TI 2018-061/III
Tinbergen Institute Discussion Paper

Earthquake risk embedded in property prices: Evidence from five Japanese cities

Masako Ikefuji[1]
Roger J. A. Laeven[2]
Jan R. Magnus[3]
Yuan Yue[4]

[1] School of Social and International Studies, University of Tsukuba, Japan
[2] Amsterdam School of Economics, University of Amsterdam, CentER, and EURANDOM
[3] Vrije Universiteit Amsterdam and Tinbergen Institute
[4] Amsterdam School of Economics, University of Amsterdam

Earthquake risk embedded in property prices: Evidence from five Japanese cities*

Masako Ikefuji

School of Social and International Studies, University of Tsukuba

Roger J. A. Laeven

Amsterdam School of Economics,

University of Amsterdam, CentER, and EURANDOM

Jan R. Magnus

Vrije Universiteit Amsterdam and Tinbergen Institute

Yuan Yue

Amsterdam School of Economics, University of Amsterdam

July 11, 2018

*Earlier versions of this paper were presented at the Technische Universität Wien, at the Universities of Oxford, Madrid and Fudan, and at the 2018 Econometric Society Australasian Meeting in Auckland and the 2018 China Meeting of the Econometric Society in Shanghai. We are grateful to the participants and to Yoshitsugu Kanemoto for helpful comments. This research was supported in part by the Japan Society for the Promotion of Science (JSPS) under grant 16K03565 (Ikefuji) and the Netherlands Organization for Scientific Research (NWO) under grant VIDI (Laeven).

Abstract: We analyze the impact of short-run and long-run earthquake risk on Japanese property prices. We exploit a rich panel data set of property characteristics, ward attractiveness information, macroeconomic variables, seismic hazard data, and historical earthquake occurrences, supplemented with short-run earthquake probabilities that we generate from a seismic excitation model. We design a hedonic property price model that allows for probability weighting, employ a multivariate error components structure, and develop associated maximum likelihood estimation and variance computation procedures. We find that distorted short-run and long-run earthquake probabilities have a significantly negative impact on property prices. Our approach enables us to identify the total compensation for earthquake risk embedded in property prices and to decompose this into pieces stemming from short-run and long-run risk, and to further decompose this into objective and distorted risk components.

JEL Classification: R20; C33; D81; Q51

Keywords: Earthquake risk; House price; Seismic excitation; Probability weighting; Hedonic pricing; Multivariate error components.

Corresponding author:
Roger J. A. Laeven
University of Amsterdam
Economics and Business
Amsterdam School of Economics
Department of Quantitative Economics
PO Box 15867
1001 NJ Amsterdam
The Netherlands

Phone: +31 20 525 4219/4252
Email: `R.J.A.Laeven@uva.nl`

The flow of the river is ceaseless
and its water is never the same.
The bubbles that float in the pools,
now vanishing, now forming,
are not of long duration:
so in the world are man and his dwellings.

Kamo no Chomei (1153–1216)

1 Introduction

After having witnessed a number of natural and personal catastrophes, in particular the great earthquake of 1185, the Japanese author and poet Kamo no Chomei decided to live as a hermit, in the forest outside the capital Kyoto. His famous essay Hojoki ('An Account of My Hut') opens with the displayed poetic lines (in the translation of Donald Keene), in which he puts the catastrophes in Buddhist perspective. More specifically, the essay argues that when a catastrophe strikes, one tends to reflect on the impermanence of property and the evil and futility of attachment — at least, in the short-run. In the long-run, however, one forgets these views and goes back to life as before.

In the spirit of Chomei's essay but with modern economics and statistics of risk, this paper analyzes the subjective evaluation of both short-run and long-run catastrophic risk, earthquake risk in particular, embedded in Japanese property prices.

It is well-known that earthquakes tend to occur in clusters rather than in isolation. These seismic clusters may take the form of foreshocks and aftershocks anticipating and following a major earthquake or of a collection of major earthquakes triggering one another by causing frictions that put strain on neighboring faults. There is therefore objective predictive content embedded in the occurrence of earthquakes. This phenomenon is known as seismic excitation and there exists a large literature in statistics aimed at capturing it.

In a different strand of the literature in economics, several papers analyze the impact of natural catastrophes on property prices. Most commonly, this literature incorporates the prevailing binary state of the world, depending on whether or not a catastrophe has occurred, into a hedonic house price model of the Rosen (1974) type, which has become the benchmark model in analyzing property prices. Within a typical hedonic price model, the characteristics of a property are viewed as detachable components that each contribute to a part of the property price. The selection of components range from traditional house attributes such as square footage, location and building age, to external factors such as macroeconomic effects. The negative effect coming from hazardous environmental events, such as flood, hurricane and earthquake, has been addressed by various researchers; see, among others, Brookshire *et al.* (1985), Kawawaki and Ota (1996), Beron *et al.* (1997), Yamaga *et al.* (2002), Bin and Polasky (2004), Nakagawa *et al.* (2007, 2009), Daniel *et al.* (2009), Naoi *et al.* (2009), Gu *et al.* (2011), Bin and Landry (2013), Hanaoka *et al.* (2015), and Hidano *et al.* (2015).

In recent years, a large body of literature has documented empirically that people do typically not treat objectively given probabilities linearly, but rather tend to overweight small probability events and underweight large probability events. This is particularly relevant when evaluating catastrophic events that are often of a low-probability high-impact nature. Various modern theories of decision under risk, such as rank-dependent utility theory and prospect theory, feature a probability weighting function that 'distorts' objective probabilities.

We contribute to this literature by introducing into a hedonic price model an objective measure of seismic excitation, next to a more conventional measure of long-run earthquake risk, while allowing for probability weighting in the spirit of the non-expected utility theories of rank-dependent utility and prospect theory. We use a rich panel data set containing property characteristics, ward attractiveness information, macroeconomic variables, seismic hazard data, and historical earthquake occurrences. We design a hedonic price model with a multivariate error component structure (Baltagi, 1980, 2008; Magnus, 1982) for which we develop associated maximum likelihood estimation and variance computation procedures. By exploiting the matrix form of the error components, we are able to estimate the model while pooling properties of different types together, in spite of the very large dimension of the variance matrix and the fact that each property type corresponds to different features and total price levels. Our approach allows to isolate the total compensation for earthquake risk embedded in Japanese property prices, and to decompose this into pieces stemming from short-run risk and long-run risk, and a further decomposition into objective and distorted risk components.

We can summarize our main findings as follows. First, we find that objective long-run earthquake risk has a significant negative impact on property prices, and increasingly so at higher risk levels. Second, given that long-run risk matters for property prices, we find that the additional impact of objective short-run earthquake risk on property prices, while estimated at negative values, is not significantly different from zero. Upon allowing for probability weighting, however, the distorted short-run earthquake probabilities do have a significantly negative effect on property prices. Third, the probability weighting function for short-run earthquake risk is found to be S-shaped, thus underweighting small probabilities and overweighting larger probabilities, contrary to the inverse-S shaped probability weighting function found in many experiments. This remarkable finding may be explained by the fact that the background arrival rate of earthquakes is positive rather than zero, in particular in Tokyo where the short-run earthquake probabilities never drop below 35% in the period that we analyze. Therefore, people may tend to evaluate and overweight temporary deviations of the short-run earthquake probabilities from the background seismicity caused by seismic excitation not with respect to zero but with respect to a positive reference probability level. In an extension of our base model, we also analyze probability distortions of long-run time-invariant earthquake probabilities. In this case, we find that small probabilities tend to be overweighted and large probabilities tend to be underweighted, in accordance with conventional wisdom.

Most of the studies on the interplay between property prices and environmental hazards investigate the risks of floods or earthquakes in the USA or Japan. In the case of the USA, Brookshire *et al.* (1985) analyze a hedonic house price model in an expected utility framework, examine self-insurance for earthquake hazards in Los Angeles and San Francisco, and show that buyers pay less for houses within a relatively risky area if they possess adequate information about earthquake hazards. Bin and Polasky (2004) estimate and compare the effects of flood hazards on property prices before and after Hurricane Floyd (the 1999 flooding in North Carolina), and show that the market price of a property located within a flood plain gets discounted by more than a property located outside the flood plain. Re-examining these findings, Bin and Landry (2013) estimate hedonic property prices for the same location with two major flooding events, and show that the implicit risk premia disappear rapidly.

In the case of Japan, Nakagawa *et al.* (2009), using the 1998 Tokyo hazard map, show strong negative impacts of earthquake risks on land prices. Gu *et al.* (2011), using an updated Tokyo hazard map, find that in previously safe areas, a decrease in risk rankings (even more safety) has a positive impact on relative land prices, while in previously dangerous areas, an increase in risk rankings (even more risk) has a negative effect. Naoi *et al.* (2009) estimate individuals' valuation of earthquake risk, based on nation-wide panel data of earthquake hazard information and records of observed earthquakes. They show that after a big earthquake people discount house prices and house rents within the earthquake area. Hidano *et al.* (2015) examine the effect of seismic hazard risk information on properties in Tokyo, and find that the price of properties in low-risk zones is higher than the prices in high-risk zones, but that for new more earthquake-resistant properties the influence of seismic hazard risk information is limited.

We also mention two survey-data studies on how risk preferences of households have changed after the Tohoku earthquake (the Great East Japan earthquake) in 2011. Naoi *et al.* (2012) find that although respondents seemed to be more prepared for natural disasters after the Tohoku experience, actual (costly) mitigation activities depend on household income. Hanaoka *et al.* (2015) examine whether risk preferences of men and women have changed, and if so whether they changed in a different way, after the Tohoku earthquake. There is some evidence that men have become more risk tolerant, while women have become more risk averse. Finally, our work is also related to the financial econometrics literature on the estimation of risk and financial excitation premia embedded in asset and derivative prices; see Aït-Sahalia *et al.* (2014, 2015), Bauer and Kramer (2016), Boswijk *et al.* (2016), and Sperna Weiland *et al.* (2018).

The rest of this paper proceeds as follows. Section 2 explains our treatment of objective seismic excitation and of probability weighting. Section 3 describes the data set. Section 4 lays out our hedonic house price model with multivariate error components and Section 5 develops the procedures for estimation. Section 6 presents the estimation results. Section 7 analyzes the influence of each component to the total property prices and calculates the implied premia for earthquake risk. Section 8 discusses the robustness of our estimation results. Section 9 concludes. Supplementary material, including a detailed data description, is contained in an online appendix; see Ikefuji *et al.* (2018).

2 Seismic Excitation and Probability Weighting

In this section we consider short-run earthquake probabilities as objective measures of seismic excitation, and develop a regression design that allows for probability weighting.

2.1 Short-Run Earthquake Probabilities

Our approach estimates an Epidemic Type Aftershock Sequence (ETAS) model and generates a panel of model-implied short-run earthquakes probabilities which vary per quarter and per city to be used in our regression design. These probabilities can be viewed as objective measures of short-run earthquake risk, summarizing publicly available information per time period and per city.

The occurrence of major earthquakes have served previously in hedonic price models with regression discontinuity design as natural exogenous events to elicit causal pricing effects. Limitations of this conventional approach include the somewhat rudimentary binary nature of this treatment, which does not reflect the multiplicity of the events, the time elapsed since the last event, and the severity of the events. By contrast, our approach relies on a continuous-time predictive earthquake intensity that depends on all previous earthquakes, with recent ones being more important than older ones, and explicitly accounts for the severity of the events. Furthermore, the earthquake intensity can be translated into objective short-run probabilities enabling us to analyze probability weighting.

The ETAS model was introduced by Ogata (1988) and has since been widely used to capture the quiescence and activation of seismic dynamics. The basic idea of the model is that each earthquake can trigger a sequence of aftershocks like 'epidemics' in that the occurrence of an earthquake makes future earthquakes more likely and that the impact of the trigger event diminishes over time (and distance). Despite the existence of several space-time extensions, we choose the temporal version of the ETAS model as described in the following, which we estimate separately for each of the five cities. Because we consider five cities this treatment is natural and simpler than first estimating a space-time version to a large area that covers all five cities and next isolating the city effects.

Formally, the ETAS model is a path-dependent marked point process and a special case of a Hawkes self-exciting process. Given observations of earthquake occurrences at times $t_1, t_2, \ldots, t_n$ over an interval $[0, T]$ $(T \geq t_n)$, the associated counting process N_t is defined as $N_t = \sum_{i=1}^{n} \mathbb{1}_{t_i \leq t}$. Denoting by $\mathcal{F}_t$ the information filtration up to time t, the corresponding left-continuous $\mathcal{F}_t$-conditional jump intensity process λ_t describes the mean jump rate per unit of time,

$$\lambda_t = \lambda(t|\mathcal{F}_t) = \lim_{h \downarrow 0} \frac{1}{h} \ \Pr\left[N_{t+h} - N_t > 0 | \mathcal{F}_t\right].$$

In the temporal ETAS model, the conditional intensity function may be written as

$$\lambda_t = \lambda_\infty + \sum_{t_i < t} c(m_i, m_c) g(t - t_i),$$

where $\lambda_\infty > 0$ (measured in number of jumps per time unit) is the background seismicity, $g(t - t_i)$ is the aftershock decay (i.e., time response) function, and the weight assigned to the aftershock decay is a function $c(m_i, m_c)$ of the magnitude of the earthquake m_i and a cut-off (i.e., threshold) magnitude m_c. Thus, the earthquake intensity depends on the background intensity and a weighted sum of all aftershock decays, where the sum is taken over all earthquakes that have occurred before time t. (In the ETAS model, g takes the form of a so-called modified Omori law and c takes an exponential form.)

We estimate the ETAS model for each of the five cities that we consider, based on the earthquake catalog of five areas covering the five cities, over the period January 1, 1970, to December 31, 2015. Next, the estimated intensities are used to generate by simulation 90-days probabilities of an earthquake exceeding a magnitude threshold of 5.5, for each city. Our simulation method follows Ogata (1981). Further details about the parameterization, estimation and simulation within the ETAS model are contained in online supplementary material (Ikefuji *et al.*, 2018).

In Figures 1 and 2 we plot the earthquake intensities along with the corresponding short-run probability series for two of the five cities, Tokyo and Nagoya. The probabilities spike up immediately after a large earthquake and die out gradually until another major earthquake occurs. The Tohoku earthquake of Friday 11 March 2011 was the most powerful earthquake ever recorded in Japan. The spike is visible in 2011/Q2 (rather than in 2011/Q1) because the short-run probabilities are simulated based on actual earthquakes up to and including the previous quarter.

The objective measure of seismic excitation given by the 90-days earthquake probabilities is included in our regression design. The rationale is that, in addition to the long-run earthquake risk that people may take into consideration when purchasing a property, news from a recent nearby earthquake may also temporarily affect property prices. Just like objective seismic excitation generated by a self-exciting process, the impact of such bad news on people's perception of risk peaks right after the event and dies out as time proceeds.

2.2 Probability Weighting

To account for probability weighting our regression design furthermore allows for a parametric probability weighting function. There is a large literature on probability weighting. Probability weighting is an important ingredient of prospect theory (Kahneman and Tversky, 1979; Tversky and Kahneman, 1992), and of the related decision theories given by the dual theory of choice under risk (Yaari, 1987) and rank-dependent utility (Quiggin, 1982), which are building blocks of prospect theory.

We shall consider two canonical one-parameter families of probability

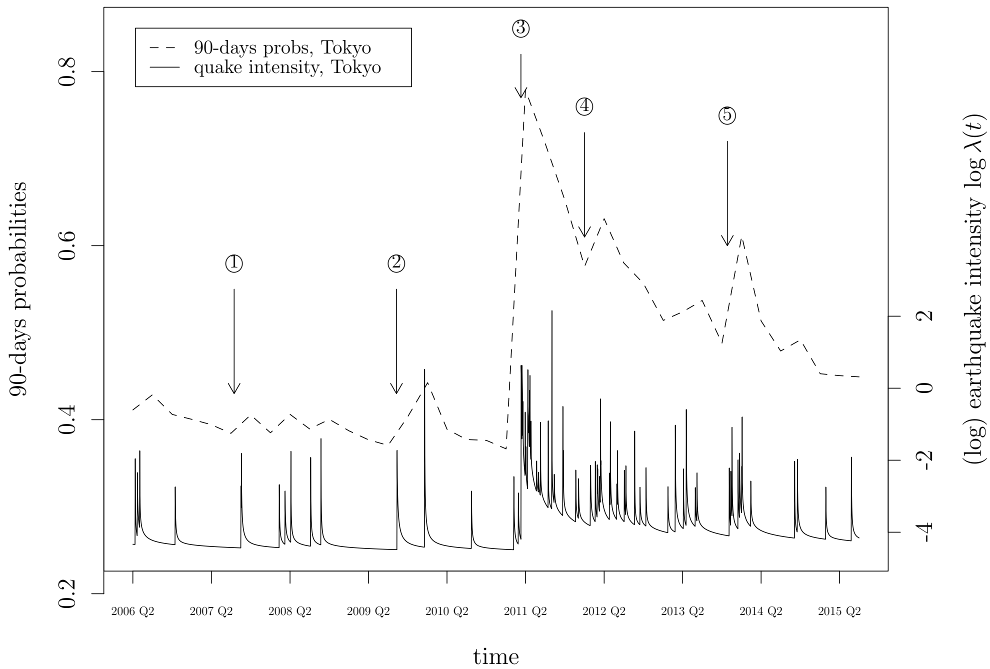

Figure 1: Short-run earthquake risk: Simulated short-run earthquake probabilities and the logarithm of the earthquake intensity series for Tokyo. Events marked in the graph: ①: 2007-07-16 Chuetsu Offshore earthquake, $M6.8$. ②: 2009-08-09 Izu Islands earthquake, $M6.8$ and 2009-08-11 Shizuoka earthquake, $M6.5$. ③: 2011-03-11 Tohoku earthquake, $M9.0$. ④: 2012-01-01 Izu Islands, $M7.0$. ⑤: 2013-10-26 Fukushima-ken oki earthquake, $M7.1$. (Source: Japan Meteorological Agency.)

weighting functions, proposed by Tversky and Kahneman (1992) and Prelec (1998), respectively. The Tversky-Kahneman function — see also Wu and Gonzalez (1996) — is given by

$$w(p) = \frac{p^\psi}{(p^\psi + (1-p)^\psi)^{1/\psi}}, \tag{1}$$

while the Prelec function is given by

$$w(p) = e^{-(-\log p)^\psi}. \tag{2}$$

The parameter ψ is restricted to be positive. When $0.279 < \psi < 1$ the Tversky-Kahneman function is inverse S-shaped, while the Prelec function is inverse S-shaped for $0 < \psi < 1$; when $\psi = 1$ both functions reduce to

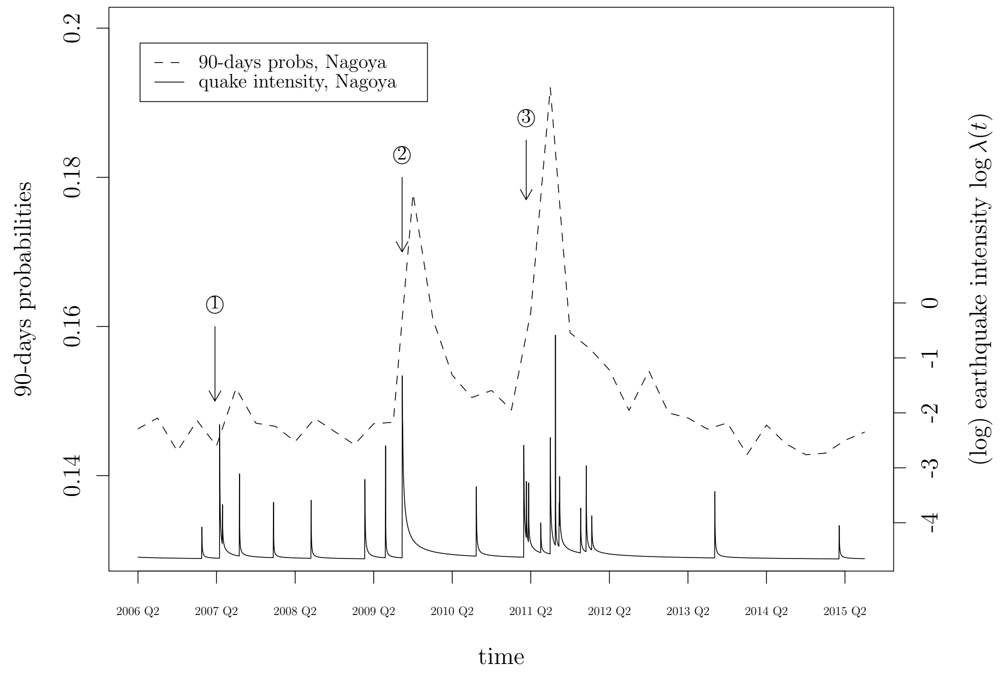

Figure 2: Short-run earthquake risk: Simulated short-run earthquake probabilities and the logarithm of the earthquake intensity series for Nagoya. Events marked in the graph: ①: 2007-03-25 Noto Hanto earthquake, $M6.9$. ②: 2009-08-11 Shizuoka earthquake, $M6.5$. ③: 2011-03-11 Tohoku earthquake, $M9.0$. (Source: Japan Meteorological Agency.)

$w(p) = p$; and when $\psi > 1$ both functions are initially S-shaped, but (only) the Tversky-Kahneman function becomes convex for large values of ψ.

In laboratory experiments (see Wu and Gonzalez, 1996, and Abdellaoui, 2000) the probability weighting function is often found to be inverse S-shaped, first concave and then convex. An inverse S-shape captures the phenomenon that people tend to become less sensitive to changes in objective probabilities as these probabilities move further away from the reference point 0 and become more sensitive as they get closer to the reference point 1. The inverse S-shape is consistent with a positive third derivative of the probability weighting function. The interpretation of the signs of the successive derivatives of the probability weighting function was recently provided by Eeckhoudt *et al.* (2017). Note that contrary to the Tversky-Kahneman function the Prelec function has an invariant fixed point and inflection point at $p = 1/e = 0.37$, which implies that it can never be globally convex or concave.

3 The Data

The data collection process for this project has been complex and elaborate, and in this section we provide a brief summary. Full details and references to sources are available in an online appendix (Ikefuji *et al.*, 2018). We are interested in the impact of earthquake risk on property prices in major cities in Japan, and we have selected five cities for our purpose. Each city is divided into wards and each ward is divided into districts. (In the original data set the word 'area' is used. We prefer 'district' to avoid confusion with other uses of the word 'area'.) Certain information that can affect (and explain/predict) the attractiveness of buying a property is available per ward. For example, population characteristics, information about schools and medical facilities, shopping, safety, etc. We distinguish between three types of properties: 'residential land (land and building)', 'residential land (land only)', and 'pre-owned condominiums' (hereafter, condos). Sales prices and property characteristics are available for each of these types in each of the five cities. We do not know the exact location of a property, but we do know in which district the property lies and we also know the distance to the nearest station and the name of that station. Some macro variables are relevant and affect house prices nationally. Finally, we have information on historical earthquake data and on earthquake risk data.

Cities. Japan has twelve cities with a population of more than one million people. Almost 100 million people, or 78% of the country's total population of 127.4 million, live in urban areas. The total population of Japan's largest 103 cities amounts to 63.9 million or just over half of all the country's residents. Tokyo, with almost nine million inhabitants, is by far the largest Japanese city. (Strictly speaking, Tokyo is not a city — it is a prefecture, but we shall call it a city.) With a population of 3.7 million, Yokohama, south of Tokyo, is Japan's second largest city. Osaka and Nagoya are Japan's third and fourth cities, each with a population of over two million. Eight cities have between one and two million inhabitants: Sapporo, Kobe, Fukuoka, Kyoto, Kawasaki, Saitama, Hiroshima, and Sendai.

From these twelve cities we selected five: Tokyo, Osaka, Nagoya, Fukuoka, and Sapporo. This choice guarantees that each of the three major metropolitan areas is represented: the greater Tokyo area (Tokyo, Yokohama, Kawasaki, Saitama) by Tokyo, the Kansai region (Osaka, Kobe, Kyoto) by Osaka, and the Chukyo metropolitan area by Nagoya. To obtain a representative geographical spread we added Sapporo, the largest city in the North, and Fukuoka, the second largest city in the West after Osaka. Data limitations prevented us from including Hiroshima, while Sendai was not included because it is too close to Fukushima where the 2011 nuclear disaster took place following the Tohoku earthquake.

Wards. A designated city is a Japanese city that has a population greater than 500,000 and has been designated as such by order of the Cabinet of Japan. Designated cities are delegated many of the tasks normally performed by prefectural governments, such as public education, social welfare, sanitation, business licensing, and urban planning. Designated cities are required to subdivide themselves into wards ('ku'), each of which has a ward office conducting various administrative functions for the city government. The 23 special wards of Tokyo are not part of this system, as Tokyo is a prefecture, and its wards are effectively independent cities. The five cities together contain 80 wards (regular and special together): 23 in Tokyo, 24 in Osaka, 16 in Nagoya, 7 in Fukuoka, and 10 in Sapporo.

When considering to buy a property in a given city, one is likely to be interested in certain characteristics of these wards. The original data set contains one hundred characteristics divided into eleven categories. Since many of these are highly correlated we first select eleven of these divided into six categories: two from population; three from schools, culture and welfare; one from medical facilities; one from safety; two from shopping facilities; and two from employment. Only four of these appear in our base model, but extensive sensitivity analyses will be conducted in Section 8 to assess how adding more characteristics may affect the results.

Districts. Within each city there are wards, and within each ward there are districts (usually 'cho', sometimes 'machi'). An average ward in Nagoya contains 86 districts, an average ward in Osaka only 23. The number of districts ranges from 318 in Fukuoka to 1383 in Nagoya (1379 after prescreening). In total there are 3714 districts (3710 after prescreening) in the five cities together.

Property types. In a given district i we have observations on three types of (residential) properties: land and buildings, land only, and condos. Most properties are condos (45.1%), followed by land and buildings (34.1%) and land only (20.8%). We have observations over $T = 38$ quarters, from 2006/Q2 to 2015/Q3.

Records with obvious errors have been excluded. Also excluded are records where the walking time to the nearest station is longer than thirty minutes or the nearest station is unknown; records with a living area larger than 2000 square meters; and properties built before the war (1945). After

Table 1: Distribution of properties over cities, wards, and districts

City	Ward	District	Land & building	Land only	Condo	Station
Tokyo	23	898	57,568	33,991	92,518	482
Osaka	24	564	21,064	6,901	21,855	220
Nagoya	16	1,379	14,640	13,110	11,029	159
Fukuoka	7	318	7,847	5,660	12,475	75
Sapporo	10	551	11,763	9,461	11,461	86
Total	80	3,710	112,882	69,123	149,338	1,022

applying the above criteria we are left with $N = 3710$ districts in total. The number of wards, districts, properties of each type, and stations in each city is displayed in Table 1.

Property prices and characteristics. We work with sales prices rather than with rental prices, because sales are more permanent than rentals and we would therefore expect that the effect of earthquake risk on choosing a property will be more informative.

Nakagawa *et al.* (2009) use land prices over various years (from 1980 onwards) and describe the data in their Section 3 (for the Tokyo area). Their data are based on the Koji-Chika data set published by the Ministry of Land, Infrastructure, Transport, and Tourism (MLIT). The Koji-Chika data set provides *fictional* sales prices (as produced by 'experts') and they are only available at annual intervals, which we consider to be too long for our purpose. We use a different data set, which provides self-reported *transaction* prices at three-months intervals. This data set, also provided by the MLIT, is known as the 'real estate transaction-price information'; see

`www.land.mlit.go.jp/webland_english/servlet/MainServlet`.

The information in this data set is based on the results of a questionnaire survey of persons involved in real estate transactions conducted by MLIT, compiled and published quarterly. We thus know the transaction price and the transaction date (quarter), and also in which district the property lies and the name of the nearest station. In addition, many property characteristics are provided, of which we shall only consider: total area in square meters, total floor area in square meters, distance to nearest station measured in walking minutes, age of the building (if applicable), building structure (reinforced concrete, steel, or wood), purpose of city planning in the urban control area, maximum building coverage ratio, and maximum floor area ratio. Different types may have different regressors. For example, the equation for land only does not have 'building structure' or 'building age' as a regressor; and the equation for condos does not use 'building structure' as a regressor.

Economic indicators. Property prices are affected by general economic conditions. In order to incorporate possible effects of these economic conditions, we have selected two national macroeconomic indicators: GDP and CPI.

Long-run earthquake risk. We consider two measures of earthquake risk: short-run risk (i.e., seismic excitation; see Section 2.1) and long-run risk. Long-run earthquake risk is defined as the probability of an earthquake exceeding certain intensity thresholds in the next 30 years in a given area, provided by the Japan Seismic Hazard Information Station (JSHIS). We select the threshold intensities '5-lower' (medium risk) and '6-lower' (high risk) in our analysis. The JSHIS probabilities are provided in various mesh sizes, varying from one square km to 250 square meters. For each district we identify its center and then define the risk of that district as the JSHIS risk associated with the smallest available mesh in which this center lies. Although the JSHIS exceedance probabilities are updated every one or two years, we take the average of the JSHIS risk data over all available years, thus obtaining a time-invariant measure of long-run risk for each district. These probabilities are included as objective measures of long-run earthquake risk in our regression design, at the district level. Choosing a district of relative safety may be viewed as a form of self-insurance. Therefore, provided this information, which is publicly available, is known among consumers, we would expect higher property prices in relatively safe areas all else being equal.

If the intensity is '5 lower' then, according to the Japan Meteorological Agency, many people will be frightened and feel the need to hold on to something stable. Hanging objects (such as lamps) will swing violently, books may fall from bookshelves, and unstable furniture may topple over. Windows may break and fall down, electricity poles may move, and roads may sustain damage. There may be cracks in the walls of wooden properties. If the intensity is '6 lower' then the effects will be more severe. It will be difficult to remain standing, unsecured furniture will move and topple over, and cracks in walls, crossbeams, and pillars will appear not only in wooden properties but also in properties built from reinforced concrete.

Summary statistics are shown in Table 2. It is clear from Table 2 that Tokyo, Nagoya, and Osaka are high-risk cities with respect to 'small' earthquakes. In fact, it is almost certain that an earthquake will occur in Tokyo

Table 2: Seismic hazard probabilities per city, averaged over districts and time (2005–2014)

City	mean	min	25%	50%	75%	max	sd
Exceeding intensity level '5 lower'							
Tokyo	1.00	0.99	0.99	1.00	1.00	1.00	0.00
Osaka	0.93	0.90	0.92	0.94	0.95	0.97	0.02
Nagoya	0.96	0.91	0.94	0.97	0.98	0.98	0.02
Fukuoka	0.39	0.06	0.30	0.42	0.48	0.56	0.12
Sapporo	0.33	0.05	0.21	0.33	0.44	0.51	0.12
Exceeding intensity level '6 lower'							
Tokyo	0.35	0.16	0.22	0.28	0.49	0.59	0.13
Osaka	0.37	0.22	0.30	0.39	0.44	0.52	0.09
Nagoya	0.56	0.21	0.41	0.61	0.67	0.77	0.14
Fukuoka	0.03	0.00	0.02	0.03	0.03	0.05	0.01
Sapporo	0.01	0.00	0.01	0.01	0.02	0.03	0.01

with an intensity more severe than '5 lower' within the next 30 years. Regarding the occurrence of 'severe' earthquakes ('6 lower'), Nagoya is more exposed than Tokyo and Osaka, and much more exposed than Fukuoka and Sapporo. The variation in probabilities of severe earthquakes in Tokyo, Osaka, and Nagoya is also much larger than in the other two cities. Fukuoka and Sapporo are not likely to have severe earthquakes, but there is still considerable probability (and variation) of smaller earthquakes. This suggests that it is important to use both thresholds, 5-lower and 6-lower, in characterizing the distribution of long-run earthquake risk for our purpose. In particular, this will guarantee sufficient variation of long-run probabilities in the hedonic price model discussed in Section 4.

Location information. The location information from our property data set (district and nearest station) needs to be linked to the location information of the risk data. We use the Google Maps API and a web scraping script to automatically search the coordinates of the center of a given district or station.

4 The Model

The dependent variable is log-property price, and we denote the h-th observation of type k in district i during quarter t as $y_{it}^{(h,k)}$. The most common method of modeling the property market is hedonic pricing, pioneered by Rosen (1974) who argued that an item's total price can be thought of as the sum of the price of each of its homogeneous characteristics, so that the effect of each characteristic on the price can be determined by regressing (log)price on these characteristics.

We shall follow the hedonic approach. In our case the (log)price is determined by characteristics of the property itself (size, age, etc.), the surrounding environment (location, crime rate, schools, etc.), earthquake risk factors, and macroeconomic influences.

The district i determines the city $c(i)$, which takes values $1, \dots, 5$ depending on the city in which district i is situated. Also, the time variable t determines in which quarter $q(t)$ the transaction took place, taking values $1, \dots, 4$ depending on whether t refers to the first, second, third, or fourth quarter. The number of observations varies per district, type and quarter, and this affects the precision. We let $H_{it}^{(k)}$ denote the number of observations on each type $k = 1, 2, 3$ in district i during quarter t.

We model the difference between cities by a shift $\alpha_{c(i)}$ in the intercept term, but we assume that all other parameters are the same between cities. The difference between cities is thus completely captured by the $\alpha_{c(i)}$.

Our model can now be written as

$$\begin{aligned} y_{it}^{(h,k)} &= \alpha_0^{(k)} + \alpha_{c(i)} + \gamma_{q(t)} + x_{i\cdot}^{(k)\prime}\beta_1 + x_{\cdot t}^{(k)\prime}\beta_2 + x_{it}^{(h,k)\prime}\beta_3 \\ &\quad + r_{it}(\psi)'\beta_4 + u_{it}^{(h,k)}, \end{aligned} \tag{3}$$

where $x_{i\cdot}$ denotes a variable that is constant over time, but varies over districts (attractiveness variables), $x_{\cdot t}$ denotes a variable that is constant over districts, but varies over time (economic indicators), x_{it} denotes a variable that varies over districts and over time (property characteristics), and r_{it} denotes the risk data (same for each type k) given by the (distorted) short- and long-run earthquake probabilities. The reference dummies are the city dummy for Tokyo and the quarter dummy for Q4; these are set to zero.

Although the model appears to be linear in the parameters this is not completely the case, because the risk variable r_{it} is a non-linear function of one or more ψ's which appear in the probability weighting function $w(p)$ discussed in Section 2. This complicates the estimation, and we shall discuss this issue in the next section.

In order to obtain a (balanced) panel we average over h, and obtain

$$\begin{aligned} \bar{y}_{it}^{(k)} &= \alpha_0^{(k)} + \alpha_{c(i)} + \gamma_{q(t)} + x_{i\cdot}^{(k)\prime}\beta_1 + x_{\cdot t}^{(k)\prime}\beta_2 + \bar{x}_{it}^{(k)\prime}\beta_3 \\ &\quad + r_{it}(\psi)'\beta_4 + \bar{u}_{it}^{(k)}, \end{aligned} \tag{4}$$

where we average over $H_{it}^{(k)}$ items, which thus depends on how many properties of type k there are in a given district.

Next we combine the three types of property into one 3×1 vector:

$$\bar{y}_{it} = \alpha_0 + (\alpha_{c(i)} + \gamma_{q(t)})\,\imath + X_{i\cdot}^{*}\beta_1 + X_{\cdot t}^{*}\beta_2 + X_{it}^{*}\beta_3 + \imath r_{it}(\psi)'\beta_4 + \bar{u}_{it}, \qquad (5)$$

where $\imath = (1,1,1)'$, which we write more succinctly as

$$\bar{y}_{it} = \bar{X}_{it}\beta + \bar{u}_{it} \qquad (i = 1, \ldots, N;\ t = 1, \ldots, T), \qquad (6)$$

where $\bar{y}_{it}$ is a $p \times 1$ vector of random observations, explained by (non-random) regressors $\bar{X}_{it} = \bar{X}_{it}(\psi)$, an unknown parameter vector β, and random errors $\bar{u}_{it}$ $(p \times 1)$. In our case $p = 3$.

5 Estimation Method

The errors are assumed to follow a p-variate three-error components structure,

$$\bar{u}_{it} = \zeta_i + \eta_t + \epsilon_{it}, \qquad (7)$$

a sum of three independent components each of which is iid with zero means and variances

$$\mathrm{var}(\zeta_i) = \Sigma_\zeta, \qquad \mathrm{var}(\eta_t) = \Sigma_\eta, \qquad \mathrm{var}(\epsilon_{it}) = \Sigma_\epsilon, \qquad (8)$$

where Σ_ζ and Σ_η are positive semidefinite, and Σ_ϵ is positive definite, all of order $p \times p$. Multivariate two-error components were first employed by Chamberlain and Griliches (1975) using maximum likelihood techniques. Multivariate three-error components were first considered by Avery (1977) who derived a feasible Aitken estimator, which is however not maximum likelihood and turns out to be asymptotically inefficient. Baltagi (1980) derived an alternative estimator, also not maximum likelihood, which is asymptotically efficient. Magnus (1982) discussed the estimation and testing of the multivariate two- and three-error components models in a maximum likelihood context.

Our error structure implies that

$$\mathrm{E}(\bar{u}_{it}\,\bar{u}_{js}') = \begin{cases} \Sigma_\zeta + \Sigma_\eta + \Sigma_\epsilon & \text{if } i = j \text{ and } t = s, \\ \Sigma_\zeta & \text{if } i = j \text{ and } t \neq s, \\ \Sigma_\eta & \text{if } i \neq j \text{ and } t = s, \\ 0 & \text{if } i \neq j \text{ and } t \neq s. \end{cases} \qquad (9)$$

Let

$$Y = \begin{pmatrix} \bar{y}_{11} & \bar{y}_{12} & \dots & \bar{y}_{1T} \\ \bar{y}_{21} & \bar{y}_{22} & \dots & \bar{y}_{2T} \\ \vdots & \vdots & & \vdots \\ \bar{y}_{N1} & \bar{y}_{N2} & \dots & \bar{y}_{NT} \end{pmatrix}, \qquad U = \begin{pmatrix} \bar{u}_{11} & \bar{u}_{12} & \dots & \bar{u}_{1T} \\ \bar{u}_{21} & \bar{u}_{22} & \dots & \bar{u}_{2T} \\ \vdots & \vdots & & \vdots \\ \bar{u}_{N1} & \bar{u}_{N2} & \dots & \bar{u}_{NT} \end{pmatrix}, \tag{10}$$

and

$$\bar{X}_{(t)} = \begin{pmatrix} \bar{X}_{1t} \\ \bar{X}_{2t} \\ \vdots \\ \bar{X}_{Nt} \end{pmatrix}, \qquad X = \begin{pmatrix} \bar{X}_{(1)} \\ \bar{X}_{(2)} \\ \vdots \\ \bar{X}_{(T)} \end{pmatrix}. \tag{11}$$

Then we can write (6) in stacked form as

$$y = X\beta + u, \tag{12}$$

where $y = \operatorname{vec} Y$ and $u = \operatorname{vec} U$. We shall assume that y is normally distributed with mean $\mu = X\beta$ and variance $\Omega(\theta)$, so that β refers to the mean parameters and θ to the variance parameters. One complication lies in the fact that the non-random matrix X depends on a parameter (vector) as well, so that $X = X(\psi)$.

Under normality, the loglikelihood takes the form

$$L(\beta, \psi, \theta) = \text{constant} - (1/2) \log |\Omega| - (1/2)(y - X\beta)'\Omega^{-1}(y - X\beta). \tag{13}$$

Maximizing L with respect to β and θ is assumed to be (relatively) easy, while maximization with respect to ψ is more difficult. Upon differentiating μ we obtain

$$d\mu = X d\beta + (dX)\beta = X d\beta + (\beta' \otimes I_n) Z d\psi, \tag{14}$$

where

$$Z = \partial \operatorname{vec} X / \partial \psi'. \tag{15}$$

Differentiating the loglikelihood then gives

$$\begin{aligned} dL = &-(1/2)\operatorname{tr}(\Omega^{-1} d\Omega) + (1/2)(y - X\beta)'\Omega^{-1}(d\Omega)\Omega^{-1}(y - X\beta) \\ &+ (y - X\beta)'\Omega^{-1} X d\beta + (y - X\beta)'\Omega^{-1}(dX)\beta, \end{aligned} \tag{16}$$

and hence

$$\begin{aligned} d^2L = {}&(1/2)\operatorname{tr}(\Omega^{-1} d\Omega)^2 - (y - X\beta)'\Omega^{-1}(d\Omega)\Omega^{-1}(d\Omega)\Omega^{-1}(y - X\beta) \\ &- (d\mu)'\Omega^{-1}(d\mu) - 2(y - X\beta)'\Omega^{-1}(d\Omega)\Omega^{-1}(d\mu) + (y - X\beta)'\Omega^{-1}(d^2\mu) \\ &- (1/2)\operatorname{tr}(\Omega^{-1} d^2\Omega) + (1/2)(y - X\beta)'\Omega^{-1}(d^2\Omega)\Omega^{-1}(y - X\beta). \end{aligned}$$

Minus the expectation of the second differential takes the simple form

$$-\,\mathrm{E}(d^2L) = (1/2)\,\mathrm{tr}(\Omega^{-1}d\Omega)^2 + (d\mu)'\Omega^{-1}(d\mu), \tag{17}$$

which implies that the information matrix will be block-diagonal in (β, ψ) and θ. This shows that we don't have to take the variance of the maximum likelihood (ML) estimator $\hat{\theta}$ into account when calculating the variance of the ML estimators $(\hat{\beta}, \hat{\psi})$. Thus, writing

$$(d\mu)'\Omega^{-1}(d\mu) = \begin{pmatrix} d\beta \\ d\psi \end{pmatrix}' \begin{pmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{pmatrix} \begin{pmatrix} d\beta \\ d\psi \end{pmatrix}, \tag{18}$$

where

$$V_{11} = X'\Omega^{-1}X, \qquad V_{12} = V_{21}' = X'\left(\beta' \otimes \Omega^{-1}\right)Z, \tag{19}$$

and

$$V_{22} = Z'\left(\beta\beta' \otimes \Omega^{-1}\right)Z, \tag{20}$$

we obtain estimators of the variances of $\hat{\beta}$ and $\hat{\psi}$ as

$$\widehat{\mathrm{var}}(\hat{\beta}) = V_{11}^{-1} + V_{11}^{-1}V_{12}\left(V_{22} - V_{21}V_{11}^{-1}V_{12}\right)^{-1}V_{21}V_{11}^{-1} \tag{21}$$

and

$$\widehat{\mathrm{var}}(\hat{\psi}) = \left(V_{22} - V_{21}V_{11}^{-1}V_{12}\right)^{-1}, \tag{22}$$

where the parameters in the V_{ij} matrices are replaced by their estimators.

It follows from (16) that the first-order conditions are

$$\begin{aligned} (y - X\beta)'\Omega^{-1}X\,d\beta &= 0, \\ (y - X\beta)'\Omega^{-1}(d\Omega)\Omega^{-1}(y - X\beta) &= \mathrm{tr}(\Omega^{-1}d\Omega), \\ (y - X\beta)'\Omega^{-1}(dX)\beta &= 0, \end{aligned} \tag{23}$$

for β, θ, and ψ, respectively. This implies that $\hat{\beta}$ takes the simple form

$$\hat{\beta}(\theta, \psi) = (X'\Omega^{-1}X)^{-1}X'\Omega^{-1}y, \tag{24}$$

so that we can concentrate the likelihood with respect to β. The concentrated loglikelihood is

$$L^* = L(\theta, \psi) = \text{constant} - (1/2)\log|\Omega| - (1/2)\hat{u}'\Omega^{-1}\hat{u}, \tag{25}$$

where

$$\hat{u} = y - X\hat{\beta} = y - X(X'\Omega^{-1}X)^{-1}X'\Omega^{-1}y.$$

For fixed ψ we have $d\psi = 0$ and

$$\begin{aligned} dL^* &= -(1/2)\,\mathrm{tr}(\Omega^{-1}d\Omega) + (1/2)\hat{u}'\Omega^{-1}(d\Omega)\Omega^{-1}\hat{u} \\ &\quad - \hat{u}'\Omega^{-1}X(X'\Omega^{-1}X)^{-1}X'\Omega^{-1}(d\Omega)\Omega^{-1}\hat{u}, \end{aligned} \tag{26}$$

using the fact that

$$\begin{aligned} d\hat{\beta} &= [d(X'\Omega^{-1}X)^{-1}]X'\Omega^{-1}y + (X'\Omega^{-1}X)^{-1}d(X'\Omega^{-1}y) \\ &= -(X'\Omega^{-1}X)^{-1}X'\Omega^{-1}(d\Omega)\Omega^{-1}\hat{u}. \end{aligned} \tag{27}$$

The variance matrix $\Omega = \mathrm{var}(u)$ is of a very large dimension, but the error components structure allows us to write it in a convenient form, allowing simple expressions for its inverse and determinant; see Proposition A.1 in the Appendix. We also need simple expressions for quadratic forms like $v'\Omega^{-1}v$ and $X'\Omega^{-1}X$. These are provided in Proposition A.2.

Estimation of the parameters then proceeds as follows. For given ψ we maximize the concentrated likelihood (25) with respect to the variance parameters θ, where using the explicit expression (26) for the gradient will speed up the optimization. Performing a grid search on ψ we obtain the ML estimates $\hat{\theta}$ and $\hat{\psi}$. Then we find $\hat{\beta}$ from (24). Finally, the estimated variances of $\hat{\beta}$ and $\hat{\psi}$ are obtained from (21) and (22).

6 Estimation Results

Our primary interest is in earthquake risk and its impact on property prices. More specifically, we wish to answer three questions:

(1) Do objective long-run earthquake probabilities have an effect on property prices?

(2) If so, do objective short-run earthquake probabilities have an effect on property prices, in addition to the effect of long-run probabilities?

(3) And do potentially distorted short-run earthquake probabilities have an effect on property prices, in addition to the effect of long-run probabilities?

Before we answer these questions and comment on our estimates in Table 3, we explain our econometric modelling strategy. This strategy is based on two ingredients. First, we aim for parsimony. We want the smallest model that captures the essence of our story. This means that sometimes regressors have been deleted from our model even when the associated parameters

Table 3: Results under various risk assumptions

	variable	LR only	LR and objective SR	Base model
	land & building	3.7592	4.5593	4.3812
intercepts	land only	3.5949	4.3940	4.2155
	condo	3.1025	3.9024	3.7244
	Osaka	−0.2273	−0.2625	−0.2615
city dummies	Nagoya	−0.3801	−0.4100	−0.4139
	Fukuoka	−0.8770	−0.9133	−0.9108
	Sapporo	−1.2050	−1.2458	−1.2388
	immigrants	6.7245	6.7224	6.7218
ward	crime	−0.0437	−0.0436	−0.0436
attractiveness	unemployment	−4.3360	−4.3395	−4.3399
	executives	3.3426	3.3447	3.3464
economic	log(GDP)	0.5606	0.5220	0.5229
indicators	log(CPI)	1.5347	1.4687	1.5030
	area (m^2)	0.0025	0.0025	0.0025
	floor area (m^2)	0.0006	0.0006	0.0006
	distance to nearest station	−0.0145	−0.0145	−0.0145
	age	−0.0121	−0.0121	−0.0121
	built 1981–2000	0.1674	0.1658	0.1652
property	built after 2000	0.4136	0.4126	0.4123
characteristics	structure: reinf. concrete	0.4348	0.4344	0.4343
	structure: steel	0.1867	0.1867	0.1867
	structure: wood	−0.1264	−0.1266	−0.1266
	urban control	−0.8972	−0.8967	−0.8967
	max building coverage ratio	−0.0019	−0.0019	−0.0019
	max floor area ratio	0.0004	0.0004	0.0004
	long run 45–55	−0.1433	−0.1427	−0.1427
risk	long run 55+	−0.5037	−0.5039	−0.5041
	short run	—	$-0.0915^{\ddagger}$	−0.0514
	$\hat{\psi}$	—	—	$3.74^{\dagger}$
$\Delta \log L$		−68.5	−15.8	—

are ‘significant’. Significance does not imply importance, and importance is what interests us. Second, we make a distinction between focus and auxiliary regressors. The focus regressors are the effects that we are interested in or are part of the minimum set that would make up a credible model, while the auxiliary regressors are only in the model because they improve the estimation of the focus parameters.

Since we have many observations, most estimates are likely to be significant at the usual 1.96 level. We provide more information about the results by strengthening the significance requirement on the t-values. Thus, a ‡ will indicate that $|t| \leq 1.96$, which we interpret as not significant, while † indicates significance with $1.96 < |t| \leq 4.00$. Estimates without superscript are therefore significant with $|t| > 4.0$. The choice of 4.0 is somewhat arbitrary and chosen *a posteriori* in order to provide more information about the precision of our estimates, in particular our parameter estimates pertaining to the risk variables. (All t-values test the null hypothesis that the parameter of interest equals zero, except the t-value of $\hat{\psi}$ which tests the null that $\psi = 1$.)

Now consider the first question. The results are presented in Table 3 under the heading 'LR only' and we see that all estimates are significant, that is, their t-value (in absolute terms) exceeds 4.0. Regarding the long-run risk effects, we remark that *long run 45–55* (medium risk) indicates the JSHIS probability that an earthquake occurs in the next thirty years of higher intensity than 5-lower and lower intensity than 6-lower; and that *long run 55+* (high risk) indicates the JSHIS probability that in the next thirty years an earthquake occurs of intensity 6-lower or higher. Both medium risk and high risk appear to have a significant negative impact on property prices. The higher risk level has a more severe impact, which is intuitively reasonable. Hence, long-run risk matters. This answers the first question.

Next, we consider the second question: given that long-run risk plays a role, do objective short-run probabilities also have an effect on property prices? The results are presented in the next column of Table 3 under the heading 'LR and objective SR'. Apparently they don't: the effect of the risk variable *short run*, while negative as we would expect, is not significantly different from zero.

Finally, we consider the third question: given that long-run risk plays a role, do potentially *distorted* short-run probabilities also have an effect on property prices? The results are displayed in the final column of Table 3 under the heading 'Base model'. Apparently they do: after distortion, short-run probabilities have a significant negative effect on property prices.

The difference between distorted and objective short-run risk is that short-run probabilities are now allowed to be distorted using a probability weighting function, in this case the one-parameter weighting function (2) proposed by Prelec (1998), which yields the highest likelihood. The parameter ψ in the Prelec function is estimated to be 3.74 and is significantly different from unity, since the absolute value of its t-value lies between 1.96 and 4.00; in fact $|t| = 2.91$.

As shown in Figure 3, the estimated probability weighting function has an S-shaped pattern where small probabilities are underweighted and large probabilities are overweighted, which is in contrast to the inverse S-shaped probability weighting function often found in experiments. This contrast may be explained by the fact that with a positive background intensity of earthquakes, temporary deviations of short-run earthquake probabilities caused by seismic excitation are not evaluated (and overweighted) with respect to a

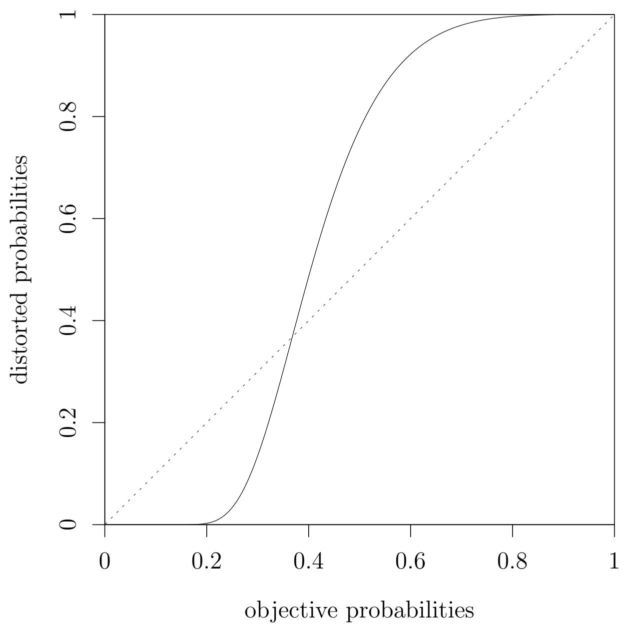

Figure 3: Estimated probability weighting of short-run probabilities, Prelec probability weighting function, $\hat{\psi} = 3.74$

reference probability of zero but with respect to a positive reference probability level. This applies in particular to Tokyo where the 90-day probability of an earthquake exceeding magnitude threshold of 5.5 never drops below 35% in the period that we analyze.

In summary: long-run risk matters, objective short-run risk does not, but distorted short-run risk does. In addition, all non-risk parameter estimates in the second and third columns are similar to the ones in the first column and all are significant (with a t-value larger than 4.0 in absolute value).

We briefly comment on these other (non-risk) parameters in the base model.

Intercept and city dummies. Tokyo, of course, is the most expensive city to buy property. If we set the property price level of Tokyo at 1.00, then the average property price levels of the other cities are 0.77 in Osaka, 0.66 in Nagoya, 0.40 in Fukuoka, and 0.29 in Sapporo. (Recall that we don't regress price but log-price on these dummies.) Also, if we set the price of land and building at 1.00, then the average price of the other types of property are 0.85 for land only and 0.52 for condos.

Quarterly effects. Estate agents sometimes tell customers that some months are better to buy or sell than others. Our results (in quarters, not months) are ambiguous, which is why we have omitted the quarter dummies from our regression. We return to this issue in our sensitivity analysis section.

Ward attractiveness. As discussed in Section 3, we selected eleven characteristics for each ward, divided into six categories. Only four of these eleven characteristics appear in our base model: percentage of immigrants (representing population); number of criminal offenses (representing safety); and unemployment ratio and percentage of executives (representing employment). Executives make a ward more attractive, while crime and unemployment make it less attractive. Immigrants too make a ward more attractive, which makes sense if we realize that the word 'immigrant' refers to somebody moving into the ward from another municipality, usually within Japan. Hence, the more people move in from other areas in Japan, the more attractive the ward apparently is.

Economic indicators. Property prices are affected by general economic conditions, and two indicators appear in Table 3 and in our base model: log(GDP) and log(CPI), both of which have a positive effect on property prices. The inclusion of log(CPI) has the additional advantage that if we wish to explain real (rather than nominal) property prices, then all results remain the same except that the effect of log(CPI) is now 0.503 rather than 1.503. Hence, CPI has a positive effect not only on nominal but also on real property prices.

Property characteristics. A large (floor) area and proximity to the nearest station contribute positively to the price. New buildings are preferred to old ones, where we have included two dummies because major changes occurred in the regulations on earthquake-resistance standards in 1981 and 2000. As a result, buyers prefer a house built between 1981 and 2000 over a house built before 1981, and they like a house built after 2000 even better. Regarding the structure, wood is not desirable, steel is desirable, but reinforced concrete is preferred. Urban control signifies restrictions on development possibilities, and this has a negative effect on prices.

For all three property types the designated maximum building coverage ratio (BCR) and the maximum floor-area ratio (FAR) are provided. These ratios are legally allowed maxima, different for each piece of land. The BCR is the percentage of the building area to the site area; the FAR is the percentage of the total floor area to the site area. We use both ratios in our regression and find a negative effect of BCR and a positive effect of FAR. Shimizu and Nishimura (2006) and Nakagawa *et al.* (2009) use only FAR and find mixed effects and a positive effect, respectively. Hidano *et al.* (2015) use both ratios (as we do) and find a negative effect of BCR and a mixed effect of FAR.

Error components. We estimated the coefficients using the multivariate three-error components structure, as described in Section 5. It turns out that

$$\mathrm{tr}(\Sigma_\epsilon) > \mathrm{tr}(\Sigma_\zeta) \gg \mathrm{tr}(\Sigma_\eta)$$

and as a result we set $\Sigma_\eta = 0$, so that we end up with a two-error components structure. The effect of this is negligible and will be discussed further in our sensitivity analysis in Section 8.

7 Importance Ordering and Premia for Earthquake Risk

Next we wish to determine an ordering of importance of the explanatory variables, in particular the importance of the risk variables, and to calculate the premia for earthquake risk embedded in property prices.

We write the prediction based on our original model (4) as

$$\hat{\bar{y}}_{it}^{(k)} = \hat{\alpha}_0^{(k)} + \hat{\alpha}_{c(i)} + \hat{\gamma}_{q(t)} + x_{i\cdot}^{(k)\prime}\hat{\beta}_1 + x_{\cdot t}^{(k)\prime}\hat{\beta}_2 + \bar{x}_{it}^{(k)\prime}\hat{\beta}_3 + r_{it}(\hat{\psi})'\hat{\beta}_4. \qquad (28)$$

In order to determine an ordering of importance of the explanatory variables, we note that the size of an estimated parameter gives no indication of the size of its influence, because this influence depends also on how the associated regressor is measured. We write (28) symbolically as

$$\log(\text{price}) = \text{intercept} + W_+ - |W_-| + M + P_+ - |P_-| - |R_{lr}| - |R_{sr}|, \qquad (29)$$

where the intercept comprises the (combined) constant term $\hat{\alpha}_0^{(k)} + \hat{\alpha}_{c(i)} + \hat{\gamma}_{q(t)}$ (positive); W_+ and W_- contain the two positive and two negative ward regressors in $x_{i\cdot}^{(k)\prime}\hat{\beta}_1$; M contains the two macro regressors in $x_{\cdot t}^{(k)\prime}\hat{\beta}_2$ (both positive); P_+ and P_- contain the seven positive and five negative property regressors in $\bar{x}_{it}^{(k)\prime}\hat{\beta}_3$, respectively; and R_{lr} and R_{sr} contain the long-run and short-run risk regressors in $r_{it}(\hat{\psi})'\hat{\beta}_4$ (all negative).

Table 4: Influences of each component for each type and city, real prices. Interquartile range between brackets.

	intercept	ward	macro	property	long-run risk	short-run risk
Type						
land & building	0.2894	0.0641	0.5394	0.0811	0.0155	0.0012
	(0.0175)	(0.0113)	(0.0335)	(0.0229)	(0.0062)	(0.0021)
land only	0.2928	0.0659	0.5626	0.0476	0.0158	0.0000
	(0.0206)	(0.0115)	(0.0371)	(0.0250)	(0.0067)	(0.0021)
condo	0.2597	0.0732	0.5747	0.0638	0.0164	0.0015
	(0.0162)	(0.0252)	(0.0265)	(0.0107)	(0.0060)	(0.0026)
City						
Tokyo	0.2628	0.0707	0.5741	0.0664	0.0160	0.0022
	(0.0300)	(0.0160)	(0.0302)	(0.0186)	(0.0062)	(0.0013)
Osaka	0.2760	0.0739	0.5482	0.0680	0.0182	0.0000
	(0.0303)	(0.0270)	(0.0318)	(0.0180)	(0.0038)	(0.0000)
Nagoya	0.2995	0.0616	0.5482	0.0693	0.0225	0.0000
	(0.0248)	(0.0124)	(0.0315)	(0.0281)	(0.0066)	(0.0000)
Fukuoka	0.3094	0.0650	0.5424	0.0666	0.0047	0.0000
	(0.0248)	(0.0244)	(0.0312)	(0.0277)	(0.0015)	(0.0000)
Sapporo	0.3320	0.0628	0.5299	0.0673	0.0025	0.0000
	(0.0200)	(0.0111)	(0.0343)	(0.0330)	(0.0024)	(0.0000)

Some categories (the ward characteristics W and the property characteristics P) contain both positive and negative influences. Simple addition would then be misleading since two opposite forces would hide possibly important influences. Hence we calculate the influences by first defining

$$A = \text{intercept} + W_+ + |W_-| + M + P_+ + |P_-| + |R_{lr}| + |R_{sr}|, \tag{30}$$

where all items are positive (by construction). Influences can then be decomposed into contributions from various categories by using A as the common denominator, that is, by computing intercept$/A$, $(W_+ + |W_-|)/A$, *et cetera*.

Table 4 presents the median of the relative influences for each component, by type and by city, using log *real* property prices as the dependent variable. Macroeconomic indicators are very important, contributing around 56%. The intercepts for type and city are also important, contributing around 28%. Location matters as well, as ward attractiveness takes up around 7% of the influence, while individual property characteristics add up to another 7%. This leaves around 2% for long-run and short-run risk. The influence of long-run risk is almost the same for all property types, but it differs substantially among different cities. Fukuoka and Sapporo, where earthquakes are relatively rare, are not much influenced by long-run risk, while Nagoya is the most influenced. Regarding short-run risk, only Tokyo is influenced and the importance of short-run risk in Tokyo is about one-tenth of the influence of long-run risk.

While the macro variables are by far the most relevant in explaining *median* house prices, they may be less relevant in explaining the *dispersion around* the median. To consider this aspect, Table 4 also displays the interquartile ranges (in brackets) of the relative influences. They reveal that the macro variables are still important, but all other variables (including the risk variables) are also quite important. More specifically, we see that individual property characteristics and intercepts for type and city are relevant in explaining dispersion in property prices (0.020 and 0.018 on average, respectively), still surpassed by macroeconomic variables (0.032), and quite closely followed by ward characteristics (0.011) and risk variables (0.009). Remarkably, the risk variables thus almost stand on equal footing with ward characteristics in explaining dispersion in property prices.

We can also compute these influences per quarter, in particular the quarter after the Tohoku earthquake (2011/Q2). The median influences of each component are essentially the same in that quarter with the exception of short-run risk in Tokyo, which is 0.22% overall but 0.35% in 2011/Q2. Large earthquakes have an important short-run effect in Tokyo. The influence of long-run risk remains the same.

We now investigate the influence of long-run and short-run risk in more detail, by decomposing the premia for earthquake risk. More precisely, we calculate and compare the predictions from four models. In model $\mathcal{M}_0$ there are no risk variables, either long-run or short-run; in model $\mathcal{M}_1$ we only have the two (objective) long-run risk variables; in model $\mathcal{M}_2$ we have long-run plus objective short-run risk variables; and in model $\mathcal{M}_3$ we have long-run plus distorted short-run risk variables (our base model).

Table 5: Decomposition of the premia for earthquake risk per type and city

type	city	median log-price	median premium $m_1 - m_0$	$m_2 - m_1$	$m_3 - m_2$
land & building	Tokyo	17.7275	−0.2620	−0.0246	−0.0092
	Osaka	17.2495	−0.2783	−0.0049	0.0049
	Nagoya	17.4264	−0.3393	−0.0076	0.0076
	Fukuoka	17.2812	−0.0630	−0.0043	0.0043
	Sapporo	17.0736	−0.0558	−0.0016	0.0016
land only	Tokyo	17.7073	−0.2409	−0.0241	−0.0087
	Osaka	17.2167	−0.2691	−0.0048	0.0048
	Nagoya	17.1113	−0.3293	−0.0077	0.0077
	Fukuoka	16.9066	−0.0658	−0.0046	0.0046
	Sapporo	16.3805	−0.0517	−0.0016	0.0016
condo	Tokyo	17.0344	−0.2621	−0.0246	−0.0093
	Osaka	16.5881	−0.2677	−0.0051	0.0051
	Nagoya	16.5236	−0.3175	−0.0079	0.0079
	Fukuoka	16.2134	−0.0740	−0.0042	0.0042
	Sapporo	16.2134	−0.0457	−0.0015	0.0015

Table 5 contains the results of this experiment. Let us denote the median of the log-price predictions in the four models by m_0, m_1, m_2, and m_3, respectively. Then the column $m_1 - m_0$ contains the premium of including (objective) long-run risk compared to not including any risk variable; the column $m_2 - m_1$ contains the premium of including objective short-run risk (in addition to long-run risk) compared to not including short-run risk; and the column $m_3 - m_2$ contains the premium of including distorted short-run risk (in addition to long-run risk) compared to including objective short-run risk.

We see that there is not much difference between different types of property and that the premium for long-run risk (compared to no risk) is much larger than the additional premium for short-run risk. Tokyo, Osaka, and Nagoya have a substantial premium for (objective) long-run risk of about 24–34%, while in Fukuoka and Sapporo this premium is 5–7%, thus much smaller. This is consistent with their different long-run risk profile. All long-run premia are negative, which means that long-run risk is compensated for through an adjustment in property prices in all cities.

Regarding short-run risk, there is a big difference between Tokyo and the other cities. In Tokyo, property prices are compensated for objective short-run risk with a median premium of about 2.5%, and there is an additional median compensation for distorted short-run risk of about 1%, because people tend to overweight large short-run earthquake probabilities in the Tokyo property market. In the quarter after the Tohoku earthquake these median premia rise to 3.0% and 1.7%, respectively.

Outside Tokyo we see that $(m_3 - m_2) \approx -(m_2 - m_1)$, which implies that the overall effect $(m_3 - m_1)$ is almost zero. This is caused by the shape of the estimated probability weighting function. The short-run probabilities outside Tokyo are relatively small, and after probability weighting they become even smaller (bottom part of the S-curve). People thus underweight small short-run probabilities; in fact they almost ignore them altogether. This effect (or lack of effect) can be decomposed into a 'compensation' $(m_2 - m_1 < 0)$ for objective short-run risk and a 'reward' $(m_3 - m_2 > 0)$ for underweighting short-run risk.

The power of econometrics is well-illustrated by the fact that, while property prices are the highest in Tokyo, the largest compensation (that is, reduction) for short-run risk (objective and distorted) and a sizeable compensation for long-run risk is in Tokyo.

8 Sensitivity Analysis

Our base model depends on assumptions regarding which variables to include and which not, how to measure or group certain variables, the choice of functional forms, and the stochastic specification. We wish to show that our results are robust, and we shall do so by deviating from our base model in various directions. (Of course, the selected base model was, in fact, itself the result of extensive sensitivity analyses.) In each case we are interested to find out whether our focus parameters are affected by these deviations. We are less interested to find out whether the deviations themselves are 'significant' or not, since these deviations typically represent auxiliary variables and are not the primary focus of our investigation.

Our focus variables are the risk variables and, in addition, four key characteristics of the property: area (m^2), floor area (m^2), distance to the nearest station, and age of the property. We have chosen the location (distance to nearest station) and the size (area and floor area) as our focus variables, and one characteristic of the property (age).

Ward attractiveness. Our base model contains four variables which measure the attractiveness of a ward. We extend this list by adding seven ward characteristics: the percentage of foreigners, and the number of hospitals, daycare centers, kindergartens, homes for the aged, department stores, and large retail stores.

Table 6: Sensitivity — ward attractiveness and economic indicators

	Base	+Attr.	−GDP
area (m^2)	0.0025	0.0025	0.0025
floor area (m^2)	0.0006	0.0006	0.0006
distance to nearest station	−0.0145	−0.0142	−0.0145
age	−0.0121	−0.0121	−0.0122
long run 45–55	−0.1427	−0.1961	−0.1411
long run 55+	−0.5041	−0.5706	−0.5024
short run	−0.0514	−0.0519	−0.0839
$\hat{\psi}$	$3.74^{\dagger}$	$3.75^{\dagger}$	$2.63^{\dagger}$
$\Delta \log L$	—	472.9	−407.8

If we compare the column '+Attr.' with the base model ('Base') in Table 6 we see that very little changes, thus showing the robustness with regard to these ward characteristics. These additional ward characteristics are therefore omitted in view of parsimony and the fact that, while they may be significant, they are not important.

Economic indicators. In the same Table 6 we also experiment with deleting log(GDP), so that the only economic indicator is log(CPI). This has some (although not a large) effect in particular on short-run risk, so that we keep GDP in the model as a general plausible indicator of economic activity.

Property characteristics. Next we experiment with the property characteristics. We consider three deviations from the base model, reported in Table 7.

Table 7: Sensitivity — property characteristics

	Urban control	Build. Struct.	Land use
area (m^2)	0.0025	0.0025	0.0025
floor area (m^2)	0.0006	0.0009	0.0006
distance to nearest station	−0.0147	−0.0159	−0.0146
age	−0.0121	−0.0119	−0.0121
long run 45–55	−0.1060	−0.1685	−0.1387
long run 55+	−0.4661	−0.5263	−0.4767
short run	−0.0516	−0.0508	−0.0515
$\hat{\psi}$	$3.72^{\dagger}$	$3.89^{\dagger}$	$3.76^{\dagger}$
Δ log L	−786.6	−5824.4	33.9

In the first column we remove the urban control variable; in the second column we remove the three building structure dummies; and in the third column we add, in addition to urban control, three further land-use variables ('residential', 'commercial', and 'industrial'), which describe the city's intentions of the usage of the land. Again, the estimated parameters appear to be robust to these changes; inclusion of urban control and, in particular, building structure dummies appears to substantially increase the loglikelihood, which makes sense because building a property costs more when steel is used instead of wood, and even more when reinforced concrete is used.

Cities. In our base model we have selected five Japanese cities. Although our selection is based on careful considerations (geographical spread and risk variation, in particular) as discussed in Section 3, this is still somewhat arbitrary. Suppose we only had four cities. How would this affect our estimates? This is shown in Table 8. In the first column we delete Tokyo, in the second column we delete Osaka, and in the third column we delete Nagoya. The effect on the non-risk parameters (area, distance, age) is small, but the effect on the risk parameters is not so small. Deleting Tokyo has quite a large effect on the risk parameters, because the short-run risk of Osaka, Nagoya, Fukuoka and Sapporo is relatively small compared to Tokyo, and estimation

Table 8: Sensitivity — four cities

	Tokyo	Osaka	Nagoya
area (m^2)	0.0023	0.0024	0.0025
floor area (m^2)	0.0006	0.0006	0.0006
distance to nearest station	−0.0152	−0.0145	−0.0147
age	−0.0126	−0.0127	−0.0115
long run 45–55	−0.2427	−0.1124	−0.1571
long run 55+	−0.4302	−0.4759	−0.6160
short run	−0.1873‡	−0.0627	−0.0525
$\hat{\psi}$	1.9‡	4.04†	4.11†

is less accurate when there is less variation in the risk variables. Deleting Osaka or Nagoya only affects the risk estimates marginally. Deleting Fukuoka or, in particular, Sapporo leads to unreliable results for the long-run risk parameters, probably caused by the fact that without these cities there is insufficient variation in the long-run risk variables leading to inaccurate estimation results. They are therefore omitted from the table. (Notice that we do not show the difference in loglikelihood in this table since the numbers of observations are different with different subsets of the sample.)

Time dimension. Our observations are per quarter and we could include quarter dummies to capture the idea that buying or selling in one quarter is more advantageous than in another.

Table 9: Sensitivity — quarters and Tohoku dummy

	Base	Q123	Q4	Tohoku
area (m^2)	0.0025	0.0025	0.0025	0.0025
floor area (m^2)	0.0006	0.0006	0.0006	0.0006
distance to nearest station	−0.0145	−0.0145	−0.0145	−0.0145
age	−0.0121	−0.0120	−0.0120	−0.0121
long run 45–55	−0.1427	−0.1415	−0.1406	−0.1426
long run 55+	−0.5041	−0.5033	−0.5025	−0.5040
short run	−0.0514	−0.0162†	−0.0208†	−0.0562
$\hat{\psi}$	3.74†	4.56‡	3.89‡	3.27†
$\Delta \log L$	—	1091.3	1007.8	6.3

Our base model does not include quarter dummies and in Table 9 we experiment with three possible extensions, namely adding three quarter dummies, adding one dummy for the fourth quarter (because there are relatively few earthquakes in the fourth quarter), and adding one dummy for the quarter following the Tohoku earthquake, respectively. In the cases Q123 and Q4 the likelihood increases substantially, but the key estimates don't change much, although the short-run risk parameters now become less significant. In the case of Tohoku even the likelihood does not increase much. Because the quarter dummies and the short-run risk are both time effects, which are likely to interact with each other, the results are ambiguous. This is why we prefer to exclude quarter dummies, thus making the interpretation easier and more transparent.

Stochastics. In our base model we have estimated two variance matrices:

$$\Sigma_\zeta = 0.129 \begin{pmatrix} 0.16 & 0.10 & -0.00 \\ 0.10 & 0.18 & -0.04 \\ -0.00 & -0.04 & 0.66 \end{pmatrix}, \quad \Sigma_\epsilon = 0.407 \begin{pmatrix} 0.31 & 0.01 & 0.00 \\ 0.01 & 0.33 & 0.00 \\ 0.00 & 0.00 & 0.36 \end{pmatrix},$$

while we set $\Sigma_\eta = 0$. This is because when we estimate the full three-error components model, we find

$$\Sigma_\zeta = 0.129 \begin{pmatrix} 0.16 & 0.11 & -0.00 \\ 0.11 & 0.18 & -0.04 \\ -0.00 & -0.04 & 0.66 \end{pmatrix}, \quad \Sigma_\epsilon = 0.406 \begin{pmatrix} 0.31 & 0.01 & 0.00 \\ 0.01 & 0.33 & 0.00 \\ 0.00 & 0.00 & 0.36 \end{pmatrix},$$

while

$$\Sigma_\eta = 0.002 \begin{pmatrix} 0.32 & 0.35 & 0.00 \\ 0.35 & 0.44 & -0.06 \\ 0.00 & -0.06 & 0.24 \end{pmatrix}.$$

The matrices Σ_ζ and Σ_ϵ are thus hardly affected and Σ_η is about one hundred times smaller than the other two.

Table 10: Sensitivity — stochasticity and station versus district

	Base	3-errors	station
area (m^2)	0.0025	0.0025	0.0026
floor area (m^2)	0.0006	0.0006	0.0006
distance to nearest station	−0.0145	−0.0146	−0.0137
age	−0.0121	−0.0121	−0.0115
long run 45–55	−0.1427	−0.1448	−0.1378†
long run 55+	−0.5041	−0.5067	−0.5742
short run	−0.0514	−0.0443	−0.0548
$\hat{\psi}$	3.74†	3.52†	3.56†
$\Delta \log L$	—	735.2	

In Table 10, column 2 we see that the key parameters are also hardly affected, although the likelihood (with six additional parameters) increases substantially. A formal test (not trivial in this case) may indicate that the hypothesis $\Sigma_\eta = 0$ is rejected in favor of $\Sigma_\eta > 0$, but we opt — in line with current ideas about the theory of applied econometrics (Angrist and Pischke, 2009; Magnus, 2017) — for parsimony and importance rather than for significance.

Station versus district. We know a lot about each property from the data, but not its exact location. We know in which district the property lies and we also know the name of the nearest station. In our setup we use districts as our location reference and there are 3,710 districts in our data set. But we could also use the nearest station as our location reference. There are 1,022 stations, so the district measure should be more precise. In fact, as Table 10, column 3 shows, the results are amazingly similar, demonstrating that the precise method of approximating the location is not so important.

Probability weighting functions: an extension. In our base model we use objective long-run probabilities and distorted short-run probabilities based on the Prelec probability weighting function. This raises various questions. First, one could argue that we should allow long-run probabilities to be distorted too; and second, we could experiment with different probability weighting functions.

Table 11: Sensitivity and extension — probability weighting functions

	Base	dist. SR TK	dist. LR Prelec	dist. LR TK
area (m^2)	0.0025	0.0025	0.0025	0.0025
floor area (m^2)	0.0006	0.0006	0.0006	0.0006
distance to nearest station	−0.0145	−0.0145	−0.0143	−0.0143
age	−0.0121	−0.0121	−0.0121	−0.0121
long run 45–55	−0.1427	−0.1427	−0.8644	−0.4856
long run 55+	−0.5041	−0.5039	−1.3838	−1.5028
short run	−0.0514	−0.0733‡	−0.0517	−0.0518
$\hat{\psi}$	3.74†	1.40‡	3.78†	3.77†
$\hat{\gamma}$	—	—	0.17	0.32
$\Delta \log L$	—	−14.6	152.8	167.6

In Table 11 we experiment with an alternative functional form for the short-run risk variable, namely the weighting function (1) introduced by Tversky and Kahneman (1992). In particular, in column 2 (dist. SR, TK) we replace the Prelec function applied to the short-run earthquake probabilities with the Tversky-Kahneman probability weighting function. The estimation results are similar to the base model, but somewhat less precise, and the loglikelihood decreases. The Tversky-Kahneman probability weighting function is found to be S-shaped, just like the Prelec function, confirming the robustness of this finding.

Next we also allow long-run risk to be distorted using both the Prelec and the Tversky-Kahneman weighting functions. The model contains two related time-invariant long-run probabilities and we quite naturally assume that these two probabilities share the same weighting function with the same parameter γ. (In particular, *distorted long run 45–55* is computed as *distorted long run 45+* minus *distorted long run 55+*, consistent with Choquet integration.) In columns 3 and 4 of Table 11 we allow both long-run risk and short-run risk to be distorted.

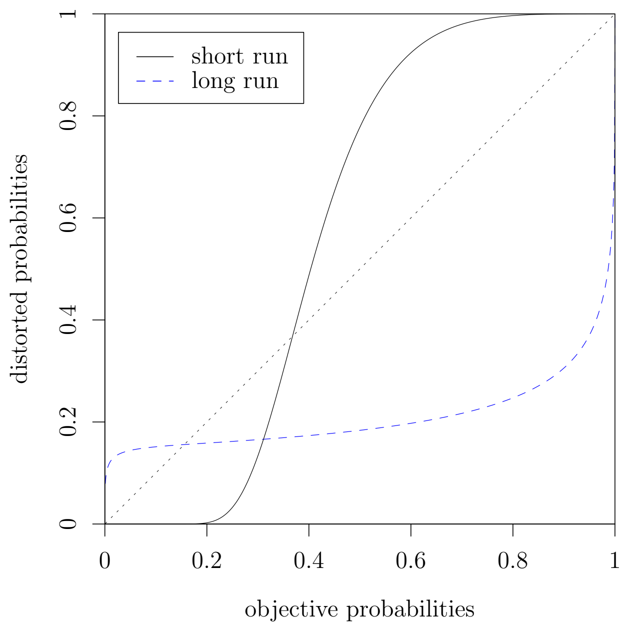

Figure 4: Implied probability weighting functions of long-run and short-run earthquake risk

The model with the higher likelihood is the one with an inverse S-shaped Tversky-Kahneman weighting function for long-run risk and an S-shaped Prelec weighting function for short-run risk, as shown in Figure 4. We note that the Prelec function for long-run risk, although yielding a lower loglikelihood than the Tversky-Kahneman weighting function, is also found to be inverse S-shaped, which is reassuring for the robustness of our results. Thus, in an extension of our base model that allows for distortion of time-invariant long-run earthquake probabilities we find evidence of a conventional inverse

S-shaped probability weighting function. This means that when purchasing property in Japan, people tend to overweight small long-run probabilities and underweight large long-run probabilities.

Summarizing, we have conducted extensive sensitivity analyses on our base model, always moving *one* step away from our base model. The base model proved to be remarkably robust in most directions. In some cases, however, one could argue that the base model should have been adjusted. The reason why we have not done so and prefer the current base model is twofold. First, we aim for parsimony; we prefer a simpler model over a more complex model. Second, if we were to change our base model, we would need to do (and we have done) the sensitivity analysis again for all cases, now based on the new base model. Then there will be other directions that prove to be sensitive. It is unlikely that there exists a model that is insensitive in every direction.

9 Conclusion

We have studied the impact of earthquake risk on Japanese property prices using a rich panel data set. We have not only allowed for time-invariant long-run earthquake probabilities to impact property prices, but we have also analyzed the impact of short-run earthquake probabilities generated from a seismic excitation model. We have designed a hedonic property prices model that accommodates probability weighting, employing a multivariate error components structure, and have developed the associated maximum likelihood and variance computation procedures.

We have shown that long-run earthquake probabilities negatively impact property prices and increasingly so at higher risk levels. We have also shown that short-run earthquake probabilities have a negative impact on property prices, and that this effect becomes statistically significant only after we allow for probability weighting.

The probability weighting function associated with short-run earthquake probabilities is found to be S-shaped. That stands in contrast to the familiar inverse S-shaped probability weighting functions predominantly found in experiments. The shape we find may be explained by the fact that in our setting there is a non-negligible positive background arrival rate of earthquakes. People may therefore tend to evaluate earthquake probabilities, and overweight their temporary deviations under seismic excitation, not with respect to zero but with respect to a positive reference probability level. This remarkable finding calls for the development of reference-dependent models for *probabilities* to augment the large literature on reference-dependent models for changes in *wealth* levels.

Appendix: Multivariate three-error components

Given the error components structure proposed in Section 5, we show that the $(NTp) \times (NTp)$ variance matrix of the error term u in (12) takes a particularly convenient form, allowing an easy way to calculate its inverse and determinant:

Proposition A.1 *Let $\imath_T$ and $\imath_N$ denote vectors containing only ones, of orders T and N, respectively, and let $J_T = \imath_T \imath_T'/T$ and $J_N = \imath_N \imath_N'/N$. Then,*

$$\Omega = \mathrm{var}(u) = V_1 \otimes \Delta_1 + V_2 \otimes \Delta_2 + V_3 \otimes \Delta_3 + V_4 \otimes \Delta_4,$$

where

$$V_1 = J_T \otimes J_N, \qquad V_2 = J_T \otimes (I_N - J_N),$$
$$V_3 = (I_T - J_T) \otimes J_N, \qquad V_4 = (I_T - J_T) \otimes (I_N - J_N),$$

and

$$\Delta_1 = \Sigma_\epsilon + T\Sigma_\zeta + N\Sigma_\eta, \qquad \Delta_2 = \Sigma_\epsilon + T\Sigma_\zeta,$$
$$\Delta_3 = \Sigma_\epsilon + N\Sigma_\eta, \qquad \Delta_4 = \Sigma_\epsilon.$$

In addition,

$$\Omega^{-1} = V_1 \otimes \Delta_1^{-1} + V_2 \otimes \Delta_2^{-1} + V_3 \otimes \Delta_3^{-1} + V_4 \otimes \Delta_4^{-1}$$

and

$$|\Omega| = |\Delta_1|\, |\Delta_2|^{N-1}\, |\Delta_3|^{T-1}\, |\Delta_4|^{(N-1)(T-1)}.$$

Proof: We write

$$\begin{aligned}\Omega = \mathrm{var}(u) &= \imath_T \imath_T' \otimes I_N \otimes \Sigma_\zeta + I_T \otimes \imath_N \imath_N' \otimes \Sigma_\eta + I_T \otimes I_N \otimes \Sigma_\epsilon \\ &= J_T \otimes I_N \otimes T\Sigma_\zeta + I_T \otimes J_N \otimes N\Sigma_\eta + I_T \otimes I_N \otimes \Sigma_\epsilon \\ &= V_1 \otimes \Delta_1 + V_2 \otimes \Delta_2 + V_3 \otimes \Delta_3 + V_4 \otimes \Delta_4.\end{aligned}$$

We note that the V_i are idempotent matrices, that $V_iV_j = 0$ $(i \neq j)$, and that $\sum_i V_i = I_{NT}$. The results now follow from Baltagi (1980), Magnus (1982, Lemma 2.1), and Abadir and Magnus (2005, Exercise 8.73). $\|$

In the special case where $\Sigma_\zeta = 0$ we have

$$\Delta_1 = \Delta_3 = \Sigma_\epsilon + N\Sigma_\eta, \qquad \Delta_2 = \Delta_4 = \Sigma_\epsilon, \tag{A.1}$$

and

$$\Omega = I_T \otimes J_N \otimes \Delta_1 + I_T \otimes (I_N - J_N) \otimes \Delta_2. \tag{A.2}$$

In the special case where $\Sigma_\eta = 0$ we have

$$\Delta_1 = \Delta_2 = \Sigma_\epsilon + T\Sigma_\zeta, \qquad \Delta_3 = \Delta_4 = \Sigma_\epsilon, \tag{A.3}$$

and

$$\Omega = J_T \otimes I_N \otimes \Delta_1 + (I_T - J_T) \otimes I_N \otimes \Delta_3. \tag{A.4}$$

Both are examples of a multivariate two-error components structure. Notice that we employ two idempotent matrices when there are two components, but that we need four (rather than three) when there are three components.

Given (25), we can obtain the ML estimates of the unknown parameters under normality by minimizing

$$L^* = \log|\Omega| + (y - X\beta)'\Omega^{-1}(y - X\beta). \tag{A.5}$$

Given the special structure of Ω this function also takes a convenient form:

Proposition A.2 *We have*

$$\begin{aligned}
L^* = {} & \log|\Delta_1| + (N-1)\log|\Delta_2| + (T-1)\log|\Delta_3| + (N-1)(T-1)\log|\Delta_4| \\
& + (1/N)(1/T)(\sum_{i,t} v_{it})'(\Delta_1^{-1} - \Delta_2^{-1} - \Delta_3^{-1} + \Delta_4^{-1})(\sum_{i,t} v_{it}) \\
& + (1/T)\sum_i(\sum_t v_{it})'(\Delta_2^{-1} - \Delta_4^{-1})(\sum_t v_{it}) \\
& + (1/N)\sum_t(\sum_i v_{it})'(\Delta_3^{-1} - \Delta_4^{-1})(\sum_i v_{it}) + \sum_{i,t} v_{it}'\Delta_4^{-1}v_{it},
\end{aligned}$$

where $v_{it} = \bar{y}_{it} - \bar{X}_{it}\beta$. *In addition, we have*

$$\begin{aligned}
X'\Omega^{-1}X = {} & (1/N)(1/T)(\sum_{i,t} X_{it})'(\Delta_1^{-1} - \Delta_2^{-1} - \Delta_3^{-1} + \Delta_4^{-1})(\sum_{i,t} X_{it}) \\
& + (1/T)\sum_i(\sum_t X_{it})'(\Delta_2^{-1} - \Delta_4^{-1})(\sum_t X_{it}) \\
& + (1/N)\sum_t(\sum_i X_{it})'(\Delta_3^{-1} - \Delta_4^{-1})(\sum_i X_{it}) + \sum_{i,t} X_{it}'\Delta_4^{-1}X_{it}.
\end{aligned}$$

Proof: Let $e_i^{(N)}$ denote the ith column of I_N and let $e_t^{(T)}$ denote the tth column of I_T. Then, writing

$$v = \sum_{i=1}^{N}\sum_{t=1}^{T} e_t^{(T)} \otimes e_i^{(N)} \otimes v_{it}, \qquad X = \sum_{i=1}^{N}\sum_{t=1}^{T} e_t^{(T)} \otimes e_i^{(N)} \otimes X_{it},$$

and

$$\begin{aligned}\Omega^{-1} &= J_T \otimes J_N \otimes \Delta_1^{-1} + J_T \otimes (I_N - J_N) \otimes \Delta_2^{-1} \\ &\quad + (I_T - J_T) \otimes J_N \otimes \Delta_3^{-1} + (I_T - J_T) \otimes (I_N - J_N) \otimes \Delta_4^{-1},\end{aligned}$$

we obtain

$$\begin{aligned}v'\Omega^{-1}v &= \sum_{i,j,s,t}(1/T)(1/N)v_{it}'\Delta_1^{-1}v_{js} + \sum_{i,j,s,t}(1/T)(\delta_{ij} - 1/N)v_{it}'\Delta_2^{-1}v_{js} \\ &\quad + \sum_{i,j,s,t}(\delta_{st} - 1/T)(1/N)v_{it}'\Delta_3^{-1}v_{js} \\ &\quad + \sum_{i,j,s,t}(\delta_{st} - 1/T)(\delta_{ij} - 1/N)v_{it}'\Delta_4^{-1}v_{js},\end{aligned}$$

where δ_{ij} and δ_{st} denote the Kronecker δ, that is, $\delta_{ij} = 1$ if $i = j$ and zero otherwise; and $\delta_{st} = 1$ if $s = t$ and zero otherwise. Hence,

$$\begin{aligned}v'\Omega^{-1}v &= (1/T)(1/N)\sum_{i,j}\sum_{t,s} v_{it}'\left(\Delta_1^{-1} - \Delta_2^{-1} - \Delta_3^{-1} + \Delta_4^{-1}\right)v_{js} \\ &\quad + (1/T)\sum_{i}\sum_{t,s} v_{it}'\left(\Delta_2^{-1} - \Delta_4^{-1}\right)v_{is} \\ &\quad + (1/N)\sum_{i,j}\sum_{t} v_{it}'\left(\Delta_3^{-1} - \Delta_4^{-1}\right)v_{jt} + \sum_{i}\sum_{t} v_{it}'\Delta_4^{-1}v_{it}.\end{aligned}$$

The result for $X'\Omega^{-1}X$ follows in a similar manner. ‖

References

Abadir, K. M. and Magnus, J. R. (2005). *Matrix Algebra*, Volume 1 of *Econometric Exercises* (Eds. K. M. Abadir, J. R. Magnus, and P. C. B. Phillips). Cambridge University Press: New York.

Abdellaoui, M. (2000). Parameter-free elicitation of utility and probability weighting functions. *Management Science*, 46, 1497–1512.

Aït-Sahalia, Y., Laeven, R. J. A., and Pelizzon, L. (2014). Mutual excitation in Eurozone sovereign CDS. *Journal of Econometrics*, 183, 151–167.

Aït-Sahalia, Y., Cacho-Diaz, J. A., and Laeven, R. J. A. (2015). Modeling financial contagion using mutually exciting jump processes. *Journal of Financial Economics*, 117, 585–606.

Angrist, J. D. and Pischke, J.-S. (2009). *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press: Princeton.

Avery, R. B. (1977). Error components and seemingly unrelated regressions. *Econometrica*, 45,199–209.

Baltagi, B. H. (1980). On seemingly unrelated regressions with error components. *Econometrica*, 48, 1547–1551.

Baltagi, B. H. (2008). *Econometric Analysis of Panel Data*. John Wiley: New York.

Bauer, D. and Kramer, F. (2016). The risk of a mortality catastrophe. *Journal of Business & Economic Statistics*, 34, 391–405.

Beron, K. J., Murdoch, J. C., Thayer, M. A., and Vijverberg, W. P. M. (1997). An analysis of the housing market before and after the 1989 Loma Prieta earthquake. *Land Economics*, 73, 101–113.

Bin, O. and Landry, C. E. (2013). Changes in implicit flood risk premiums: Empirical evidence from the housing market. *Journal of Environmental Economics and Management*, 65, 361–376.

Bin, O. and Polasky, S. (2004). Effects of flood hazards on property values: Evidence before and after Hurricane Floyd. *Land Economics*, 80, 490–500.

Boswijk, H. P., Laeven, R. J. A., and Lalu, A. (2016). Asset returns with self-exciting jumps: Option pricing and estimation with a continuum of moments. Mimeo, University of Amsterdam.

Brookshire, D. S., Thayer, M. A., Tschirhart, J., and Schulze, W. D. (1985). A test of the expected utility model: Evidence from earthquake risks. *Journal of Political Economy*, 93, 369–389.

Chamberlain, G. and Griliches, Z. (1975). Unobservables with a variance-components structure: Ability, schooling, and the economic success of brothers. *International Economic Review*, 16, 422–449.

Daniel, V. E., Florax, R. J. G. M., and Rietveld, P. (2009). Flooding risk and housing values: An economic assessment of environmental hazard. *Ecological Economics*, 69, 355–365.

Eeckhoudt, L. R., Laeven, R. J., and Schlesinger, H. (2017). Risk apportionment: The dual story. Technical report, arXiv. org.

Gu, T., Nakagawa, M., Saito, M., and Yamaga, H. (2011). On asymmetric effects of changes in regional risk rankings on relative land prices in the Tokyo Metropolitan area: A test of implications implied by prospect theory using market equilibrium prices (in Japanese). *Journal of Behavioral Economics and Finance*, 4, 1–19.

Hanaoka, C., Shigeoka, H., and Watanabe, Y. (2015). Do risk preferences change? Evidence from panel data before and after the Great East Japan earthquake. NBER Working Paper No. w21400.

Hidano, N., Hoshino, T., and Sugiura, A. (2015). The effect of seismic hazard risk information on property prices: Evidence from a spatial regression discontinuity design. *Regional Science and Urban Economics*, 53, 113–122.

Ikefuji, M., Laeven, R. J. A., Magnus, J. R., and Yue, Y. (2018). Earthquake risk embedded in property prices: Evidence from five Japanese cities — Online Appendix (available from `http://www.rogerlaeven.nl`).

Kahneman, D. and Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, 47, 263–292.

Kawawaki, Y. and Ota, M. (1996). The influence of the Great Hanshin-Awaji earthquake on the local housing market. *Review of Urban and Regional Development Studies*, 8, 220–233.

Magnus, J. R. (1982). Multivariate error components analysis of linear and nonlinear regression models by maximum likelihood. *Journal of Econometrics*, 19, 239–285.

Magnus, J. R. (2017). *Introduction to the Theory of Econometrics.* VU University Press: Amsterdam.

Nakagawa, M., Saito, M., and Yamaga, H. (2007). Earthquake risk and housing rents: Evidence from the Tokyo Metropolitan area. *Regional Science and Urban Economics*, 37, 87–99.

Nakagawa, M., Saito, M., and Yamaga, H. (2009). Earthquake risks and land prices: Evidence from the Tokyo Metropolitan area. *The Japanese Economic Review*, 60, 208–222.

Naoi, M., Seko, M., and Ishino, T. (2012). Earthquake risk in Japan: Consumers' risk mitigation responses after the Great East Japan earthquake. *Journal of Economic Issues*, 46, 519–529.

Naoi, M., Seko, M., and Sumita, K. (2009). Earthquake risk and housing prices in Japan: Evidence before and after massive earthquakes. *Regional Science and Urban Economics*, 39, 658–669.

Ogata, Y. (1981). On Lewis' simulation method for point processes. *IEEE Transactions on Information Theory*, 27, 23–31.

Ogata, Y. (1988). Statistical models for earthquake occurrences and residual analysis for point processes. *Journal of the American Statistical Association*, 83, 9–27.

Prelec, D. (1998). The probability weighting function. *Econometrica*, 66, 497–527.

Quiggin, J. (1982). A theory of anticipated utility. *Journal of Economic Behavior and Organization*, 3, 323–343.

Rosen, S. (1974). Hedonic prices and implicit markets: Product differentiation in pure competition. *Journal of Political Economy*, 82, 34–55.

Shimizu, C. and Nishimura, K. G. (2006). Biases in appraisal land price information: The case of Japan. *Journal of Property Investment and Finance*, 24, 150–175.

Sperna Weiland, R. C., Laeven, R. J. A., and de Jong, F. C. J. M. (2017). Feedback between credit and liquidity risk in the US corporate bond market. Mimeo, University of Amsterdam and Tilburg University.

Tversky, A. and Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, 5, 297–323.

Wu, G. and Gonzalez, R. (1996). Curvature of the probability weighting function. *Management Science*, 42, 1676–1690.

Yaari, M. E. (1987). The dual theory of choice under risk. *Econometrica*, 55, 95–115.

Yamaga, H., Nakagawa, H., and Saito, M. (2002). Earthquake risks and land pricing: The case of the Tokyo Metropolitan area (in Japanese). *Journal of Applied Regional Science*, 7, 51–62.