

RISE | RICE INITIATIVE *for the* STUDY *of* ECONOMICS

RISE Working Paper 15-013

“Pricing Characteristics: An Application of Shephard’s Dual Lemma”

by Rolf Färe, Shawna Grosskopf, Chenjun Shang and Robin Sickles



RICE

Department of Economics

Baker Hall, MS22

6100 Main Street, Houston, Texas 77005

<https://economics.rice.edu>

Pricing Characteristics: An Application of Shephard's Dual Lemma*

Rolf Färe
Oregon State University

Shawna Grosskopf
Oregon State University

Chenjun Shang
Rice University

Robin Sickles
Rice University

revised, May 1, 2015

Abstract

The recent housing bubble has provided impetus for revisiting indicators of housing price inflation and property characteristics. Diewert (2011, 'Alternative Approaches to Measuring Housing Price Inflation,' paper presented at the Economic Measurement Group Workshop, 2011, UNSW, Australia) for example has provided a comparison of various methods of constructing property price indices using index number and hedonic regression methods, which he illustrates using data from a small Dutch town over a number of quarters. We provide an alternative approach based on Shephard's dual lemma and apply it to the same data used by Diewert. This method avoids the multicollinearity problem associated with traditional hedonic regression, and the resulting prices of property characteristics show smoother trends than Diewert's results. We also revisit the Diewert and Shimizu (2013) study that employed hedonic regressions to decompose the price of residential property in Tokyo into land and structure components and that constructed constant quality indexes for land and structure prices respectively. We use three models from Diewert and Shimizu (2013) to fit our real estate data from town "A" in Netherlands, and

*We are grateful to W.E. Diewert for his comments and for sharing his data.

also construct the price indices for land and structure, which are compared with our results derived above.

Keywords: Hedonic price index, duality theory, distance function, housing prices

JEL Classification Codes: C2, C23, C43, D12, E31, R21.

1 Introduction

During the recent international financial distresses that led to the Great Recession the balance sheets of banks were (and continue to be) under severe oversight by banking regulators. Basel II requires that bank holding companies have combined Tier 1 and Tier 2 capital ratios of at least 8%. These capital ratios reflect the percentages of a bank's capital to its risk-weighted assets. One of the major assets of a bank is its portfolio of residential mortgages, which typically make up between 20% and 25% of a bank's total assets. As of June 2012 the total residential loans outstanding in the US were roughly US\$ 3.5 trillion.

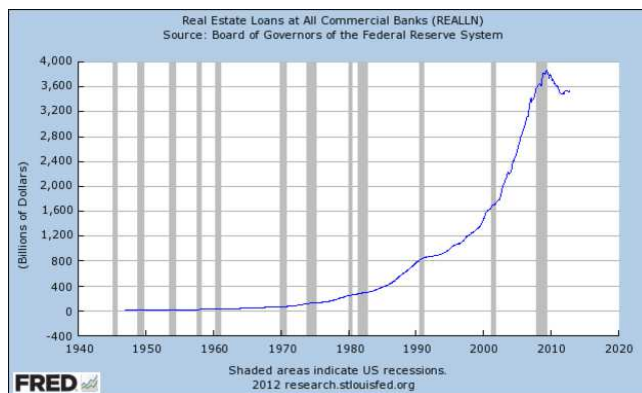


Figure 1: <http://research.stlouisfed.org/fred2/series/REALLN?cid=100>

The precipitous drop in the value of such holdings, measured in the valuations from mark to market accounting instead of the more conventional book valuations, was the key component in causing the Great Recession.

There are a set of reasonable questions to ask about the way in which banks, banking regulators, and rating agencies establish the valuation of a bank's real estate holdings and thus the bank's solvency.

1. What methods do banks use to value their real estate holdings?
2. Are these methods the same across real estate holdings in different regions of a country?
3. Are these methods the same across real estate holdings in different countries?
4. Are these methods transparent and easily calculated?
5. Are these methods consistent with the economic assumptions underlying the value at risk paradigm used to set reserve benchmarks?

Our paper addresses these questions by providing a transparent and easily implementable methodology for constructing real estate price indices based on economic assumptions in keeping with the other economic assumptions underlying many of the regulatory criteria used by banking regulators in, for example, the assessment of the value at risk paradigm that provides banks with the rules for setting reserve benchmarks. We use an input distance function to describe the value generating process of residential properties (also referred to as the "dwelling unit"), which is a euphemism for the output of a production process whose price is the price of the residence. The inputs into the production process are a set of characteristics that a buyer demands, proxied in our empirical analysis by the square footage of the structure, the amount of land on which the structure sits, and the age of the structure. The specific form of the input distance function is translog and the shadow prices are derived based on duality theory. We use these shadow prices to construct an imputed residential properties (dwelling unit) price index and compare it to those generated by more conventional hedonic methods. These methods and their advantages and disadvantages are examined by Good, Sickles, and Weiher (2008), among many others. We implement our modeling approach using a single-output, multi-input distance function. The standard method to estimate the parameters of such an input distance function is to normalize the regression model by moving one of the input variables to the left hand side and to treat it as the dependent variable and

view the unknown distance as a right hand side error that is combined with the normal idiosyncratic error in the regression model. Of course this former source of error is bounded due to the bounded nature of the distance function itself. We utilize several methods to address this aspect of the composed error and compare our imputed residential price index across different approaches. The methods are 1) corrected ordinary least squares with White-type robust standard errors; 2) time dummy least squares regression with White-type robust standard errors; 3) stochastic frontier model (Belotti, et al., 2013).

The paper uses data from Diewert (2011) to construct price indices for residential properties in a small Dutch town using quarterly data from 2005 II to 2008 II. We compare results from our approach, which involves an application of Shephard’s Dual Lemma (Shephard, 1953), with methods employed in Diewert and Shimizu (2013) that utilize stratification techniques and various hedonic treatments. Compared with these hedonic regression approaches, our empirical models can simultaneously estimate the shadow prices of the main property characteristics without suffering from typical problems of collinearity among the quality characteristics. The residential property (dwelling unit) price indices that we derive from our estimations show similar trends to Diewert’s results but appear to be less volatile.

2 The Theoretical Model

It is known from mathematics that a gradient vector of a function belongs to the dual space of its variables. In economics, a classic example is Shephard’s lemma, which says that the derivative of the cost function with respect to a price is an input quantity, i.e., the derivative takes us from the price space to the quantity space. In this paper, we use Shephard’s dual lemma (Shephard, 1953), which says that the derivative of the input distance function with respect to an input is an input price. Next, we apply the dual lemma and use it to derive shadow prices of property characteristics.

Assume that a good is endowed with $z = (z_1, \dots, z_N)$ characteristics. These characteristics in turn generate a value of the good equal to $p \geq 0$.¹ We model this relation with an input correspondence

$$L(p) = \{z \in \mathfrak{R}_+^N : z \text{ generates value } p\}, p \geq 0. \quad (1)$$

¹The value of the good is $p = wz$, where $w = (w_1, \dots, w_N) \in \mathfrak{R}_+^N$ are the unknown prices of the characteristics.

This correspondence can in turn be given a functional representation via Shephard's (1953) input distance function

$$D_i(p, z) = \sup\{\lambda : z/\lambda \in L(p)\}, \quad (2)$$

which, with some mild assumptions on $L(p)$ ² provides a complete characterization of the input correspondence, i.e.,

$$D_i(p, z) \geq 1 \Leftrightarrow z \in L(p). \quad (3)$$

Dual to the input distance is the cost function

$$C(p, w) = \min\{wz : z \in L(p)\} \quad (4)$$

where $w \in \mathfrak{R}_+^N$ are the (unknown) prices of the characteristics. From the duality between $C(p, w)$ and $D(p, z)$ we find the shadow price vector of the characteristics to be

$$w^s = \frac{p \cdot \nabla_z D_i(p, z)}{D_i(p, z)}, \quad (5)$$

i.e., the unknown price vector w can be derived from observed data (p, z) . (See Appendix for the proof.)

To parameterize the distance function (2), we begin by choosing the broad family including generalized quadratic (Chambers, 1988) or transformed quadratic (Diewert, 2002) since these functions are linear in their parameters and provide a second-order Taylor approximation. In addition, if these functions are homogeneous of degree +1 (like the input distance function is in inputs), they take two specific functional forms (Färe and Sung, 1986). Either they are quadratic means of order ρ (Denny, 1974; Diewert, 1976) or translog (Christensen et al., 1971). The former function has only second-order parameters, while the translog has both second- and first-order parameters. Having no zeros in our data, we choose to estimate the translog formulation of the distance function (2). From these estimates, by applying (5), we can derive the desired input-characteristics shadow price vector.

²See Färe and Primont (1995).

To relate our model to the single-output production models by Thorsness (1997) and McMillen (2003), assume that the technology $L(p)$ exhibits constant returns to scale, i.e.,

$$L(\lambda \cdot p) = \lambda \cdot L(p), \lambda > 0. \quad (6)$$

Then and only then can the input distance function be written as

$$D_i(p, z) = \frac{1}{p} D_i(1, z), \quad (7)$$

noting that p is a scalar. Assuming that z belongs to the isoquant of $L(p)$ so that

$$D_i(p, z) = \frac{1}{p} D_i(1, z) \geq 1, \quad (8)$$

then our distance function formulation takes the traditional single-output production function expression,

$$p \leq D_i(1, z). \quad (9)$$

Thus the Cobb-Douglas model by McMillen and the CES model of Thorsness are special cases of (9) with a pricing formula (5) now given by

$$w^s = \frac{p \cdot \nabla_z D_i(1, z)}{D_i(1, z)}. \quad (10)$$

Upon applying(9) we find that

$$w^s = \nabla_z D_i(1, z) \quad (11)$$

3 Data Description

In this section we make use of data kindly provided to us by Professor W. Erwin Diewert and analyzed in Diewert (2011). The data consist of 2289 observations on quarterly sales of detached houses (what we label residential property or dwelling unit) over 14 quarters for the town of ‘A’ in the Netherlands. This is a small city (roughly 60,000 inhabitants) and its exact location and of course the name has been masked by Statistics Netherlands. Transactions on dwelling units begin in the first quarter of 2005 and end the second quarter of 2008.

4 Empirical Model

4.1 General Specification

To justify our choice of the translog model we use in our empirical analyses below, we first note that the input distance function is homogeneous of degree +1 in inputs. It is known, see Färe and Sung (1986), that a homogeneous generalized quadratic function with arguments x_1, x_2 such as

$$\begin{aligned}\varphi^{-1}(F(x_1, x_2)) = & a_0 + a_1 f(x_1) + a_2 f(x_2) + b_1 f(x_1) f(x_1) \\ & + b_2 f(x_2) f(x_2) + b_3 f(x_1) f(x_2)\end{aligned}\quad (12)$$

takes either a translog or generalized mean of order ρ functional form. As the latter function has only second-order parameters, the translog is the preferred choice in our empirical analysis. In the application below we specify the arguments as characteristics of property services, which we have denoted as $z_i, i = 1, \dots, n$.

In the case of detached residential properties, we treat each dwelling unit as the output whose price is influenced by a number of characteristics. The main variables used are

- p : value of the residential dwelling unit;
- L : land area of the property;
- S : floor space area of the structure;
- A : age of the structure.

The land area L , floor space area S , and structure age A are treated as the three input characteristics ($z_i, i = 1, 2, 3$), and p is the value of output. The translog input distance function is specified as below:

$$\begin{aligned}\ln D_i = & \alpha_0 + \alpha_1 \ln p + \frac{1}{2} \alpha_{11} (\ln p)^2 + \beta_1 \ln S + \beta_2 \ln L + \beta_3 \ln A + \frac{1}{2} \beta_{11} (\ln S)^2 \\ & + \frac{1}{2} \beta_{22} (\ln L)^2 + \frac{1}{2} \beta_{33} (\ln A)^2 + \beta_{12} \ln S \ln L + \beta_{13} \ln S \ln A \\ & + \beta_{23} \ln L \ln A + \gamma_1 \ln S \ln p + \gamma_2 \ln L \ln p + \gamma_3 \ln A \ln p.\end{aligned}\quad (13)$$

The assumptions of homogeneity of degree +1 in inputs and symmetry imply the following restrictions on the parameters:

$$\beta_1 + \beta_2 + \beta_3 = 1 \quad (14)$$

$$\sum_{l=1}^3 \beta_{kl} = 0, \quad k = 1, 2, 3 \quad (15)$$

$$\gamma_1 + \gamma_2 + \gamma_3 = 0 \quad (16)$$

$$\beta_{kl} = \beta_{lk}, \quad k, l = 1, 2, 3. \quad (17)$$

Utilizing these restrictions, the distance function can be rewritten as

$$\begin{aligned} \ln D_i &= \alpha_0 + \alpha_1 \ln p + \frac{1}{2} \alpha_{11} (\ln p)^2 + \beta_1 \ln \frac{S}{A} + \beta_2 \ln \frac{L}{A} + \ln A \\ &+ \frac{1}{2} \beta_{11} (\ln \frac{S}{A})^2 + \frac{1}{2} \beta_{22} (\ln \frac{L}{A})^2 + \beta_{12} \ln(\frac{S}{A}) \ln(\frac{L}{A}) \\ &+ \gamma_1 \ln \frac{S}{A} \ln p + \gamma_2 \ln \frac{L}{A} \ln p \end{aligned} \quad (18)$$

From the model, we know that the shadow price vector is

$$w^s = \frac{p \cdot \nabla_z D_i(p, z)}{D_i(p, z)} \quad (19)$$

where p is the value of output, and z is the vector of all inputs. Denoting $w_S, w_L,$ and w_A as the shadow prices of the structure, the land, and the age of the structure respectively, we can then derive the explicit expressions for the shadow prices of the input characteristics as:

$$w_S = \frac{p}{S} (\beta_1 + \beta_{11} \ln S + \beta_{12} \ln L + \beta_{13} \ln A + \gamma_1 \ln p) \quad (20)$$

$$w_L = \frac{p}{L} (\beta_2 + \beta_{22} \ln L + \beta_{12} \ln S + \beta_{23} \ln A + \gamma_2 \ln p) \quad (21)$$

$$w_A = \frac{p}{A} (\beta_3 + \beta_{33} \ln A + \beta_{13} \ln S + \beta_{23} \ln L + \gamma_3 \ln p). \quad (22)$$

4.2 Specification with Constant>Returns-to-Scale Assumption

To make our setting comparable with the traditional hedonic regression model, we consider the assumption of constant returns to scale (CRS). As shown in section 2, we can derive the relation of output price and the input

distance function under CRS assumption as in 9. The inequality is captured by the distance to the frontier by $D_i(1, z)$. Plug the translog functional form into (9), we can obtain the regression equation as below,

$$\begin{aligned} \ln p &= \alpha_0 + \beta_1 \ln S + \beta_2 \ln L + \beta_3 \ln A + \frac{1}{2}\beta_{11}(\ln S)^2 + \frac{1}{2}\beta_{22}(\ln L)^2 \\ &+ \frac{1}{2}\beta_{33}(\ln A)^2 + \beta_{12} \ln S \ln L + \beta_{13} \ln S \ln A + \beta_{23} \ln L \ln A + \ln D_i \end{aligned} \quad (23)$$

We can see that it is the same as the general specification when one sets $\alpha_{11} = \gamma_1 = \gamma_2 = 0$. This model with CRS assumption is, in essence, the same as the hedonic regression with S , L and A as the characteristics.

Again, we impose the assumption of homogeneity of degree 1. Compared with the general case, now there are no terms on the right-hand side involving $\ln p$, thus we do not have the γ coefficients any more. The homogeneity condition now implies equations (??), (??) and (??). With these three constraints, we can rewrite the equation (24) as

$$\begin{aligned} \tilde{\ln p} &= \alpha_0 + \beta_1 \ln \frac{S}{A} + \beta_2 \ln \frac{L}{A} + \frac{1}{2}\beta_{11}(\ln \frac{S}{A})^2 \\ &+ \frac{1}{2}\beta_{22}(\ln \frac{L}{A})^2 + \beta_{12} \ln \frac{S}{A} \ln \frac{L}{A} \end{aligned} \quad (24)$$

where $\tilde{\ln p} = \ln p - \ln A$. Under this new specification, the shadow prices of the three characteristics are

$$w_S = \frac{1}{S}(\beta_1 + \beta_{11} \ln S + \beta_{12} \ln L + \beta_{13} \ln A) \quad (25)$$

$$w_L = \frac{1}{L}(\beta_2 + \beta_{22} \ln L + \beta_{12} \ln S + \beta_{23} \ln A) \quad (26)$$

$$w_A = \frac{1}{A}(\beta_3 + \beta_{33} \ln A + \beta_{13} \ln S + \beta_{23} \ln L) \quad (27)$$

4.3 Construction of the price index

For comparison purposes, we use the matched model Fisher index discussed in Diewert (2011). Diewert constructs price indices for land and for structures that make up the dwelling unit of a particular age and we do likewise. Dwelling units are grouped into 45 cells consisting of 3 categories for land area (small, medium, large), 3 categories for structures (small, medium, large) and 5 groups for age. The break points for the size of land and structure are chosen in a way that about 50% of the units fall in the medium group, and roughly 25% units are in small and large group respectively. The break

points for land area are $L_1 = 160 m^2$ and $L_2 = 300 m^2$, and the break points for structure size are $S_1 = 110 m^2$ and $S_2 = 140 m^2$. Age of the structure is identified by when the structure was built and ranges from 1960 to 2008. For houses built in 2000-2008, $A = 2$; $A = 3$ for 1990-1999 and so on³. Using the structure (or land) prices derived from the above model, we define w_n^t to be the average structure (land) price for properties in cell n that were sold in period t ,

$$w_n^t = \frac{\sum_{i \in n} w_i^t z_i^t}{\sum_{i \in n} z_i^t} = \frac{\sum_{i \in n} w_i^t z_i^t}{z_n^t} \quad (28)$$

where z_i^t and w_i^t represent the structure (land) area and its corresponding shadow price in cell n . As there is no transaction in some cells across two compared periods, we define $S(s, t)$ to be the set of cells n that have at least one transaction in each of the quarters s and t . The indices are then computed over these matched components. The Laspeyres (L) and Paasche (P) indices for periods s and t are:

$$w^L(s, t) = \frac{\sum_{n \in S(s, t)} w_n^t z_n^s}{\sum_{n \in S(s, t)} w_n^s z_n^s} \quad (29)$$

$$w^P(s, t) = \frac{\sum_{n \in S(s, t)} w_n^t z_n^t}{\sum_{n \in S(s, t)} w_n^s z_n^t}. \quad (30)$$

Diewert constructs the (ideal) Fisher index by taking the geometric mean of the above two indices:

$$W^F(s, t) = [w^L(s, t)w^P(s, t)]^{\frac{1}{2}} \quad (31)$$

Two sets of indices are constructed for both structure and land prices. One is a fixed Fisher index, which uses the first quarter as the base period. The other is a chained index: we construct the Fisher index for every two consecutive periods, and the chained index for period t is computed as:

$$I_F^t = W^F(1, 2)W^F(2, 3) \cdots W^F(t - 1, t)$$

³In Diewert (2011), the range of age A is 0 to 4. To accommodate our use of the translog distance function, we shift the range of A to 2 to 6.

5 Regression and Results

5.1 General Specification

We utilize three different regression methods to estimate the input distance function (13). The methods are 1) corrected ordinary least squares with White-type robust standard errors; 2) time dummy least squares regression with White-type robust standard errors; 3) stochastic frontier model (Belotti, et al., 2013). The results are shown in Table 1.

The input distance function has long been utilized in theoretical papers to measure the technical efficiency level of a production process. The input distance is bounded from below by unity, which represents a technically efficient level of production. The use of the distance function in empirical work can be traced back to Färe, Grosskopf and Lovell (1985) wherein linear programming was used to estimate non-parametric distance functions and measure technical efficiency. Starting in the 1990's, researchers also considered parametric functions and used econometric methods for estimation. Lovell, Richardson, Travers and Wood (1994) specified a translog distance function, and used OLS to estimate the parameters. The translog functional form was also used in Coelli and Perelman (1996, 2000) and also was estimated using OLS.

In our application, the land, structure and age of a detached property are regarded as the inputs, which are used to “produce” this property, and input distance D_i gives us an estimate of the (in)efficiency level compared to the efficiency frontier. As discussed in previous sections, we choose translog functional form for the input distance. In most empirical studies using translog input distance functions with m outputs and k inputs, the negative of the logged k th input is treated as the dependent variables and is regressed upon the rest terms, and the negative of the logged input distance is treated as an error term. Employing this method, the equation to be estimated is rearranged as:

$$\begin{aligned} -\ln A &= \alpha_0 + \alpha_1 \ln p + \frac{1}{2}\alpha_{11}(\ln p)^2 + \beta_1 \ln \frac{S}{A} + \beta_2 \ln \frac{L}{A} \\ &+ \frac{1}{2}\beta_{11}(\ln \frac{S}{A})^2 + \frac{1}{2}\beta_{22}(\ln \frac{L}{A})^2 + \beta_{12} \ln(\frac{S}{A}) \ln(\frac{L}{A}) \\ &+ \gamma_1 \ln \frac{S}{A} \ln p + \gamma_2 \ln \frac{L}{A} \ln p - \ln D_i \end{aligned} \quad (32)$$

Note that by the definition of the input distance function, $D_i \geq 1$, thus $-\ln D_i \leq 0$ and can be interpreted as a one-sided error. Now the objective function (32) fits into the production frontier models, and we can utilize frontier models for estimation, such as those developed in Stata (Belotti, et al., 2013).

OLS can be used to estimate the coefficients of the distance function. After obtaining the estimates of all coefficients, we correct the estimated intercept by adding the largest positive residual such that the adjusted function bounds all the observed points from below. This gives us the corrected OLS (COLS) estimates. Some researchers suggest that there exists potential simultaneous equation bias, as one of the input variables (here “ $-\ln A$ ”) is assumed to be endogenous and there are ratios of inputs in the right-hand-side (“ $\ln(S/A)$ ” and “ $\ln(L/A)$ ”). However, as shown in Coelli (2000), under the cost minimization assumption, OLS provides consistent estimates of the parameters. Though our observations are detached properties with access to basically the same amenities, we still use robust errors in the regression to account for possible heteroskedasticity. The parameter estimates are shown in Table 1, and the R^2 is 0.8726.⁴

The data begins in the first quarter of 2005 and ends in the second quarter of 2008. Considering that property prices might be affected by the date of the transaction due to changing market conditions and other factors proxied by time, we add dummy variables to account for yearly effects and reestimate the model by COLS. The adjusted R^2 improves a bit to 0.878.

As noted in Coelli and Perelman (1996), both the linear programming technique (see, for example, Färe et. al., 1993) and the COLS method assume the distance to full efficiency is due entirely to technical inefficiency. To account for the effect of data noise, we can employ stochastic frontier methods. Adding a pure noise term to equation (32), we now have a composite error $\epsilon_i = v_i - \ln D_i = v_i - u_i$, where u_i is the idiosyncratic error

⁴We have also utilized a modified IV estimator to address the potential for endogeneity in the log ratio’s of the other input characteristics and the age characteristic as well as considered other normalizations, for example normalizing with respect to the price of the residential dwelling unit. However, the input distance function we utilize is consistent with cost minimization, holding fixed output and input prices, and thus one may question why price should be considered endogeneous. As one would expect, the renormalization made little difference in our results. Unfortunately, instrumenting the right-hand-side variables that interact with the left-hand-side variable $\ln(A)$ was not possible due to a dearth of potential instruments.

assumed to be $i.i.d.N(0, \sigma_v)$; u_i is our logged input distance, representing technical inefficiency. Here we assume u_i follows the half-normal distribution, i.e., $u_i = |z_i|$, $z_i \sim N(0, \sigma_u)$, and use standard ML techniques for these estimates. These and other stochastic frontier models can be estimated using Stata (Belotti et. al., 2013). The stochastic frontier estimates from equation (32) are reported in Table 1.

5.2 Comparisons with the methods of Diewert and Shimizu (2013)

Diewert and Shimizu (2013) employed hedonic regression techniques to decompose the price of residential property in Tokyo into land and structure components, and constructed constant quality indices for land and structure prices respectively. In this section we use three different models from Diewert and Shimizu (2013) to fit our real estate data from town “A” in Netherlands. We also construct the price indices for structures and land and compared these results with those derived above.

In traditional hedonic regression models, the price of one unit of commodity under study is assumed to depend on a function of its characteristics. Diewert (2003), among others, provides the microeconomic support for this method. If one assumes that an agent can consume an hedonic commodity Z with a set of characteristics $z = (z_1, \dots, z_k)$ and other commodity X , then the consumption of Z units of the hedonic commodity gives a subutility of $f(z_1, \dots, z_k)$, and the consumption of Z and X together generates utility $u = U(Z, X)$. Denote p^t and w^t to be the prices of the general commodity and hedonic commodity at period t , respectively. The consumer then faces a standard cost minimization problem:

$$\min_{X,Z} \{p^t X + w^t Z : U(X, Z) = u^t\}$$

Under some regulatory conditions, the price of the hedonic commodity can be expressed as:

$$w^t = \rho_t f(z_1, \dots, z_t)$$

That is, the price of the hedonic good is the product of some time-dependent effects and the utility a consumer gets from its characteristics. The hedonic regression methods are widely used in real estate studies. It explains the property value based on actual choices of people, and it can be comfortably

modified to take into account the interaction between the characteristics of the property itself and the effects of the surrounding environment. Some limitations also exist with the classic hedonic pricing model as they do in our approach that utilizes the input distance function. For example, the model assumes that the price of the residential dwelling unit can change immediately after the change in one or some of its characteristics, whereas in reality, there may be a substantial time lag. The model assumes that there are a variety of properties in the market so that consumers can choose the one with the desired combination of characteristics, which is only possible if the market is deep. Another problem is the multicollinearity among characteristics, which we will encounter below. A consumer may find that some properties have all the good characteristics, while some alternatives are inferior in all aspects.

The basic paradigm in Diewert and Shimizu (2013) is referred to as a builder's model, which is based on the assumption that the value of a residential property is the sum of the value of the land on which it is built and the construction cost of its structure. Considering that the structure's price usually falls as the structure ages, they assume the constant quality structure to be a function of its age and a constant depreciation rate over all time periods. Following the notation used above, we can specify the model as below,

$$p_{it} = w_{L,t}L_{it} + w_{S,t}(1 - \delta A_{it})S_{it} + \epsilon_{it} \quad i = 1, \dots, N(t); t = 1, \dots, 14 \quad (33)$$

where $N(t)$ is the number of properties sold in period t . One concern about this model is the multicollinearity between the land size and structure size: we would expect a larger house structure to be built with a larger land area. Our data shows that the correlation between land size and structure size is 0.6278. Thus the coefficient estimates of land and structure may not be reliable in the sense that small variations in the data may result in erratic changes in the estimates. To deal with this multicollinearity, Diewert and Shimizu (2013) assumed the price of a constant quality structure was proportional to a property construction cost index published by the relevant authority. This method was also employed in Diewert (2011) as one approach to exploring the price change of residential properties, in which the index used was the New Dwelling Output Price Index (NDOPI) published by the Central Bureau of Statistics of Netherlands. The resulting land price index from Diewert (2011) will be included as part of our comparison.⁵ Though

⁵Compared to what we describe here, Diewert (2010) used a different method to con-

Diewert examined the same dataset for town “A” as our paper, the focus of the research was on the price index construction of the residential properties rather than their characteristics. Thus we use applicable methods from Diewert and Shimizu (2013) to make comparisons, which focused more on the decomposition of property price into land and structure components.

We set the constant quality structure price to be proportional to the New Dwelling Output Price Index (NDOPI) mentioned above, the same index used in Diewert (2011) as we examine the same dataset. As Diewert assumes that the price of the structure is proportional to the construction cost index, we follow this assumption in order to provide results as comparable as possible to the hedonic methods he uses and set $w_{S,t} = w_S P_{C,t}$, where P_{C_t} represents the exogenous cost index. The model can then be written as

$$p_{it} = w_{L,t}L_{it} + w_{S,t}P_{C_t}(1 - \delta A_{it})S_{it} + \epsilon_{it} \quad i = 1, \dots, N(t); t = 1, \dots, 14 \quad (34)$$

We denote this model as DS0, corresponding to the basic builder’s model in Diewert and Shimizu (2013).⁶ Coefficients of this nonlinear model are estimated by minimizing the mean squared error of the residual term. $w_{L,t}$ is interpreted as a suitable constant quality land price for all residential properties sold in period t , and the constant quality land price index for quarter t is defined by Diewert and Shimizu to be

$$I_{L,1t} = w_{L,t}/w_{L,1}. \quad (35)$$

The age-adjusted constant quality structure is defined to be $(1 - \delta A_{it})S_{it}$ and the corresponding structure price index for quarter t is defined by Diewert and Shimizu to be

$$I_{S,1t} = (w_S P_{C_t})/(w_S P_{C_1}) = P_{C_t}/P_{C_1}. \quad (36)$$

The second model employs splines on both the land size and structure age. Empirical evidence indicates that the growth rate of the property land prices vary with land size. To model the possible changes in land prices as land

struct the land price index.

⁶In Diewert and Shimizu (2013), the basic builder’s model also included the 21 dummy variables indicating different wards in Tokyo, to account for possible differences in land prices. In our dataset, all observations are detached residential properties with access to basically the same amenities, thus the difference in locations has little effect on property prices. The R^2 is quite satisfactory in our regression.

area increases, Diewert and Shimizu (2013) divided all observations into 3 groups based on the land size, and assumed that the land price in each group was linear in land size. Diewert (2011), which analyzed the same dataset as we do, also considered the possibility of changing land prices over different land area ranges in one of his approaches to measure the property price. The method used in Diewert (2011) was the same as Diewert and Shimizu (2013) wherein all observations were grouped into 3 categories based on their land sizes and land price was assumed to be piecewise linear in land areas. To make our results comparable with that from Diewert (2011), we divide our data into the same 3 groups, with break points at $L_1 = 160$ and $L_2 = 300$. This generates a grouping with approximately 50% of the properties in the middle group and 25% in the lower and upper groups. The piecewise linear relative land value function is thus specified as

$$f_L(L_{it}) = \text{DL}_{it,1}\gamma_1 L_{it} + \text{DL}_{it,2}(\gamma_1 L_1 + \gamma_2(L_{it} - L_1)) \quad (37) \\ + \text{DL}_{it,3}(\gamma_1 L_1 + \gamma_2(L_2 - L_1) + \gamma_3(L_{it} - L_2))$$

where $\text{DL}_{it,j}$, $j = 1, 2, 3$ are the land dummy variables with $\text{DL}_{it,j} = 1$ indicating that the property falls into category j and $\text{DL}_{it,j} = 0$ meaning that observation i is not in category j . γ_k , $k = 1, 2, 3$ are unknown parameters to be estimated.

We also group all the observations into three categories based on the age of the structure in order to be consistent with Diewert as he points out that depreciation rates will not be the same for structures of different ages. The break points are chosen to be $A_1 = 1$ and $A_2 = 2$, and the categorical dummy variables for ages are denoted as $\text{DA}_{it,j}$, $j = 1, 2, 3$, with $\text{DA}_{it,j} = 1$ indicating that the property is in category j . This is intended to ensure that the three groups have roughly the same size. However, due to the categorical nature of the data and the unequal number of properties in each of the five age categories (ages are indicative of the building decade and range from 0 to 4 with structures with ages new to 10 years a much larger number than the other 4 categories) the actual number of properties in the three groups are unequal and set at 1052, 481, and 474, respectively. The piecewise linear depreciation function of the structure's age is defined as

$$g_A(A_{it}) = 1 - (\text{DA}_{it,1}\delta_1 A_{it} + \text{DA}_{it,2}(\delta_2 A_1 + \delta_2(A_{it} - A_1))) \quad (38) \\ + \text{DA}_{it,3}(\delta_1 A_1 + \delta_2(A_2 - A_1)) + \delta_3(A_{it} - A_2))$$

where δ_k , $k = 1, 2, 3$ are the unknown parameters modeling the depreciation schedule of different structure ages.

Now the new model with generalization on land size and structure age can be defined as

$$p_{it} = w_{L,t}f_L(L_{it}) + w_S P_{C_t} g_A(A_{it}) S_{it} + \epsilon_{it} \quad (39)$$

We denote this model as DS1. This is also a nonlinear regression model, and we can see that the 3 land relative value parameters (γ 's) and the 14 land time parameters ($w_{L,t}$'s) cannot all be identified unless we impose some normalization condition. Thus, we make the normalization that $\gamma_1 = 1$. Note that in this extended model, the marginal land prices for each category, the γ 's, are assumed to be the same over all time periods, while $w_{L,t}$ represents the time change in land price for properties in all groups. Thus the constant quality land price index is again defined as in (35) for period t , while the constant quality structure price index is again defined as in (36).

Besides these main characteristics such as the land and structure size, the price of a residential property is also affected by other factors concerning the design of the structure and use of the land space. The model is further extended to adjust for number of rooms, which affect the quality of the structures. To model the effects from number of rooms, we utilize the same technique that first divides all observations into 3 groups and then define the piecewise linear function of number of rooms. In our data, the number of rooms, denoted as N_{it} , ranges from 2 to 10. We first transform the variable to be $R_{it} = N_{it} - 2$, which ranges from 0 to 8, and then divide the properties into three groups based on R_{it} . The break points for R_{it} are chosen to be $R_1 = 2$ and $R_2 = 3$. Let $DR_{it,j}$ be the dummy variable for the number of rooms, and the piecewise linear function of R_{it} is defined as:

$$g_B(R_{it}) = \theta_1 + DR_{it,1}R_{it} + DR_{it,2}(\theta_2 R_1 + \theta_3(R_{it} - R_1)) \\ + DR_{it,3}(\theta_2 R_1 + \theta_3(R_2 - R_1) + \theta_4(R_{it} - R_2)) \quad (40)$$

Then model incorporating this adjustment of structure quality is specified as:

$$p_{it} = w_{L,t}f_L(L_{it}) + w_S P_{C_t} g_A(A_{it}) g_B(R_{it}) S_{it} + \epsilon_{it} \quad (41)$$

where of course w_S does not vary with time in order to address the high collinearity between structure and land sizes. As the case in DS1, not all parameters can be identified unless some normalizations are made. We thus add two normalization conditions that $\gamma_1 = 1$ and $\theta_1 = 1$, and denote this model as DS2. The constant quality price indices for land and structure are defined the same as in the previous two models.

The land price indices constructed from different models are plotted in Figure 2. We can see that the three indices derived using our model exhibit similar trends, while the indices derived from the DS models show somewhat parallel but shifted temporal patterns. The upturns and downturns basically occur at the same time period, but indices generated from our model appear less variable.

Note that the land price indices derived using the DS models climb in the third quarter in 2007, and then fall sharply in the next two quarters, while indices from our model move rather smoothly and exhibit only a moderate fall in 2008 quarter 1. To further explore the difference, we need to jointly look at the property prices and the implied prices of the land and the structure. Using the method described in section 4.1, we can similarly construct the Fisher fixed-base and chained indices for the properties, which are shown in Table 4. We can see that the property price index also peaks in 2007 quarter 3, and then falls back a little. When decomposing this temporal change of property price into components of land price and structure price, the DS models assume the structure prices to be proportional to the exogenous construction cost index (NDOPI), which increases all the way over 2007 to the first quarter in 2008. Thus all the decrease in the property price in 2008, quarter 1 is attributed to change in the land price, resulting in a substantial fall of the land price index as we observed. In our model, this decrease in the residential property price index is explained by both land and structure price. Seen from Figure 3, the structure price indices derived from our model also fall moderately in the first quarter in 2008.⁷

The use of NDOPI as the structure price index in the DS models is adopted to deal with multicollinearity problems, and cannot be supported within the model. With the same idea, we now introduce exogenous information on land price and derive the structure prices from the DS models. The time-varying parameter $w_{L,t}$ now is set to be $w_L P_{L,t}$, proportional to some land price index, and the structure price parameter $w_{S,t}$ is allowed to freely change over each period. With all other settings remain the same, we still denote the models as DS0, DS1 and DS2. The exogenous land price index is chosen to be the one generated from our SFA model. The resulting structure price indices are shown in Table 7 and plotted in Figure 4. We can see that with the land price index derived from our model as exogenous information,

⁷In all DS models, the structure price indices are equal to the construction cost index (NDOPI), thus we only use DS0 as a representation.

the structure price indices generated from the DS models show basically the same trend as our structure price indices. Thus, the reliability of land or structure prices generated using the DS models heavily depends on how well the exogenous information reflects the true changing pattern of the price of the other component. The method we propose here can avoid the problem of multicollinearity in hedonic regression and construct the land and structure price indices at the same time. Our methods also tend to smooth the price fluctuations and are less sensitive to market changes compared to the DS models.

5.3 Specification with Constant>Returns-to-Scale Assumption

Compared with the general specification, the setting under CRS assumption is a constrained regression with $\alpha_{11} = \gamma_1 = \gamma_2 = 0$. We conduct the joint F-test for these constraints, and the F-statistic is 42.5, with a p-value smaller than 10^{-4} . Thus we would reject the assumption of constant-returns-to-scale based on the data. Nevertheless, we impose the CRS constraint here to make our setting comparable with the hedonic regression model.

Assuming the white noises are i.i.d with zero mean and constant variance, and taking into account the one-sided log-distance term $\ln D_i$, again we have composite errors in the regression. From the results of the general specification, we can see that the slope coefficients estimates are very close with each other under different regression methods, and only the intercept needs to be corrected for the one-sided error. As only the estimates of slope coefficients are used in the formula generating the shadow prices of characteristics, the choice of which regression method to use does not lead to much difference. Thus I will use robust OLS method for simplicity. To take into account the possible year effects and seasonality, we add dummies for both years and quarters in our regression. The estimation results are shown in Table 3. Using formulas (27), we can then construct the price indexes of structure and land, which are shown in Table 8. As we add dummies to pick up the yearly and quarterly effects in the regression, we compare the results under the CRS assumption with those obtained in general specification with time dummy variables, and denote as TD-CRS and TD correspondingly. The plots of the indexes are shown in Figure 5 and Figure 6. As we can see from the figures, compared with the general specification, the structure price indexes

under CRS assumption are almost the same, while the land price indexes are lower in all periods. To uncover where this difference comes from, we go back to check the coefficients estimates with and without the CRS assumption. In the general specification with time dummy variables, the coefficient of the cross term of structure and property price, $\ln S \ln p$, is considerably smaller, and the corresponding t-value is near zero, indicating that the coefficient is insignificant. As a contrast, the coefficient of the cross term of land and property price, $\ln L \ln p$, is much larger and significant. In the specification with the CRS assumption, both cross terms are deleted, and their coefficients are no longer in the calculation of the shadow price of corresponding characteristics. Thus the lack of $\ln S \ln p$ has nearly no influence on the shadow price calculation of structures, while the lack of $\ln L \ln p$ results in obvious change in the calculated shadow prices of lands. Despite the decrease of the levels of the land price indexes, the changing trends over the studied period are basically the same with or without the CRS assumption.

6 Conclusions

We are optimistic about the potential usefulness of our new approach to construct residential property price indices. It has reasonable theoretical underpinnings and is parsimonious in terms of required data. Rather than relying on exogenous information to circumvent problems of multicollinearity between different property characteristics, our method can estimate the shadow price of each characteristic with little computational burden. As this model is less sensitive to actual market fluctuations, it also can be combined with traditional hedonic regression methods to provide bounds on residential property prices based on mark-to-market adjustment and less volatile consumer preferences.

7 References

- Belotti, F., Daidone, S., Ilardi, G, and Atella, V. (2013) “Stochastic frontier analysis using Stata,” *Stata Journal*. Forthcoming.
- Chambers, R. G. (1988) *Applied production analysis: a dual approach*. Cambridge University Press, Cambridge.
- Christensen, L. R., Jorgenson, D. W., and Lau, L. J. (1971) “Conjugate duality and the transcendental logarithmic production function,” *Econometrica*. 39, 255-265.
- Coelli, T. (2000) “On the econometric estimation of the distance function representation of a production technology,” *Center for Operations Research & Econometrics, Université Catholique de Louvain*.
- Coelli, T and Perelman, S. (1996) “Efficiency measurement, multiple-output technologie and distance functions: with application to European railways,” *Centre de Rechercheen Économie Publique et de la Population (C.R.E.P.)*, Liège, Belgium.
- Coelli, T. and Perelman, S. (2000) “Technical efficiency of European railways: a distance function approach,” *Applied Economics*. 32(15), 1967-1976.
- Denny, M. C. (1974) “The relationship between functional forms for the production system,” *Canadian Journal of Economics*. 7, 21-31.
- Diewert, E. W. (1976) “Exact and superlative index numbers,” *Journal of Econometrics*. 9, 481-507.
- Diewert, E. W. (2002) “The quadratic approximation lemma and decomposition of superlative indexes,” *Journal of Economic and Scocial Measurement*. 28, 63-88
- Diewert, E. W. (2003) “Hedonic regressions: a consumer theory approach,” *Scanner data and price indexes*. 317-348, University of Chicago Press.
- Diewert, E. W. (2011) “Alternative approaches to measuring house price inflation,” *Discussion Paper 10-10, Economics, The University of British Columbia*.

- Diewert, E. W., de Haan, J. and Hendriks, R. (2010) “The decomposition of a house price index into land and structures components: a hedonic regression approach,” *Discussion Paper 10-01, Economics, The University of British Columbia*.
- Diewert, E. W. and Shimizu, C. (2013) ”A conceptual framework for commercial property price indexes,” *Discussion Paper 13-11, School of Economics, The University of British Columbia*.
- Epple, D., Gordon, B., and Sieg, H. (2010) “A new approach to estimating the production function for housing,” *American Economic Review*. 100, 905-924.
- Färe, R., Grosskopf, S., Lovell, C.A.K., and Yaisawarng, S. (1993) “Derivation of shadow prices for undesirable outputs: a distance function approach,” *The Review of Economics and Statistics*. 75, 374–380.
- Färe, R., Grosskopf, S., and Lovell, C.A.K.. (1985) *The measurements of efficiency of production*. Vol. 6. Springer.
- Färe, R. and Sung, K.J. (1986) “On second order Taylor’s series approximations and linear homogeneity,” *Aequationes Mathematicae*. 30, 180-186.
- Färe, R. and Primont, D. (1995) *Multi-output production and duality: theory and applications*. Norwell, MA: Kluwer Academic.
- Good, D. H., Sickles, R. C., and Weiher, J. (2008) “A hedonic price index for airline travel,” *Review of Income and Wealth*. 54, 438-465.
- Lovell, C.A.K., Travers, P., Richardson, S., and Wood, L. (1994) *Resources and functionings: a new view of inequality in Australia*. 787-807, Springer Berlin Heidelberg.
- McMillen, D. P. (2003) “The return of centralization to Chicago: using repeat sales to identify changes in house price distance gradients,” *Regional Science and Urban Economics*. 33, 287-304.
- Shephard, R. W. (1953), *Cost and production functions*. Princeton NJ: Princeton University Press.

Thorsnes, P. (1997) "Consistent estimates of the elasticity of substitution between land and non-land inputs in the production of housing," *Journal of Urban Economics*. 42, 98-108.

Appendix

We prove the shadow price expression (5) in two ways, both when z adjusts to w and when w adjusts to z .

Define the cost function as

$$C(p, w) = \min_z wz - \mu(D_i(p, z) - 1) \quad (\text{A.1})$$

then

$$F.O.C : \quad w - \mu^* \nabla_z D_i(p, z) = 0. \quad (\text{A.2})$$

Also note that

$$\mu^* = C(p, w). \quad (\text{A.3})$$

This follows from

$$\begin{aligned} \tilde{C}(p, w, \alpha) &= \min_z wz - \mu(D_i(p, z) - \alpha) \\ &= \alpha C(p, w), \end{aligned} \quad (\text{A.4})$$

and

$$\frac{\partial \tilde{C}}{\partial \alpha} = \mu = C(p, w). \quad (\text{A.5})$$

Thus, by the *FOC* we have,

$$w = C(p, w) \cdot \nabla_z D_i(p, z) \quad (\text{A.6})$$

and by Euler's theorem we have

$$p = wz = C(p, w) D_i(p, z) \quad (\text{A.7})$$

or

$$C(p, w) = p / D_i(p, z) \quad (\text{A.8})$$

and thus

$$w^s = \frac{p \cdot \nabla_z D_i(p, z)}{D_i(p, z)}. \quad (\text{A.9})$$

Next consider the dual optimization problem

$$\begin{aligned} D_i(p, z) &= \min_w wz - \lambda(C(p, w) - 1) \\ &= w^* z - \lambda^*(C(p, w^*) - 1) \end{aligned} \quad (\text{A.10})$$

As above one can show that

$$D_i(p, z) = \lambda^* \tag{A.11}$$

thus

$$D_i(p, z) = w^* z - D_i(p, z)(C(p, w^*) - 1) \tag{A.12}$$

and

$$D_i(p, z) = \frac{w^* z}{C(p, w^*)} \tag{A.13}$$

Therefore

$$\nabla_z D_i(p, z) = \frac{w^*}{C(p, w^*)} - \frac{w^* \cdot \nabla_w C(p, w^*)}{C(p, w^*)^2}. \tag{A.14}$$

Now since $C(p, w^*) = 1$, a constant function,

$$w^* \cdot \nabla_w C(p, w^*) = 0 \quad \text{and}$$

$$\nabla_z D_i(p, z) = \frac{w^*}{C(p, w^*)}. \tag{A.15}$$

Again, using the Euler's theorem and noting that $w^* z = p$, we have

$$w^{s,*} = \frac{p \cdot \nabla_z D_i(p, z)}{D_i(p, z)} \tag{A.16}$$

Tables and Figures

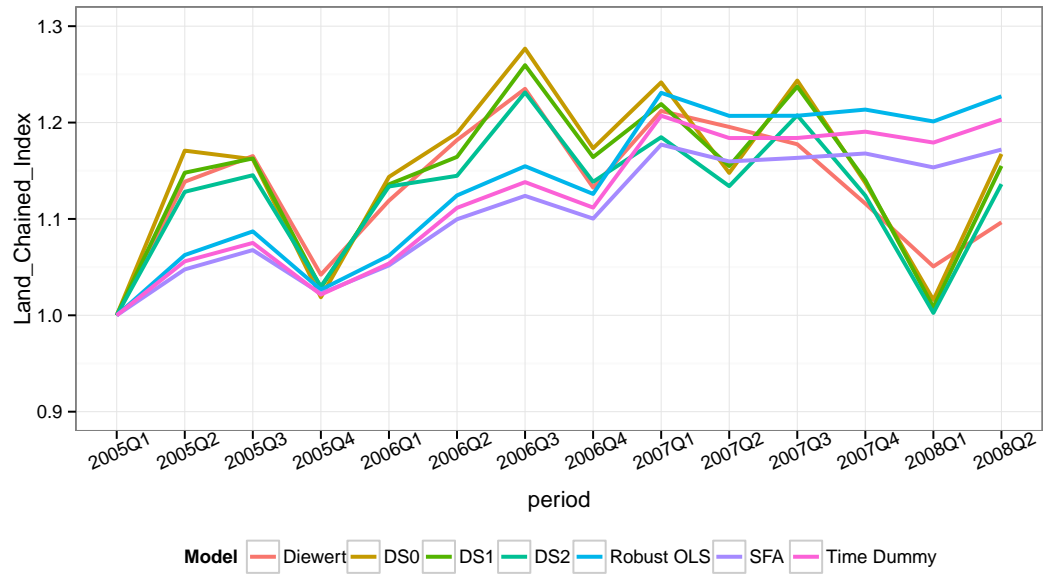


Figure 2: Land Price Index Comparison

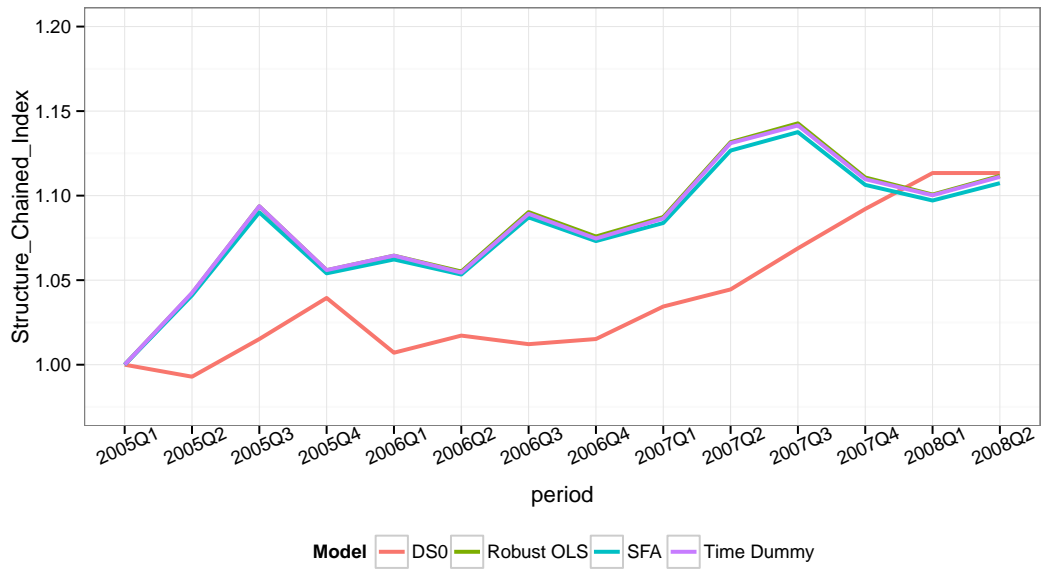


Figure 3: Structure Price Index Comparison

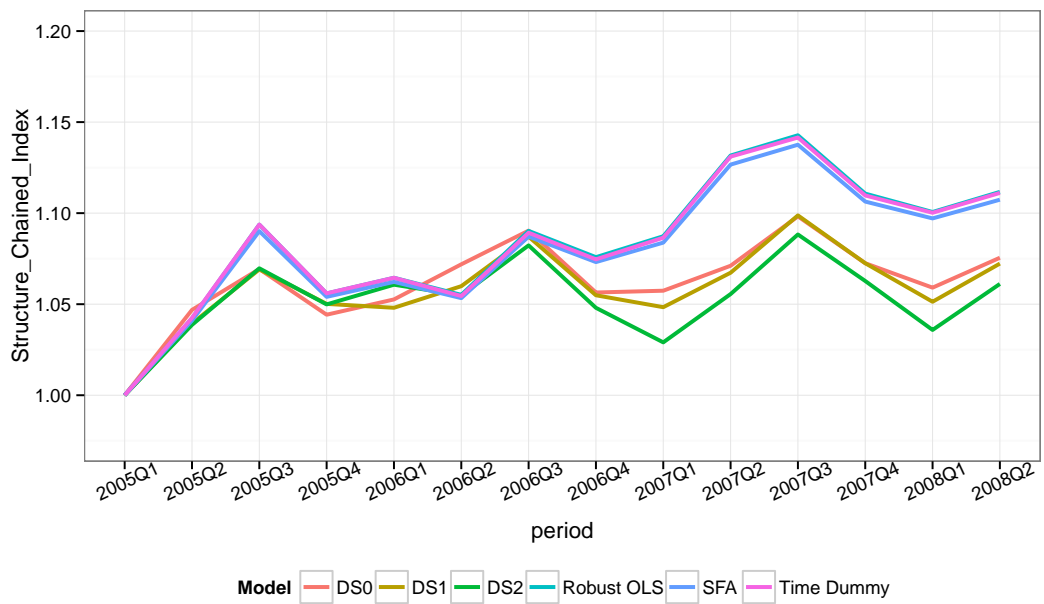


Figure 4: Structure Price Index Comparison with Exogenous Land Price

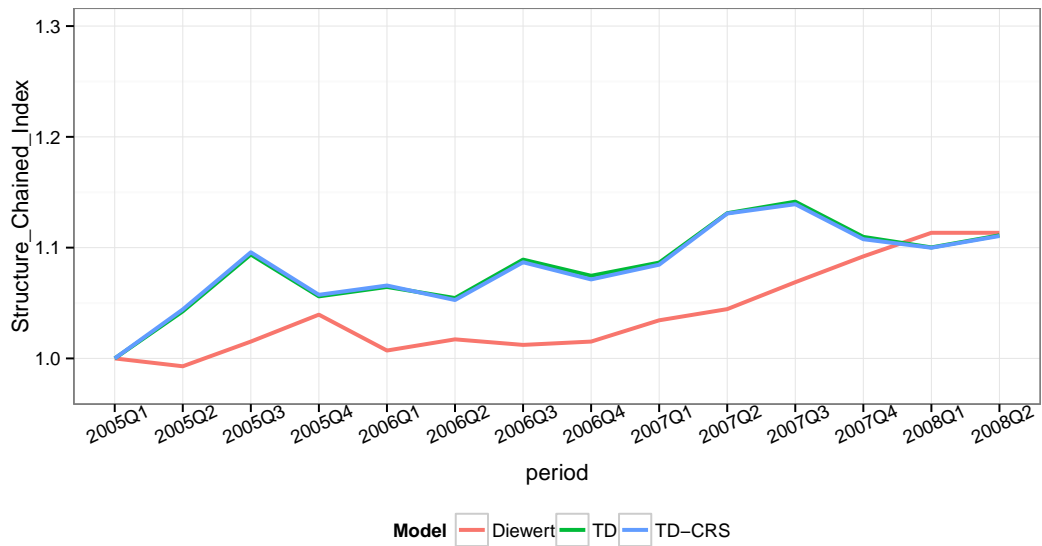


Figure 5: Structure Price Index Comparison with CRS

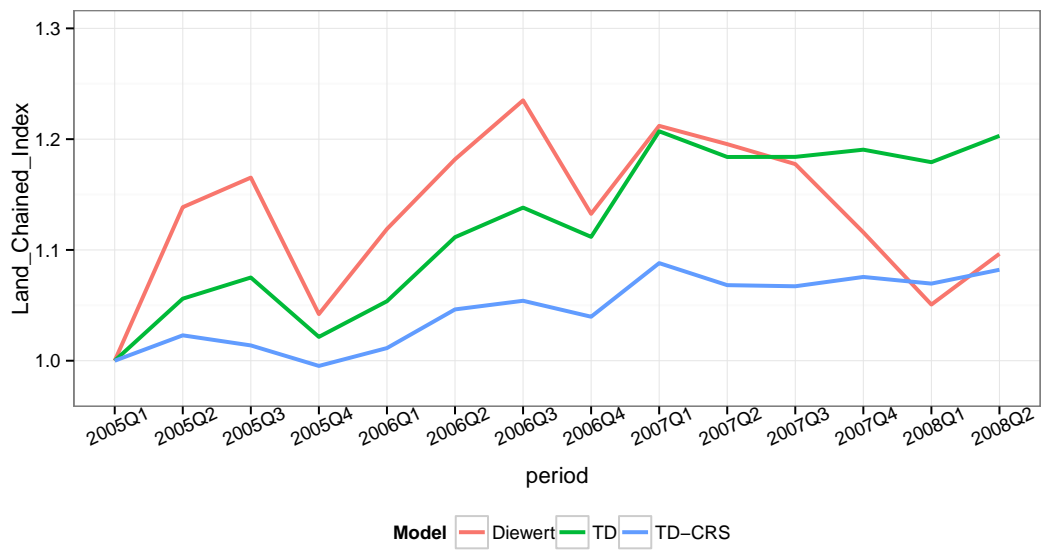


Figure 6: Land Price Index Comparison with CRS

Parameter	COLS			Time Dummy			SFA		
	Coefficient	Std.Err.	t	Coefficient	Std.Err.	t	Coefficient	Std.Err.	t
α_0	-8.654262	0.9184771	-9.42				-7.886361	0.533136	-14.79
α_1	2.999968	0.6205874	4.83	2.519516	0.2859612	8.81	2.506335	0.2969787	8.44
α_{11}	-0.432247	0.1122404	-3.85	-0.3851873	0.042537	-9.06	-0.339389	0.0458892	-7.40
β_1	0.4251001	0.1919702	2.21	0.5175217	0.1231534	4.20	0.5230654	0.1267647	4.13
β_2	-0.7420501	0.2855724	-2.6	-0.6834723	0.1417875	-4.82	-0.5483521	0.1469673	-3.73
$\frac{1}{2}\beta_{11}$	0.0491277	0.0205348	2.39	0.0358318	0.0168336	2.13	0.0550541	0.0168423	3.27
$\frac{1}{2}\beta_{22}$	-0.0471125	0.0296489	-1.59	-0.0423398	0.0150257	-2.82	-0.0146476	0.0160396	-0.91
β_{12}	0.0074992	0.0490074	0.15	0.0073379	0.0249217	0.29	0.0101349	0.0255948	0.40
γ_1	-0.0105265	0.0651377	-0.16	-0.0098862	0.0375809	-0.26	-0.0390997	0.0388763	-1.01
γ_2	0.2456549	0.1031222	2.38	0.2307362	0.0425736	5.42	0.1550929	0.0459129	3.38

Table 1: Regression Results of Different Methods

	Diewert	Ours
Num. of observations	2289	2280
Ave. sale price	190.13	189.76
Ave. land area	257.6	258.98
Ave. structure area	127.2	127.09

Table 2: Summary Statistics Comparison

	Coef.	Std. Err.	t
α_0	-0.257646	0.114862	-2.24
β_1	0.9153355	0.075584	12.11
β_2	0.1219993	0.0539718	2.26
$\frac{1}{2}\beta_{11}$	-0.018214	0.0189258	-0.96
$\frac{1}{2}\beta_{22}$	0.029778	0.0107326	2.77
β_{12}	-0.000395	0.0241278	-0.02

Table 3: Coefficients of Specification with CRS

Time Period	Num. of Obs.	NDOPI	Fisher Fixed-Base House Price Index	Fisher Chained House Price Index
1	157	1.0000	1.0000	1.0000
2	155	0.9929	1.0240	1.0240
3	154	1.0152	1.0682	1.0784
4	155	1.0395	1.0490	1.0408
5	163	1.0071	1.0444	1.0408
6	175	1.0172	1.0668	1.0575
7	157	1.0122	1.0731	1.0734
8	152	1.0152	1.0768	1.0671
9	159	1.0344	1.0683	1.0895
10	194	1.0445	1.1189	1.1148
11	137	1.0688	1.1220	1.1247
12	187	1.0921	1.1132	1.1048
13	148	1.1134	1.1107	1.1045
14	187	1.1134	1.1058	1.1119

Table 4: House Price Index and NDOPI

Period	DS0	DS1	DS2	COLS	SFA	Time Dummy	Diewert
1	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000
2	1.17085	1.14797	1.12816	1.06255	1.04764	1.05601	1.13864
3	1.16199	1.16277	1.14536	1.08708	1.06768	1.07512	1.16526
4	1.01873	1.02859	1.02990	1.02655	1.02264	1.02149	1.04214
5	1.14362	1.13589	1.13368	1.06191	1.05178	1.05369	1.11893
6	1.18897	1.16439	1.14474	1.12440	1.09966	1.11148	1.18183
7	1.27672	1.25955	1.23127	1.15472	1.12393	1.13820	1.23501
8	1.17338	1.16416	1.13854	1.12592	1.10034	1.11176	1.13257
9	1.24168	1.21921	1.18481	1.23086	1.17709	1.20706	1.21204
10	1.14775	1.15482	1.13401	1.20693	1.15966	1.18394	1.19545
11	1.24358	1.23731	1.20753	1.20707	1.16338	1.18398	1.17747
12	1.13746	1.14025	1.12421	1.21345	1.16790	1.19053	1.11588
13	1.01596	1.00828	1.00247	1.20116	1.15347	1.17922	1.05070
14	1.16739	1.15499	1.13613	1.22721	1.17198	1.20303	1.09648

Table 5: Land Price Indexes Comparison

Period	DS0	COLS	SFA	Time Dummy
1	1.00000	1.00000	1.00000	1.00000
2	0.99291	1.04218	1.04103	1.04250
3	1.01518	1.09368	1.09011	1.09375
4	1.03947	1.05588	1.05404	1.05597
5	1.00709	1.06450	1.06224	1.06452
6	1.01721	1.05513	1.05333	1.05448
7	1.01215	1.09026	1.08712	1.08924
8	1.01518	1.07582	1.07312	1.07462
9	1.03441	1.08732	1.08382	1.08652
10	1.04453	1.13165	1.12666	1.13105
11	1.06883	1.14276	1.13757	1.14152
12	1.09211	1.11073	1.10636	1.10971
13	1.11336	1.10063	1.09711	1.10022
14	1.11336	1.11168	1.10738	1.11111

Table 6: Structure Price Index Comparison

Period	DS0	DS1	DS2	COLS	SFA	Time Dummy
1	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000
2	1.04687	1.03938	1.03864	1.04218	1.04103	1.04250
3	1.06929	1.06910	1.06970	1.09368	1.09011	1.09375
4	1.04422	1.05016	1.04987	1.05588	1.05404	1.05597
5	1.05261	1.04803	1.06071	1.06450	1.06224	1.06452
6	1.07189	1.05990	1.05488	1.05513	1.05333	1.05448
7	1.09040	1.08710	1.08229	1.09026	1.08712	1.08924
8	1.05641	1.05490	1.04801	1.07582	1.07312	1.07462
9	1.05740	1.04835	1.02903	1.08732	1.08382	1.08652
10	1.07110	1.06728	1.05556	1.13165	1.12666	1.13105
11	1.09823	1.09876	1.08828	1.14276	1.13757	1.14152
12	1.07249	1.07250	1.06274	1.11073	1.10636	1.10971
13	1.05908	1.05134	1.03586	1.10063	1.09711	1.10022
14	1.07550	1.07229	1.06113	1.11168	1.10738	1.11111

Table 7: Structure Price Indexes Comparison with Exogenous Land Price

TD-CRS	TD	Diewert	TD-CRS	TD	Diewert
1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
1.0442	1.0425	0.9929	1.0229	1.0560	1.1386
1.0958	1.0938	1.0152	1.0138	1.0751	1.1653
1.0574	1.0560	1.0395	0.9953	1.0215	1.0421
1.0658	1.0645	1.0071	1.0114	1.0537	1.1189
1.0527	1.0545	1.0172	1.0463	1.1115	1.1818
1.0868	1.0892	1.0122	1.0541	1.1382	1.2350
1.0713	1.0746	1.0152	1.0397	1.1118	1.1326
1.0846	1.0865	1.0344	1.0881	1.2071	1.2120
1.1308	1.1311	1.0445	1.0682	1.1839	1.1955
1.1392	1.1415	1.0688	1.0672	1.1840	1.1775
1.1076	1.1097	1.0921	1.0756	1.1905	1.1159
1.0999	1.1002	1.1134	1.0696	1.1792	1.0507
1.1105	1.1111	1.1134	1.0821	1.2030	1.0965

(a) Structure Price Indexes Comparison with CRS

(b) Land Price Indexes Comparison with CRS

Table 8: Price Indexes Comparison with CRS