

# **Discussant remarks on Chihiro Shimizu, Kiyohiko G Nishimura and Tsutomu Watanabe's paper “House prices from magazines, realtors, and the Land Registry”**

Yongheng Deng<sup>1</sup>

## **Introduction**

This is a very important and timely paper addressing issues regarding appropriateness in measuring property asset prices in the housing market. In doing empirical research in the field of finance and economics, including the housing financial market, the researcher has to tackle the problem of model as well as the problem of data. For example, current US Federal Reserve Chairman Ben Bernanke commented that "... this originate-to-distribute model appears to have contributed to the breakdown in underwriting standards, as lenders often found themselves able to pass on the credit risk without much resistance from the ultimate investors" (Ben Bernanke, 14 March 2008 remarks on the mortgage crisis). Others believe that the models are fine, but that they have an input problem. It becomes a number that the researcher plucks out of the air (a comment made by a former Citigroup banker, Satyajit Das). Former Fed Chairman Alan Greenspan remarked on 23 October 2008 that "the whole intellectual edifice, however, collapsed in the summer of last year because the data inputted into the risk management models generally covered only the past two decades – a period of euphoria". An et al (2011) examine the model stability and data input issues using the subprime mortgage crisis as a natural experiment.

In this paper, Shimizu, Nishimura and Watanabe carefully address both modelling issues and data challenges in constructing the housing price index that can appropriately measure the dynamics of asset prices at various stages of the housing market. The research is carefully designed and the paper is well written and easy to read. The study contributes to the existing housing price literature through carefully examining the distribution of the empirical data from different timing of the price formation and marketing periods, as well as from various data sources. The study provides valuable insight into how to tackle challenges due to the data limitations, a challenge faced not only by housing economists, but also by many others who want to conduct rigorous empirical research in broadly defined finance and economics fields.

## **The challenges of housing price index modelling**

The literature on housing price research can be traced back to almost half a century ago, when Bailey et al (1963) proposed a weighted repeat sales house price index model that is estimated on the basis of paired sales of properties which have been sold more than once, so that the characteristics of the properties are standardised with reference only to themselves. Kain and Quigley (1970) propose a hedonic technique to account for the important non-temporal determinants of price variation. They illustrate the complexity of the

---

<sup>1</sup> Provost's Chair Professor and Director, Institute of Real Estate Studies, National University of Singapore. This is an edited version of comments delivered at the MAS-BIS Workshop on Property Markets and Financial Stability, 5 September 2011, Singapore.

bundle of residential services and the importance of residential quality as well as neighbourhood effects that affect housing price. Case and Shiller (1989) apply the repeat sales model to study the efficiency of the market for single-family homes. Englund et al (1999) and Shimizu et al (2010) compare various methodologies of computing housing price indexes with respect to temporal aggregation and sample definition. However, it has been well documented that existing housing price modelling approaches suffer from various limitations, for example, the strict assumption with regard to the constant quality sampling pair applied in the repeat sales approach, etc. Deng and Quigley (2008) document the magnitude and bias of the historical house price index due to the underlying data revisions, and analyse their systematic effects on the settlement prices in housing options markets.

On the data side, Deng et al (2011) report that the distribution of sale prices shifted much farther to the right at high prices than at lower prices for 1995–2010, and this pattern is particularly evident in the boom periods of 1996 and 2005–2007. The variance of the sale price distribution increased significantly during boom periods. These results imply that the variance of sale prices increased markedly during times of rapidly appreciating sale prices. More recently, China's statistics agency stopped publishing the country's much-watched official index of national property prices, owing to the underlying data accuracy being questioned (The Wall Street Journal Asia, 17 February 2011).

To handle the limited availability of housing price data, Shimizu, Nishimura and Watanabe in this paper cleverly compile a unique sample by merging four different data sets for house prices collected from different timing of the price formation and marketing periods and from various resources:

1. Initial asking price (listed in a residential information weekly magazine);
2. Final asking price (listed in a residential information weekly magazine);
3. Contract price (collected by realtors); and
4. Registry price (from the Land Registry and the Ministry of Land, Infrastructure, Transport and Tourism).

Shimizu, Nishimura and Watanabe carefully test the distributions of four different data sets, which are collected from different timing of the marketing periods. Given the timing gap between these marketing periods can be different by as much as 30 weeks in between, relying upon any single one of them (as most existing empirical research works would have done due to data unavailability) may lead to potential distributional truncation bias.

To address the modelling limitation discussed above, Shimizu, Nishimura and Watanabe adopt a quantile hedonic approach following McMillen (2008) to conduct the data quality adjustment. They find that after quality adjustment, the difference among distributions becomes smaller when two sets of prices come from closer stages. However, the p-values associated with the three tests suggest rejecting the null hypothesis that these prices are from a common distribution.

The findings reported above, however, seemingly contradict the statement in the abstract of the paper, where the authors state that “once quality differences are controlled for, there remain only small differences between the price distributions”.

## Final comments

The above conflict may be attributed to the fact that the quality controls used in the current study are limited to the spatial and hedonic characteristics of the building, i.e., floor space, age of building, distance to the nearest station, and travel time to the terminal station. There are other, potentially omitted quality variations missing in the current model. For example, there may have been some market variations during the 10–31 weeks' lag between the data

points from different pricing data sets; the heterogeneity in the seller's listing strategy may also lead to price distribution vitiation; furthermore, heterogeneity in negotiation between buyers and sellers in different market situations can also lead to distributional differential. A more careful specification with regard to quality adjustment, such as controlling for market characteristics and buyer and seller strategy/behaviour, is warranted here; at minimum, these can serve as robustness tests.

In addition, Shimizu, Nishimura and Watanabe can consider a revised model that follows Heckman's approach to developing a strategy to handle potential bias due to truncating distribution arising from market characteristics, the buyer and seller's heterogeneous behaviour, etc. They may also consider adopting a propensity score approach to controlling quality adjustment, and testing which quality adjustment tools/variables explain the price distributional variation.

To conclude: Shimizu, Nishimura and Watanabe have written a careful paper that improves our understanding of the housing price dynamics. I am certain this paper will play an important role in the housing price literature.

## References

- An, X., Deng, Y., Rosenblatt, Eric, and Yao, V., 2011. Model stability and the subprime mortgage crisis. *Journal of Real Estate Finance and Economics*, forthcoming.
- Bailey, M., Muth, R., and Nourse, H., 1963. A regression method for real estate price index construction. *Journal of the American Statistical Association* 58, 933–942.
- Case, K., and Shiller, R., 1989. The efficiency of the market for single-family homes. *The American Economic Review* 79, 125–137.
- Deng, Y. McMillen, D., and Sing T., 2011. Private residential price indices in Singapore: A matching approach. *Regional Science and Urban Economics*, forthcoming.
- Deng, Y., and Quigley, J. M.; 2008. Index revision, house price risk, and the market for house price derivatives. *Journal of Real Estate Finance and Economics* 39, 191–209.
- Englund, P., Quigley, J. M., and Redfearn, C. L., 1999. The choice of methodology of computing housing price indexes: comparisons of temporal aggregation and sample definition. *Journal of Real Estate Finance and Economics* 19, 91–112.
- Kain, J., and Quigley, J. M., 1970. Measuring the value of housing quality. *Journal of the American Statistical Association* 65, 532–548.
- McMillen, D., 2008. Changes in the distribution of house prices over time: structural characteristics, neighborhood or coefficients? *Journal of Urban Economic* 64, 573–589.
- Shimizu, C., Nishimura, K. G., and Watanabe, T. 2010. Housing prices in Tokyo: A comparison of hedonic and repeat-sales measures. *Journal of Economics and Statistics* 230, 792–813.