Methodology of JRPPI:
Japan Residential Property Price Index

August 2020

Real Estate and Construction Economy Bureau
Ministry of Land, Infrastructure, Transport and Tourism

Contents

1. Outline of the Japan Residential Property Price Index (JRPPI)

The Residential Property Price Index is an index of prices of residential property (residential plots of land and unit ownership buildings) nationwide that is calculated based on the data accumulated through the System to Provide Real Estate Transaction Price Data (Land General Information System) operated by the MLIT, with the quality of each property adjusted by using the Hedonic Approach.

The outline of the Japan Residential Property Price Index is as follows.

Table: Outline of the Residential Property Price Index

Index name (Use)	・Residential property { ・Residential Land ・Detached House ・Condominiums （mainly used）
Regions	・Nationwide(Japan) ・Regions（Hokkaido, Tohoku, Kanto, Hokuriku, Chubu, Kinki, Chugoku, Shikoku, Kyushu-Okinawa, Total 9 regions） ・Metropolitan Areas（Tokyo including suburbs, Nagoya including suburbs, Osaka including suburbs） ・Prefectures (Tokyo, Aichi, Osaka)
Nature of transactions	・Transactions between private bodies. Excludes transactions with public bodies.
Term	・From April, 2008 (Tokyo, Aichi and Osaka Residential property and Condominiums : From April, 2007 Residential Land and Detached House : From April, 1984)
Base year	・The calendar year 2010 is used as the base year (Arithmetic mean of 2010 equals 100)
Periodicity	・Monthly
Method of calculation	・Hedonic Approach(time dummy variable method)
Method of seasonal adjustment	・X-12-ARIMA Seasonal adjustment method
Data	・Transaction price by questionnaire
Lag between transaction period and publication	・Approximately three months (the preliminary index shall be updated for three months and then fixed)

Regions: Hokkaido: Hokkaido. Tohoku: Aomori, Iwate, Miyagi, Akita, Yamagata, Fukushima, and Niigata. Kanto: Ibaraki, Tochigi, Gunma, Saitama, Chiba, Tokyo, Kanagawa, and Yamanashi. Hokuriku: Toyama, Ishikawa, and Fukui. Chubu: Nagano, Shizuoka, Gifu, Aichi and Mie. Kinki: Shiga, Kyoto, Osaka, Hyogo, Nara, and Wakayama. Chugoku: Tottori, Shimane, Okayama,

Hiroshima and Yamaguchi. Shikoku: Tokushima, Kagawa, Ehime, and Kochi. Kyushu-okinawa: Fukuoka, Saga, Nagasaki, Kumamoto, Oita, Miyazaki, Kagoshima and Okinawa.
Metropolitan areas: Tokyo including suburbs: Saitama, Chiba, Tokyo and Kanagawa. Nagoya including suburbs: Gifu, Aichi, and Mie. Osaka including suburbs: Kyoto, Osaka, and Hyogo.

2. Data Generating Process

Data used for compiling the Residential Property Price Index are the data accumulated through the System to Provide Real Estate Transaction Price Data (Land General Information System) operated by the MLIT (hereinafter referred to as the "Transaction Case Data").

(1) System to Provide Real Estate Transaction Price Data

The System to Provide Real Estate Transaction Price Data aims to enhance the credibility and transparency of the real estate market, thereby facilitating and vitalizing real estate transactions. Under this system, information on real estate prices in actual transactions, which is compiled based on the questionnaire survey targeting transaction parties, is provided to the general public.

This system was commenced in some areas in FY2005, and since FY2008, the survey has been conducted nationwide. Such information as the transaction price, time of the transaction, address, floor area, age, and nearby stations, etc., is accumulated for each transaction, and is published after processing for secrecy, on a quarterly basis, on a webpage titled "Land General Information System"[1] on the MLIT website.

Prices in the Transaction Case Data are those of land or land with buildings that were determined in actual transactions. Transactions in the market involve various circumstances that trigger aggressive buying or delay of buying decisions or otherwise affect the course of transactions, and rights and interests are sometimes attached to land and buildings. The Transaction Case Data are characterized as being crude information reflecting all these factors.

(2) Data Used for JRPPI

The Transaction Case Data are compiled in three steps, based on (i) information on changes in registry, (ii) questionnaire answers, and (iii) on-site surveys. What is actually published on the webpage titled "Land General Information System" is information after

[1] https://www.land.mlit.go.jp/webland_english/servlet/MainServlet

processing for secrecy.

JRPPI figures are based on the data from (i) information on changes in registry and (ii) questionnaire answers, while past figures are based on all data, additionally including information from (iii) on-site surveys.

(3) Definition of Area

Factors to formulate prices of residential property vary by the usage and location, etc., of the property. Therefore, in order to compile price indices properly, stratification is required by classifying residential property into appropriate subpopulations.

For example, residential property can be classified by usage into residential zones, commercial zones, and industrial zones, etc., or can be stratified by prefecture, region, or metropolitan area.

Since factors to formulate prices of real estate vary significantly depending on the usage and location of the property, indices by usage and location are made through stratification of residential property into homogeneous strata when compiling the Residential Property Price Index.

Regarding locations, residential property is classified by region and by metropolitan area. Regions are divided as follows, in light of geographical and economic segmentation and in consideration of the necessity to secure a certain number of stable samples for calculating indices.

Table: Stratification by Location

Nationwide index	Index by region	Index by metropolitan area	Prefecture name
Nationwide(Japan)	Hokkaido		Hokkaido
	Tohoku		Aomori
			Iwate
			Miyagi
			Akita
			Yamagata
			Fukushima
			Niigata
	Kanto		Ibaraki
			Tochigi
			Gunma
		Tokyo including suburbs	Saitama
			Chiba
			Tokyo
			Kanagawa
			Yamanashi
	Hokuriku		Toyama
			Ishikawa
			Fukui
	Chubu		Nagano
			Shizuoka
		Nagoya including suburbs	Gifu
			Aichi
			Mie
	Kinki		Shiga
		Osaka including suburbs	Kyoto

Nationwide index	Index by region	Index by metropolitan area	Prefecture name
			Osaka
			Hyogo
			Nara
			Wakayama
	Chugoku		Tottori
			Shimane
			Okayama
			Hiroshima
			Yamaguchi
	Shikoku		Tokushima
			Kagawa
			Ehime
			Kochi
	Kyushu-Okinawa		Fukuoka
			Saga
			Nagasaki
			Kumamoto
			Oita
			Miyazaki
			Kagoshima
			Okinawa

Note) Indices colored in light gray are published.

(4) Definition of Usage

Transaction Case Data include not only residential properties, but also commercial properties such as office, retail and industrial properties.

In the JRPPI figures, which focus on the residential properties without commercial properties, we publish the three series: “residential land”, “detached house”, and “condominiums.”

Usage of land in the Residential Property Price Index is defined as follows.

Table: Definition of Usage

Usage		Definition
Residential Property	Residential Land	・ Property whose transfer was registered under the land category “residential land” ・ Property was found to be “land” or “limited proprietary right of land” through the on-site surveys or the questionnaire survey. ・ Property was found to be “residential land” through the on-site surveys or GIS, or the questionnaire survey.
	Detached House	・ Property whose transfer was registered under the land category “residential land” ・ Property was found to be “detached house” through the on-site surveys or the questionnaire survey. ・ Property was found to be “residential detached house” through the on-site surveys or the questionnaire survey.
	Condominiums	・ Property whose transfer was registered as a “unit ownership building” or a “site for a unit ownership building” ・ Property was found to be “residential” through the registered data.

Note: We use the on-site survey data for the properties transacted before March 2014. After April 2014, we mainly use the questionnaire survey data.

(5) Data Cleaning

As the Transaction Case Data are compiled based on information from the questionnaire survey, in addition to information on changes in registry, attribute information (transaction prices, the area, distance from the nearest station, etc.) of each estate may contain errors or missing parts. Therefore, data in which specific attribute information is missing and data that contain values beyond the minimum and maximum ranges of the outliers predetermined by prefecture are eliminated from the set of data for calculations.

However, in order to maintain the largest number of data samples to the extent possible, elimination of missing and abnormal values is limited to the minimum necessary.

With regard to abnormal values, data that contain values beyond the minimum and maximum ranges of the outliers (maximum and minimum values) predetermined by prefecture are eliminated from the set of data for calculations.

Table: Attributes Whose Abnormality Causes Elimination from Calculation Data

Usage	Targeted item	Minimum	Maximum
Residential Land	Unit price (yen/m^2)	One-half of the record low of "Land Market Value Publication" in each prefecture.	Double of the record high of "Land Market Value Publication" in each prefecture.
	The area (m^2)	One-half of the record low of "Land Market Value Publication" in each prefecture.	Double of the record high of "Land Market Value Publication" in each prefecture.
Detached House	Unit price (yen/m^2)	One-half of the record low of "Land Market Value Publication" in each prefecture.	Double of the record high of "Land Market Value Publication" in each prefecture.
	The land area (m^2)	One-half of the record low of "Land Market Value Publication" in each prefecture.	Double of the record high of "Land Market Value Publication" in each prefecture.
	Age (years)	0	60
	The floor area (m^2)	25(1)	-
	The floor area (m^2) / The land area (m^2)	-	3(2)
Condominium	Unit price (yen/m^2)	10,000	5,000,000
	The area (m^2)	10	500

	Age (years)	0	50
	Total number of floors	1	60
	Number of floors	1	Total number of floors

Note: The record level of “Land Market Value Publication” is the minimum or maximum in each prefecture from 1984.

(1) We assume the minimum floor area for the detached house is single household’s minimum living level.

(2) When “The floor area (m^2) / The land area (m^2)” is over three (300%), we assume that the property is “commercial property” such as the apartment building.

3. Calculation Model

In the calculation of the Residential Property Price Index, the hedonic regression method, or in particular, the time dummy variable model is adopted. Furthermore, in order to minimize the necessity to revise time series data, the rolling window method (or the moving window method) is adopted.

(1) Concept in Calculating the Residential Property Price Index

Each house has different attributes concerning location, equipment, specification, age, etc. There are no two properties with completely identical attributes. Additionally, the quality of houses changes rapidly due to technological innovation, etc., and transactions of houses are characteristically sporadic. In other words, residential property has the uniqueness of being inhomogeneous and its market is highly specific.

Therefore, in order to compare the value of transaction prices at different points in time, fluctuations in housing prices need to be broken down into fluctuations due to changes in attributes and sheer price fluctuations.

One of the methods of compiling a quality adjusted price index is the Hedonic Approach.

(2) Outline of Residential Property Price Index Calculation Models

The following is the explanation of the calculation model of the Residential Property Price Index.

First of all, the transaction price of **Real Estate** n at **Point in Time** t is set to be p_n^t. Each **Real Estate** n is presumed to have K number of **Attributes** z_{nk}^t. Supposing that there are data over the **Periods** $1, 2, \cdots, T$, **Time Range** τ starting from **Period** r is indicated as $[r, r+\tau-1]$. **Time Range** τ is here called as "window time range" and τ represents the length of the window[2].

With regard to **Time Ranges** $[1, \tau], [2, \tau+1], \cdots, [r, r+\tau-1], \cdots, [T-\tau+1, T]$, the following model is applied sequentially.

[2] For example, supposing the length of **Time Range** τ to be 12 months and the first period to be April 2008, calculation is first made for **Time Range [1,12]** (from April 2008 to March 2009), and then for **Time Range [2,13]** (from May 2008 to April 2009).

$$\ln p_n^t = \beta_0 + \sum_{t=1}^{\tau} \delta^t D^t + \sum_{k=1}^{K} \beta_k z_{nk}^t + \epsilon_n^t$$

Variables are as follows:

p_n^t : Transaction price of **Real Estate** n at **Point in Time** t

β_0 : Constant term

δ^t : Time dummy parameter at **Period** t

D^t : Time dummy variable ("1" at the time of the transaction and "0" at any other time (the base point in time is "0"))

β_k : Parameter for **Attribute** k of a house (supposing that it does not change during **Time Range** τ)

z_{nk}^t : Attribute value of **Attribute** k of **Real Estate** n at **Point in Time** t

ϵ_n^t : Error term

"**ln**" represents a natural logarithm, and "t" represents "$1, 2, \cdots, \tau$" (periods numbered sequentially from the first one in **Time Range** τ that is included in the **Periods** $1, 2, \cdots, T$, for which data are available).

Next, **Time Dummy Parameter** $\hat{\delta}^t$ obtained by applying the aforementioned model to **Time Range** $\tau([r, r+\tau-1])$ starting from **Period** r out of all time ranges (**Periods** $1, 2, \cdots, T$) is indicated as $\hat{\delta}_{(r)}^r, \hat{\delta}_{(r)}^{r+1}, \cdots, \hat{\delta}_{(r)}^{r+\tau-1}$ by explicitly showing **Period** r.

Table: Time Dummy Parameters by Window

r \ t	1	2	3	⋯	τ	$\tau+1$	$\tau+2$	⋯	$T-\tau+1$	⋯	T
1	$\hat{\delta}_{(1)}^1$	$\hat{\delta}_{(1)}^2$	$\hat{\delta}_{(1)}^3$	⋯	$\hat{\delta}_{(1)}^{\tau}$						
2		$\hat{\delta}_{(2)}^2$	$\hat{\delta}_{(2)}^3$	⋯	$\hat{\delta}_{(2)}^{\tau}$	$\hat{\delta}_{(2)}^{\tau+1}$					
3			$\hat{\delta}_{(3)}^3$	⋯	$\hat{\delta}_{(3)}^{\tau}$	$\hat{\delta}_{(3)}^{\tau+1}$	$\hat{\delta}_{(3)}^{\tau+2}$				
⋯				⋯	⋯	⋯	⋯	⋯	⋯	⋯	
$T-\tau+1$									$\hat{\delta}_{(T-\tau+1)}^{T-\tau+1}$	⋯	$\hat{\delta}_{(T-\tau+1)}^{T}$

When the log price index during **Time Range** $[1, \tau]$ (the first time range), i.e., $q^r \quad (r = 1, 2, \cdots, \tau)$, is defined as $q^1 = 0,\ q^2 = \hat{\delta}_{(1)}^2, \quad \cdots,\ q^{\tau} = \hat{\delta}_{(1)}^{\tau}$, the subsequent **Log Price Index** $q^{\tau+1}$ can be indicated as $q^{\tau+1} = q^{\tau} + (\hat{\delta}_{(2)}^{\tau+1} - \hat{\delta}_{(2)}^{\tau})$.

Specifically, **Quality Adjusted Price Index** $P^{\tau+1}$ converted from log to antilog is as follows.

$$\frac{P^{\tau+1}}{P^{1}} = \exp\left(\hat{\delta}^{\tau}_{(1)}\right) \times \frac{\exp\left(\hat{\delta}^{\tau+1}_{(2)}\right)}{\exp\left(\hat{\delta}^{\tau}_{(2)}\right)}$$

Therefore, when supposing the first period as the base point, the difference between price at **Period 1** and price at **Period $\tau+1$** can be obtained based on the time dummy parameter calculated for the last period of the first window time range (**Point in Time τ**) and the time dummy parameters calculated for the last and the second to the last periods of the next window time range (**Point in Time τ** and **Point in Time $\tau+1$**).

By sequentially conducting the aforementioned calculations for all window time ranges, quality adjusted price indices can be obtained for all time ranges by assuming the base point in time to be "1."

In calculating the Residential Property Price Index, **Window Time Range τ** is set to be 12 months. Window time ranges may be set free from any theoretical constraint, but in general, the window time range for the Residential Property Price Index should be no shorter than one year (12 months) because transactions in the housing market vary seasonally, being active at the fiscal year end (February and March) and becoming slow thereafter. However, on the other hand, when a window time range is too long, it will become difficult to ascertain changes in the market structure appropriately. Therefore, the window time range at the stage of trial operation of the Residential Property Price Index is set to be 12 months.

(3) Explanatory Variables

Attribute information available under data acquisition constraints that contains few missing values, affects prices significantly, and is also commonsensical in terms of sign condition is selected as explanatory variables for calculating the Residential Property Price Index.

Specifically, the following variables are adopted with regard to attributes that define the basic qualities of real estate, such as size, closeness, newness, location and transaction terms.

Table: Explanatory Variables Adopted in Calculating the Residential Property Price Index

Attributes	Explanatory variables	Residential Land	Detached House	Condominiums
Size	The area	X	X	X

Attributes	Explanatory variables	Residential Land	Detached House	Condominiums
	On what floor			X
	Total floor area		X	
	Total number of floors			X
Closeness	Distance from the nearest station	X	X	X
	Distance from the main station	X	X	X
Newness	Age		X	X
	Renovated			X
Location	City	X	X	X
	Zoning	X	X	X
	Facing south			X
Transaction Terms	Transaction party	X	X	X

(4) Basis for Calculation of the Index

The Residential Property Price Index is calculated by assuming that the monthly indexation by average for the period between January 2010 and December 2010 equals 100. The time series assuming the values for first month to be 100 are obtained through the aforementioned Hedonic Approach based on the time dummy variable model using the rolling window method. Therefore, based on the time series thus obtained, the arithmetic mean for 2010 is assumed to be 100 and is used as the calculation basis.

Since the survey on the original case data was begun nationwide in FY2008 (from April 2008 onward), the calculation of the Residential Property Price Index also starts from April 2008.

(5) Methods of Integration into Superior Indices

In addition to the Residential Property Price Index by region and by usage, the General Residential Property Price Index (integrating usage) and the Nationwide General Residential Property Price Index (integrating regions) are also calculated as their superior indices.

By metropolitan area, the index for the Tokyo including suburbs area, index for the Nagoya including suburbs area and the index for the Osaka including suburbs area are

calculated. Those indices are calculated by integrating indices for prefectures belonging to the respective areas.

In order to integrate the Residential Property Price Index by region, by usage and by prefecture into the nationwide index, the general index, and the index by metropolitan area, respectively, weighted averages are calculated based on monthly transaction amount.

In particular, we use the Fisher index, which is calculated as the geometric mean of Paasche and Laspeyres indices.

(6) Methods of Seasonal Adjustment

The time series data of the Residential Property Price Index released monthly are likely to have seasonality that could increase or decrease the index value depending on the season and the month. Therefore, it seems necessary to remove the seasonal effect inherent in the time series data.

As shown in the equation below, time series data are usually decomposed into: Tt (the trend component at time t); Ct (the cyclical component at time t); St (the seasonal component at time t); and It (the irregular (or error) component at time t.

Seasonal adjustment is generally carried out by removing the seasonal (S) and irregular (I) components from data series based on moving averages.

$$Yt = Tt + Ct + St + It$$

The decomposition of time series data can be achieved in the following way: (1) It is possible to remove the seasonal and irregular components by determining the seasonal fluctuation cycle of the time series data (Yt) and applying an appropriate moving average to the fluctuation cycle (for example, if the seasonal fluctuation cycle is four quarters, take a four-quarter moving average to smooth out the highs and lows observed among the periods); and (2) If the moving-average smoothing method in (1) was successful, the trend (T) and cyclical (C) components should remain in the time series. The cyclical (C) component can be obtained by removing the trend (T) component by utilizing data graphs and other materials.

For the Residential Property Price Index, the X-12-ARIMA seasonal adjustment method developed by the US Census Bureau was employed. X-12-ARIMA is based on the X-11 method which was commonly used for seasonal adjustment in the past.

The Residential Property Price index without seasonal adjustment has also been published.

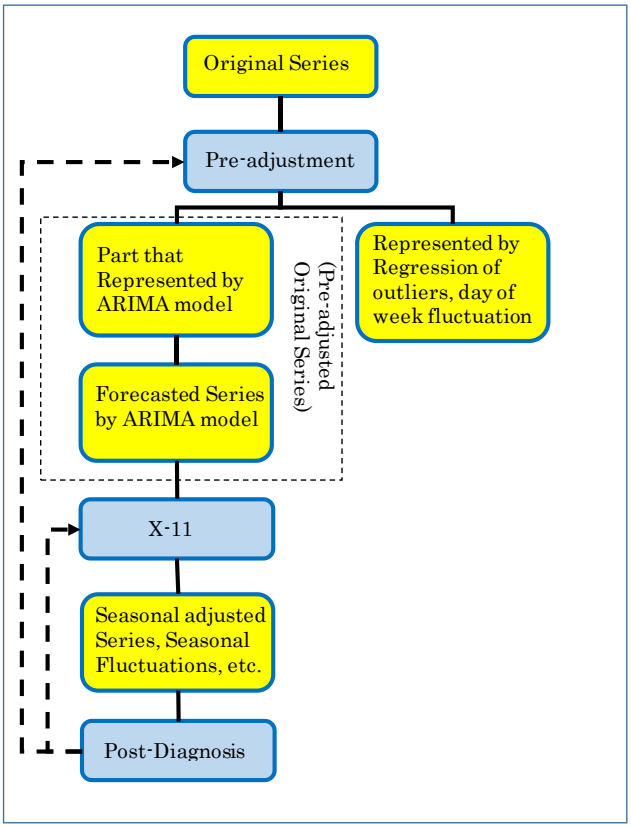
1) X-11 seasonal adjustment method

The X-11 seasonal adjustment method is a technique developed by the US Census Bureau in 1965. It repeatedly applies moving averages to an original time series in order to decompose it. Its seasonal component should be separated from the trend and irregular components. The X-11 procedure involves centered moving averages, for which each moving average is centered at the middle of the values being averaged. In terms of some latest values in a time series, however, there are not enough future values to calculate their centered moving averages. For those latest values, backward moving averages, which utilize values up to the newest value for calculating its average, are used instead. When another value of the most recent period is added to the time series in the future, the previously calculated averages of the latest values could be revised significantly.

2) X-12-ARIMA seasonal adjustment method

X-12-ARIMA was developed by the US Census Bureau to improve the X-11 and to adjust outliers and calendar effects. X-12-ARIMA uses the REGARIMA model that divides an original time series into the following two parts: (a) a portion that can be represented by a regression analysis of calendar effects and outliers; and (b) the remaining part that can be explained by the ARIMA model (Auto-Regressive Integrated Moving Average Model). Then, by making use of Item (b) above, the REGARIMA model generates a pre-adjusted original series that complements the original data with out-of-range values. Then, X-12-ARIMA makes seasonal adjustment to the pre-adjusted original series using the X-11 seasonal adjustment method. In this process, centered moving averages can be calculated for all values, including the most recent value, and outliers and calendar effects can be also adjusted. In addition, because X-12-ARIMA requires the analyst to carefully consider which model to represent the original series, it provides a post-diagnosis function to examine seasonally adjusted values. The figure below shows the process of seasonal adjustment by the X-12-ARIMA model.

Figure　Seasonal Adjustment Process by X-12-ARIMA



Source: Bank of Japan, "X-12-ARIMA Manual," Feb. 1997.

(Reference) Definition of Explanatory Variables

Definitions of the variables used in preliminary figures and revised figures are as follows.

(i) The area

In the case of condominiums, the area means as follows.

The area = Registered area of the exclusively-owned area

In the case of land and detached houses, the area means as follows.

The area = Size of the plot (registered area)

(ii) On what floor

On what floor = On what floor the exclusively-owned unit is located

(iii) Total floor area

Total floor area = Total floor area of the building

(iv) Total number of floors

Total number of floors = Total number of floors of the condominium

(v) Distance from the nearest station

The linear distance from the center of the town section, etc., where the estate is located to the nearest station is used.

Distance from the nearest station = Linear distance from the center of the town section, etc., where the estate is located to the nearest station

(vi) Distance from the prefectural capital

The linear distance from the center of the town section, etc. where the estate is located to the capital of the prefecture where the estate is located (the central station of the prefectural capital) is used.

Distance from the prefectural capital = Linear distance from the center of the town section, etc. where the estate is located to the capital of the prefecture where the estate is located

(vii) Age

The age of buildings is calculated as follows.

Age = (Date of causes of registration – Date of construction) / Number of days

Data in which the date of construction is missing are processed as follows.

- **When only the year of construction is known, the date of construction is considered to be January 1 of said year.**
- **When the month of construction is known, the date of construction is considered to be the first day of said month.**

(viii) Renovated or not

In the case of condominiums, information on whether or not the renovation has been done is used.

Renovation dummy = "1" when the renovation has been done, and "0" when not

(ix) City

The City dummy is set as follows.

Prefecture dummy = "1" when the transaction is made within the city, and "0" when not

(x) Zoning Code

Zoning Code dummy = "1" when the estate is located in each zoning area, and "0" when not

(xi) Facing south or not

South-facing dummy = "1" when the estate faces south, and "0" when not

(xii) Transaction party

Dummy variables to show which party is the seller and which is the purchaser in a real estate transaction are used. Sellers and purchasers are categorized as follows.

Categories of sellers and purchasers = 1: Individual
2: Private corporation
3: Local government
4: National government

Since we exclude the data which the public entity engages in JRPPI, the following dummy variables are set for the attributes of sellers and purchasers.

Corporation seller dummy = "1" when the seller is categorized into 2, and "0" when not

Corporation purchaser dummy = "1" when the purchaser is categorized into 2, and "0" when not