitate further analysis but is outside of the scope of this study.

Based on our analysis, we argue that increasing average house price growth rates before the housing market collapse created an optimistic environment in a way that weakening responsiveness of individuals’ action on house price growth rate movements (higher  $T$  compared to crash period) before the collapse is one of the reasons that caused system to build up a period of housing boom/bubble.

This may be because glamorous expectations for the future of housing market deteriorated individuals' capability of processing the market signals. Individuals' responsiveness reached its maximum during the crisis when the collapse showed itself as a fact rather than just being a possibility.

In addition to weak responsiveness of individuals to market outcomes (higher  $T$ ) between 2004–2006, strong willingness of individuals to sell/buy in the same period seems to represent "wrong (unfulfilled)" expectations which leads to the speculative market behaviors before the crisis. Along with a weakened negative feedback mechanism between actions and house price changes (lower  $\beta$ ), strong willingness to transact in the market seems to aggravate further increases in house price growth rates triggering the housing bubble during the pre-crisis periods, although the probability of selling (supply) was higher than the probability of buying (demand). This period, therefore, seems to create a market punishment for the buyers as they kept buying at a "too high prices" when the prices were appreciating at an increasing rate.

Between 2007–2009, even though the feedback mechanism recovered during the first half of the crisis (2007–June 2008) (higher  $\beta$ ), the market could not escape from the collapse due to the fact that high responsiveness and deteriorated willingness of selling to declining price growth rates (lower  $T$  and "unfulfilled" expectations—broadening gap between  $\mu$  and  $\xi$ ) prevented selling actions to take place. As a result, the forces keeping the system in statistical equilibrium became unsustainable, and the system-dynamics collapsed leading to dis-equilibrium in the market between 2007 and 2008. Finally, after the crash, average actual growth rate of the market and estimated parameters start to return to their pre-crisis values revealing a recovery of statistical equilibrium after the collapse.

Overall, this study is the first application of the QRSE models on the US housing markets in which we are able to capture and explain different characteristics and behaviors throughout the most recent boom-bust cycle by distinguishing out-of-equilibrium from times of market equilibrium. Modeling highly complex social systems such as housing markets is difficult because their dynamics evolve through the interactions among a large number of heterogeneous individuals under the influence of each others' decisions and possible external impacts (e.g., policy changes, natural events etc.). This is mostly because market interactions and individuals actions often are not observable to the researcher and the parameter space necessary to model them has high dimensionality, which is not possible to fully account. Moreover, factors such as the lack of high quality data to represent unique non-repetitive social events and the computational inabilities make it even harder. Given the existing difficulties confronted in the literature, we present a simple-alternative method which has a high potential to improve modeling techniques of complex social systems for future economic research problems.

A further extension of the QRSE housing market model can include other economically important variables, such as interest rates, mortgage rates, credit and loan options and debt-to-GDP ratios of different regions—which play significant roles in individual's buying/selling behavior in order to provide a more economically detailed analysis of housing market dynamics.

One difficulty with QRSE models, on the other hand, is that they require discretization of the data in the form of coarse-grained bins in order to recover statistical equilibrium joint distribution of action  $A$ , and outcome  $x$ . Estimation of the model parameters may be affected by the bin size chosen by the modeler. This topic is of interest for future research.

**Funding:** This research received no external funding.

**Acknowledgments:** I would like to thank to Duncan K. Foley for detailed discussions and his help and support throughout this study. I am also grateful to Mark Setterfield, Paulo dos Santos, Lance Taylor, Ellis Scharfenaker, Jangho Yang and Luis Daniel Torres Gonzales for their detailed comments. Finally, I would like to thank to the New School for Social Research administration for their support.

**Conflicts of Interest:** The author declare no conflict of interest.

## Appendix A. Data

Freddie Mac House Price monthly index, FMHPI, which is available for DC area and 367 Metropolitan statistical areas (MSAs) throughout the US, and the second one is monthly S&P

Case-Shiller Index for 20 largest US metropolitan cities called as “Superstar” cities in the literature [39]. Both data sets employ the repeat sales method of index calculation. Compared with S&P Case-Shiller Home Price index, FMHPI has a larger sample size, and covers more geographic areas. FMHPI is based on transactions involving single family detached and town-home properties purchased by Fannie Mae or Freddie Mac, and is a weighted index, meaning that it measures average price changes in repeat sales or refinancing on the same properties. The S&P index is calculated monthly using a 3 month-moving average algorithm for the data on single-family housing stocks that have sold at least twice in order to capture the appreciation in the values of each sales unit. The most important difference between the two is that the S&P Case-Shiller index is value-weighted which means that price trends for more expensive houses have greater influence on estimated price changes than other homes, whereas FMHPI’s purchase only index weights price trends equally for all properties.

## Appendix B. The Maximum Entropy Program of the QRSE Model for the Housing Market

This section presents the solution to the maximum entropy program, which is used to predict the marginal frequencies,  $f[x]$ , subject to three constraints; the constraint on the difference between the weighted conditional expected outcomes, Equation (4), a constraint on the mean value of the outcome,  $\int f[x]xdx = \xi$ , and the assumption that the conditional action is a logit function with parameters  $\mu$  and  $T$ .

Incorporating the assumption of quantal response, constraint on the difference between the weighted conditional expected outcomes, Equation (4) with the conditional action constraint takes a more compact form which is a hyperbolic tangent function that is:

$$\begin{aligned} f[sell]E[x|sell] - f[buy]E[x|buy] &= \int (f[sell|x]f[x] - f[buy|x])f[x]dx \\ &= \int \left( \frac{1}{1 + e^{-\frac{(x-\mu)}{T}}} - \frac{e^{-\frac{(x-\mu)}{T}}}{1 + e^{-\frac{(x-\mu)}{T}}} \right) f[x]dx \\ &= \int \text{Tanh}\left[\frac{(x-\mu)}{2T}\right] f[x]dx \leq \delta \end{aligned}$$

Since the data on actions are not available, the entropy of the joint distribution can be written in terms of the marginal distribution  $f[x]$ , as:

$$H = - \int f[x] \text{Log}[f[x]]dx + \int H_{\mu,T}[x]dx$$

where  $H_{\mu,T}$  is the binary entropy function:

$$H_{\mu,T}[x] = - \left( \int f[x] \text{Log}\left[\frac{1}{1 + e^{-\frac{(x-\mu)}{T}}}\right] + \int f[x] \text{Log}\left[\frac{e^{-\frac{(x-\mu)}{T}}}{1 + e^{-\frac{(x-\mu)}{T}}}\right] \right)$$

Therefore, the maximum entropy program that determines the marginal distribution  $f[x]$  given the parameters  $T, \mu, \delta, \xi$  can be expressed as:

$$\begin{aligned} \underset{f[x] \geq 0}{\text{Max}} \quad & - \int [f[x] \text{Log}[f[x]]dx + \int f[x]H_{\mu,T}[x]dx] \\ \text{subject to} \quad & \int f[x]dx = 1 \\ & \int f[x]xdx = \xi \\ & \int \text{Tanh}\left[\frac{(x-\mu)}{2T}\right] f[x]dx \leq \delta \end{aligned}$$

The Lagrangian to this maxent program, hence is:

$$\begin{aligned}\mathcal{L}[f[x], \lambda, \beta] = & - \int [f[x] \log[f[x]] dx + \int f[x] H_{\mu, T}[x] dx - \lambda \left( \int f[x] dx - 1 \right) \\ & - \gamma \left( \int f[x] x dx - \xi \right) - \beta \left( \int \tanh \left[ \frac{(x - \mu)}{2T} \right] f[x] x dx - \delta \right)\end{aligned}$$

The associated first order conditions for maximizing entropy of the joint frequencies and conditional frequency require:

$$\frac{\partial \mathcal{L}}{\partial f[x]} = -\log[f[x]] - 1 - \lambda + H_{\mu, T}[x] - \gamma \mu - \beta \tanh \left[ \frac{(x - \mu)}{2T} \right] x = 0$$

Solving for  $f[x]$  gives the maximum entropy (recovered) marginal distribution  $\hat{f}[x]$  with normalization factor:

$$\hat{f}[x] = \frac{e^{H_{\mu, T}[x]} e^{-\gamma x} e^{\beta \tanh \left[ \frac{x - \mu}{2T} \right] x}}{\int e^{H_{\mu, T}[x]} e^{-\gamma x} e^{\beta \tanh \left[ \frac{x - \mu}{2T} \right] x} dx}$$

### Appendix C. Posterior Probability Estimates

As mentioned above, predicted marginal distribution  $\hat{f}[x]$  from Equation (5) is a Kernel to the maximum entropy program, and together with the parameters  $T, \mu, \beta, \gamma$ , provides a multinomial distribution for the model. In a Bayesian sense, observed marginal distribution,  $\bar{f}[x]$  is just a sample distribution from the multinomial distribution. We use the Kullback-Leibler divergence as an approximation to the log posterior probability for the multinomial model since it allows us to make posterior inferences about the parameter estimations (See [10] for more details).

Kullback-Leibler divergence measures the discrepancy between the observed marginal frequencies,  $\bar{f}[x]$  and predicted marginal frequencies  $\hat{f}[x; T, \mu, \beta, \gamma]$  inferred from the maximum entropy Kernel as:

$$D_{KL}[\hat{f}[x], \bar{f}[x]] = \sum \hat{f}[x] \log \left[ \frac{\hat{f}[x]}{\bar{f}[x]} \right] \quad (\text{A1})$$

As a result, KL divergence provides us with a tool to compare the observed marginal frequency distribution with the predicted marginal distribution. If  $\hat{f}[x] = \bar{f}[x]$ ,  $D_{KL}$  becomes zero indicating that two distribution are the same. Therefore, smaller the KL divergence is closer the observed distribution to the predicted one and better the fit.

Model parameters  $T, \mu, \beta, \gamma$  are estimated jointly by minimizing the  $D_{KL}$  function between the observed and inferred marginal frequencies. To measure the closeness of the model fit, we used information distinguishability criteria (ID) introduced by Soofi and Retzer [40]. ID measure shows approximately how much of the informational content of the observed frequencies to be captured from the results of the maximum entropy program.

Finally, to make inferences about the posterior parameters and their credible intervals, conditional distributions of each parameter, holding all others at their maximum posterior probability estimates  $\bar{\Omega}$ , (where the model parameters are  $\Omega = \{T, \mu, \beta, \gamma\}$ ), are calculated.

### References

1. Gleaser, E.L.; Nathanson, C.G. Housing Bubbles. *NBER* **2014**, *99*, 1–45.
2. Shiller, R.J. Turning Points in Real Estate. *East. Econ. J.* **2008**, *34*, 1–13. [[CrossRef](#)]
3. Wheaton, W.C. Vacancy Search and Prices in a Housing Market Matching Model. *J. Polit. Econ.* **1990**, *98*, 1270–1292. [[CrossRef](#)]

4. Diaz, A.; Jerez, B. House Prices Sales, and Time on the Market: A search Theoretic Framework. *Int. Econ. Rev.* **2013**, *54*, 837–872. [[CrossRef](#)]
5. Kouwenberg, R.; Zwinkels, R.C.J. Forecasting the U.S. housing market. *Int. J. Forecast.* **2014**, *30*, 415–425. [[CrossRef](#)]
6. Bolt, W.; Demertzis, M.; Diks, C.; Hommes, C.; van der Leij, M. Identifying Booms and Busts in House Prices under Heterogeneous Expectations. *Eur. Comm. Econ. Pap.* **2014**, *540*. [[CrossRef](#)]
7. Arthur, W.B. Out-of Equilibrium Economics and Agent Based Modeling. *Handb. Comput. Econ.* **2006**, *2*, 1551–1564.
8. Erlingsson, E.; Roberto, M.; Steffansson, H.; Sturluson, J.T. Integrating the housing market into an agent-based economic model. In *Managing Market Complexity: The Approach of Artificial Economics*; Springer: Berlin/Heidelberg, Germany, 2012; Volume 662.
9. Axtell, R.; Farmer, D.; Geanokoplos, J.; Howitt, P. An Agent-Based Model of the Housing Market Bubble in Metropolitan Washington, D.C. In Proceedings of the Deutsche Bundesbank’s Spring Conference on Housing Markets and the Macroeconomy: Challenges for Monetary Policy and Financial Stability, Eltville, Germany, 5–6 June 2014; Working Paper.
10. Scharfenaker, E.; Foley, D.K. Quantal Response Statistical Equilibrium in Economic Interactions: Theory and Estimation. *Entropy* **2017**, *19*, 444. [[CrossRef](#)]
11. Boltzmann, L. Über das Wärmegleichgewicht zwischen mehratomigen Gasmolekülen. *Wien. Ber.* **1871**, *63*, 397–418.
12. Gibbs, J.W. *Elementary Principles in Statistical Mechanics*; Scribner: New York, NY, USA, 1902.
13. Maxwell, J.C. Illustrations of the Dynamical Theory of Gases. *Philos. Mag.* **1860**, *20*, 21–37. [[CrossRef](#)]
14. Jaynes, E.T. Information Theory and Statistical Mechanics. *Phys. Rev.* **1957**, *160*, 620–630. [[CrossRef](#)]
15. Cho, M. House Price Dynamics: A Survey of Theoretical and Empirical Issues. *J. Hous. Res.* **1996**, *7*, 145–172.
16. Case, K.E.; Shiller, R.J. The efficiency of the market for single-family homes. *Am. Econ. Rev.* **1989**, *79*, 125–137.
17. Gleaser, E.L.; Gyourko, J.; Morales, E.; Nathanson, C.N. Housing Dynamics: An urban approach. *J. Urban Econ.* **2014**, *81*, 145–156. [[CrossRef](#)]
18. Fama, E.F. *Random Walks in Stock Market Prices*; Selected Papers; University of Chicago: Chicago, IL, USA, 1965.
19. Farjoun, F.; Machover, M. *Laws of Chaos: A Probabilistic Approach to Political Economy*; Verso: London, UK, 1983.
20. Foley, D.K. Statistical Equilibrium in Economics: Method, Interpretation and an Example. In Proceedings of the Workshop on General Equilibrium: Problems, Prospects and Alternatives, Siena, Italy, 4–11 July 1999.
21. Scharfenaker, E.; Semieniuk, G. A Statistical Equilibrium Approach to the Distribution of Profit Rates. *Metroeconomica* **2017**, *68*, 465–499. [[CrossRef](#)]
22. Smith, A. *The Wealth of Nations*; W. Strahan and T. Cadell: London, UK, 1904.
23. Marx, K. *Capital: Volume III*; Penguin Group: London, UK, 1981.
24. Jaynes, E.T. *Probability Theory: The Logic of Science*; Cambridge University Press: Cambridge, MA, USA, 2003.
25. Foley, D.K. Statistical equilibrium in a simple labor market. *Metroeconomica* **1996**, *47*, 125–147. [[CrossRef](#)]
26. dos Santos, P.L.; Scharfenaker, E. Capital-market competition, informational performance, and the distribution of tobin’s q. *New School Soc. Res.* **2016**, *6*, 1–24.
27. Yang, J. A Quantal Response Statistical Equilibrium Model of Induced Technical Change in an Interactive Factor Market: Firm-Level Evidence in the EU Economies. *Entropy* **2018**, *20*, 156. [[CrossRef](#)]
28. Shannon, C.E. A Mathematical Theory of Communication. *Bell Syst. Tech. J.* **1948**, *27*, 379–423. [[CrossRef](#)]
29. Jaynes, E.T. *Where Do We Stand on Maximum Entropy?* MIT Press: Cambridge, MA, USA, 1978.
30. Muth, J.F. Rational expectations and the theory of price movements. *Econometrica* **1961**, *29*, 315–335. [[CrossRef](#)]
31. Friedman, M. *Essays in Positive Economics*; University of Chicago Press: Chicago, IL, USA, 1953.
32. Lucas, R.E. Asset prices in an exchange economy. *Econometrica* **1978**, *46*, 1429–1446. [[CrossRef](#)]
33. Foley, D.K. Unfulfilled Expectations: One Economist’s History. *Guggenheim Prize Talk 2017*, Author’s unpublished notes.
34. Suppes, P.; Krantz, D.H.; Luce, R.D.; Tversky, A. *Foundations of Measurement Volume II: Geometrical, Threshold, and Probabilistic Representations*; Dover Publications: New York, NY, USA, 1989.
35. McFadden, D.L. Quantal Choice Analysis: A Survey. *Ann. Econ. Soc. Meas.* **1976**, *5*, 363–390.
36. Sims, C.A. Implications of Rational Inattention. *J. Monet. Econ.* **2003**, *50*, 665–690. [[CrossRef](#)]

37. Krainer, J. A Theory of Liquidity in Residential Real Estate Markets. *J. Urban Econ.* **2001**, *49*, 32–53. [[CrossRef](#)]
38. Case, K.; Shiller, R.J. Is there a bubble in the housing market? *Brook. Pap. Econ. Act.* **2003**, *2*, 299–362. [[CrossRef](#)]
39. Gyourko, J.; Mayer, C.; Sinai, T. Superstar Cities. *Am. Econ. J.* **2013**, *5*, 167–199.
40. Soofi, E.S.; Retzer, J.J. Information Indices: Unification and Applications. *J. Econ.* **2002**, *107*, 17–40. [[CrossRef](#)]



© 2018 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

Copyright of Entropy is the property of MDPI Publishing and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.