Considerations in the Design & Construction of Investment Real Estate Research Indices

David Geltner & David Ling

(Forthcoming: Journal of Real Estate Research)

Presentation for National University of Singapore

Department of Real Estate

David Geltner

July, 2006

"In summary, we argue that the NCREIF Index is ready to evolve into two more specialized successor families of index products: one tailored for fundamental asset class research support, and the other tailored for investment performance evaluation benchmarking and performance attribution."

-- From: D.Geltner & D.Ling, <u>Benchmarks & Index Needs in the U.S. Private Real Estate Investment Industry: Trying to Close the Gap</u> (A RERI Study for the Pension Real Estate Association), October 17, 2000.

• Development of large-scale electronic databases of commercial property transaction prices (e.g., *RCA*, *CoStar*, *NCREIF*, *IPD*).

• Advance of econometric techniques for index construction from such data (e.g., *BMN*, *Griliches*, *Rosen*, *Case-Shiller*, *Clapp-Giacotto*, *Gatzlaff-Haurin*, *FGGH*, *Goetzmann*, *etc.*).

➔ **Time is ripe for developing transactions-based indices of commercial property price and investment performance.**

Four Determinants of Index *Dynamic Quality*:

1. Index return reporting frequency *(m* = index reporting periods per year)*;*

2. Frequency of revaluation observations per property (f = valuation obs/yr per property);

3. Number of properties in the underlying population tracked by the index (p): ➔ Data density = Number of valuation obs per index reporting period = $n = pf / m$;

4. Index construction "technology," or methodology used to construct the index from the underlying valuation observations.

Exhibit 1a: The General Trade-off Between Index Statistical Quality Per Period, and the Index Reporting Frequency
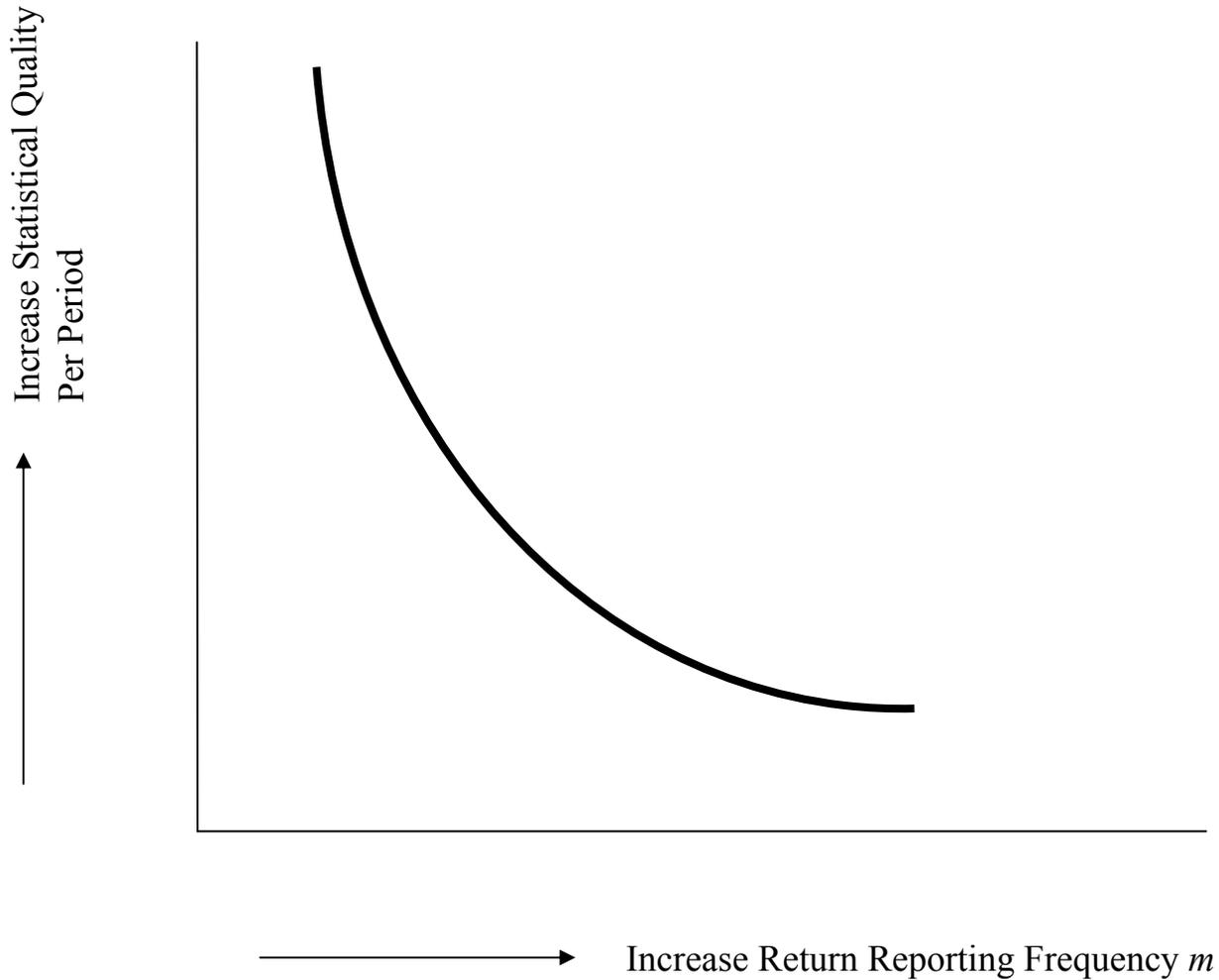
Exhibit 1b: Index Utility Isoquants: Index Utility Increases With Both Statistical Quality Per Period and Reporting Frequency: $U_0 < U_1 < U_2$; With Diminishing Marginal Substitutability (convex isoquants)…
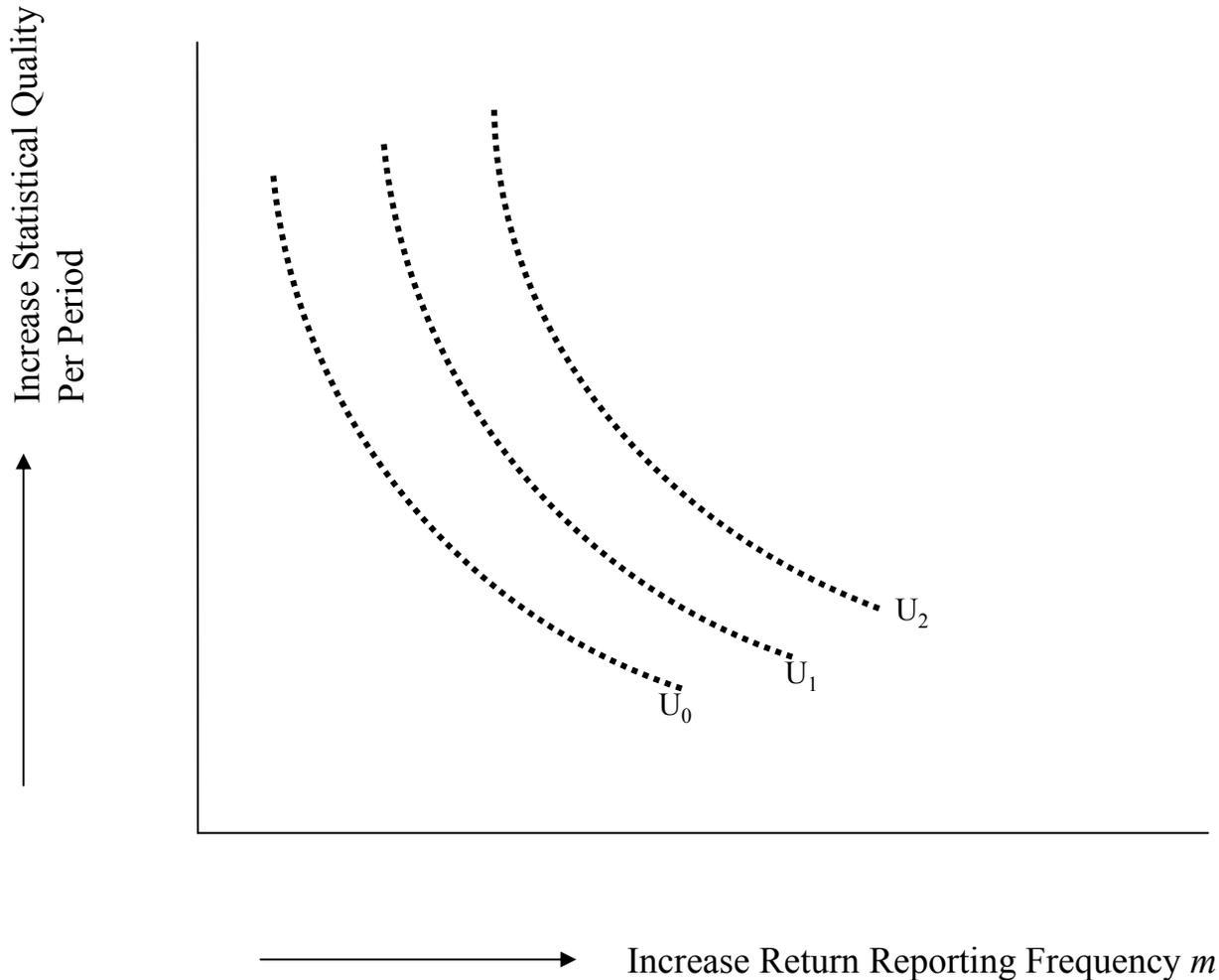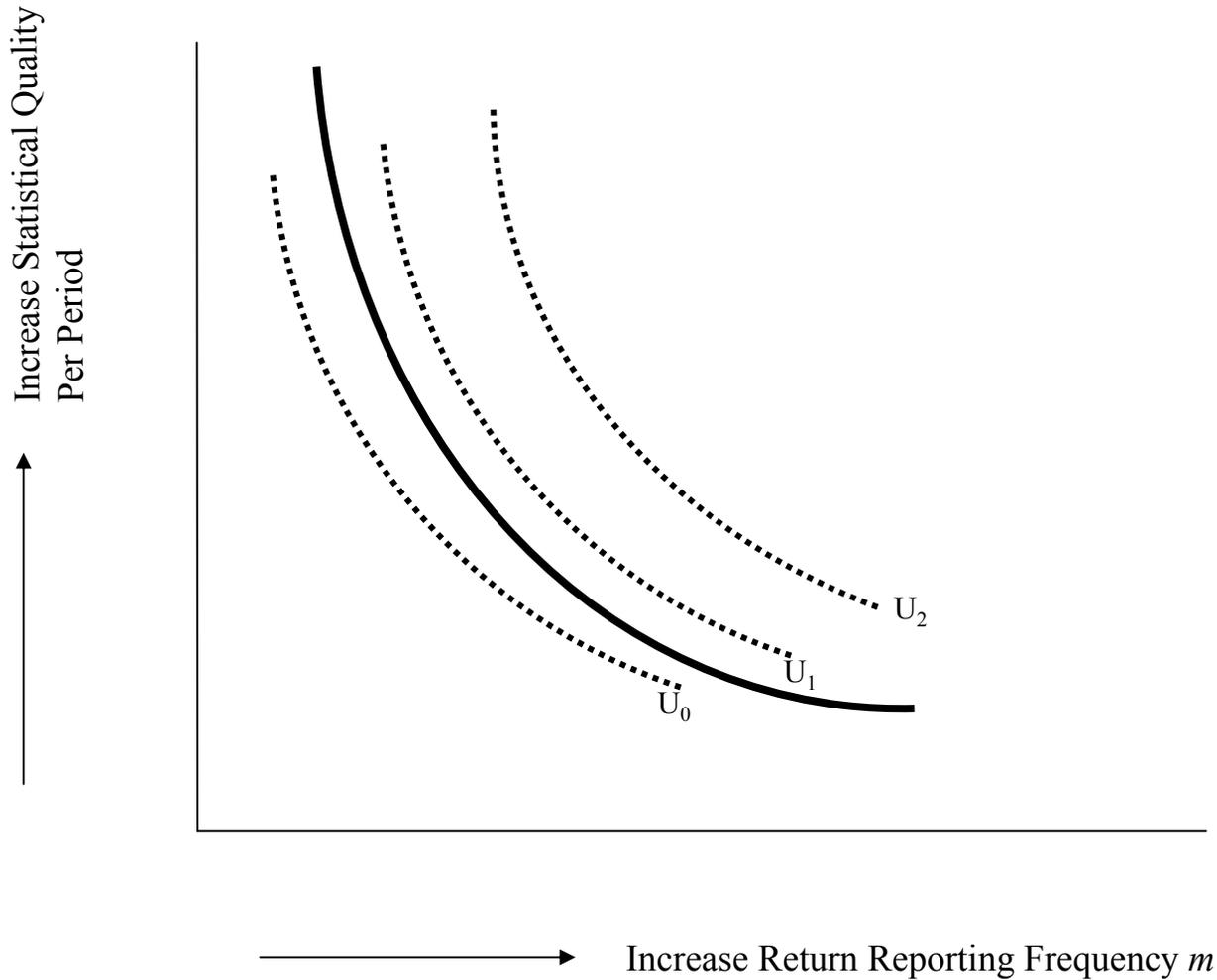
Exhibit 1c: Optimal Index Reporting Frequency Is Ambiguous or Not Unique . . .

Increase Statistical Quality Per Period

U_2

U_1

U_0

Increase Return Reporting Frequency m

$P_{i,t}$ = Transaction price of property i in index reporting period t.

i = 1, 2, . . ., n ; where n = Number of properties sold in each index reporting period.

Each property is "representative" of population (identical).

$$P_{i,t} = V_t + \varepsilon_{i,t}$$

V_t = True (population) market value in period t.

$\varepsilon_{i,t}$ = Random error (iid normal)

$\hat{V}_t[N]$ = Arithmetic average across N of the transaction prices starting with transactions that are among the n that occur within period t ;

N / n = an integer $>= 1$.

$$\hat{V}_t[N] = \frac{1}{N} \sum_{i=1}^{N} P_{i,t}$$

$$E[\hat{V}_t[N]] = \frac{1}{N/n} \sum_{s=0}^{N/n-1} V_{t-s} = \frac{1}{L+1} \sum_{s=0}^{L} V_{t-s}$$

A simple stylized model of optimal traditional (individual property) appraisal & optimal aggregate (transactions-based) index estimation

$$MSE = E\left[\left(\hat{V}_t[N] - V_t\right)^2\right]$$

$$= E\left[\left(\frac{1}{N}\left(\sum_{i=1}^{n}\left(P_{i,t} - V_t\right) + \sum_{i=n+1}^{2n}\left(P_{i,t-1} - V_t\right) + \ldots + \sum_{i=N-n+1}^{N}\left(P_{i,t-L} - V_t\right)\right)\right)^2\right]$$

$$= \frac{1}{N^2}\left(\sum_{i=1}^{N}\varepsilon_i^2 + n^2\sum_{s=1}^{N/n-1}\left(V_{t-s} - V_t\right)^2\right) = \frac{1}{(n(L+1))^2}\left(\sum_{i=1}^{n(L+1)}\varepsilon_i^2 + n^2\sum_{s=1}^{L}\left(V_{t-s} - V_t\right)^2\right)$$

$$MSE = \frac{VAR[\varepsilon_i]}{N} + \frac{n^2 VAR[r_t]}{N^2}\sum_{s=1}^{N/n-1}s^2 = \frac{\sigma_\varepsilon^2}{N} + \frac{n^2\sigma_r^2}{N^2}\sum_{s=1}^{N/n-1}s^2 = \frac{\sigma_\varepsilon^2}{n(L+1)} + \frac{\sigma_r^2}{(L+1)^2}\sum_{s=1}^{L}s^2$$

Where L = Maximum lag in value estimation sample.

MSE as a function of lag: $\longrightarrow$

$$\frac{\Delta MSE}{\Delta L} = \left(\frac{\sum_{s=1}^{L+1} s^2}{L+2} - \frac{\sum_{s=1}^{L} s^2}{L+1} \right) \sigma_r^2 - \left(\frac{1}{n(L+1)(L+2)} \right) \sigma_\varepsilon^2$$

Optimal valuation methodology rule (minimizing *MSE*): $\longrightarrow$

$$\frac{\Delta MSE}{\Delta L} = 0, \Rightarrow$$

$$\left(\frac{\sum_{s=1}^{L+1} s^2}{L+2} - \frac{\sum_{s=1}^{L} s^2}{L+1} \right) \sigma_r^2 = \left(\frac{1}{n(L+1)(L+2)} \right) \sigma_\varepsilon^2, \Rightarrow$$

$$\left(\frac{\sum_{s=1}^{L+1} s^2}{L+2} - \frac{\sum_{s=1}^{L} s^2}{L+1} \right) (L+1)(L+2) = \frac{\sigma_\varepsilon^2}{n \sigma_r^2}$$

Optimal lag is:

• Increasing function of the price dispersion σ_ε , and

• Decreasing function of both the transaction density n and the market return volatility σ_r

• A function only of the <u>*ratio*</u> of the two dispersion parameters: $\sigma_\varepsilon / \sigma_r$, not of the absolute values of these parameters.

This model is essentially a discrete time version of Quan and Quigley's (1991) continuous time optimal appraisal model, except that the Quan-Quigley model was conditioned on the existence of a prior value estimate by the same appraiser (optimal updating), whereas the present model is an unconditional optimization, not assuming any prior value estimation.

By working in discrete time, we can address important index policy issues, such as index reporting frequency and transaction price sample size.

Our model here is also very similar to, and our results here are consistent with, those in Giaccotto and Clapp (1992) and Geltner (1997b). Giaccotto & Clapp explored a broader range of appraisal rules and allowed for serial correlation in the market value returns. Geltner (1997b) employed a numerical simulation rather than an analytical model, had declining appraiser weights on comps with the age of the comp, and also relaxed the random walk assumption of the underlying true real estate returns.

A simple stylized model of optimal traditional (individual property) appraisal & optimal aggregate (transactions-based) index estimation

MIT CENTER FOR REAL ESTATE

To determine the optimal lag, we employ a numerical algorithm that starts with $L = 0$, and increments one lag at a time as long as the MSE is decreasing. *e.g.*:

Exhibit 2: Optimal Lag (L) as Function of Dispersion Ratio & Transaction Density: *(Optimal Lag in Number of Periods)*				
	Dispersion Ratio ($\sigma_\varepsilon/\sigma_r$):			
Transaction Density:	4	2	1	0.5
$n = 1$	6	3	1	0
$n = 2$	4	2	0	0
$n = 10$	1	0	0	0
$n = 35$	0	0	0	0

Optimal lag diminishes rapidly as transaction density increases and as $(\sigma_\varepsilon / \sigma_r)$ decreases.

Optimal lag = 0 for transaction densities as low as $n = 2$ & $(\sigma_\varepsilon / \sigma_r)$ as high as 1.

A simple stylized model of optimal traditional (individual property) appraisal & optimal aggregate (transactions-based) index estimation

Empirical evidence ➜ σ_ε and σ_r range from 5 percent to 10 percent per year.

➜ Dispersion ratio range between 1/2 and 2 for annual frequency periods, dispersion ratio between 1 and 4 for quarterly periods.
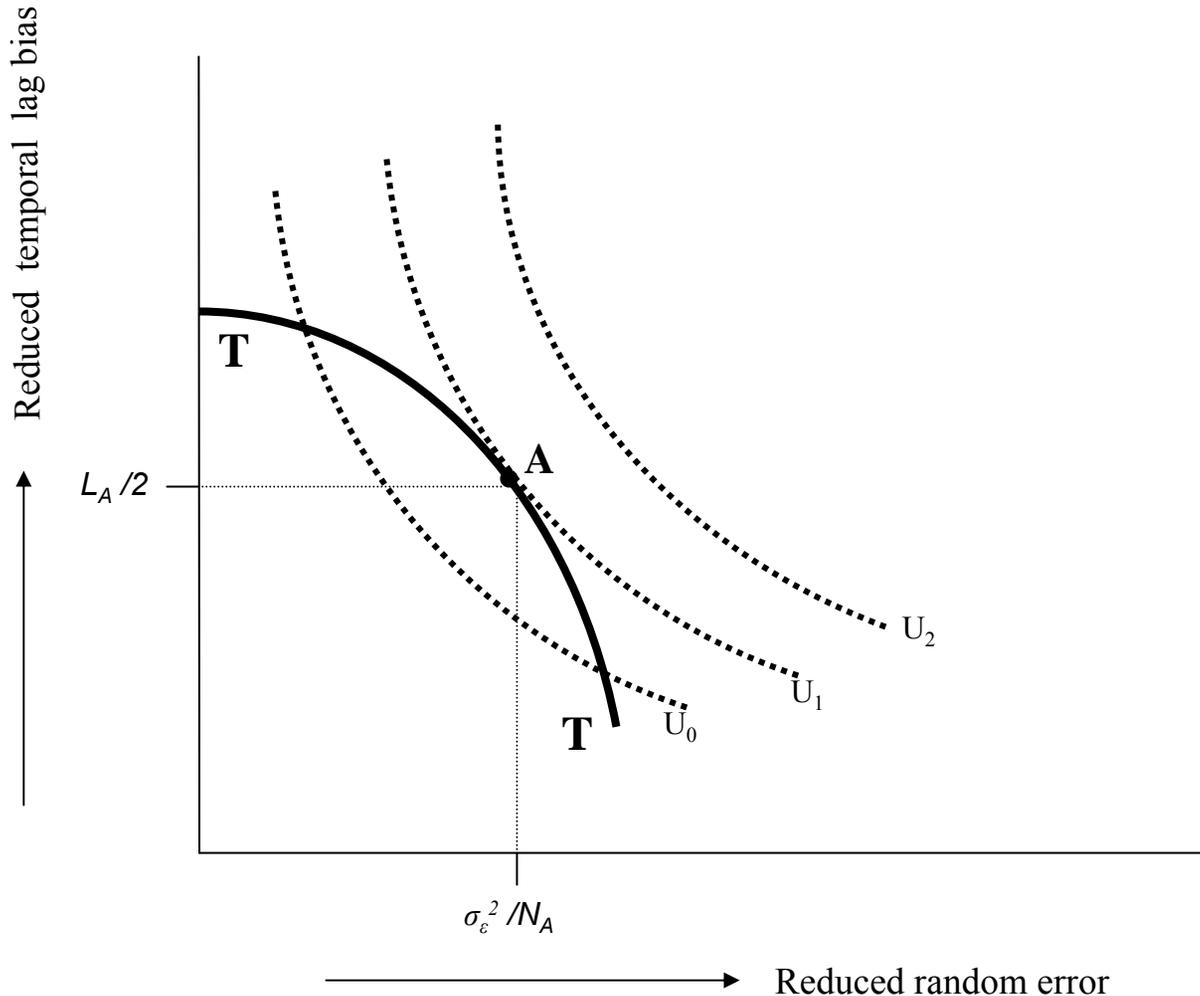
Assuming $\sigma_\varepsilon = 10\%$, we have Exhibit 3…

Exhibit 3: RMSE at Optimal Lag (*L*, & Optimal Sample Size *N*) as a Function of Transaction Density and Market Volatility (both per period). (Assuming σ_ε = 10%, σ_r = 10%/year)			
Transaction Density per Period:	Volatility per Period, σ_r :		
	(Monthly Index) $2.9\% = 10\% / \sqrt{12}$	(Quarterly Index) $5.0\% = 10\% / \sqrt{4}$	(Annual Index) $10.0\% = 10\% / \sqrt{1}$
n = 1	5.42% (*L*=5, *N*=6)	6.85% (*L*=3, *N*=4)	8.66% (*L*=1, *N*=2)
n = 3	3.95% (*L*=3, *N*=12)	4.79% (*L*=1, *N*=6)	5.77% (*L*=0, *N*=3)
n = 12	2.50% (*L*=1, *N*=24)	2.89% (*L*=0, *N*=12)	2.89% (*L*=0, *N*=12)

Exhibit 4a: The Noise vs Lag Trade-off Frontier and Optimal Property Value Estimation Methodology

Exhibit 4b: Greater Utility for an Up-to-Date Valuation:

Reduced temporal lag bias

L = Lag of oldest observation in estimation sample for period t.

N = Number of observations in estimation sample for period t.

$$L_B < L_A$$

$$N_B < N_A$$

T

$L_B/2$

B

U_2

U_1

U_0

T

σ_ε^2/N_B

Reduced random error
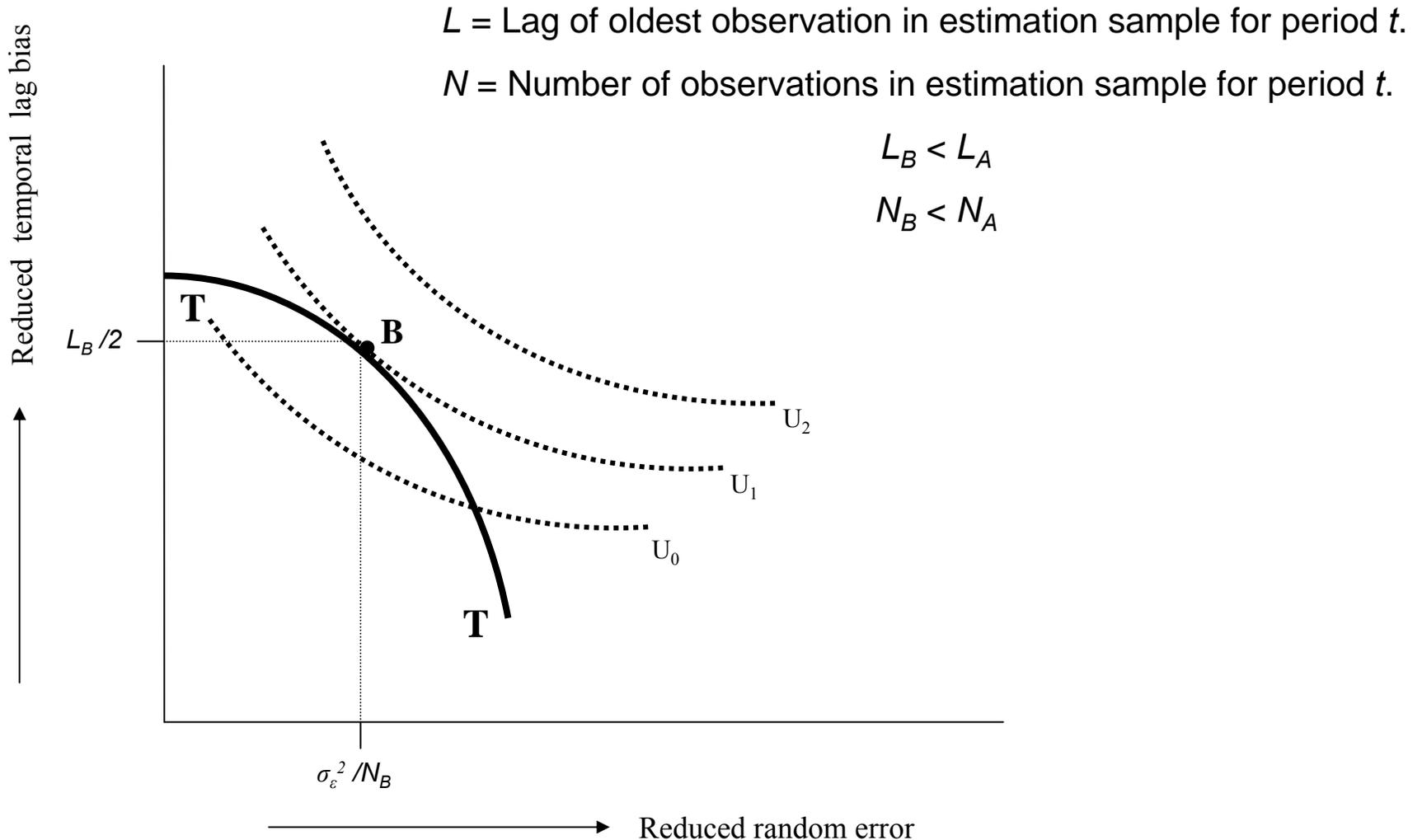
Example with typical numbers…

Exhibit 5: Optimal Valuation Methodology for Traditional Appraisal:
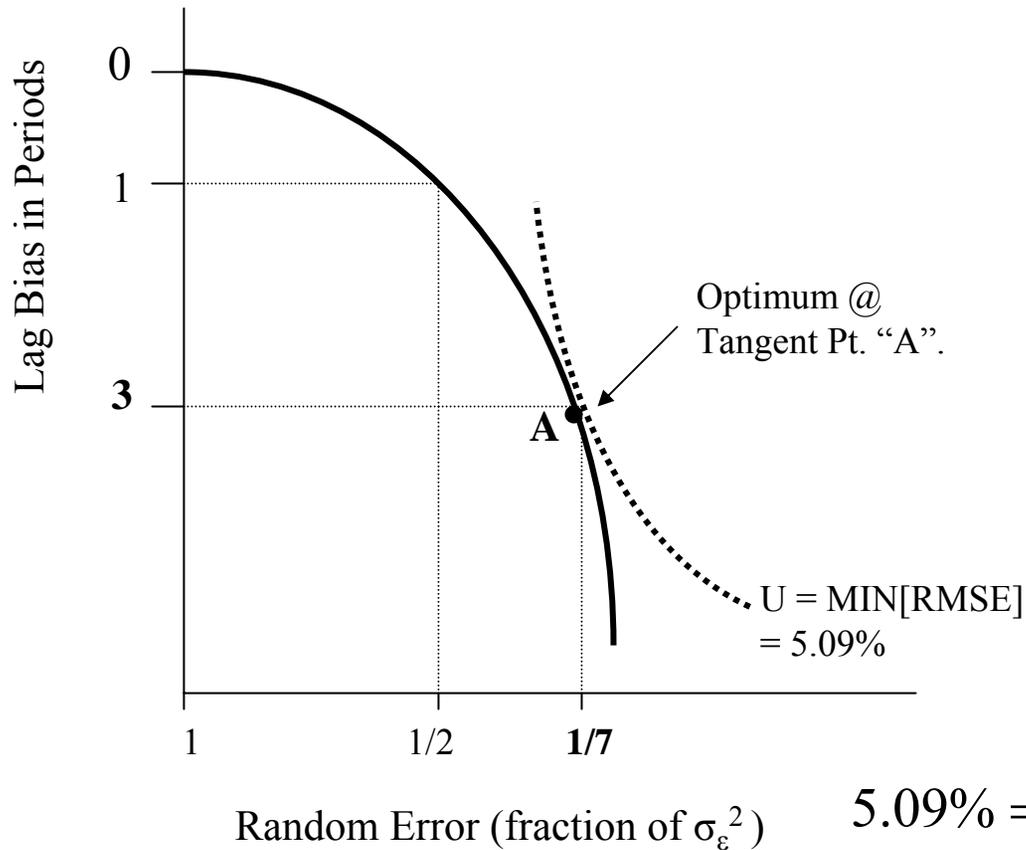
$U = MIN[MSE]$, $n = 1$, $(\sigma_\varepsilon / \sigma_r) = 4$:

➔ $L_A = 6$, $N_A = 7$.

➔ $E[lag] = 3$ periods; *Random Error Variance* $= \sigma_\varepsilon^2 / 7$.

➔ $RMSE = 5.09\%$ if $\sigma_\varepsilon = 10\%$.

n = Number of available transaction observations per period.

N = Number of transaction observations used in estimation of value as of period t.

L = Maximum lag of transaction observations in estimation sample.

Lag Bias in Periods

0

1

3

A

Optimum @ Tangent Pt. "A".

U = MIN[RMSE] = 5.09%

1   1/2   **1/7**

Random Error (fraction of σ_ε^2)

$$5.09\% = \sqrt{\frac{.10^2}{7} + \left(\frac{.025}{7}\right)^2 \sum_{s=1}^{6} s^2}$$

Exhibit 6: Optimal Methodology for a Transactions-Based Index:

$U = MIN[\text{MSE}], n = 35, (\sigma_\varepsilon / \sigma_r) = 4$:

➔ $L_C = 0, N_C = 35.$

➔ $E[lag] = 0$ period; *Random Error Variance* $= \sigma_\varepsilon^2 / 35.$

➔ $RMSE = 1.69\%$ if $\sigma_\varepsilon = 10\%.$

Suppose regression-based ("mass appraisal") index expands scope of estimation sample 35-fold compared to traditional individual property appraisal ("comps").

"Corner Solution" Optimum @ Zero Lag: Point "C".

U_C: RMSE=1.69%

U_B: RMSE=3.47%

U_A: RMSE=5.09%

Lag Bias in Periods

Random Error (fraction of σ_ε^2)
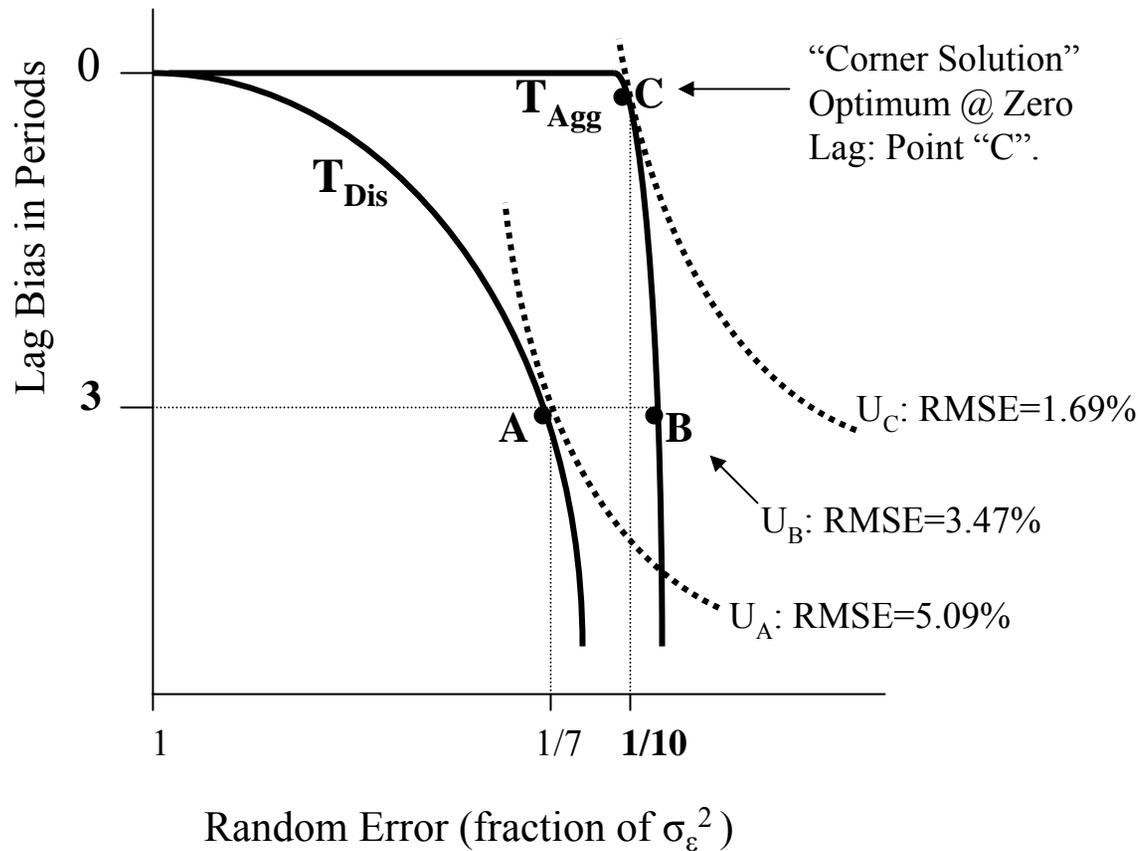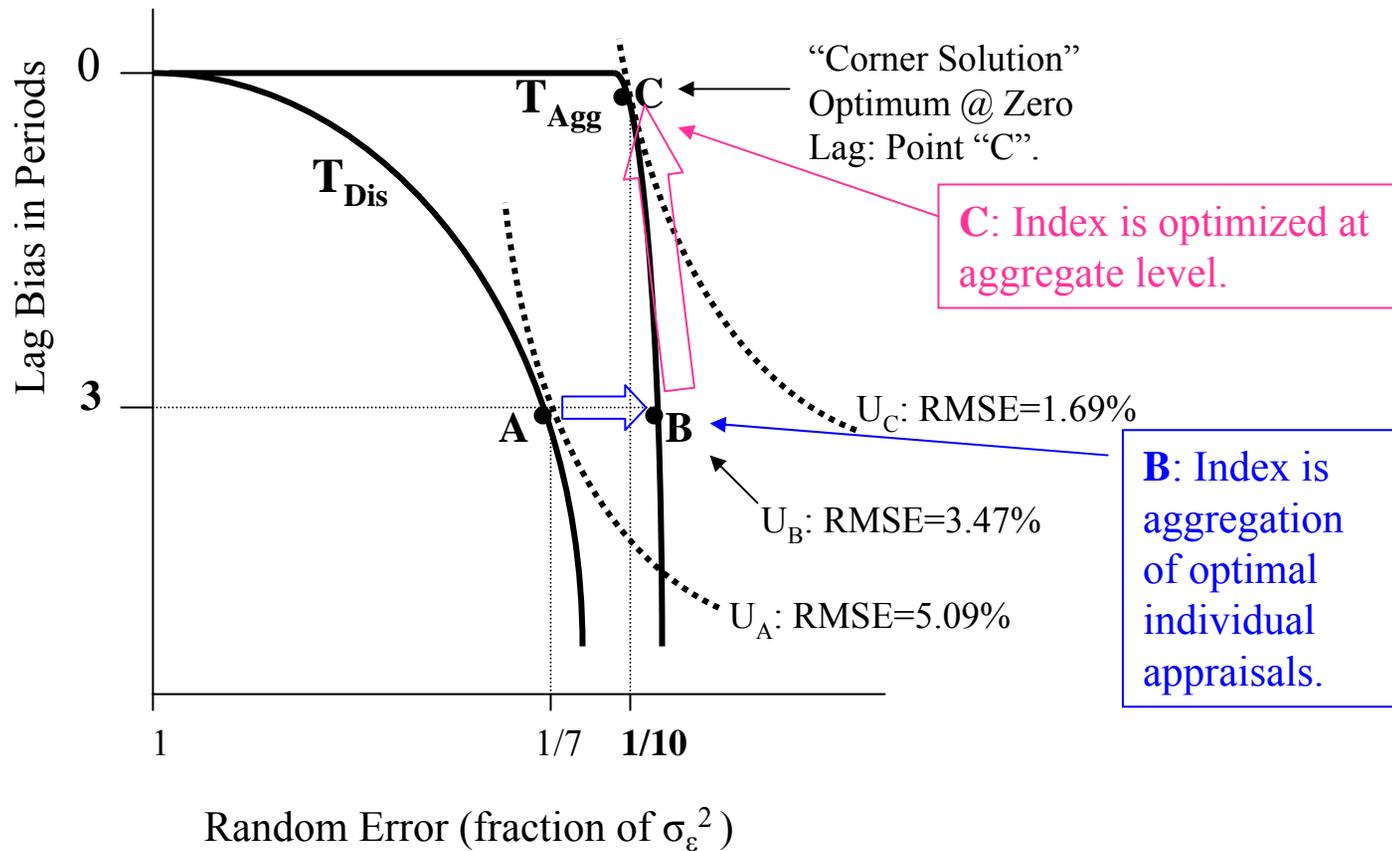
Example with typical numbers…

Exhibit 6: Optimal Methodology for a Transactions-Based Index:

$U = MIN[\text{MSE}]$, $n = 35$, $(\sigma_\varepsilon/\sigma_r) = 4$:

➔ $L_C = 0$, $N_C = 35$.

➔ $E[lag] = 0$ period; *Random Error Variance* $= \sigma_\varepsilon^2 / 35$.

➔ *RMSE* = 1.69% if $\sigma_\varepsilon = 10\%$.

"Corner Solution" Optimum @ Zero Lag: Point "C".

C: Index is optimized at aggregate level.

U_C: RMSE=1.69%

U_B: RMSE=3.47%

U_A: RMSE=5.09%

B: Index is aggregation of optimal individual appraisals.

Numerical Example Comparison of Appraisal-Based vs Transactions-Based Aggregate Property Value Index. *(Error components measured in square root of variance.)*		
Error Type:	Appraisal-Based Index:	Transactions-Based Index:
Purely Random (noise):	0.64%	1.69%
Temporal Lag Bias:	3.41%	0
Total:	3.47%	1.69%

$$\sqrt{\frac{.10^2}{7*35}} = 0.0064 \qquad \sqrt{\left(\frac{.025}{7}\right)^2 \sum_{s=1}^{6} s^2} = 0.0341 \qquad 0.0347 = \sqrt{\frac{.10^2}{7*35} + \left(\frac{.025}{7}\right)^2 \sum_{s=1}^{6} s^2} \qquad 0.0169 = \sqrt{\frac{0.10^2}{35}}$$

Lag Bias in Periods

0

T_{Agg} C   "Corner Solution" Optimum @ Zero Lag: Point "C".

T_{Dis}

3

A   B

U_C: RMSE=1.69%

U_B: RMSE=3.47%

U_A: RMSE=5.09%

1   1/7   **1/10**

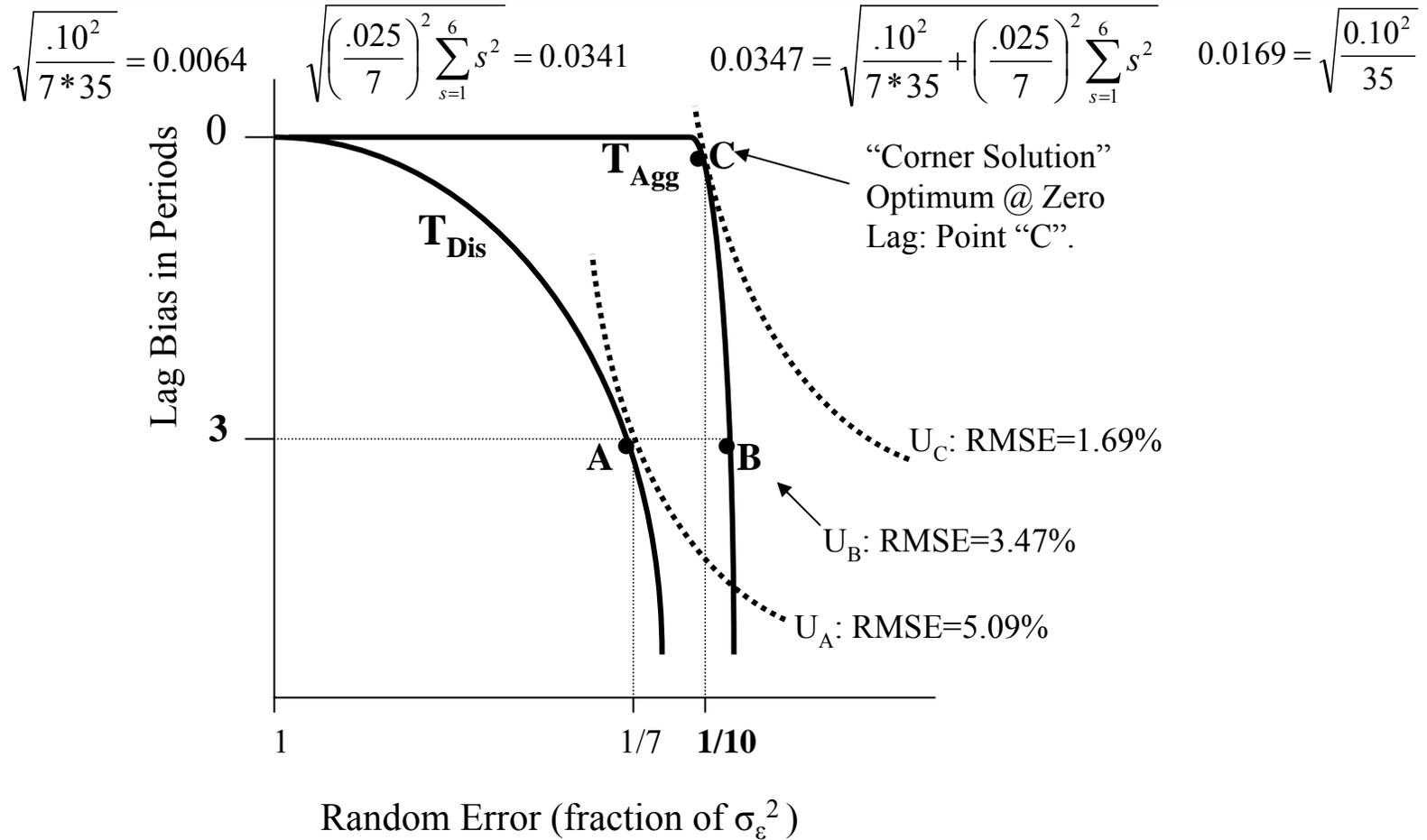Random Error (fraction of σ_ε^2)

Exhibit 7: Noise-vs-Lag Trade-off Frontiers with Disaggregate (Traditional Appraisal) and Aggregate (Transactions Based Regression) Valuation Methodologies

N_A = Disaggregate (Appraisal) Optimal Sample Size (# comps) = $n_A(L_A+1)$, n_A = comps density/period;
N_C = Aggregate (Transactions Based Regression) Optimal Sample Size (obs per period);
Q = Number of Market Segments in the Index Population (as multiple of number used by appraiser).
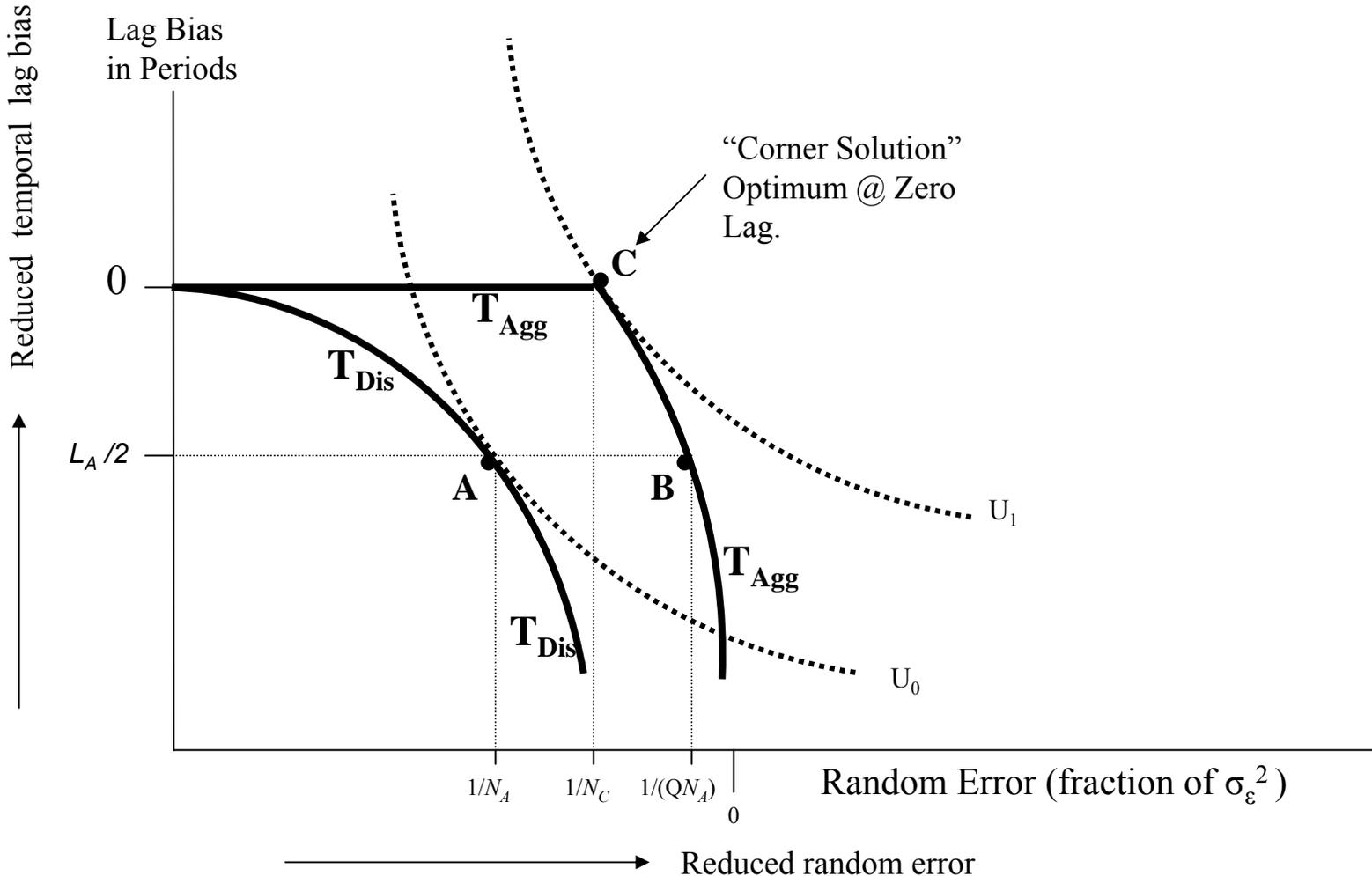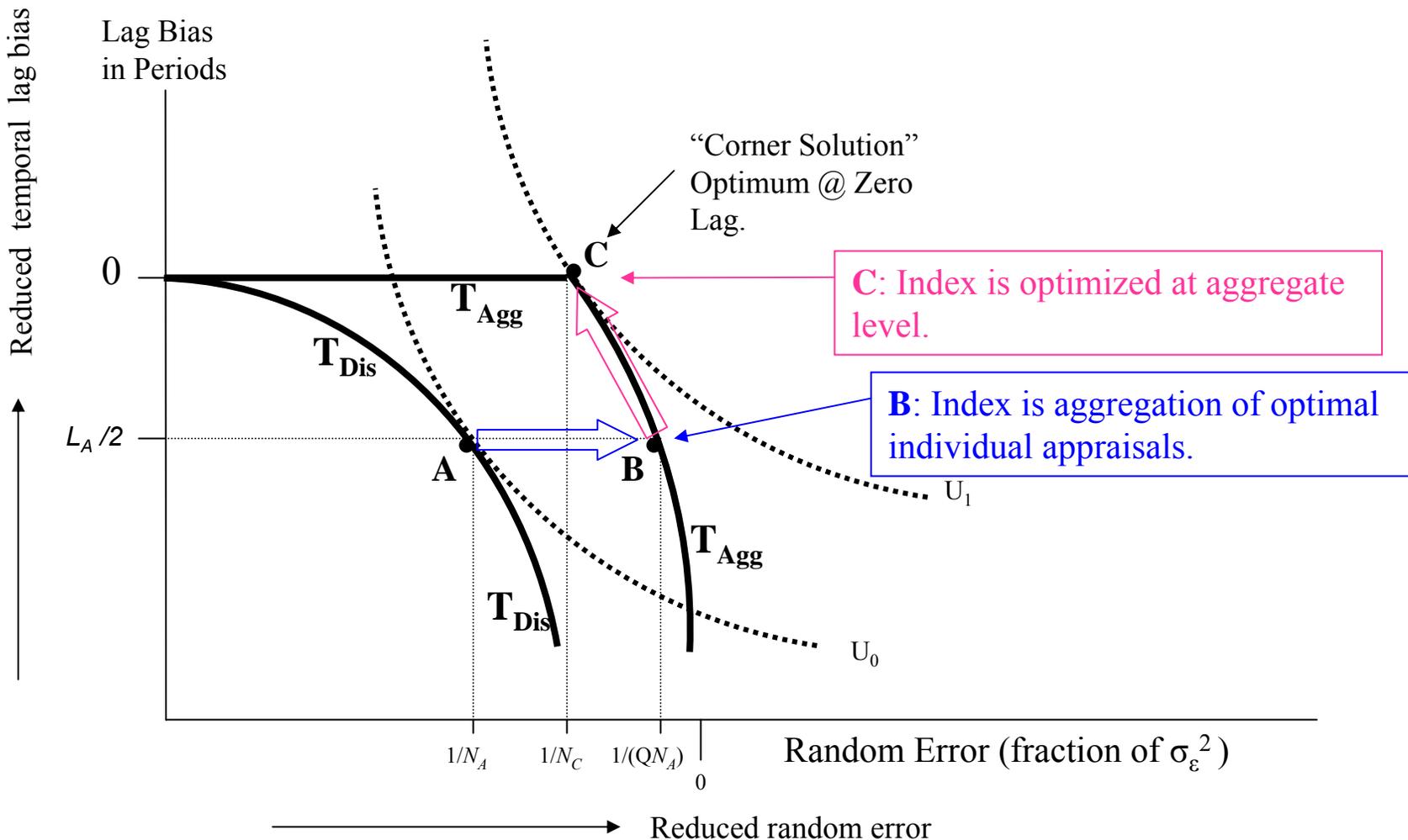
Exhibit 7: Noise-vs-Lag Trade-off Frontiers with Disaggregate (Traditional Appraisal) and Aggregate (Transactions Based Regression) Valuation Methodologies

N_A = Disaggregate (Appraisal) Optimal Sample Size (# comps) = $n_A(L_A+1)$, n_A = comps density/period;
N_C = Aggregate (Transactions Based Regression) Optimal Sample Size (obs per period);
Q = Number of Market Segments in the Index Population (as multiple of number used by appraiser).

Lag Bias in Periods

Reduced temporal lag bias

"Corner Solution" Optimum @ Zero Lag.

C

T_{Agg}

T_{Dis}

0

C: Index is optimized at aggregate level.

B: Index is aggregation of optimal individual appraisals.

$L_A/2$

A

B

U_1

T_{Agg}

T_{Dis}

U_0

$1/N_A$ $1/N_C$ $1/(QN_A)$ 0

Random Error (fraction of σ_ε^2)

Reduced random error

Optimal development of a commercial property price index useful for investment and urban economic *research* purposes requires:
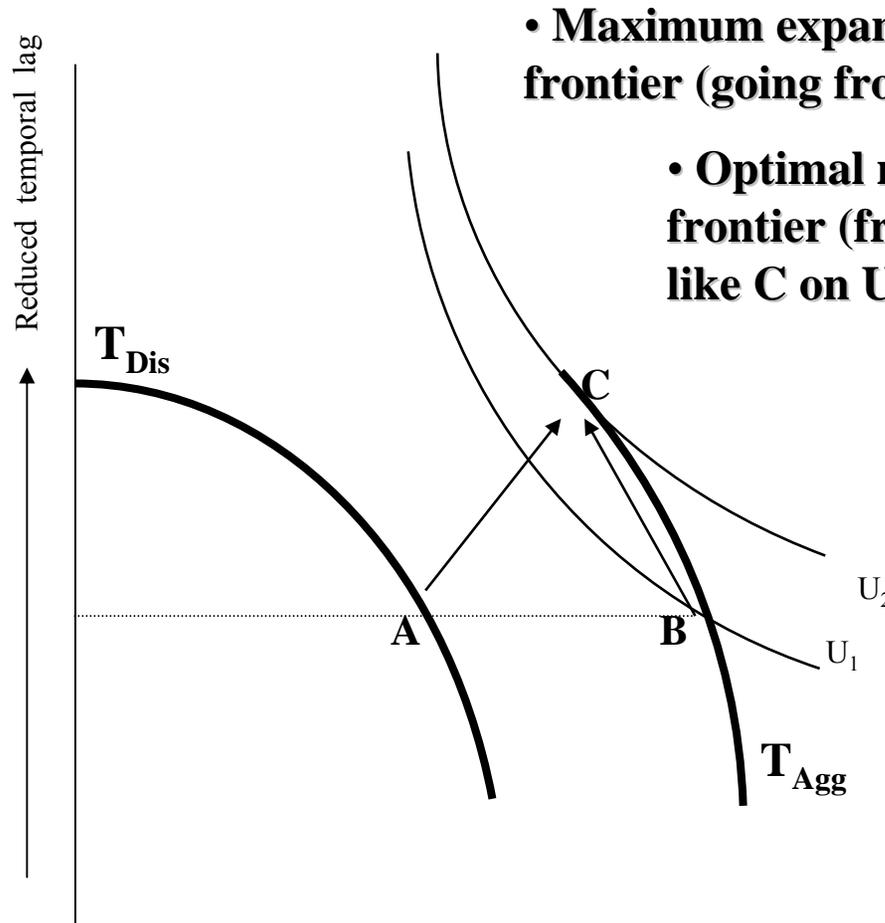
- **Maximum expansion of the accuracy trade-off frontier (going from T_{Dis} to T_{Agg});**

- **Optimal movement along the trade-off frontier (from points like B on U_1 to points like C on U_2).**

- **This in turn requires use of:**

  - **State-of-the-art statistical methodology;**

  - **Maximum possible quantity of empirical valuation data;**

  - **Maximum possible quality of data.**

Reduced temporal lag

T_{Dis}

C

U_2

A   B U_1

T_{Agg}

Reduced random noise

- For **research that is highly sensitive to temporal lag bias**, but less sensitive to random error in the index returns, **transactions-based indices are preferable** to appraisal-based indices, because of the temporal lag bias inherent in appraisal-based returns.

- For **research that is highly sensitive to random error**, but not very sensitive to temporal lag bias, an **appraisal-based index may be preferable** to a transactions-based index because of the greater frequency of the appraisals in the population and (possibly) because the appraisals may be less noisy than the transaction prices.

• For **research that is equally sensitive** to lag bias error and random error (e.g., as represented by the MSE-minimization objective for the index), **transactions-based indices are preferable** to appraisal-based indices.

● Except for indices tracking small populations of properties where transaction density is less than two or three dozen observations per index reporting period, transactions-based indices minimizing the MSE criterion can be produced with no temporal lag bias.