The HKU Scholars Hub   The University of Hong Kong   香港大學學術庫

Title	Adjusting for non-linear age effects in the repeat sales index
Author(s)	Chau, KW; Wong, SK; Yiu, CY
Citation	Inaugural Singapore/Kong Hong International Real Estate Research Symposium, Singapore, 18-19 July, 2003. In Journal of Real Estate Finance and Economics, 2005, v. 31 n. 2, p. 137-153
Issued Date	2005
URL	http://hdl.handle.net/10722/81857
Rights	Creative Commons: Attribution 3.0 Hong Kong License

ADJUSTING FOR NON-LINEAR AGE EFFECTS IN THE REPEAT SALES INDEX

by

K.W. Chau,[1] S.K. Wong,[1] and C.Y. Yiu[2]

[1]Department of Real Estate and Construction, The University of Hong Kong
[2]Department of Building and Real Estate, The Polytechnic University of Hong Kong

*Please send comments to:
Professor K.W. Chau
Chair of Real Estate and Construction
Department of Real Estate and Construction
The University of Hong Kong
Pokfulam Road
Hong Kong

Tel : (852) 2859 - 2128
Fax: (852) 2559 - 9457
E-mail : HRRBCKW@HKU.HK

1

ADJUSTING FOR NON-LINEAR AGE EFFECTS IN THE REPEAT SALES INDEX

K.W. Chau, S.K. Wong, and C.Y. Yiu

ABSTRACT:

A true constant quality real estate price index should measure the general change in price level free from any change in quality over time. In recent years, the repeat-sales method has been widely used to construct constant quality property price indices. Since buildings depreciate over time, a simple repeat-sales index would underestimate the growth in property prices. The major problem of controlling the effects of age constant in a repeat-sales model arises from the exact multicollinearity between the age variable and the time dummy variables. In this study, we derive a solution that is theoretically sound and practical by allowing the age effects to be non-linear. In case of leasehold properties, we further incorporated interest rates into the model because the effects of age on real estate prices depend theoretically on interest rates. A sample of residential units in Hong Kong sold more than once from Quarter 2 of 1991 to Quarter 1 of 2001 (more than 11,000 repeat sales pairs) are used for the empirical analysis.

KEYWORDS:

Age effects, depreciation, duration, multicollinearity, repeat-sales index

1. INTRODUCTION

An ideal price index measures the general change in the price level of a group of commodities over time. It is "general" in the sense that the measured change only reflects the price effects common to all commodities in the group, provided that such common effects exist. This raises two difficulties in real assets. First, due to their infrequent trading and heterogeneous characteristics, the properties transacted in one period are unlikely to be the same as those transacted in another period. Second, even if the properties transacted remain the same between periods, their qualities are likely to change over time. These two problems may undesirably cause the price index to measure quality variations in addition to general price changes.

Two approaches have been commonly adopted to eliminate quality differences in constructing real estate price indices. One is the hedonic pricing model that controls for the varying characteristics of properties using regression (see, for example, Palmquist, 1980; Greenlees, 1982; and Mark and Goldberg, 1984). The price index is then computed either from the coefficients of time dummies in a single regression or from the values of a standard property through regressions for each time period. The major

criticisms of this approach, however, are the assumption on functional form[1] and the difficulty in specifying and measuring a correct and complete set of attributes (e.g. Case, *et al.*, 1991). Both problems may lead to biased estimates of the coefficients.

The other approach is the repeat-sales model originally proposed by Bailey, *et al.* (1963) and later extended by Case and Shiller (1987). The idea of this approach is to control quality variations by utilizing the transacted prices of the same property in different time periods. Provided that property characteristics and their implicit prices do not change between sales, the price differences can be solely explained by time dummies,[2] and thus the problems of specifying the functional form and attributes, as in the hedonic method, can be avoided. The price index is then obtained directly from the coefficients of time dummies. However, this approach has often been criticized for its inefficiency (e.g. Case, *et al.*, 1991), sample selection bias (Clapp and Giaccotto, 1992; Meese and Wallace, 1997), and incapability to control for age effects (Bailey, *et al.*, 1963).

This paper provides a solution to the age effects problem in the repeat-sales model. Age effects refer to the change of property value when it ages over time, *ceteris paribus.* There are several possible sources of age effects. First, physical deterioration and functional obsolescence give rise to higher maintenance costs and/or lower productivity. Second, older buildings will be sold at a discount if they are regarded as lemons. Third, for leasehold land,[3] building age also bears a negative relationship with the residual length of leasehold interests (i.e. the term to maturity). The shortened future income stream for older buildings provides an additional source of negative age effects on property prices. In the presence of these age effects, the assumption of no quality change between sales of the same property in the repeat-sales model obviously does not hold. If they are ignored, the model will produce biased estimates of time dummy coefficients (omitted variable bias), and thus an erroneous price index. Unlike other occasional quality changes such as renovations, we cannot simply drop the age effects from our sample or control them by sample selection because they apply to all transaction data. As a result, the only solution is to model the age effects explicitly in the repeat-sales model.

However, incorporating an age variable directly into the model will result in exact collinearity with time dummies. Several researchers have attempted to overcome this problem. The earliest attempts to solve a similar problem in the hedonic model were Cagan (1971) and Hall (1971). They suggested imposing an *a priori* restriction on the coefficients. Palmquist (1980, 1982) used an independent estimate of depreciation to adjust for the age effects occurred between sales, but his method required the assumption of a stable rate of depreciation over time, as well as the estimation of the hedonic pricing models. Another line of approach involved the 'hybrid' technique, which combines the

---

[1] The functional form problem can be alleviated by using a more flexible parameterisation such as Box-Cox transformation or non-parametric techniques (Meese and Wallace, 1991).
[2] When there are changes in property characteristics between sales (e.g. renovations), they can be either eliminated from the sample or explicitly taken into account in the repeat-sales model. See Section 2 for details.
[3] For our purpose, leasehold land refers to land owned for a fixed period of time, whereas freehold land is owned perpetually.

repeat-sales model and the hedonic pricing model (Case and Quigley, 1991; Quigley, 1995; Hill, *et al.,* 1997; Englund, *et al.,* 1998). Since the data required to estimate a hybrid model is often difficult to obtain, the hybrid model approach is often not practical or very expensive. Recently, Clapp and Giaccotto (1998) proposed adding the interaction terms of age variables and time dummies, whereas Cannaday, *et al.* (1998) proposed using age dummies. Both tried to avoid putting an age variable directly into the model. We will show that their methods are erroneous in the next section.

In this paper, we argue that the exact multicollinearity in estimating age and time effects concurrently in the repeat-sales model can be resolved by dropping the assumption of a constant depreciation rate over time. This is an assumption being made for convenience, and has no theoretical justification. We developed a repeat-sales model that allows depreciation to be non-linear, so that age and time effects can be disentangled.[4] We further generalized our model to cover properties under a leasehold tenure where age effects are theoretically interest-rate-dependent. Therefore, our model is applicable to both freehold and leasehold properties.

The organization of this article is as follows. Section 2 critically reviews the previous attempts to solve the multicollinearity problem in the repeat-sales model. Section 3 discusses the reasons for non-linear age effects and establishes the mathematical exposition. Section 4 presents the resulting age-adjusted repeat-sales model. An empirical analysis using transactions of residential properties in Hong Kong is shown in Section 5. The conclusion is given in the final section.

2. PREVIOUS ATTEMPTS AT SOLVING THE MULTICOLLINEARITY PROBLEM

2.1 The Exact Multicollinearity Problem

Assuming the characteristics and their implicit prices of the same property do not change between the first sale (t_1) and second sale (t_2), Bailey, *et al.* (1963) introduced the following repeat-sales model (the BMN model):

$$\ln\left(\frac{P_{it_2}}{P_{it_1}}\right) = \sum_{t=1}^{T} \alpha_t D_{it} + \varepsilon_{it_1 t_2} \tag{1}$$

where P_{it} denotes the sales price of property i in period t ($i=1,\ldots,n$; $t=0,\ldots,T$); α_t denotes the estimated coefficient for time dummy D_{it}; and $\varepsilon_{it_1 t_2}$ is the error term. The time dummies are set to take the value -1 if $t = t_1$, $+1$ if $t=t_2$, and zero if otherwise. Note that α_0 has been set to zero so as to normalize the price index at time period 0.

---

[4] The paper only aims to separately identify the age and time effects. The separation of various components of age-related effects such as time-to-lease-expiration and physical deterioration is beyond the scope of this study. We regard the age-related depreciation estimate to be net of maintenance efforts.

4

When the assumption of unchanged property characteristics was violated, Bailey, *et al.* showed how to modify Equation (1) to take into account any quality changes that occurred between sales:

$$\ln\left(\frac{P_{it_2}}{P_{it_1}}\right) = \sum_{t=1}^{T}\alpha_t D_{it} + \sum_{k=1}^{K}\beta_k\left(Q_{kit_2} - Q_{kit_1}\right) + \varepsilon_{it_1t_2} \tag{2}$$

where β_k ($k=1,\ldots,K$) denotes the estimated coefficient for the change in property characteristic Q_{kit} from first sale t_1 to second sale t_2. This specification, however, is unable to deal with linear age effects because the difference in building age between two sales is an exact linear combination of the time dummies.

This collinear relationship can be shown mathematically. Assuming the only quality change is due to age effects for which age (A) acts as a proxy, we can substitute A for Q in Equation (2). Since age difference and time difference between two sales of the same property are always identical, we can write for all i that:

$$A_{it_2} - A_{it_1} \equiv \sum_{t=1}^{T} t D_{it} \tag{3}$$

This will lead to the problem of exact multicollinearity and make the estimation of coefficients impossible.

2.2 Restrictions on Coefficients

Early attempts at solving the multicollinearity problem focused mainly on separating the effects of age, time, and vintage (completion date of the building) in the hedonic model, which bears a very close relationship with the age-time effects in the repeat-sales model. Cagan (1971) and Hall (1971) suggested imposing some restrictions on the coefficients based on *a priori* information. For example, if it is known that the vintages of two consecutive models are identical, equality restrictions can be imposed on the coefficients representing their respective vintage. One simple way to set the restrictions is to normalize the coefficients of both periods. However, the disadvantages of this method are that it is not applicable when there is no *a priori* knowledge on vintages, and all the estimated coefficients will become biased if the restrictions are incorrect. Moreover, in most cases, this method only turns the problem of exact multicollinearity artificially into high collinearity, making the estimated coefficients highly unreliable.

2.3 Independent Estimates of Depreciation Rates

The first literature that dealt with the multicollinearity problem in the repeat-sales model was Chinloy (1977), but his method of estimating depreciation has been theoretically shown to be incorrect (Palmquist 1979). Emphasizing the need for external information to separate the time and age effects, Palmquist (1980, 1982) proposed to

adjust the price index manually by an independent estimate of depreciation from the hedonic model. Although this method appears arbitrary, it is reminiscent of the recently emerged hybrid technique, which simultaneously estimates the repeat-sales and hedonic regression models (Case and Quigley, 1991; Quigley, 1995; Hill, *et al.,* 1997; Englund, *et al.,* 1998).

2.3 Age Dummies

Cannaday, *et al.* (1998) also tried to tackle the problem in the repeat-sales model by breaking down the continuous age variable ranging from 0 to U years old into a series of age dummies. Their model can be written as:

$$\ln\left(\frac{P_{it_2}}{P_{it_1}}\right) = \sum_{t=1}^{T}\alpha_t D_{it} + \sum_{a=1}^{U}\delta_a Y_{ia} + \varepsilon_{it_1t_2} \tag{4}$$

where δ_a ($a=0,\ldots,U$) denotes the estimated coefficient for age dummy Y_{ia}. The age dummies are set to take the value –1 and +1 if a is the age of the property at the time of first sales and second sales, respectively, but zero if otherwise. Note that δ_0 has been set to zero so as to normalize the age index at age 0. Their original idea to avoid multicollinearity was to perform a non-linear transformation on age. This should be a feasible direction. Unfortunately, when they implemented their idea, the fact that multicollinearity is a problem of explanatory variables, but not their coefficients, was overlooked. Exact multicollinearity, therefore, cannot be avoided. Mathematically, the exact collinear relationship is:

$$\sum_{a=1}^{U}aY_{ia} \equiv \sum_{t=1}^{T}tD_{it} \tag{5}$$

2.2 Time-Age Interaction

Clapp and Giaccotto (1998) attempted to separate time and age by interacting age with time dummies. Their model can be written as:

$$\ln\left(\frac{P_{it_2}}{P_{it_1}}\right) = \sum_{t=1}^{T}\alpha_t D_{it} + \sum_{t=0}^{T}\gamma_t D_{it} A_{it} + \varepsilon_{it_1t_2} \tag{6}$$

where γ_t denotes the estimated coefficient for the interaction term. Their model in Equation (6) simply introduces a set of time-varying age coefficients, but not a transformation of the age variables. As a result, this method has the same problem as Cannaday, *et al.* (1998), and hence still results in exact multicollinearity. Mathematically, the exact collinear relationship of their model is as follows:

6

$$\sum_{t=0}^{T} D_{it} A_{it} \equiv \sum_{t=1}^{T} t D_{it} \tag{7}$$

3. NON-LINEAR AGE EFFECTS

3.1 Reasons for Non-Linearity

Estimation of the housing depreciation rate has commanded intensive research. Malpezzi, et al. (1987) carried out a comprehensive survey on 12 previous estimates (1942-1986) of depreciation rates. They ranged from 0.38% to 1.6%. However, different definitions of depreciation, models, and specifications are involved in these 12 studies. For example, Chinloy (1979) adopted semi-log hedonic models on housing units in Canadian cities, and found a 1% depreciation rate. Grether and Mieszkowski (1974) used the linear hedonic model on single family home sales in New Haven, CT, and found a 0.5% rate of depreciation. Cannaday and Sunderman (1986) considered the log-linear hedonic model on single family home sales in Champaign, IL and got 0.38-0.75%. Malpezzi, et al. (1987) also estimated the depreciation rates of housing in 59 metropolitan areas in the U.S. by way of the semi-log hedonic model with an age-squared variable. Their model catered for non-linear age effects, which allowed them to find depreciation rates of 0.43 – 0.93% per annum.

There are several reasons why age might affect property values in a non-linear manner. Physical deterioration itself may not be linearly related with age, especially when maintenance and renovation are ignored. Chinloy (1980) studied the impact of maintenance on housing depreciation rates, and found that the total depreciation rate was 1.52%, of which 37% was accounted for by maintenance for housing units in Canadian cities. Since maintenance expenditure is not independent of the age structure, but increases with a house's age, this implies a non-linear depreciation rate.

Chinloy (1978) earlier suggested that the more rapid depreciation rate in earlier years of a building should be observed in transacted properties because of the dominance of 'lemons' (poor quality products) in the transaction market. Hulten and Wykoff (1981) and Dixon, et al. (1999) also recognized that 'lemons' can pose a condition of non-linearity in depreciation measurement.

Functional obsolescence, or vintage effect, is another commonly-agreed upon reason for non-linearity. Randolph (1988), Wykoff (1989), and Yiu (2002) studied theoretically and empirically on this issue. Gallimore, et al. (1996), Waddell, et al. (1996), So, et al. (1997), and Tse and Love (2000) also explained their non-linear depreciation rate estimates by means of vintage effects. Goodman and Thibodeau (1997), on the other hand, argued that obsolescence results in heteroskedasticity. They, after controlling the dwelling-age-related heteroskedasticity, obtained depreciation rates of 1.9 – 4.05% for housing units in Dallas. Clapp and Giaccotto (1998) developed a model that explicitly captures the effects of obsolescence, and their estimated depreciation rates ranged from -0.30 – 0.10% for single family residential properties in Fairfax County, VA.

The above three reasons for non-linear age effects apply no matter if the land is freehold or leasehold. Yet, in dealing with leasehold land, there is an additional source of non-linearity: sensitivity to interest rates. Based on the duration analysis of bonds, we know that a longer-term bond is more price-sensitive to interest rate changes than a shorter-term bond (see Appendix for the derivation). Since leasehold properties have a definite expiration date, the repeat sales of a property must involve a change in the term to maturity (or time to lease expiration). Given that building age is negatively-related to the term to maturity of a leasehold interest, a younger building should be more price-sensitive to interest rate changes than an older building. Therefore, on top of the arguments on physical deterioration, functional obsolescence, and the lemon effect, we further add that interest rates should be taken into account when doing repeat sales analysis for leasehold properties. The mathematical exposition is given in the next section.

3.2 Mathematical Exposition

Although the mathematical exposition below appears to be based primarily on leasehold properties, there is nothing to prevent it from applying to freehold properties. This is because freehold properties can be considered conceptually as leasehold properties with infinite lease terms. Consequently, the implications drawn from a generalized exposition for leasehold properties should be directly applicable to the restrictive case of freehold properties.

Consider a leasehold property of age A. Given that the length of the lease of the property is T, where $T > A \geq 0$, the difference between T and A is the term to maturity.[5] Initially, we assume the net real income (R) of the property remains unchanged over the leasehold period, so as to isolate the effects of physical deterioration, functional obsolescence, and the lemon effect from the effects of shortened future income stream. This assumption is made for the ease of exposition, and will be relaxed later. Let the real rate of return be r. The (log) present value (V) of the property is the discounted sum of its expected net future income:[6]

$$V = \ln \int_0^{T-A} R \cdot e^{-rt} dt = \ln(R/r) + \ln\left[1 - e^{-r(T-A)}\right] \tag{8}$$

Since the partial derivative of V with respect to A is negative (Equation 9), the value of a property will decline as it ages over time (i.e. comes closer to the expiration of its lease), *ceteris paribus*. The expression in Equation (9) represents the *marginal* effect of age purely due to the shortening of future income stream as the property ages, which is not related to physical deterioration, functional obsolescence, or the lemon effect. This

---

[5] It is assumed that when the lease ends, the existing building will be scrapped. For freehold properties, T approaches infinity.
[6] The last term becomes zero for freehold properties.

marginal effect of age on property value (pure income stream effect) is a function of r, A, and T, as given by $g(r, A, T)$:[7]

$$\frac{\partial V}{\partial A} = \frac{r}{1 - e^{r(T-A)}} = g(r, A, T) \leq 0 \tag{9}$$

The overall marginal effects of age must take into account the physical deterioration, functional obsolescence, and the lemon effect as the building ages. Let this negative marginal age effect be a function of age only, given by $h(A) < 0$. It can be added to $g(.)$ to obtain the overall marginal age effects, $f(r, A, T)$, as shown in Equation (10). The overall marginal age effects are negative because both $g(.)$ and $h(.)$ are negative.

$$f(r, A, T) = g(r, A, T) + h(A) < 0 \tag{10}$$

Two important points can be inferred from Equation (10). First, the combined depreciation tends to be non-linear in age because the overall marginal age effects, $f(.)$, are a function of A. This is true even if the depreciation due to physical deterioration, functional obsolescence, and the lemon effect is linear (i.e. $h(.)$ is a constant).[8] Equation (11a) gives the partial derivative of $f(.)$ with respect to A. The sign of $g'(A)$ is negative, but that of $h'(A)$ is unknown. Consequently, the second-order effect of age on the property value has no *a priori* sign.

$$\frac{\partial f}{\partial A} = g'(A) + h'(A) = \frac{-r^2 e^{r(T-A)}}{\left[1 - e^{r(T-A)}\right]^2} + h'(A) \tag{11a}$$

The second point drawn from Equation (10) is that for leasehold properties, the combined marginal age effects are not simply a function of age, but also of r and T. Equations (11b) and (11c) denote the partial derivative of $f(.)$ with respect to r and T, respectively. They show that the combined marginal age effects increase as r or T increases, and hence their joint effects with age are positive. *Fig. 1* depicts the relationship between property values, age, and real rates of return graphically. It indicates that the value difference between two properties (one property of age $A+\Delta A$ and the other of age A) is smaller when the real rate of return is higher.

---

[7] $g(r, A, T) = 0$ for freehold properties. This means freehold property values are not affected by the pure income stream effect.

[8] However, in case of freehold properties, $h(.)$ must not be a constant, but a function of age in order to give rise to non-linear age effects.

$$\frac{\partial f}{\partial r} = \frac{\partial g}{\partial r} = \frac{1 + [r(T-A) - 1]e^{r(T-A)}}{[1 - e^{r(T-A)}]^2} > 0 \qquad (11b)$$

$$\frac{\partial f}{\partial T} = \frac{\partial g}{\partial T} = \frac{r^2 e^{r(T-A)}}{[1 - e^{r(T-A)}]^2} > 0 \qquad (11c)$$

Fig. 1: The Relationship between Property Values, Age, and Real Rates of Return
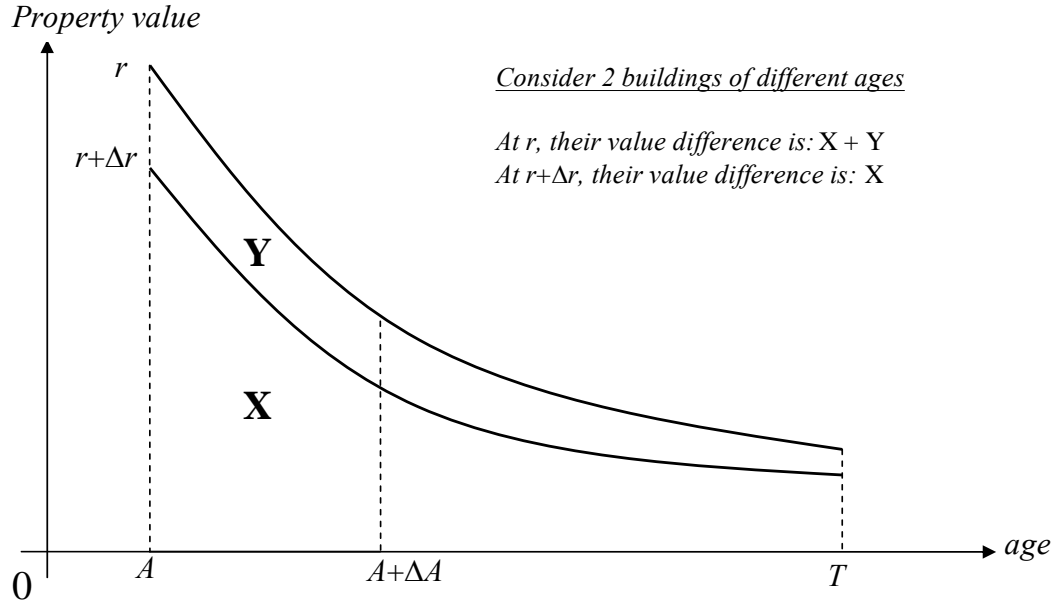
4. THE AGE-ADJUSTED REPEAT-SALES MODEL

4.1 Empirical Estimation of Age Effects

In the previous section, we argued that the age effects should be non-linear irrespective of the type of land tenure. Specifically, we showed that building age not only accounts for physical deterioration, functional obsolescence, and the lemon effect, but also for the residual income stream of a leasehold property. The mathematical exposition further suggests that the real rate of return and term to maturity should produce positive joint effects with age. Based on these theoretical arguments, it is possible to specify a simpler empirical function for age effects for estimation. Let the age affects be $F(r, A, T)$, which is a primitive function of $f(.)$. The non-linearity of age effects can be modelled and approximated by highly flexible Box-Cox functions, $y^{(\lambda)} = (y^{\lambda}-1)/\lambda$, where y can be any strictly positive variable. The joint effect between age and the real rate of return (or term to maturity) can be represented by their interaction term. For

10

simplicity, we assume that any other higher orders' compounded marginal effects with age are negligible, and the terms in $F(.)$ are additive. This produces the following empirical function for the age effects:

$$F(r, A, T) = \beta_0 + \beta_1 A^{(\lambda_1)} r^{(\theta_r)} + \beta_2 A^{(\lambda_2)} + \beta_3 A^{(\lambda_3)} T^{(\theta_T)} \tag{12a}$$

where θ_r, θ_T, λ_i ($i=1,2,3$), and β_j ($j=0,1,2,3$) are coefficients. Flexibility is maintained by applying Box-Cox transformation to each variable. It is expected that β_1 and β_3 are positive (from Equations 11b and 11c), and β_2 is negative (from Equation 10 and assuming positive β_1 and β_3). There is no *a priori* prediction for the sign of the Box-Cox parameters (i.e. θ_r, θ_T, λ_1, λ_2, and λ_3). For freehold properties, β_1 and β_3 should be set to zero:

$$F(A) = \beta_0 + \beta_2 A^{(\lambda_2)} \tag{12b}$$

4.2 The Repeat-Sales Model

It is useful to start with the hedonic pricing model because the repeat-sales model can be derived from the hedonic pricing model (Clapp and Giaccotto, 1992). Incorporating the empirical function for the age effects in Equation (12a) into the hedonic pricing model gives rise to the following specification for leasehold properties:

$$\ln P_{it} = c + \sum_{j=1}^{J} \delta_j Q_{jit} + \sum_{t=1}^{K} \alpha_t D_{it} + \beta_1 A_{it}^{(\lambda_1)} r_t^{(\theta_r)} + \beta_2 A_{it}^{(\lambda_2)} + \beta_3 A_{it}^{(\lambda_3)} T_i^{(\theta_T)} + \varepsilon_{it} \tag{13a}$$

where P_{it} ($i=1,\ldots,n$; $t=0,\ldots,K$) is the transaction price for property i at time t; Q_{jit} ($j=1,\ldots,J$) are the property attributes except age; D_{it} are the time dummies that are set to 1 at the time of sale, and 0 if otherwise; A_{it} is the property age; r_t is the real rate of return; T_i is the length of leasehold interests; c is the intercept (including β_0); and δ_j, α_t, β_1, β_2, β_3, θ_r, θ_T, λ_1, λ_2, and λ_3 are coefficients.

There are two possible variations for Equation (13a). First, if the length of leasehold interests are the same for all properties (i.e. $T_i = T$ for all i), then the interaction term between A_{it} and T_i is reduced to a function of age only. In such a case, it can be reasonably assumed that its effect can be absorbed by the Box-Cox transformed age term, so that it need not be included. This produces Equation (13b). The original coefficient of the Box-Cox transformed age term, β_2, is changed to β_{23}, which is expected to be negative according to Equation (10).

11

$$\ln P_{it} = c + \sum_{j=1}^{J} \delta_j Q_{jit} + \sum_{t=1}^{K} \alpha_t D_{it} + \beta_1 A_{it}^{(\lambda_1)} r_t + \beta_{23} A_{it}^{(\lambda_{23})} + \varepsilon_{it} \qquad (13b)$$

Second, if the properties are freehold, then the coefficient β_1 should be set to zero, resulting in:

$$\ln P_{it} = c + \sum_{j=1}^{J} \delta_j Q_{jit} + \sum_{t=1}^{K} \alpha_t D_{it} + \beta_{23} A_{it}^{(\lambda_{23})} + \varepsilon_{it} \qquad (13c)$$

The repeat-sales model is simply a differenced version of Equation (13), in which the order of difference is determined by the period between each repeat pair of transactions (t_2-t_1). Assume that all property characteristics other than age remain unchanged between the repeat sales of a property (i.e. $Q_{jit} = Q_{ji}$ for all t), and the repeat-sales specifications for varying lease terms (T_i), for constant lease terms (T), and for infinite lease terms are shown in Equations (14a), (14b), and (14c), respectively. As we can see, an exact collinear relationship between age and time does not exist as long as $\lambda_2 \neq 1$ in Equation (14a) or $\lambda_{23} \neq 1$ in Equations (14b) and (14c).[9]

$$\ln\left(\frac{P_{it_2}}{P_{it_1}}\right) = \sum_{t=1}^{K} \alpha_t D_{it} + \beta_1\left(A_{it_2}^{(\lambda_1)} r_{t_2}^{(\theta_r)} - A_{it_1}^{(\lambda_1)} r_{t_1}^{(\theta_r)}\right) + \beta_2\left[A_{it_2}^{(\lambda_2)} - A_{it_1}^{(\lambda_2)}\right] +$$
$$\beta_3 T_i^{(\theta_T)}\left(A_{it_2}^{(\lambda_3)} - A_{it_2}^{(\lambda_3)}\right) + \varepsilon_{it_1 t_2} \qquad (14a)$$

$$\ln\left(\frac{P_{it_2}}{P_{it_1}}\right) = \sum_{t=1}^{K} \alpha_t D_{it} + \beta_1\left(A_{it_2}^{(\lambda_1)} r_{t_2}^{(\theta_r)} - A_{it_1}^{(\lambda_1)} r_{t_1}^{(\theta_r)}\right) + \beta_{23}\left[A_{it_2}^{(\lambda_{23})} - A_{it_1}^{(\lambda_{23})}\right] + \varepsilon_{it_1 t_2} \qquad (14b)$$

$$\ln\left(\frac{P_{it_2}}{P_{it_1}}\right) = \sum_{t=1}^{K} \alpha_t D_{it} + \beta_{23}\left[A_{it_2}^{(\lambda_{23})} - A_{it_1}^{(\lambda_{23})}\right] + \varepsilon_{it_1 t_2} \qquad (14c)$$

Since the coefficients to be estimated are not linear, ordinary least squares (OLS) estimation is not directly applicable. However, the maximum likelihood estimation (MLE) technique can be applied to obtain consistent estimates for the unknown parameters. Alternatively, a two-stage estimation procedure can be used. First, the non-linear model is estimated by the MLE technique to select the Box-Cox parameters that fit

---

[9] In fact, Equations (13a), (13b), and (13c) can also provide a solution for estimating the vintage effects in the hedonic pricing model. However, a large number of property attributes is required, and so we have to leave this application for future research.

the data best. Then, by treating the Box-Cox parameters as given, the model becomes linear in coefficients and is re-estimated by the OLS technique.

5. EMPIRICAL DATA AND RESULTS

5.1 Hong Kong Data

The residential property market in Hong Kong has been very active over the last decade. This provides us with a rich source of data. We selected a residential district called North Point to carry out the empirical analysis. Over 11,000 pairs of repeat transactions of buildings, ranging from Q2-1991 through Q1-2001, were included in our sample, which is a very large sample size compared to all previous repeat-sales studies. Virtually all lands in Hong Kong are leasehold.

Table 1 shows the descriptive statistics of the data used. The average log price change of repeated transactions is about 12.6%.[10] The ages of the transacted properties ranged from 0.25 years to 43.75 years.[11] The *ex post* long term real rate of interest, which is the yield on 10-year Exchange Fund notes,[12] less the inflation rate,[13] is considered a proxy for the real rate of return. Since the real rate of interest is not strictly positive, Box-Cox transformation is not applied to r, and θ_r is assumed to be one. The tenures of 72 lots located in the sampled district were surveyed, and it was found that the distribution of the lease tenure of the lots was well-confined to about 99 years. Therefore, the lease term can be regarded as fixed in this sample, and the age-adjusted repeat-sales model in Equation (14b) will be adopted. A repeat-sales index based on the BMN model in Equation (1) will also be estimated for comparison purposes.

Table 1. Descriptive Statistics of the Data Used

Description	Variable	Mean	Standard Deviation	Minimum	Maximum
Log price change	$\ln(P_{it2}/P_{it1})$	12.6%	32.7%	-150.8%	178.4%
Age (in years)	A_{it1}	13.19	7.65	0.25	43.75
Real rate of return	r_t	4.6%	5.3%	-2.2%	14.6%

---

[10] The distortion of the repeat-sales price index due to the correlation between rent and price growth can be ignored in this sample because the correlation was found to be very low (r=-0.19).

[11] In estimation, the age variable is measured in quarters, not years.

[12] Since the 10-year Exchange Fund notes of Hong Kong were only available after 1996, the yields before this period are estimated from U.S. long term interest rates. Under the pegged currency exchange system with the U.S. Dollar, the extrapolation is based on the 10-year US Treasuries, with an adjustment made for the average gap between the yields in HK and the U.S. during the period 1997-2002.

[13] The inflation rate is the year-by-year change of the implicit price deflator of the GDP of Hong Kong.

5.2 Discussion of Results

To estimate Equation (14b) by the MLE technique, we assume the error term $\varepsilon_{it_1 t_2}$ is independent, drawing from a normal distribution with a mean of zero and a variance of σ^2. The estimated results of λ_1, λ_{23}, β_1, and β_{23} are shown in *Table 2*. The positive sign of β_1 implies that the joint effect of age and the real rate of return is positive. This result is consistent with the theoretical relationship that real rates of return are positively-related to the marginal age effect. The negative sign of β_{23} confirms that age has a negative effect on property prices. Moreover, the values of λ_1 and λ_{23} suggest that age effects are non-linear, and are approximately a square-root function of age.

Table 2. MLE Results of the Age-Adjusted Repeat-Sales Model

	Coefficient	Standard Error	*p*-value
λ_1	0.667	1.062	0.530
λ_{23}	0.515	0.394	0.191
β_1	0.009	0.041	0.827
β_{23}	-0.010	0.005	0.055
log-likelihood value: 5,251.9			

While β_1 and β_{23} have expected signs, and λ_1 and λ_{23} have plausible magnitudes, they were not significant at the 5% level. This is due to the loss in degrees of freedom for a highly flexible model. Since the coefficients of primary interest are β_1 and β_{23}, it is worthwhile to place restrictions on λ_1 and λ_{23}. Moreover, if λ_1 and λ_{23} are restricted to some specific values, the model becomes linear and can be estimated by the standard OLS technique.

In view of the above motivations, some special case models (i.e. λ_1=-1, 0, 0.5, 1, 2 and λ_{23}=-1, 0, 0.5, 2) are tested against the above optimal age-adjusted model using log-likelihood ratio (*LR*) test. If a special case model is not significantly different from the optimal Box-Cox model, it can be used as a good approximation for practical purposes. The test statistics is:

$$LR = 2[MLO - ML(\lambda)] \tag{17}$$

where *MLO* is the maximum log-likelihood value of the optimal model, and *ML(λ)* is the maximum log-likelihood value of a special case model. Since only two parameters are to be restricted, *LR* follows a Chi-squared distribution with two degrees of freedom, and the critical value at the 5% level is 5.991. The results of the *LR* test are reported in *Table 3*.

Table 3. LR Test Results of the Age-Adjusted Repeat-Sales Model

Special Case λ_1	Special Case λ_{23}	Log-Likelihood Value $ML(\lambda)$	Test Statistics LR
-1	-1	5,235.79	32.16
-1	0	5,246.06	11.62
-1	0.5	5,249.78	4.19*
-1	2	5,242.41	18.93
0	-1	5,244.17	15.41
0	0	5,248.77	6.21
0	0.5	5,250.95	1.85*
0	2	5,246.37	11.01
0.5	-1	5,245.96	11.82
0.5	0	5,250.32	3.10*
0.5	0.5	5,251.81	0.12*
0.5	2	5,246.62	10.51
1	-1	5,245.20	13.35
1	0	5,250.36	3.02*
1	0.5	5,251.71	0.33*
1	2	5,245.20	13.35
2	-1	5,241.23	21.28
2	0	5,248.58	6.59
2	0.5	5,250.52	2.71*
2	2	5,242.49	18.76

* The constrained model cannot be rejected at the 5% level

In *Table 3*, several constrained models are not significantly different from the optimal model at the 5% level. Since the log-likelihood value at λ_1=0.5 and λ_{23}=0.5 is the highest (with *LR*=0.12), a square-root function is chosen to approximate the non-linear effect of age. We can then estimate a linear form of the age-adjusted repeat-sales model using OLS as if the values of λ_1 and λ_{23} were known:

$$\ln\left(\frac{P_{it_2}}{P_{it_1}}\right) = \sum_{t=1}^{K}\alpha_t D_{it} + \beta_1\left(A_{it_2}^{0.5}r_{t_2} - A_{it_1}^{0.5}r_{t_1}\right) + \beta_{23}\left[A_{it_2}^{0.5} - A_{it_1}^{0.5}\right] + \varepsilon_{it_1 t_2} \qquad (15)$$

Table 4. OLS Results of the Age-Adjusted Repeat-Sales Model with λ_1=0.5 and λ_{23}=0.5

	Coefficient	Standard Error	*p*-value
β_1	0.036	0.017	0.041
β_{23}	-0.020	0.006	0.000
R-squared value: 0.785			

The OLS results of β_1 and β_{23} in Equation (15) are presented in *Table 4*. In line with the MLE results, the signs of β_1 and β_{23} are consistent with theoretical predictions. More importantly, their values become significant at the 5% level. The resulting age-adjusted repeat-sales index, α_t, is plotted in *Fig.2*. Apart from the age-adjusted index, two other indices are also included for comparison. One is the price index published by the Rating and Valuation Department (RVD) of the Hong Kong SAR Government (the R&V index)[14]. It is an index for the entire residential market of Hong Kong, rather than for North Point only. The other price index is generated by the traditional BMN repeat-sales model in Equation (1) (the BMN index). The price levels for all indices in the second quarter of 1991 are set to 1.

From *Fig. 2*, the age-adjusted index and the R&V index share a very similar trend. This confirms that the estimates of the age-adjusted index are reasonable vis-à-vis the overall residential property market in Hong Kong. Most importantly, when the age-adjusted index is compared to the BMN index, it is found that the BMN index is persistently lower than the age-adjusted index. This implies that the age-adjusted repeat-sales model makes upward adjustments for negative age effects, which the BMN model cannot handle. The gap, therefore, represents the underestimation of property returns by the BMN index. *Table 5* provides the return statistics of each index. It shows that the magnitude of underestimation averages about 0.6% per annum.

---

[14] The R&V indices are the "official" set of real estate indices. These indices control quality by a mass assessment of the net annual rental income (Rateable Value). See Chau (1997) and Brown and Chau (1997) for a more detail description of the indices.

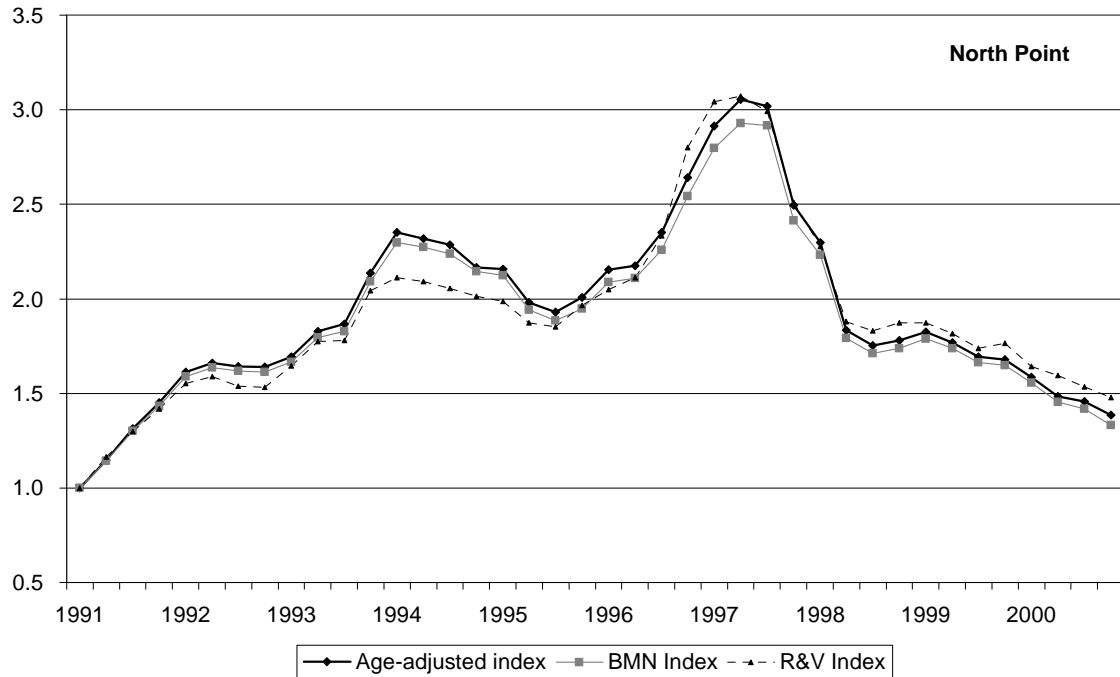*Fig. 2. Residential Property Price Indices for North Point*

Table 5. Annualized Returns of the Residential Property Market

	Mean	**Standard Deviation**
R&V Index	9.06%	34.15%
BMN Index	7.94%	32.45%
Age-adjusted Index	8.53%	33.04%

6. CONCLUSIONS

The problem of multicollinearity between time and age difference in the traditional repeat-sales model has been addressed and an age-adjusted solution has been proposed. Our solution is founded on the recognition that the age effects are non-linear in general and age bears a close relationship with the term to maturity for leasehold properties in particular. Through analysing the residential market in Hong Kong, we have demonstrated how the proposed method can adjust for the age effects, which are ignored in the traditional BMN method. The estimated depreciation rate of 0.6% per annum in this study falls within the range of most of the relevant studies.

Due to data availability, the empirical test in this study is limited to residential leasehold interests only. Further tests on our model can be extended on one hand to

freehold interests, and on the other hand to the commercial or office sectors. The theoretical model also relies on the validity of the assumptions of income approach in property valuation. This study, because of the lack of information, also ignored quality changes due to maintenance and renovation. In other words, the estimate is an age-related depreciation net of maintenance efforts.

The resulting age-adjusted index should be useful in providing a clearer picture of the performance of a property market. The same approach can be applied to the hedonic model to estimate vintage effects, although an empirical test for this is best left to further studies.

ACKNOWLEDGEMENTS

REFERENCES

Bailey, M.J., Muth, R.F. and Nourse, H.O. (1963). "A regression model for real estate price index construction", *Journal of the American Statistical Association,* **58**, 933-942.

Cagan, P. (1971). "Measuring quality changes and the purchasing power of money: an exploratory study of automobiles", in *Price Indexes and Quality Change*, (ed. Z. Griliches), Harvard University Press, Cambridge. pp. 215-239.

Cannaday, R.E. and Sunderman, M.A. (1986). "Estimation of Depreciation for Single-Family Appraisals", *AREUEA Journal*, **14**, Summer, 255-273.

Cannaday, R.E., Munneke, H.J. and Yang, T.T. (1998). "An age-adjusted repeat sales house price index", *Discussion Paper*, Department of Finance, University of Illinois, Urbana.

Case, B. and Quigley, J.M. (1991). "The dynamic of real estate prices", *Review of Economics and Statistics,* **22**, 50-58.

Case, B., Pollakowski, H.O. and Wachter, S.M. (1991). "On choosing among house price index methodologies", *AREUEA Journal,* **19**(3), 286-307.

Case, K.E. and Shiller, R.J. (1987). "Prices of single family homes since 1970: new indexes for four cities", *New England Economic Review,* **Sept/Oct**, 45-56.

Chau, K. W. (1997). "Political uncertainty and the real estate risk premiums in Hong Kong", *Journal of Real Estate Research*, **13** (3), 297-315.

Brown, G. R. and Chau, K. W. (1997). "Excess returns in the Hong Kong commercial real estate market", *Journal of Real Estate Research*, **14** (2), 91-105.

Chinloy, P.T. (1977). "Hedonic price and depreciation indexes for residential housing: a longitudinal approach", *Journal of Urban Economics,* **4**, 469-482.

Chinloy, P.T. (1978). "Depreciation, Adverse Selection and Housing Markets", *Journal of Urban Economics*, **5**, 172-187.

Chinloy, P.T. (1979). "The Estimation of Net Depreciation Rates on Housing", *Journal of Urban Economics*, **6**, 432-443.

Chinloy, P.T. (1980). "The Effect of Maintenance Expenditures on the Measurement of Depreciation in Housing", *Journal of Urban Economics*, **8**, 86-107.

Clapp, J.M. and Giaccotto, C. (1992). "Estimating price indices for residential property: a comparison of repeat sales and assessed value methods", *Journal of the American Statistical Association,* **87**(418), 300-306.

Clapp, J.M. and Giaccotto, C. (1998). "Residential hedonic models: a rational expectations approach to age effects", *Journal of Urban Economics,* **44**, 415-437.

Dixon, T.J., Crosby, N. and Law, V.K. (1999). "A Critical Review of Methodologies for Measuring Rental Depreciation applied to UK Commercial Real Estate", *Journal of Property Research*, **16**(2), 153-180.

Englund, P., Quigley, J.M. and Redfearn, C.L. (1998). "Improved price indexes for real estate: measuring the course of Swedish housing prices", *Journal of Urban Economics,* **44**, 171-196.

Gallimore, P.M., Fletcher, M. and Carter, M. (1996). "Modelling the influence of location on value", *Journal of Property Valuation and Investment*, **14**(1), 6-19.

Goodman, A.C. and Thibodeau, T.G. (1995). "Age-related Heteroskedasticity in Hedonic House Price Equations", *Journal of Housing Research*, **6**(3), 25-42.

Greenlees, J.S. (1982). "An empirical evaluation of the CPI Home Purchase Index, 1973-1978", *AREUEA Journal,* **10**, 1-24.

Grether, D.M. and Mieszkowski, P. (1974). "Determinants of Real Values", *Journal of Urban Economics*, **1**, 127-145.

Hall, R.E. (1971). "The measurement of quality change from vintage price data", in *Price Indexes and Quality Change*, (ed. Z. Griliches), Harvard University Press, Cambridge. pp. 240-271.

Hill, R.C., Knight, J.R. and Sirmans, C.F. (1997). "Estimating capital asset prices", *Review of Economics and Statistics,* **79**, 226-233.

Hopewell, M.H. and Kaufman, G.G. (1973). "Bond price volatility and term to maturity: a generalized respecification", *American Economic Review,* **63**(4), 749-753.

Hulten, C.R. and Wykoff, F.C. (1981). "The Estimation of Economic Depreciation using Vintage Price Assets", *Journal of Econometrics*, **15**, 367-396.

Macaulay, F.R. (1938). *Some theoretical problems suggested by the movement of interest rates, bond yields, and stock prices in the United States since 1856,* National Bureau of Economic Research, New York.

Malpezzi, S., Ozanne, L. and Thibodeau, T.G. (1987). "Microeconomic estimates of housing depreciation", *Land Economics,* **63**(4), 372-375.

Mark, J.H. and Goldberg, M.A. (1984). "Alternative housing price indices: an evaluation", *AREUEA Journal,* **12**(1), 30-49.

Meese, R.A. and Wallace, N.E. (1991). "Nonparametric estimation of dynamic hedonic price models and the construction of residential housing price indices", *AREUEA Journal,* **19**, 308-332.

Meese, R.A. and Wallace, N.E. (1997). "The construction of residential housing prices indices: a comparison of repeat-sales, hedonic-regression, and hybrid approaches", *Journal of Real Estate Finance and Economics,* **14**, 51-73.

Palmquist, R.B. (1979). "Hedonic price and depreciation indexes for residential housing: a comment", *Journal of Urban Economics,* **6**, 267-271.

Palmquist, R.B. (1980). "Alternative Techniques for developing real estate price indexes", *Review of Economics and Statistics,* **66**, 394-404.

Palmquist, R.B. (1982). "Measuring environmental effects on property values without hedonic regressions", *Journal of Urban Economics,* **11**, 333-347.

Quigley, J.M. (1995). "A simple hybrid model for estimating real estate price indices", *Journal of Housing Economics,* **4**, 1-12.

Randolph, W.C. (1988). "Estimation of Housing Depreciation: Short-term Quality Change and Long-term Vintage Effects", *Journal of Urban Economics*, **23**, 162-178.

So, H.M., Tse, R.Y.C. and Ganesan, S. (1997). "Estimating the influence of transport on house prices: evidence from Hong Kong", *Journal of Property Valuation and Investment*, **15**(1), 40-47.

Tse, R.Y.C. and Love, P.E.D. (2000). "Measuring residential property values in Hong Kong", *Property Management*, **18**(5), 366-374.

Waddell, P., Berry, B.J.L. and Chung, K.S. (1996). "Variations in Housing Price Depreciation: The Taste for Newness Across Heterogeneous Submarkets", *Urban Geography*, **17**(3), 269-280.

Wykoff, F.C. (1989). "Economic Depreciation and the User Cost of Business-leased Automobile", in Jorgenson, D.W. and Landau, R. (eds), *Technology and Capital Formation*, The MIT Press, Cambridge, London.

Yiu, C.Y. (2002). "The Effects of Age on Housing Prices in Hong Kong", *Unpublished PhD. Thesis*, Department of Real Estate and Construction, The University of Hong Kong, Hong Kong.

APPENDIX – DURATION ANALYSIS

It is well established that duration (D) measures the sensitivity of a bond's price (P) to a small change in the interest rate (r) (Macaulay, 1938). In contrast to the term to maturity (T), duration calculates the time structure of a bond by considering the overall cash flow pattern instead of only the time when the last cash flow occurs (Hopewell and Kaufman, 1973).[15] It weighs the time periods by the present value of the corresponding cash flow (C) relative to the price. Equation (A1) shows how the duration of a bond is related to its price, cash flow pattern, term to maturity, and the interest rate.

$$\frac{dP/P}{dr} = -D = -\frac{\int_0^T tCe^{-rt}dt}{P} \tag{A1}$$

There are several implications from Equation (A1). First, since duration is positive, a rise in the interest rate (cost of capital) will depress the bond price. Second, the larger the duration, the more sensitive is the bond price to interest rate movements. Third, for a given interest rate and cash flow pattern, duration bears a positive (though not linear) relationship with the term to maturity. It follows that a longer-term bond is more price-sensitive to interest rate changes than a shorter-term bond (Equation A2).

$$\frac{\partial}{\partial T}\left|\frac{\partial P/P}{\partial r}\right| = \frac{\partial D}{\partial T} > 0 \tag{A2}$$

---

[15] A zero coupon bond is a special case that the duration equals the term to maturity.