Diewert and Shimizu: “Hedonic Regression Models for Tokyo Condominium Sales” – Discussion

Dr Jens Mehrhoff*, Head of Section Business Cycle, Price and Property Market Statistics

* This presentation represents the author’s personal opinions and does not necessarily reflect the views of the Deutsche Bundesbank or its staff.
Structure of the presentation

1. Introduction
2. The Tokyo Condominium Data
3. The Basic Builder’s Model
4. – 10. Explaining Variations in the Imputed Land Values
11. Using the Selling Price as the Dependent Variable
12. Adjusting the Quality of the Structure
14. Constructing Land, Structure and Overall Property Price Indexes
15. Comparison with other Condo Price Indexes
16. Conclusion

“A condominium is a form of multifurcated real property tenure. Specified units of the property are separately owned, and the remainder of the property is collectively owned. Unlike apartments, which are leased by their tenants, condominium units are owned outright.” Source: Wikipedia, The Free Encyclopedia.
1. Introduction

- It is important to not only have estimates for the value of the housing stock but to decompose the overall value into (additive) land and structure components and then to further decompose these value aggregates into constant quality price and quantity components.

- The selling price values the sum of the structure and land components: 
  \[ \text{Property value} = \text{Structure value} + \text{Land value}. \]

- The problem of obtaining constant quality price components is further complicated by the fact that housing units are almost always unique assets: 
  \[ \text{Value change} = \text{Price change} + \text{Quality change}. \]
2. The Tokyo Condominium Data

- **Sales of condominium units**
  - located in nine wards in the **central area of Tokyo**
  - from the first quarter of 2000 to the first quarter of 2015 (**61 quarters**).

- In addition to the sales prices, **various characteristics of the properties**. (NB: “Sales price” here means listed price in the final week before a successful deal.)

- A total of **3,232 observations** (after range deletions to account for sparse tails of the variables and outliers in the data).
### Definitions of the variables and their units of measurement

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
<th>Unit of Measurement</th>
</tr>
</thead>
<tbody>
<tr>
<td>V</td>
<td>Value of the condo</td>
<td>in 10,000 yen</td>
</tr>
<tr>
<td>S</td>
<td>Floor space of the condo</td>
<td>in square metres</td>
</tr>
<tr>
<td>TS</td>
<td>Total floor space of the entire building</td>
<td>in square metres</td>
</tr>
<tr>
<td>TL</td>
<td>Total lot area for the entire structure</td>
<td>in square metres</td>
</tr>
<tr>
<td>A</td>
<td>Age of the structure</td>
<td>in years</td>
</tr>
<tr>
<td>H</td>
<td>Number of the story of the condo</td>
<td></td>
</tr>
<tr>
<td>TH</td>
<td>Total number of stories in the entire building</td>
<td></td>
</tr>
<tr>
<td>NB</td>
<td>Number of bedrooms in the condo</td>
<td></td>
</tr>
<tr>
<td>TW</td>
<td>Walking time to the nearest subway station</td>
<td>in minutes</td>
</tr>
<tr>
<td>TT</td>
<td>Subway running time to Tokyo station</td>
<td>in minutes</td>
</tr>
<tr>
<td>SCR</td>
<td>Reinforced concrete construction dummy variable</td>
<td>1 if reinforced</td>
</tr>
<tr>
<td>South</td>
<td>South facing dummy variable, 1 if condo faces</td>
<td>1 if condo faces south</td>
</tr>
</tbody>
</table>
3. The Basic Builder’s Model

Consider a property developer who builds a structure on a particular property: The total cost of the property after the structure is completed will be equal to

- the floor space area of the structure, say $S$ square meters, times the building cost per square metre, $\beta_t$ during quarter $t$,

- plus the cost of the land, which will be equal to the cost per square metre, $\alpha_t$ during quarter $t$, times the area of the land site, $L$.

This leads to the following hedonic regression model for period $t$ where the $\alpha_t$ and $\beta_t$ are the parameters to be estimated in the regression:

1. $V_{tn} = \alpha_t L_{tn} + \beta_t S_{tn} + \varepsilon_{tn}, \ t = 1, \ldots, 61; \ n = 1, \ldots, N_t,$

- where the error terms $\varepsilon_{tn}$ are assumed to be independently normally distributed with zero means and constant variances.
3. The Basic Builder’s Model

- Assuming a geometric depreciation model, a more realistic hedonic regression model than that defined by (1) above is the following basic builder’s model:

\[
V_{tn} = \alpha_t L_{tn} + \beta_t (1 - \delta_t)^{A_{tn}} S_{tn} + \epsilon_{tn},
\]

- where the parameter \( \delta_t \) reflects the net geometric depreciation rate as the structure ages one additional period, i.e. it is equal to a “true” gross structure depreciation rate less an average renovations appreciation rate.

- In order to deal with the multicollinearity problem, Diewert and Shimizu draw on exogenous information on the cost of building new condominium units from the Japanese Ministry of Land, Infrastructure, Transport and Tourism (MLIT) and they assume that the price of new structures is proportional to an official index of condominium building costs, \( p_{St} \). Thus, they replace \( \beta_t \) in (2) by \( \beta p_{St} \).
3. The Basic Builder’s Model

− Each condo in the building should be allocated a share of the total land value of the property. There are two simple methods for constructing an appropriate land share:

• (i) Use the unit’s share of floor space to total structure floor space, or

• (ii) simply use the share of the total number of units in the building.

− Thus, define the following two land imputations for unit $n$ in period $t$:

• (3a) $L_{Stn} = \left( S_{tn} / TS_{tn} \right) TL_{tn}$;

• (3b) $L_{Ntn} = \left( 1 / N_{tn} \right) TL_{tn}$;

• where $N_{tn}$ is the total number of units in the building for unit $n$ sold in period $t$. 
3. The Basic Builder’s Model

- In order to allocate land across the $N_{tn}$ units in a building, the unit shares should add up to one. However, the shares $S_{tn} / TS_{tn}$, if available for every unit in the building, would add up to a number less than one because the unit floor space areas, $S_{tn}$, if summed over all units in the building add up to privately owned floor space which is less than total building floor space $TS_{tn}$.

- An imperfect estimate of the ratio of privately owned floor space to total floor space for unit $n$ in period $t$ is $N_{tn} S_{tn} / TS_{tn}$. The sample wide average of these ratios was 0.899. Thus, to account for shared structure space, Diewert and Shimizu replaced the owned floor space variable in (2), $S_{tn}$, by $(1 / 0.899) S_{tn} \approx 1.1 S_{tn}$.

- Diewert and Shimizu assumed that the depreciation rate $\delta_t$ was equal to 0.03:

  - (4) $V_{tn} = \alpha_{St} L_{Stn} + \beta_{S} p_{St} (1 - 0.03)^{A_{tn}} (1.1 S_{tn}) + \varepsilon_{Stn}$
  - (5) $V_{tn} = \alpha_{Nt} L_{Ntn} + \beta_{N} p_{St} (1 - 0.03)^{A_{tn}} (1.1 S_{tn}) + \varepsilon_{Ntn}$
3. The Basic Builder’s Model

− The estimates for $\beta > 2$ were totally unsatisfactory because these parameters should have been close to unity.

− Moreover, the land price indexes that these regression models generated were subject to excessive volatility.

− In order to deal with the problem of too high estimates of $\beta$, Diewert and Shimizu decided to set it equal to unity rather than to estimate it.

− Moreover, they temporarily put aside the problem of jointly determining land and structure value to concentrate on determining sensible constant quality land prices.
3. The Basic Builder’s Model

− Diewert and Shimizu assume that the *structure value* for unit \( n \) in period \( t \), \( V_{Stn} \), is defined as follows:

\[
(6) \quad V_{Stn} = p_{St} (1 - 0.03)^{Atn} (1.1 S_{tn}).
\]

− Once the imputed value of the structure has been defined by (6), they define the *imputed land value* for condo \( n \) in period \( t \), \( V_{Ltn} \), by subtracting the imputed structure value from the total value of the condo unit:

\[
(7) \quad V_{Ltn} = V_{tn} - V_{Stn}.
\]

− They use \( V_{Ltn} \) as the dependent variable and *attempt to explain variations in these imputed land values* in terms of the property characteristics.
4. – 10. Explaining Variations in the Imputed Land Values

4. Introduction of Ward Dummy Variables

5. Introduction of Building Height as an Explanatory Variable

6. Introduction of the Height of the Unit as an Explanatory Variable

7. Introduction of a More General Method of Land Imputation

8. Introduction of the Number of Units in the Building as an Explanatory Variable

9. Introduction of Excess Land as an Explanatory Variable

10. Introduction of Subway Travel Times and Facing South as Explanatory Variables
### 11. Using the Selling Price as the Dependent Variable

\[ V_{tn} = \]

\begin{align*}
\alpha_t \\
\times (\sum_j \omega_j D_{W,tn,j}) & \quad \text{Ward dummy variables} \\
\times (\sum_h \chi_h D_{TH,tn,h}) & \quad \text{Building height dummy variables} \\
\times (\sum_m \mu_m D_{EL,tn,m}) & \quad \text{Excess land dummy variables} \\
\times (\phi_1 D_{s,tn,1} + \phi_2 D_{s,tn,2}) & \quad \text{Facing south dummy variables} \\
\times [1 + \gamma (H_{tn} - 3)] & \quad \text{Height of the unit} \\
\times [1 + \kappa (N_{tn} - 11)] & \quad \text{Total number of units in the building} \\
\times [1 + \eta (TW_{tn} - 1)] & \quad \text{Walking time to nearest subway station} \\
\times [1 + \theta (TT_{tn} - 12)] & \quad \text{Subway running time to Tokyo station} \\
\times L_{tn}(\lambda) & \quad \text{More general method of land imputation} \\
+ p_{St} (1 - \delta)^{Atn} (1.1 S_{tn}) & \quad \text{Structure component of the value} \\
+ \varepsilon_{tn} & \quad \text{} 
\end{align*}
12. Adjusting the Quality of the Structure

- Diewert and Shimizu introduce the **number of bedrooms** variable, $NB_{tn}$, and the **reinforced concrete construction** dummy variable, $SCR_{tn}$, as quality adjusters for the value of the structure:

\[
V_{tn} = \alpha_t \times (\sum_j \omega_j D_{W,tn,j}) \times (\sum_h \chi_h D_{TH,tn,h}) \times (\sum_m \mu_m D_{EL,tn,m})
\times (\phi_1 D_{s,tn,1} + \phi_2 D_{s,tn,2}) \times [1 + \gamma (H_{tn} - 3)] \times [1 + \kappa (N_{tn} - 11)]
\times [1 + \eta (TW_{tn} - 1)] \times [1 + \theta (TT_{tn} - 12)] \times L_{tn}(\lambda)
+ \rho_{St} \times (1 - \delta)^{A_{tn}} \times (1 + \sigma SCR_{tn}) \times (\sum_i \rho_i D_{B,tn,i}) \times (1.1 S_{tn}) + \varepsilon_{tn}.
\]

- A complete listing of the **estimated 101 coefficients** and their standard errors can be found in **Table 3 of the paper**.

- It is likely that the **model developed in section 13** (separate land price parameters for poor, medium and rich wards) is **not reliable**.
14. Constructing Land, Structure and Overall Property Price Indexes

- **Predicted land and structure values** for observation $tn$, $V_{Lt,tn}^*$ and $V_{St,tn}^*$, can be defined as follows:

  - (30) $V_{Lt,tn}^* = \alpha_t^* \times (\sum_j \omega_j^* D_{W,tn,j}) \times (\sum_h \chi_h^* D_{TH,tn,h}) \times (\sum_m \mu_m^* D_{EL,tn,m}) \times (\phi_1^* D_{s,tn,1} + \phi_2^* D_{s,tn,2}) \times [1 + \gamma^*(H_{tn} - 3)] \times [1 + \kappa^*(N_{tn} - 11)] \times [1 + \eta^*(TW_{tn} - 1)] \times [1 + \theta^*(TT_{tn} - 12)] \times L_{tn}(\lambda^*)$;

  - (31) $V_{St,tn}^* = p_{St} \times (1 - \delta^*)^{A_{tn}} \times (1 + \sigma^* SCR_{tn}) \times (\sum_i \rho_i^* D_{B,tn,i}) \times (1.1 S_{tn})$.

- It can be seen that each term on the right hand side of (30) has $\alpha_t^*$ as a common factor. Thus, Diewert and Shmizu define $\alpha_t^*$ as the **price index for land** in period $t$, $p_{Lt}$.

- Each term on the right hand side of (31) has $p_{St}$ as a common factor. Thus, it is natural to define this common factor $p_{St}$ as the **price index for structures** in period $t$. 
14. Constructing Land, Structure and Overall Property Price Indexes

![Graph showing land and structure price indices for Tokyo condominium sales.](chart)

Author's calculations.
Deutsche Bundesbank

Jens Mehrhoff, Deutsche Bundesbank, Directorate General Statistics
34th IARIW General Conference
Dresden, 21-27 August 2016
Page 16
15. Comparison with other Condo Price Indexes

- An overall constant quality property price index $p_t$ is formed as a chained Fisher (hedonic imputation) index of the land and structure components.

- Diewert and Shimizu also plot a traditional time dummy hedonic price index, $p_{TDt}$.

- In order to construct this index they run a linear regression:

\[
\begin{align*}
\ln V_{tn} &= c + (\Sigma \pi_t D_{T,tn,\tau}) + (\Sigma \omega_j D_{W,tn,j}) + \beta \ln S_{tn} + \alpha \ln L_{tn} \\
&\quad + \delta A_{tn} + \gamma H_{tn} + \rho \ln NB_{tn} + \eta TW_{tn} + \theta TT_{tn} \\
&\quad + \sigma SCR_{tn} + \phi \ln South_{tn} + \varepsilon_{tn} \\
&= 4.4850 + (\Sigma \pi_t D_{T,tn,\tau}) + (\Sigma \omega_j D_{W,tn,j}) + 0.9158 \ln S_{tn} + 0.0203 \ln L_{tn} \\
&\quad - 0.0168 A_{tn} + 0.0076 H_{tn} + 0.0034 \ln NB_{tn} - 0.0135 TW_{tn} - 0.0094 TT_{tn} \\
&\quad - 0.0027 SCR_{tn} + 0.0195 \ln South_{tn} + \varepsilon_{tn}.*
\end{align*}
\]

- $p_{TDt} = \exp \pi_t^*$. 
15. Comparison with other Condo Price Indexes

Hedonic regression price indices for Tokyo condominium sales

2000 Q1 = 100, log scale

Time dummy price index
Imputation Fisher price index

Author's calculations.
Deutsche Bundesbank
16. Conclusion

− The main focus of the paper has been to suggest a method for decomposing condominium values into structure and land components, where the structure value components are consistent with existing national statistical agency practices for the determination of apartment structure values.

− The models need to be tested on other data sets and improvements to the basic methodology should be made. However, the models presented in this paper provide a starting point for further research.

− Moreover, many of the explanatory characteristics that were used will probably be useful in explaining commercial property prices.
Discussion

− (1) \( V_{tn} = \alpha_t L_{tn} + \beta_t S_{tn} + \varepsilon_{tn} \)

• The **costs of developing a property** might not be a good proxy for its **value**, or as Oscar Wilde taught us in *The Picture of Dorian Gray*: “Nowadays people know the price of everything and the value of nothing.”

• This is because someone’s **willingness to pay** for the condominium may be determined by factors beyond the construction costs, or as *Aristotle* put it: “The whole is greater than the sum of its parts.”

− (2) \( V_{tn} = \alpha_t L_{tn} + \beta_t (1 - \delta_t)^{A_{tn}} S_{tn} + \varepsilon_{tn} \)

• How does **depreciation** fit into a “**builder’s model**”? If \( A_{tn} = 0 \) (newly built property), then (2) simplifies to (1).
Discussion

− (4) \( V_{tn} = \alpha_{St} L_{Stn} + \beta_{S} p_{St} (1 - 0.03)^{A_{tn}} (1.1 \ S_{tn}) + \varepsilon_{Stn} \)

• (i) \( \beta_{t} \) replaced by \( \beta_{pSt} \): Is the assumption that the price of new structures is proportional to an index of building costs supported by the data?

• (ii) \( S_{tn} \) replaced by \( 1.1 \ S_{tn} \): What is the variation between condos and over time of this sample wide average adjustment for land imputation, is it robust?

• (iii) \( \delta_{t} \) assumed to equal 0.03: On which basis (data?) is this number for the depreciation rate imputed? (Although it is estimated non time varyingly later.)

− (6) \( V_{Stn} = p_{St} (1 - 0.03)^{A_{tn}} (1.1 \ S_{tn}) \) (NB: No error term, since no estimation.)

• \( \beta \) set equal to 1: The data contradict this! Also, should \( p_{St} \) be a price index(?!), why should the parameter be unity?
Discussion

− (7) $V_{Lt} = V_{tn} - V_{Stn}$

• The structure value in (6) per square metre is, for a given age, proportional to an index of building costs and, hence, (7) is essentially the **residual value method**.

− Sections 4 – 10: Concentrating on explaining variations in these imputed land values and attempting to determine sensible constant quality land prices **does not solve the problem of jointly determining land and structure value**.

• The methodology and results are **highly educational**, though!

• At the end of the paper, **not much remains of the basic builder’s model**, the approach is **mainly empirical** (but rightly so!).
Discussion

− Increases in structure prices were driven by increasing construction costs due to earthquakes and other disasters. It can be seen that as construction prices spike upwards, land prices tend to spike downwards.

• Is it sensible for the value of the property stock to be affected by abnormal developments in construction costs?

• Also, is the spike downwards in land prices a direct result of this effect, and if yes, how can this be explained (e.g. did overall values remain stable)?

− All of the $R^2$s reported in this paper are equal to the square of the correlation coefficient between the dependent variable in the regression and the corresponding predicted variable.

• This is equivalent to the $R^2$ of a linear regression of the dependent variable on a constant and the predicted variable; are the results consistent in a Mincer/Zarnowitz sense, i.e. is the intercept zero and the slope unity?
Discussion

- The **formulation is designed to be useful for national income accountants** who require a decomposition of property value into structure and land components.

- True but **also from a financial stability perspective this is highly relevant.** An important indicator is the **change in values – price changes including quality changes** – of financed objects over time.

- This is because, from the banks' perspective, the **residual value of a home is of interest only should the debtor default**, since then the bank would have to sell the home on the market (possibly in a forced sale).

- The quality of the house, however, is not fixed but it is assumed to be **subject to a constant annual depreciation rate**.

- Hence, **for macroprudential purposes we need something like the age-price profile** in the SNA.