Review for the Estimation with Regression Equations for Rent Indices in Japan

Makoto Shimizu
m.shimizu@soumu.go.jp

Director for International Statistical Affairs
Office of Director-General for Policy Planning (Statistical Standards)
Ministry of International Affairs and Communications, Japan

3 August 2009

Abstract

This paper provides the follow-up for the paper "Hedonic Approach for House Rents in Japan" presented at the Ottawa Group Meeting, Neuchâtel, 27-29 May 2009 based on comments by participants there.

In the CPI of Japan, indices of rents are computed by type of housing units, using the house rents collected every month from households. The indices of owner-occupied rents are imputed from the rent indices. The rent indices, however, often fluctuate due to rented houses which enter or exit from the sample of the rent survey, particularly at the local level.

This paper introduces the result of an experiment of estimation of rent indices and imputation of missing data with the regression equation with floor space, construction year and population density in the computation of the rent index of the CPI comparing the carrying forward and backward of missing data in order to solve that problem. This experiment aims to improve rent indices in the local area with small size of samples where it is difficult to keep stable quality by taking advantage of the national rent index, rather than seeking for a shortcoming for the national rent index.

This experiment became possible with initiation of surveying construction year for the districts newly adopted conclusively in January 2008 in the rent survey.

I Regression for Rents

This chapter focuses on the log-linear regression equation using rent data as a dependent variable and

floor space, construction year and population density[1] as independent variables among methods to solve the problem. Data are micro data[2] of the RPS (Retail Price Survey) for rent data, floor space and construction year in 2008 while population density is population in the 2005 Population Census divided by inhabitable area[3] by the National Survey for Area of Prefectures and Municipalities, etc.

Regression equations

The regression equations are the followings;

$$r_k = e^a \prod_{i=1}^{3} x_{ik}^{b_i}$$

In these equations, r_k is the rent, while x_{1k} is the floor space; x_{2k} the construction year; x_{3k} the population density.

Taking of logarithms of the equations,

$$\ln r_k = a + \sum_{i=1}^{3} b_i \ln x_{ik}$$

The variables and performances for rents by type of houses are shown monthly from January to December 2008 on Table 1-1, 1-2, 1-3 and 1-4.

For the equation for small wooden houses[4] adjusted R-squared are 0.65-0.67, the highest among the four types of houses, and variables have two digit t-values and stable coefficients among months. (Table 1-1)

[1] Land price for dwelling area (results of the Land Price Survey by Prefecture) was used instead of population density as an independent variable in Makoto Shimizu [May 2009].

[2] Rent Survey Districts were enlarged in Ohshu, Ishinomaki and Hirado from October to December 2008 for the RPS. Data in enlarged area were used in Makoto Shimizu [May 2009], however are excluded in this paper since they are not used for the CPI compilation, having nothing to do with the regression equation for lack of construction year.

[3] In principle, they are calculated by municipality, but by ward in Ku-area of Tokyo, by consolidated area for Kofu and Fujikawaguchiko.

[4] Floor space for small houses is smaller than 30 square meter, that for medium houses equal to or larger than 30 square meter.

Table 1-1 Regression Equations for Rents of Small Wooden Houses

mon.	constant		floor space		construction year		population density		number of houses	adj.R^2
	coefficient	t-value	coefficient	t-value	coefficient	t-value	coefficient	t-value		
Jan.	−262.54	−42.41	0.35	15.81	35.33	43.23	0.27	41.00	1804	0.66
Feb.	−266.57	−42.60	0.35	15.62	35.86	43.41	0.27	40.37	1773	0.66
Mar.	−265.21	−42.06	0.35	15.54	35.67	42.85	0.27	39.75	1712	0.67
Apr.	−264.12	−41.84	0.35	15.65	35.53	42.63	0.27	39.76	1697	0.67
May	−257.52	−39.98	0.34	14.59	34.67	40.76	0.28	39.41	1658	0.66
Jun.	−255.89	−39.47	0.34	14.41	34.45	40.23	0.28	39.05	1641	0.65
Jul.	−250.23	−38.23	0.34	14.19	33.70	39.00	0.28	39.10	1622	0.65
Aug.	−248.18	−38.33	0.32	13.41	33.46	39.14	0.28	39.23	1619	0.65
Sep.	−246.93	−38.11	0.31	13.14	33.29	38.91	0.28	38.96	1598	0.65
Oct.	−247.58	−38.05	0.31	13.09	33.38	38.85	0.28	38.84	1586	0.65
Nov.	−246.94	−37.78	0.32	13.27	33.29	38.56	0.28	38.59	1573	0.65
Dec.	−245.79	−37.43	0.33	13.63	33.13	38.21	0.27	38.39	1544	0.65

For the equations for medium wooden houses adjusted R-squared are 0.61 or 0.62, and variables have still two digit t-values and stable coefficients among months. (Table 1-2)

Table 1-2 Regression Equations for Rents of Medium Wooden Houses

mon.	constant		floor space		construction year		population density		number of houses	adj.R^2
	coefficient	t-value	coefficient	t-value	coefficient	t-value	coefficient	t-value		
Jan.	−211.39	−61.33	0.42	35.40	28.60	62.94	0.20	60.19	5524	0.61
Feb.	−210.74	−61.64	0.41	35.50	28.52	63.28	0.19	60.44	5434	0.62
Mar.	−209.69	−60.50	0.41	34.80	28.39	62.14	0.19	60.13	5307	0.61
Apr.	−208.75	−60.34	0.41	35.15	28.26	61.96	0.19	60.12	5257	0.62
May	−203.81	−58.82	0.43	36.22	27.60	60.43	0.19	59.46	5207	0.61
Jun.	−203.99	−58.80	0.43	35.98	27.62	60.40	0.19	59.24	5169	0.61
Jul.	−203.38	−58.51	0.43	36.44	27.53	60.10	0.19	59.50	5142	0.61
Aug.	−201.55	−57.68	0.43	35.98	27.30	59.26	0.19	59.03	5084	0.61
Sep.	−202.61	−57.67	0.43	35.60	27.44	59.25	0.19	58.88	5055	0.61
Oct.	−202.37	−57.54	0.43	35.82	27.40	59.12	0.19	58.82	5024	0.61
Nov.	−201.85	−57.11	0.43	35.70	27.34	58.67	0.19	58.32	4980	0.61
Dec.	−203.05	−57.26	0.42	35.30	27.50	58.83	0.19	58.31	4944	0.61

For the equations for small non-wooden houses adjusted R-squared are 0.58, the lowest among the four types of houses, but variables have two digit t-values and stable coefficients among months. (Table 1-3)

Table 1-3 Regression Equations for Rents of Small Non-wooden Houses

mon.	constant		floor space		construction year		population density		number of houses	adj.R^2
	coefficient	t-value	coefficient	t-value	coefficient	t-value	coefficient	t-value		
Jan.	−115.54	−22.60	0.65	44.63	15.67	23.22	0.27	68.60	5160	0.58
Feb.	−106.79	−20.79	0.65	44.70	14.52	21.42	0.27	68.74	5115	0.58
Mar.	−108.48	−20.90	0.65	43.90	14.75	21.52	0.27	67.63	4995	0.58
Apr.	−108.62	−20.79	0.65	43.78	14.76	21.41	0.27	67.18	4921	0.58
May	−108.44	−20.60	0.65	43.20	14.74	21.21	0.27	66.54	4799	0.58
Jun.	−110.46	−21.02	0.64	42.20	15.02	21.64	0.27	66.33	4736	0.58
Jul.	−110.98	−20.78	0.66	43.49	15.06	21.36	0.27	65.06	4738	0.58
Aug.	−120.86	−22.41	0.65	42.33	16.38	22.99	0.27	64.81	4712	0.58
Sep.	−119.01	−21.85	0.66	42.96	16.13	22.42	0.27	64.36	4675	0.58
Oct.	−117.71	−21.47	0.66	42.91	15.95	22.04	0.27	64.00	4636	0.58
Nov.	−114.43	−20.55	0.68	43.60	15.50	21.08	0.27	63.78	4600	0.58
Dec.	−117.96	−21.06	0.67	43.38	15.96	21.58	0.27	63.54	4554	0.58

For the equations for medium non-wooden houses adjusted R-squared are 0.61 or 0.62, and variables have two or three digit t-values and stable coefficients among months. (Table 1-4)

Table 1-4 Regression Equations for Rents of Medium Non-wooden Houses

mon.	constant		floor space		construction year		population density		number of houses	adj.R^2
	coefficient	t-value	coefficient	t-value	coefficient	t-value	coefficient	t-value		
Jan.	−150.11	−42.71	0.66	82.39	20.20	43.63	0.27	100.85	11228	0.61
Feb.	−148.15	−41.68	0.66	81.41	19.94	42.59	0.27	99.40	10987	0.61
Mar.	−140.92	−39.50	0.67	82.74	18.97	40.36	0.27	99.32	10818	0.61
Apr.	−140.66	−39.07	0.67	82.29	18.93	39.93	0.28	98.78	10722	0.61
May	−141.66	−39.46	0.67	82.29	19.06	40.32	0.28	98.70	10650	0.61
Jun.	−144.43	−40.27	0.67	82.16	19.42	41.12	0.28	97.94	10608	0.61
Jul.	−144.31	−40.25	0.68	82.33	19.41	41.10	0.28	97.58	10567	0.62
Aug.	−145.73	−40.39	0.68	82.28	19.59	41.23	0.27	96.87	10472	0.62
Sep.	−147.04	−40.66	0.68	81.68	19.77	41.51	0.27	96.61	10430	0.62
Oct.	−147.16	−40.45	0.67	80.47	19.79	41.31	0.27	96.20	10360	0.61
Nov.	−147.28	−40.38	0.67	79.90	19.81	41.24	0.27	95.63	10298	0.61
Dec.	−147.83	−40.22	0.67	79.05	19.88	41.07	0.27	94.79	10214	0.61

Observed from those tables, coefficients for floor space are larger for non-wooden houses than wooden houses, conversely coefficients for construction year larger for wooden houses than non-wooden houses. It illustrates that rents of wooden houses tend to get relatively stronger influence from newness of houses whereas rents of non-wooden houses are reflected from space of houses.

Those equations can be used for quality adjustment for rent indices considering 0.58-0.67 adjusted R-squared, no less than 0.53-0.59; those for weights of estimated expenditures for owned houses in the official CPI. Those results depict that main factors influencing on rents are floor space, construction year and population density, that is, space and newness of houses and congestion in area. Based on stability of coefficients, the equations are plausible in view of viability, therefore, can be adopted for estimation of rents by region.

II Reviewing Quality Adjustment for Rents

The chapter provides estimation of missing or measured values using the regression equations shown in the above chapter in comparison with methods to carry forward and backward in order to solve the problem of fluctuations of rent indices; the massive fluctuations of rent indices owing to small sample sizes in a small municipality.

Regression equations for quality adjustment

The regression equations can be used for estimation of rent indices by region because independent variables in the equations have so high degree of significance that they can be fixed in time series.

The regression equation is beneficial for the purpose of adjusting quality for rents and extracting essential changes, by constructing functions with rents and other variables in order to estimate ideal levels for rents for vacant or newly included houses, for instance, when some sampled tenants remove into an area outside the districts, newly established houses are included, or the districts are exchanged. The regression equations are useful even for rent data for existing households without transfer because adjusted rent data based on structure in the national level would be influenced less than actual rent data that include outliers by transfer of households.

Points and methods for comparison

Prerequisites to solve the problem are the possibility of practical application for the CPI. Specifically, one is accessibility to all necessary data for the method before compilation of the CPI which has to be released for short period after the survey. Another is stability of the methodology, with which the procedure of the compilation remains the same although data change.

Based on those two points, three methods are reviewed as candidates to solve the problem. The first is the carrying backward of the data in the previous month for the data of tenants which newly entered the sample for the RPS. In the same way, the carrying forward is applied for data in the month when tenants exit from the sample. The second is the imputation with the regression equations for data in the previous month for entrance and those in the month for exit. The third is application of the estimation by the regression equations to all rent data, making it possible to adjust quality for all rent data including outliers.

This chapter compares those methods based on aggregation of municipal data, though the official CPI is not disseminated in municipalities other than prefectural governments and designated cities but in 10 areas composed of municipalities. Municipal data are total gross rents divided by total floor space for all tenants including those for the imputation.

Both for the carrying forward or backward and the regression method, imputed data should be applied for calculation merely for change from the previous month at the month of transfer, which will be chained every month for compilation of indices, instead of regarded as rents in the following months, because continuous inclusion of the same data once estimated by the method will lead indices to include excessive stability after the month. The CPI Manual [2004] also says that the carrying forward the last observed price should be avoided wherever possible and is acceptable only for a very limited number of periods. Further, whole termination of continuous adoption of the same data will bring slides between the old and the new indices at the next revision of samples for the CPI.

Estimated rents among methods

Firstly, their estimation are evaluated as signs and sizes from rent data for the official CPI compiled with averages without the imputation for missing data. Differences with bilateral signs guarantee non-bias estimation. The smaller the differences are, the better the methodology is.

In terms of the carrying forward or backward and the imputation with the regression equation, Table 2, accounting for weighted average of differences of monthly rents per square meter from the original data[5], observes that differences show both positive and negative signs for entrance and exit, meaning that rents of tenants with transfer include no stabilized bias from those of tenants without transfer as a whole. In addition, it provides information that differences tend to be larger for the carrying forward or backward than for the regression equation. The largest difference records eight yen per square meter for the carrying forward but six for the regression equation for exit, small non-wooden houses on July. In these ways, the regression equation usually provides more approximate estimation to the original rent data.

Besides, estimation by the regression equations tends to be lower consistently than indices calculated from actual rents. It implies that the regression equations are apt to estimate rents lower than actual rents, by lowering outliers.

[5] Results in this paper using data exclusively in the new districts are different from those by the original CPI because the CPI is compiled partly with data in the old districts from January to March. The regression equations in this paper cannot be applied to the old districts where construction year had not been surveyed.

Table 2 Weighted Averaged Differences of Monthly Rents per Floor Space among the Carrying Forward or Backward, the Imputation and the Estimation with the Regression Equations from the Original Results (Yen per Square Meter)

	small wooden houses					medium wooden houses					small non-wooden houses					medium non-wooden houses				
	entrance		exit			entrance		exit			entrance		exit			entrance		exit		
month	Car.	Reg.	Car.	Reg.	Hed.	Car.	Reg.	Car.	Reg.	Hed.	Car.	Reg.	Car.	Reg.	Hed.	Car.	Reg.	Car.	Reg.	Hed.
Jan.	2	1	-1	-2	-74	-1	-2	-2	-1	-49	-2	1	0	2	-25	-1	0	0	0	-74
Feb.	0	0	-1	-5	-82	1	0	-1	-2	-49	-1	2	-2	-1	-24	1	1	0	-4	-77
Mar.	0	1	2	-2	-85	-1	0	1	1	-49	1	0	0	0	-28	0	-1	1	-2	-82
Apr.	1	4	2	-1	-82	1	1	0	-1	-47	1	1	0	0	-28	0	-1	0	-1	-82
May	1	0	2	-3	-95	1	0	0	-1	-51	-1	1	1	2	-27	0	0	0	-1	-81
Jun.	1	2	5	3	-95	1	0	0	-1	-52	-1	2	0	1	-27	0	0	0	0	-80
Jul.	0	0	-5	-3	-97	1	0	0	0	-51	0	2	8	6	-24	0	0	1	-1	-81
Aug.	2	0	3	2	-93	0	-1	2	0	-50	-2	1	0	-2	-28	0	0	1	0	-80
Sep.	0	1	0	-3	-95	0	0	0	0	-49	1	1	0	0	-26	-2	-2	0	0	-79
Oct.	4	2	0	1	-97	0	0	1	0	-49	1	1	1	0	-28	0	0	-1	-2	-79
Nov.	3	0	2	-1	-96	-1	0	-1	-1	-50	0	1	3	-2	-28	0	0	0	0	-79
Dec.	-	-	1	0	-91	-	-	0	1	-48	-	-	1	-1	-31	-	-	0	-1	-79

Note:
Entrance is the case to adopt the imputation of missing data in the previous month for newly included tenants.
Exit is the case to adopt the imputation of missing data in the month for newly excluded tenants.
Car. is the case with the carrying forward or backward of missing data.
Reg. is the case with the regression equations for the imputation of missing data.
Hed. is the case with the regression equations replacing all actual data with the estimation.
Differences are those of weighted average with municipal weights of aggregated rents per aggregated floor space by municipality from the original in a month.

Monthly changes of estimated rents among methods

In the CPI, changes are more important than actual charges. This section focuses on evaluation among the three methods using monthly changes estimated from rents of the RPS.

As prerequisites to the monthly release of data, the carrying forward or backward and the regression equation, from practical view, should be applied only for imputation for unavailable monthly changes, calculated from rent data, instead of unavailable rent data themselves for transfer of tenants, because once rent data are released in the previous month when entrance of new tenants have not been noticed, they cannot be corrected by the imputation as rents in the previous month later, yet monthly changes, released for the first time in the current month can mirror correction of rent data in the previous month.

The first evaluation point of comparison among the methods is effectiveness by adjustments of excessive fluctuations of rent indices by municipality. Adjustment of outliers of the indices sometimes brings estimation bias to the national average in spite of smoothing individual municipal indices. Thus, the second evaluation point is the national level loss of estimation biases according to adjustment of indices in specific municipalities. The former can be evaluated with standard deviation of monthly changes of estimated indices, the latter with averaged differences of monthly changes between estimated indices and original indices, with municipal weights among municipalities.

As the result of the first evaluation, Table 3 illustrates comparison of averaged standard deviations in the year 2008 of monthly changes of rents per floor space by municipality among original results, the carrying forward or backward, the imputation of missing values and the estimation with the regression equation by type of houses.

The carrying forward or backward can amplify stability of data by regarding rents by entering or exiting tenants as the same as those when they exist. The table demonstrates that averaged standard deviations for the carrying forward or backward are smaller than those for the original results. On the other hand, it pictures that those by the regression equations are not always smaller than those for the original results.

The estimation with the regression equations depicts the smallest standard deviations, much lower than the original results, by adjusting all rent data based on relation with basic variables in the national level. From the result, the estimation with the regression equations provides most stable rent indices.

Table 3 Averaged Standard Deviations in 2008 of Monthly Changes of Rents per Floor Space by Municipality among the Original Results, the Carrying Forward or Backward, the Imputation and the Estimation with the Regression Equations (Percent Point)

	method of imputation	small wooden houses	medium wooden houses	small non-wooden houses	medium non-wooden houses
original results		1.7	1.1	1.7	0.7
entrance	carrying backward	1.6	0.9	1.4	0.5
	regression equation	2.0	1.1	1.7	0.7
exit	carrying forward	0.8	0.7	0.7	0.5
	regression equation	1.7	1.0	1.5	0.8
estimation with regression equations		0.3	0.2	0.2	0.1

Note:
Entrance is the case to adopt the imputation of missing data in the previous month for entering tenants in the current month.
Exit is the case to adopt the imputation of missing data in the month for exiting tenants in the current month.
Standard deviations are those among changes of aggregated rents per aggregated floor space by municipality in a month, averaged from January to December 2008.
Changes for the carrying forward or backward and the imputation with the regression equations are ratios of the original results in the current month divided by the estimation in the previous month for entrance and ratios of the estimation in the current month divided by the original results in the previous month for exit.
Changes for the estimation with the regression equations are ratios of the estimation in the current month divided by the estimation in the previous month.

As the latter evaluation, Table 4 shows biases of the estimation from the original results as weighted average of change over the previous month of aggregated rents per aggregated floor space by municipality among the three methods. It explains that estimation biases are not so large in size, less than 0.9 percent point at the largest, observing both of positive and negative signs for all the methods, and that absolute averaged estimation biases among 12 months record less than 0.1 percent point, less than 1.2 percent point annually. In particular, those for medium wooden houses with the largest weight as the owner-occupied houses are less than 0.3 percent point both monthly and annually.

It pictures in addition that estimation biases by the imputation with the regression equations tend to be

smaller than those by the carrying forward or backward and the estimation with the regression equations, for example, the largest estimation bias records 0.36 percent point by the imputation with the regression equation, 0.43 percent point by the carrying forward for exit, small non-wooden houses in July; and -0.84 percent point by the estimation with the regression equation for small non-wooden houses in May. But the differences of estimation biases do not always follow to the tendency depending on situation.

Table 4 Weighted Averaged Estimation Biases of Monthly Changes by Municipality among the Carrying Forward or Backward and the Imputation and the Estimation with the Regression Equations from the Original Results in 2008 (Percent Point)

	small wooden houses					medium wooden houses					small non-wooden houses					medium non-wooden houses				
	entrance		exit			entrance		exit			entrance		exit			entrance		exit		
month	Car.	Reg.	Car.	Reg.	Hed.	Car.	Reg.	Car.	Reg.	Hed.	Car.	Reg.	Car.	Reg.	Hed.	Car.	Reg.	Car.	Reg.	Hed.
Jan.	0.03	-0.01	-0.07	-0.12	-0.27	-0.03	-0.01	-0.12	-0.03	0.05	-0.13	0.04	-0.01	0.09	0.04	0.08	0.05	0.01	0.00	0.06
Feb.	-0.19	-0.10	-0.02	-0.14	-0.46	0.04	0.15	-0.14	-0.19	-0.07	0.08	-0.04	-0.08	-0.06	0.03	0.06	-0.02	-0.04	-0.23	-0.29
Mar.	0.03	-0.01	0.09	-0.10	-0.13	-0.08	0.00	0.07	0.06	-0.05	0.01	-0.08	0.03	0.03	-0.21	-0.06	-0.07	0.05	-0.17	-0.43
Apr.	0.01	-0.04	0.11	-0.01	0.18	0.10	-0.01	0.00	-0.03	0.19	-0.06	-0.01	0.02	0.02	-0.05	-0.02	0.07	0.05	-0.04	0.00
May	-0.09	-0.23	0.14	-0.13	-0.84	-0.10	-0.11	-0.03	-0.04	-0.21	-0.04	-0.08	0.05	0.18	-0.01	0.00	0.03	0.03	-0.06	0.03
Jun.	-0.03	0.02	0.19	0.09	-0.02	-0.07	0.02	0.06	-0.02	-0.06	0.08	-0.06	0.02	0.04	-0.03	0.03	0.04	-0.01	-0.02	0.02
Jul.	-0.03	-0.10	-0.18	-0.13	-0.08	-0.06	-0.03	-0.03	0.03	0.01	0.03	-0.10	0.43	0.36	0.31	0.02	0.01	0.04	-0.03	-0.01
Aug.	-0.01	0.01	0.10	0.11	0.19	-0.03	-0.01	0.09	0.00	0.08	0.00	-0.11	0.02	-0.03	-0.20	0.01	0.01	0.03	0.00	0.05
Sep.	-0.07	0.01	-0.01	-0.17	-0.08	0.01	0.04	0.04	0.02	0.10	0.08	-0.04	0.01	0.03	0.09	0.00	0.02	0.04	0.02	0.05
Oct.	-0.03	-0.07	-0.06	0.02	-0.15	-0.03	-0.04	0.06	0.07	0.02	-0.02	-0.04	0.03	0.02	-0.07	0.09	0.07	-0.04	-0.09	0.03
Nov.	-0.09	0.08	0.16	0.00	0.13	-0.02	-0.01	-0.04	-0.04	-0.10	-0.01	-0.06	0.16	-0.07	0.01	0.04	0.02	-0.01	0.00	0.02
Dec.	-0.10	0.00	0.02	-0.05	0.38	0.00	-0.01	0.05	0.07	0.11	-0.03	-0.04	0.04	0.00	-0.12	0.01	0.00	-0.02	-0.04	0.01
average	-0.05	-0.04	0.04	-0.05	-0.10	-0.02	0.00	0.00	-0.01	0.01	0.00	-0.05	0.06	0.05	-0.02	0.02	0.02	0.01	-0.05	-0.04
annual	-0.58	-0.42	0.48	-0.64	-1.14	-0.26	0.00	0.01	-0.10	0.06	-0.01	-0.61	0.72	0.58	-0.20	0.25	0.22	0.11	-0.65	-0.45

Note:
Entrance is the case to adopt the imputation of missing data in the previous month for entering tenants in the current month.
Exit is the case to adopt the imputation of missing data for exiting tenants in the current month.
Car. is the case with the carrying forward or backward of missing data.
Reg. is the case with the regression equations for the imputation of missing data.
Hed. is the case with the regression equations replacing all actual data with the estimation.
They are differences of weighted averages with municipal weights of monthly changes of aggregated rents per aggregated floor space by municipality from the original in a month.
Changes for the carrying forward or backward and the imputation with the regression equations are ratios of the original results in the current month divided by the estimation in the previous month for entrance and ratios of the estimation in the current month divided by the original results in the previous month for exit.
Changes for the estimation with the regression equations are ratios of the estimation in the current month by the previous month.

Lastly, influences by estimation biases to the CPI are calculated.

Table 5 estimated influences to change of the CPI by aggregating those with weights for private rented houses and imputed owned houses by type of houses in the CPI.[6] Since there is no tendency in positive or negative signs for monthly indices, the influences are apt to be smaller, aggregated among types of

[6] Weights for rents reach 17.0 percent composed of 2.8 percent for private rented houses and 14.2 percent for imputed owned houses in the 2005 base CPI. Especially, for medium wooden houses the influences are huge since 0.8 percent for private rented houses is applied to 10.9 percent for imputed owned houses. Strictly speaking, influences depend on size of rent indices as well as weights, substituted with aggregations solely with weights as rent indices by type of houses are near 100 in this paper.

houses. Averaged into annual level, the influence records the widest but only -0.03 percent point for the carrying forward for entrance. As the imputation is applied simultaneously for entrance and exit, aggregated, the influences for both sides record -0.02 percent point for both the imputation by the carrying forward and backward and with the regression equations wider than 0.00 percent point for the estimation with the regression equations. In the background of those figures, estimation bias for the estimation with the regression equations is large for any type of houses, yet becomes smaller as a whole, through offsetting those with both signs. The results by adding estimation biases for entrance and exit are simply references because the methodology to reflect on indices should be reviewed further when entrance and exit are considered and applied together in the imputation for missing values.

Table 5 Monthly Influences of Estimation Biases on Changes of the Official CPI for the Carrying Forward or Backward, the Imputation and the Estimation with the Regression Equation in 2008 (Percent Point)

	entrance		exit		exit		
month	Car.	Reg.	Car.	Reg.	Car.	Reg.	Hed.
Jan.	0.00	0.00	-0.01	0.00	-0.02	0.00	0.01
Feb.	0.00	0.02	-0.02	-0.02	-0.01	-0.01	-0.01
Mar.	-0.01	0.00	0.01	0.01	0.00	0.01	-0.01
Apr.	0.01	0.00	0.00	0.00	0.01	0.00	0.02
May	-0.01	-0.01	0.00	0.00	-0.01	-0.02	-0.03
Jun.	-0.01	0.00	0.01	0.00	0.00	0.00	-0.01
Jul.	-0.01	0.00	0.00	0.01	-0.01	0.00	0.00
Aug.	0.00	0.00	0.01	0.00	0.01	0.00	0.01
Sep.	0.00	0.00	0.01	0.00	0.01	0.01	0.01
Oct.	0.00	0.00	0.01	0.01	0.00	0.00	0.00
Nov.	0.00	0.00	0.00	0.00	-0.01	-0.01	-0.01
Dec.	0.00	0.00	0.01	0.01	0.01	0.01	0.01
average	0.00	0.00	0.00	0.00	0.00	0.00	0.00
annual	-0.03	0.00	0.01	-0.01	-0.02	-0.02	0.00

Note:
Car. is the case with the carrying forward or backward of missing data.
Reg. is the case with the regression equation for the imputation of missing data.
Hed. is the case with the regression equation replacing all actual data with the estimation.
They are differences of weighted averages with municipal weights of monthly changes of aggregated rents per aggregated floor space by municipality from the original in a month.
Changes for the carrying forward or backward and the imputation with the regression equations are ratios of original results in the current month divided by the estimation in the previous month for entrance and ratios of the estimation in the current month divided by the original results in the previous month for exit.
Changes for the estimation with the regression equations are ratios of the estimation in the current month by the previous month.

From those results, no critical influence would appear to the national CPI by methods reviewed here. It implies that those methods can avoid influences to the national CPI as well as shrinking fluctuations of rent indices by type of houses and by municipality.

In addition, among those methods, the estimation with the regression equations is excellent in lowering excessive fluctuations of municipal indices but includes the largest biases from the original indices. Differences from the original indices are, however, smaller than differences among samples and show no

long-term tendency.

As a whole, from stability of rent indices, the best method is the estimation with the regression equations, the next the carrying forward and backward, while from estimation biases, it is difficult to identify the best method. On top of that, it is difficult by a single methodology to satisfy both of stability and approximation to the original data.

Strictly speaking, it is rational to implement the above comparisons after dividing the CPI by prices on the base year because the CPI is compiled by aggregating current prices per base prices with municipal weights when influences to the CPI are estimated. However, the same estimation is impossible at present as construction year has been surveyed since 2008 although base prices should be estimated based on the same method. Therefore, this paper roughly substituted with comparisons of change over the previous month of rents per floor space.

Conclusions

The regression equation with floor space, construction year and population density can be effective and useful as the quality adjustment method for rent indices. The method can be applied as the imputation of missing values or for the estimation of all rents.

All of methods of the imputation of missing values and the estimation reviewed in this paper do not shift change of rent indices in the national or annual level although they often do that in local or monthly level.

More detailed and long-term data should be applied to scrutinize further issues for improving the estimation considering practical situation. Especially, it is necessary to estimate rent indices again after accumulation of data for construction year which exist only after 2008. Other variables should be reviewed for the regression equation for the purpose of enhancing accuracy of estimation of indices. Further various combinations should be reviewed, including those for the imputation both for entrance and exit, both for the imputation and the estimation, or whatever.

References

Makoto Shimizu, Hedonic Approach for House Rents in Japan, 11th Ottawa Group Meeting, Neuchâtel, 27-29 May 2009

ILO/ IMF/ OECD/ Statistical Office of the European Communities/ UN/ The International Bank for Reconstruction and Development/ The World Bank, Consumer Price Index Manual: Theory and Practice, 2004

Website of the Statistics Bureau, Japan; http://www.stat.go.jp