House Price Indexes and the Global Financial Crisis

Presentation at the Financial Crisis and Property Price Indexes Seminar at Reitaku University,

October 18, 2013

Erwin Diewert

University of British Columbia and

University of New South Wales

and

Chihiro Shimizu

University of British Columbia and

Reitaku University

Introduction

- **The European statistical agency, Eurostat, has recently published a Residential Property Price Index (RPPI) Handbook.**
- **This book describes some of the problems associated with constructing price indexes for residential house prices and gives advice on methods that could be used in order to construct house price indexes.**
- **Why is this important? The Global Financial Crisis has several causes but a main cause was a housing bubble in the U.S. which led banks to make mortgage loans that were based on the assumption that house prices were not rising unusually quickly.**
- **If accurate regional house price indexes for the U.S. would have been widely available to the public, it is unlikely that so many bad housing loans would have been made and perhaps the Global Financial Crisis could have been averted.**

Why Have National Statistical Agencies not Published Official Property Price Indexes?

- **It is very difficult to construct accurate house price indexes (we will explain why this is the case below) and so statistical agencies have been reluctant to allocate their scarce resources to the construction of indexes where there has not been international agreement on how exactly to construct such an index;**
- **House prices by themselves do not occupy an important position in the major statistics that countries construct; i.e., house price indexes do not appear directly in either the Consumer Price Index or in the main components of GDP.**

But House Price Indexes do play a role in the SNA:

- **In the Balance Sheet Accounts as part of constructing real wealth measures and**
- **As part of the treatment of Owner Occupied Housing.**

OOH and House Price Indexes

- **Current CPI practise either ignores owner occupied housing or prices it according to what the rental value of the house is.**
- **The first alternative is not very satisfactory and while the second alternative understates the true opportunity cost of owning a very expensive house and using its services.**
- **The true opportunity cost of using the services of an owned house is the *maximum* of what the owner could rent the house for and the financial user cost of the capital that is tied up in the house.**
- **For less expensive houses, the rent is approximately equal to the user cost in most countries but for expensive houses, the rent is only about one half of the corresponding user cost.**
- **Thus in order to calculate the true opportunity cost of using the services of an owned house, we need to construct accurate house price indexes in order to construct user costs and opportunity costs.**

Private Sector Companies do Construct House Price Indexes

- **The problem with these privately provided indexes is that they use different methodologies of varying quality and they are typically not comparable within and across countries.**
- **This heterogeneity explains why Eurostat funded the production of a Handbook that outlines various measurement methodologies and provides some guidance to countries who will probably produce official house price indexes in the near future.**
- **Japan is one such country which will produce an official Residential Property Price Index in the future.**

Why is it so difficult to construct house price indexes?

- **How are price indexes are formed for "normal" commodities like say a can of soup?**
- **The basic "matched model" methodology collects prices on the same item (sold in the same location) at two different points in time and then forms a relative price with these matched prices.**
- **These price relatives are then aggregated over say several varieties of soup and this possibly weighted average of the soup price ratios is then a soup price index.**
- **The problem with constructing a price index for house sales in a given area is that the houses that transact in two different periods are generally *different* and thus the usual "matched model" methodology that is a basic building block in the construction of a price index cannot be applied.**

The Repeat Sales Method

- **This method looks only at houses that sold in both of the periods under consideration and so at first glance, it appears that this method solves the lack of matching problem.**
- **Unfortunately, the same house at different points in time is not really the same house: in the time between the two sales, the house has depreciated (and hence is less valuable) and it may have undergone some renovations.**
- **Thus the repeat sales method generally has a (small) downward bias that will grow as the time between sales grows.**
- **Another problem is that data on repeat sales may be very sparse and hence lead to index volatility.**
- **Finally, there are problems with very quick repeat sales (the property flipping problem) and with the fact that the historical index values change as new data become available.**

Methods that Rely on Housing Characteristics Information

The first step is to determine what are the important price determining characteristics of a house. The main characteristics are:

- **The age of the house;**
- **The size of the land area of the house in square meters;**
- **The size of the structure in square meters of floor space;**
- **The type of structure (detached house, apartment, wood construction) and**
- **The location of the house (which determines its access to amenities)**

Two Methods Relying on Characteristics Information

The two methods are:

- **Stratification methods. House price sales are grouped into cells where the main characteristics of the houses in the cell are approximately similar. This method usually fails due to sparse cell data.**
- **Hedonic regression methods. Here the selling price of a house is set equal to a function of the characteristics listed on the previous slide (and other characteristics as well) plus an error term.**

The RPPI Handbook explains how the above methods work and provides some empirical illustrations of the recommended methods using sales data for the small town of "A" in the Netherlands.

The Problem of Decomposing Selling Prices into Land and Structure Components

- **For some purposes (in the Balance Sheet accounts of a country and in constructing Multifactor Productivity Accounts for a country), it is necessary to have a breakdown of property sales into structure and land components.**
- **This turns out to be a very difficult problem but the RPPI Handbook worked out a useful methodology and applied it to the town of "A".**
- **Diewert and Shimizu have also applied this methodology to Tokyo residential house sales starting at Quarter 1 of 2000 and ending in Quarter 4 of 2010.**
- **I will conclude my talk by describing our methodology and the results of our hedonic regressions for Tokyo.**

The Tokyo Data

- **Our data on sales of residential houses in Tokyo was not dense enough for us to implement our basic model for individual neighbourhoods (or Wards) of Tokyo so we used Ward dummy variables to take into account neighbourhood effects on the price of land.**
- **In the *RPPI Handbook*, it was found that information on the sales price of a house in the town of "A" along with information on the lot area, the structure area and the age of the structure was sufficient to explain about 85% of the variation in house prices. (These are the basic characteristics).**
- **In our work with the Tokyo data, the basic *RPPI* model was extended to include other house characteristics. These extra characteristics were added as spline variables in order to achieve a better description of the data.**

The Tokyo Data: the Characteristics Used

- **V = The value of the sale of the house in 10,000,000 Yen;**
- **S = Structure area (floor space area) in units of 100 meters squared;**
- **L = Lot area in units of 100 meters squared;**
- **A = Approximate age of the structure in years;**
- **NB = Number of bedrooms;**
- **WI = Width of the lot in meters;**
- **TW = Walking time in minutes to the nearest subway station;**
- **TT = Subway running time in minutes to the Tokyo station from the nearest station during the day (not early morning or night).**

The Data

- **There were a total of 5578 observations (after range deletions) in our sample of sales of single family houses in the Tokyo area over the 44 quarters covering 2000-2010.**

Table 1: Descriptive Statistics for the Variables

Name	No. of Obs.	Mean	Std. Dev	Minimum	Maximum
V	5578	6.2310	2.95420	2.0500	20
S	5578	1.0961	0.36255	0.5012	2.4789
L	5578	1.0283	0.42538	0.5001	2.4977
A	5578	14.689	8.91460	2.0140	49.7230
NB	5578	3.9518	1.04090	2	8
WI	5578	4.6987	1.26090	2.5	9
TW	5578	9.9295	4.48510	2	29
TT	5578	31.677	7.55220	4	48

The Data (cont)

- **We deleted 9.2 per cent of the observations because they fell outside our range limits for the variables V, L, S, A, NB and W.**
- **It is risky to estimate hedonic regression models over wide ranges when observations are sparse at the beginning and end of the range of each variable.**
- **The a priori range limits for these variables were as follows: $2 \le V \le 20$; $0.5 \le S \le 2.5$; $0.5 \le V \le 2.5$; $1 \le A \le 50$; ; $2 \le NB \le 8$; $2.5 \le W \le 9$.**
- **In order to eliminate the multicollinearity problem between the lot size L and floor space area S for an individual house, we assumed that the value of a new structure in any quarter was proportional to a Construction Cost Price Index for Tokyo.**

The Data (cont)

- **In addition to having the information listed in Table 1 on residential houses sold in Tokyo over 2000-2010, we also had the address for each transaction.**
- **We used this information in order to allocate each sale into one of 21 Wards for the Tokyo area.**
- **We constructed Ward dummy variables and made use of these variables in most of our regressions as locational explanatory variables.**

The Basic Builder's Model

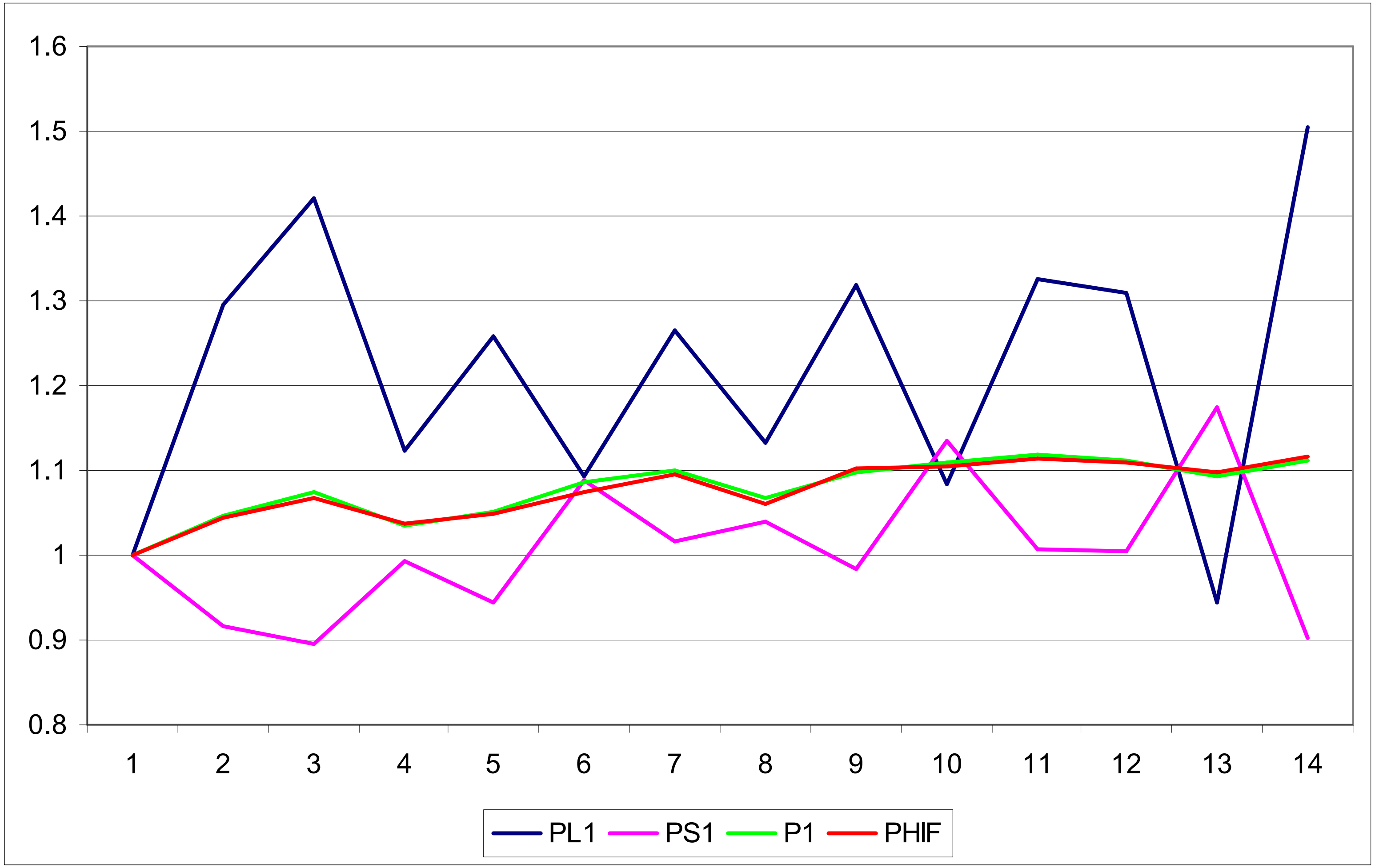
Consider the following hedonic regression model for quarter t:

(2) $V_{tn} = \alpha_t L_{tn} + \beta_t(1 - \delta_t A_{tn})S_{tn} + \varepsilon_{tn}$; $t = 1,\ldots,44$; $n = 1,\ldots,N(t)$

where the parameter δ_t reflects the *net depreciation rate* as the structure ages one additional year, the parameter α_t is the price of land in Tokyo in quarter t and β_t is the price of a new structure in period t. Note that we are assuming straight line depreciation here. To eliminate the multicollinearity problem, we combined all of the 44 quarterly regressions into a single regression and also used our exogenous price of new construction, p_{Ct} (Note that the β_t have been replaced by βp_{Ct}):

(3) $V_{tn} = \alpha_t L_{tn} + \beta p_{Ct}(1 - \delta A_{tn})S_{tn} + \varepsilon_{tn}$; $t = 1,\ldots,44$; $n = 1,\ldots,N(t)$

The Multicollinearity Problem Illustrated: Figure 9.1 from the RPPI Handbook: The Price of Land P_{L1}, the Price of Structures P_{S1}, the Overall Fisher House Price Index P_1 and the Fisher Hedonic Imputation House Price Index P_{HIF}



The Basic Builder's Model: Results

- **For the model defined by equations (3), we have 5578 degrees of freedom to estimate 44 land price parameters α_t, one structure price parameter β that determines the level of prices over our sample period and one annual straight line depreciation rate parameter δ, a total of 46 parameters.**
- **The R^2 for the resulting nonlinear regression model was only 0.5704.**
- **Thus the simple Builder's Model defined by (3) was not as satisfactory as was the corresponding Builder's Model for the small town of "A" in the Netherlands where the R^2 was 0.8703 using the same information on characteristics of the house and lot.**
- **In the case of the town of "A", the structures were all much the same and all houses in the town had access to basically the same amenities. The situation in the huge city of Tokyo is very different: different neighborhoods have access to very different amenities and so we would expect substantial variations in the price of land across the various neighborhoods.**

The Basic Builder's Model with Ward Dummy Variables

- **In order to take into account possible neighbourhood effects on the price of land, we introduced *ward dummy variables*, $D_{W,tn,j}$, into the hedonic regression (3).**
- **We now modify the model defined by (3) to allow the *level* of land prices to differ across the 21 Wards of Tokyo:**

$$(5)\ V_{tn} = \alpha_t(\textstyle\sum_{j=1}^{21} \omega_j D_{W,tn,j})L_{tn} + \beta p_{Ct}(1 - \delta A_{tn})S_{tn} + \varepsilon_{tn}\ ; \qquad t = 1,\ldots,44;\ n = 1,\ldots,N(t).$$

- **It can be seen that we have added an additional 21 *ward relative land value parameters*, $\omega_1,\ldots,\omega_{21}$, to the model defined by (3). We call the model defined by (5) and (6) Model 1.**
- **However, not all parameters can be identified so we impose the following normalization on the parameters:**

$$(6)\ \omega_{10} \equiv 1.$$

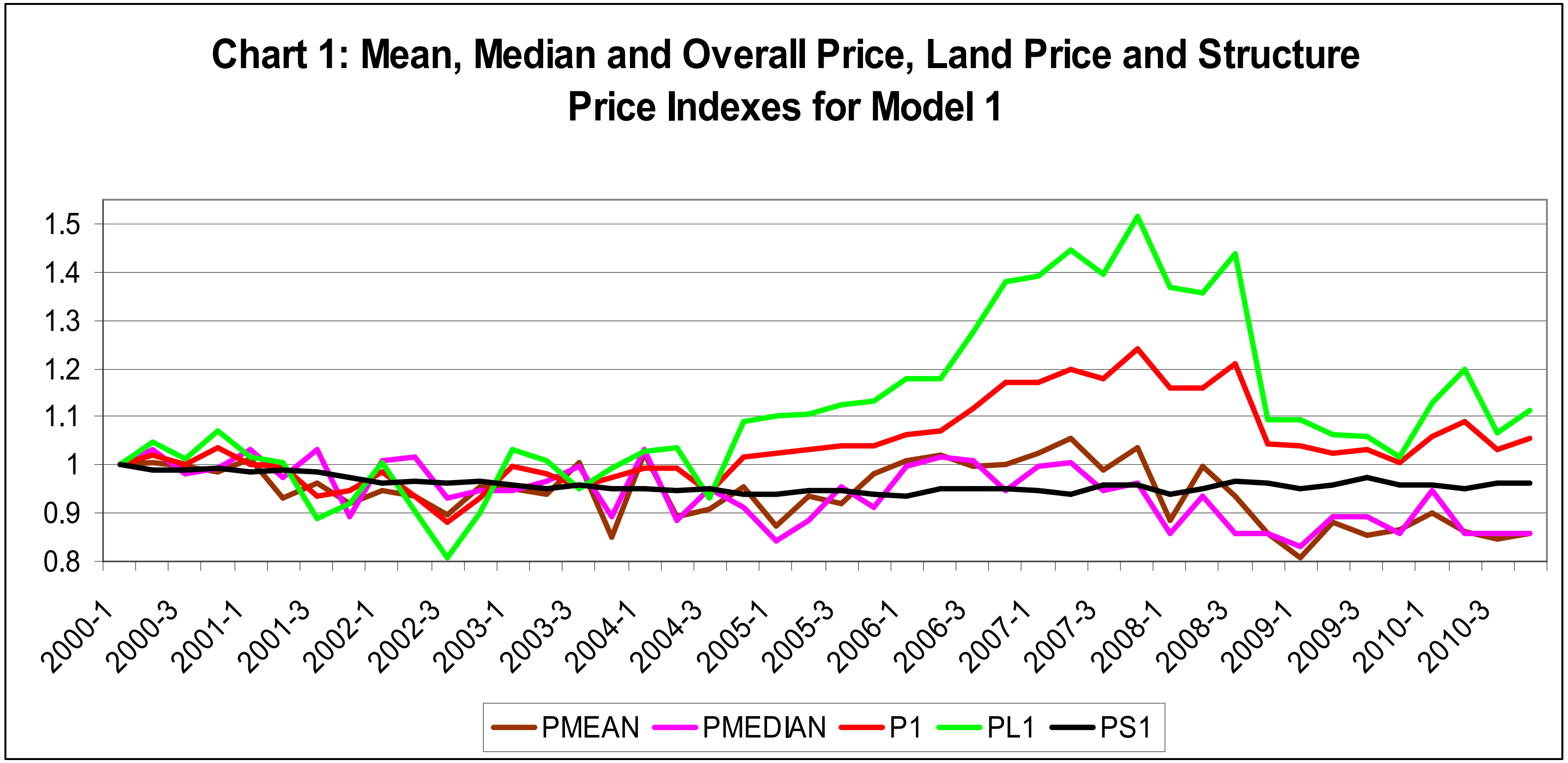
- **The tenth ward, Setagay, has the most transactions in our sample (1158 transactions over the sample period).**

The Basic Builder's Model with Ward Dummy Variables: Results

- **The R^2 for this model turned out to be 0.8168 and the log likelihood (LL) was –9233.0, a huge increase of 2270.6 over the LL of the model defined by (3).**
- **Thus the Ward variables are very significant determinants of Tokyo house prices.**
- **Note that we used only four characteristics for each house sale: the land area L, the structure area S, the age of the structure A and its Ward location.**
- **However, we also required an exogenous house construction price index to implement our method.**
- **We will omit the details on how the results of the hedonic regression (5) were used to construct separate land and structure price indexes. Chained Fisher indexes were used to aggregate the land and structure components into an overall house price index. The following Chart plots these indexes.**

The Basic Builder's Model with Ward Dummy Variables: Results

- **The overall Model 1 house price index P_{1t} as well as the land and structure price indexes P_{L1t} and P_{S1t} for Tokyo are graphed in Chart 1.**
- **We have also computed the quarterly mean and median house prices transacted in each quarter.**



Discussion of Model 1 Results

- The mean and median price indexes were plotted on the previous chart. (Brown and purple lines).
- Our preferred hedonic index is the red line and it lies well above the mean and median price lines.
- Our overall index (the red line) is a weighted average of the structure price index (the black line which is equal to a constant times the official residential house price construction index for Tokyo) and the green line which is our estimated residential land price index for Tokyo.
- In our recent paper, "Residential Property Price Indexes for Tokyo", we estimate 4 more hedonic regression models, adding additional characteristics to the previous model.
- Our final model is Model 5 where we estimate separate land prices for expensive and less expensive wards. The results for Model 5 are shown on the following slide.

Allowing for Land and Structure Price Differences Across Wards

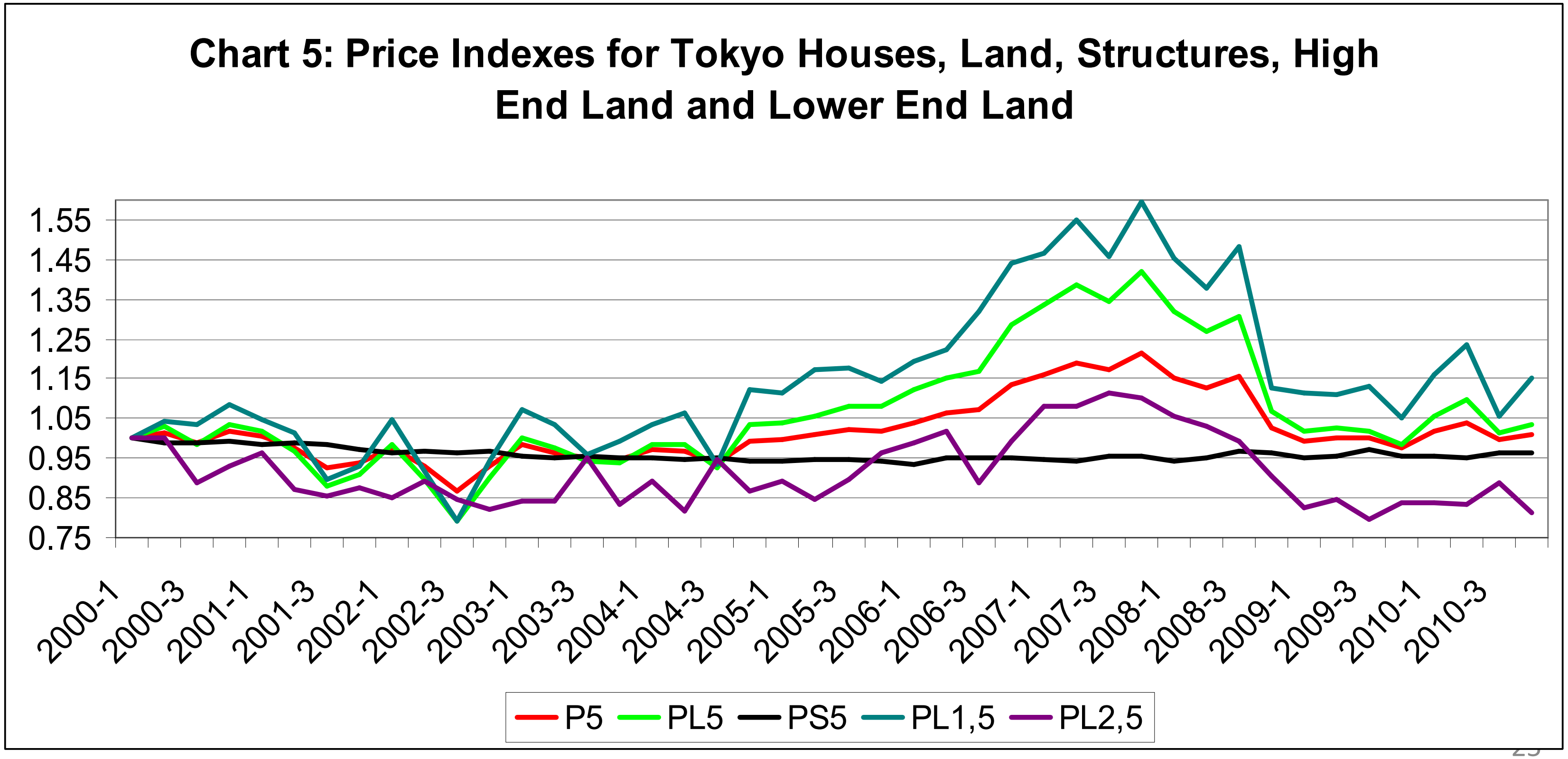
- **Usually, land price movements in high end properties are more volatile than in lower end properties.**
- **It would be preferable to have separate land price parameters (the α_t) for each Ward. However, we do not have enough degrees of freedom to accurately measure land price movements ward by ward. We do have a sufficient number of observations so that we can divide Wards into two groups based on the estimated ω_j parameters from Model 4: Group 1 Wards are those whose estimated relative land price levels ω_j exceeded 0.75 and Group 2 Wards are those whose estimated land price levels ω_j were less than 0.75.**
- **The following Wards were in Group 1 (the *expensive or high end Wards*): 1-4, 7-11, 13-14. The following Wards were in Group 2 (the *cheaper or lower end Wards*): 5, 12, 15-21. We will allow land prices to evolve over time in a completely independent manner for high and lower end Wards.**

Allowing for Land and Structure Price Differences Across Wards (cont)

- **We also allow for separate lot size quality adjustments in the high and lower end wards.**
- **Finally, we now allow the level of structure prices to differ in high and lower end wards so that the previous structure price level parameter β is now replaced by β_1 (the level of structure prices in high end wards) and β_2 (the level of structure prices in lower end wards). Our expectation is that β_2 will be less than β_1 since we would expect the quality of construction to be higher in the high end wards.**
- **Our final nonlinear hedonic regression model (Model 5) is defined by equations (38) in the paper along with the normalizations in equations (39).**
- **There are 128 unknown parameters to be estimated in Model 5.**

Allowing for Land and Structure Price Differences Across Wards: Results

- **The R^2 for Model 5 was 0.8476 and the log likelihood was −8709.9, an increase of 106.0 over the Model 4 log likelihood.**
- **The resulting overall house price index P5, the overall land price PL5, the land price indexes for high and low end Wards, PL1 and PL2 respectively are plotted on Chart 5.**

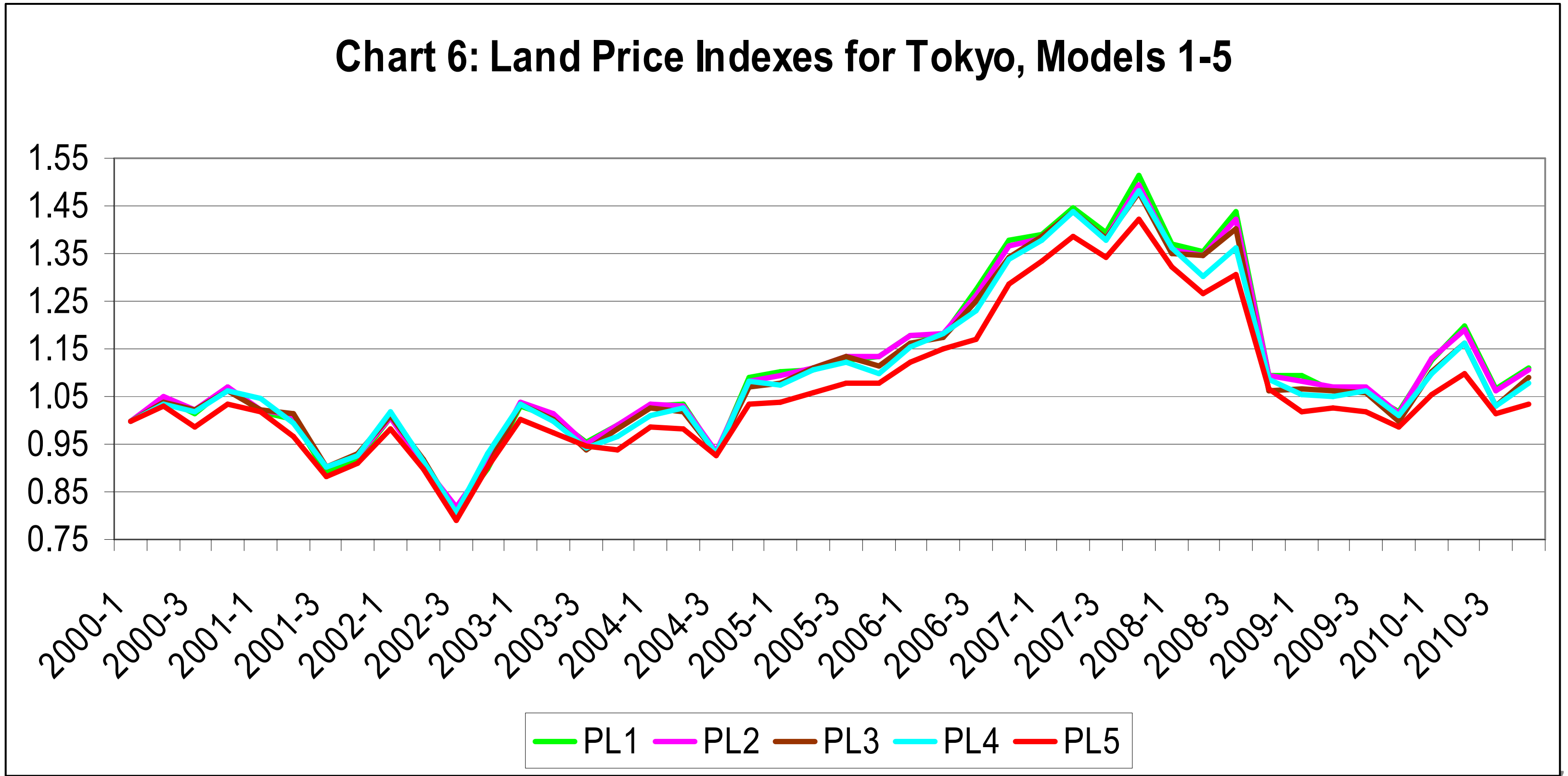


Allowing for Land and Structure Price Differences Across Wards: Results (cont)

- **As expected, the pattern of land price movements is very different in the high and low end wards.**
- **Price movements have generally been higher and more volatile in the more expensive wards; i.e., $P_{L1,5t}$ generally lies above $P_{L2,5t}$ and $P_{L1,5t}$ has a higher variance than $P_{L2,5t}$.**
- **However, it can also be seen from viewing Chart 5 that the overall land price index for Model 5, P_{L5t}, is not that different from the land price indexes from Model 1.**
- **Thus information on the main important characteristics can get us pretty close to the "truth".**
- **We compare the Model 1 to Model 5 overall land price indexes in Chart 6 on the following slide.**

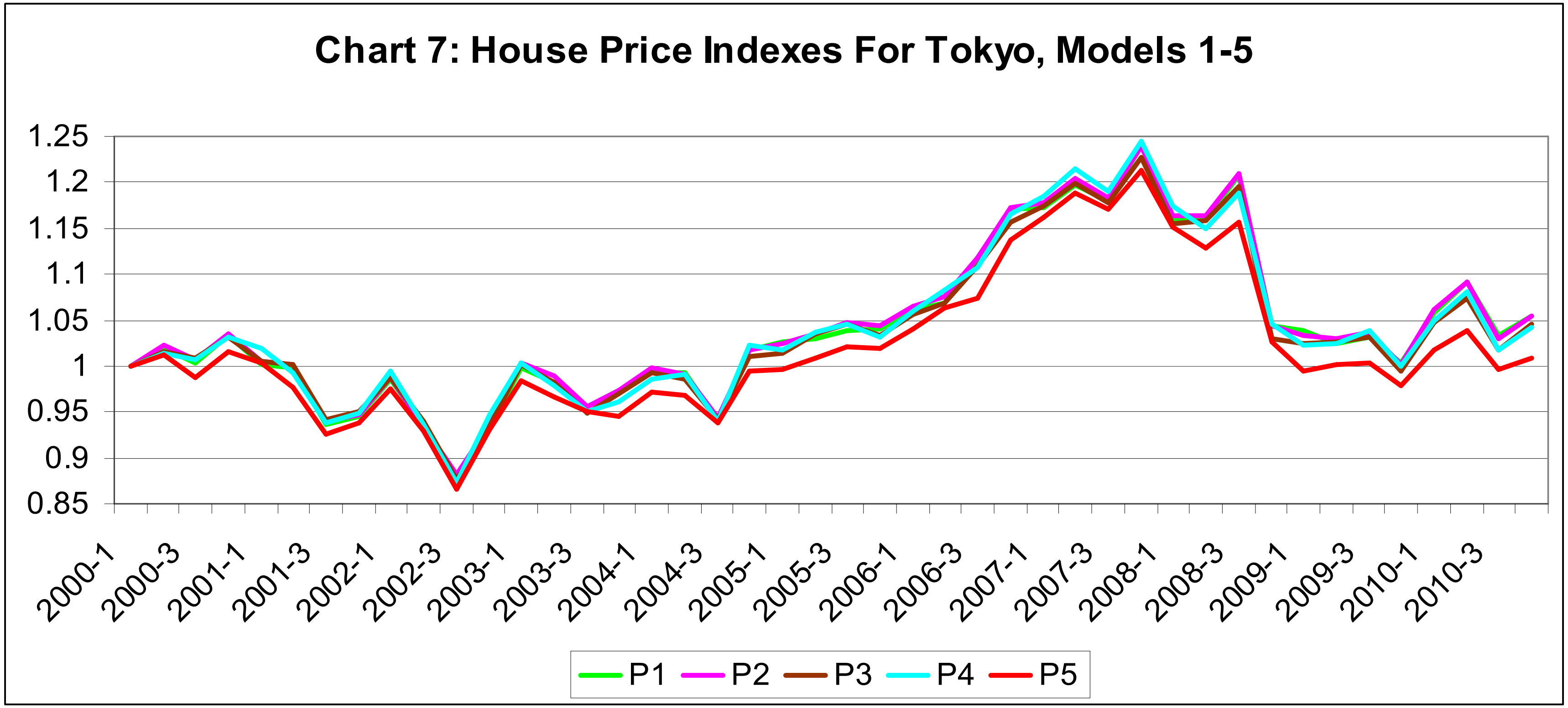
Comparison of Overall Land Price Indexes, Models 1-5

- **It can be seen that the overall land price series for Models 1-4, P_{L1t}-P_{L4t}, are generally quite close .The overall land price series for Model 5 drops a more substantial amount: about 3% on average from the other series.**



Comparison of Overall House Price Indexes, Models 1-5

- **It is also useful to compare the overall house price indexes for Models 1-5 and this is done in Chart 7 below.**
- **The Model 5 overall house price index P_{5t} is about 2% lower on average compared to the levels in the other Models.**



Conclusion

- **The Median house price index, which is widely used as a house price index, has a downward bias relative to our preferred hedonic index P5.**
- **This downward bias of about 16% over our sample period is due to the fact that the Median index cannot take into account the effects of structure depreciation and other changes in the characteristics of the houses that are sold.**
- **Over the two years, 2006-2007, overall land prices increased 27.4%, a mini bubble. Over the same two years, land prices increased about 22.9% for the less expensive wards and 27.2% for the more expensive wards.**
- **It can be seen from Chart 5 that the fluctuations in land prices are larger in the rich wards and over the entire 11 year period, land prices increased 17% in the richer wards while they decreased 25% in the poorer wards.**
- **Our conclusion is that the hedonic method works well.**