Appendix to

"Earthquake Risk Embedded in Property Prices: Evidence from Five Japanese Cities"

M. Ikefuji, R.J.A. Laeven, J.R. Magnus, Y. Yue

May 4, 2021

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)

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 statistics and economics of risk, our paper (hereafter ILMY) analyzes the subjective evaluation of both short-run and long-run catastrophic risk, earthquake risk in particular, embedded in Japanese property prices.

The length of the paper is constrained by the restrictions imposed by the journal, and in this appendix we provide supplementary material which could not be included in the paper itself. This supplementary material should be considered an integral part of the paper. All context, notation, and definitions are as in ILMY. The additional reference list only provides references to papers which are referred to in the appendix but not in ILMY.

The appendix is organized as follows. In Section A we provide various technical results relating to the maximum likelihood procedure and the multivariate three-error components model. In Section B we analyze the robustness of our estimation results. In Section C we try 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. In Section D we provide a brief literature review. In Section E we provide a map to illustrate the spatial windows used for estimation of the ETAS model.

A Maximum Likelihood

A.1 Optimization of the Loglikelihood under Normality

Consider the model $y = X\beta + u$, eq. (12) in ILMY, where $\mathrm{E}(u) = 0$, $\mathrm{var}(u) = \Omega(\theta)$, and $X = X(\psi)$. Under normality, the loglikelihood takes the form

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

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

$$d\mu = Xd\beta + (dX)\beta = Xd\beta + (\beta' \otimes I_n)Zd\psi, \tag{A.2}$$

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

Differentiating the loglikelihood then gives

$$\begin{aligned} dL = &-(1/2)\operatorname{tr}(\Omega^{-1}d\Omega) + (1/2)(y - X\beta)'\Omega^{-1}(d\Omega)\Omega^{-1}(y - X\beta) \\ &+ (y - X\beta)'\Omega^{-1}Xd\beta + (y - X\beta)'\Omega^{-1}(dX)\beta, \end{aligned} \tag{A.3}$$

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

$$-\,\mathrm{E}(d^2L) = (1/2)\,\mathrm{tr}(\Omega^{-1}d\Omega)^2 + (d\mu)'\Omega^{-1}(d\mu), \tag{A.4}$$

which implies that the information matrix will be block-diagonal in (β, ψ) and θ. This shows that we do not 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

$$(d\mu)'\Omega^{-1}(d\mu) = \begin{pmatrix} d\beta \\ d\psi \end{pmatrix}' \begin{pmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{pmatrix} \begin{pmatrix} d\beta \\ d\psi \end{pmatrix}, \tag{A.5}$$

where

$$V_{11} = X'\Omega^{-1}X, \qquad V_{12} = V_{21}' = X'\left(\beta' \otimes \Omega^{-1}\right)Z, \qquad V_{22} = Z'\left(\beta\beta' \otimes \Omega^{-1}\right)Z, \tag{A.6}$$

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

$$\widehat{\mathrm{var}}(\hat{\beta}) = V_{11}^{-1} + V_{11}^{-1}V_{12}\left(V_{22} - V_{21}V_{11}^{-1}V_{12}\right)^{-1}V_{21}V_{11}^{-1}, \tag{A.7}$$

$$\widehat{\mathrm{var}}(\hat{\psi}) = \left(V_{22} - V_{21}V_{11}^{-1}V_{12}\right)^{-1}, \tag{A.8}$$

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

It follows from (A.3) that the first-order conditions are

$$\begin{aligned} (y - X\beta)'\Omega^{-1}X\,d\beta &= 0, \\ (y - X\beta)'\Omega^{-1}(d\Omega)\Omega^{-1}(y - X\beta) &= \mathrm{tr}(\Omega^{-1}d\Omega), \\ (y - X\beta)'\Omega^{-1}(dX)\beta &= 0, \end{aligned} \tag{A.9}$$

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

$$\hat{\beta}(\theta, \psi) = (X'\Omega^{-1}X)^{-1}X'\Omega^{-1}y, \tag{A.10}$$

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

$$L^* = L(\theta, \psi) = \text{constant} - (1/2)\log|\Omega| - (1/2)\hat{u}'\Omega^{-1}\hat{u}, \tag{A.11}$$

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

$$\begin{aligned} dL^* &= -(1/2)\operatorname{tr}(\Omega^{-1}d\Omega) + (1/2)\hat{u}'\Omega^{-1}(d\Omega)\Omega^{-1}\hat{u} \\ &\quad - \hat{u}'\Omega^{-1}X(X'\Omega^{-1}X)^{-1}X'\Omega^{-1}(d\Omega)\Omega^{-1}\hat{u}, \end{aligned} \tag{A.12}$$

using the fact that

$$\begin{aligned} d\hat{\beta} &= [d(X'\Omega^{-1}X)^{-1}]X'\Omega^{-1}y + (X'\Omega^{-1}X)^{-1}d(X'\Omega^{-1}y) \\ &= -(X'\Omega^{-1}X)^{-1}X'\Omega^{-1}(d\Omega)\Omega^{-1}\hat{u}. \end{aligned} \tag{A.13}$$

A.2 Multivariate Three-Error Components

Given the error components structure proposed in Section 5 of ILMY, we show that the $(NTp) \times (NTp)$ variance matrix Ω of the error term u in eq. (12) of ILMY 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

$$\Delta_1 = \Delta_3 = \Sigma_\epsilon + N\Sigma_\eta, \qquad \Delta_2 = \Delta_4 = \Sigma_\epsilon, \tag{A.14}$$

and

$$\Omega = I_T \otimes J_N \otimes \Delta_1 + I_T \otimes (I_N - J_N) \otimes \Delta_2. \tag{A.15}$$

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

$$\Delta_1 = \Delta_2 = \Sigma_\epsilon + T\Sigma_\zeta, \qquad \Delta_3 = \Delta_4 = \Sigma_\epsilon, \tag{A.16}$$

and

$$\Omega = J_T \otimes I_N \otimes \Delta_1 + (I_T - J_T) \otimes I_N \otimes \Delta_3. \tag{A.17}$$

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 (A.11), we can obtain the ML estimates of the unknown parameters under normality by minimizing

$$L^* = \log|\Omega| + (y - X\beta)'\Omega^{-1}(y - X\beta). \tag{A.18}$$

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

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

we obtain

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

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,

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

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

B 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.

If we compare the column '+Attr.' with the base model ('Base') in Table B1 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 B1 we also experiment with deleting log(GDP), so

Table B1: 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†	3.75†	2.63†
$\Delta \log L$	—	472.9	−407.8

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 B2.

Table B2: 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†	3.89†	3.76†
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 of ILMY, this is still somewhat arbitrary. Suppose we only had four cities. How would this affect our estimates? This is shown in Table B3. In the first

Table B3: 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	$-0.1873^{\ddagger}$	−0.0627	−0.0525
$\hat{\psi}$	$1.9^{\ddagger}$	$4.04^{\dagger}$	$4.11^{\dagger}$

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 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 B4: 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	$-0.0162^{\dagger}$	$-0.0208^{\dagger}$	−0.0562
$\hat{\psi}$	$3.74^{\dagger}$	$4.56^{\ddagger}$	$3.89^{\ddagger}$	$3.27^{\dagger}$
$\Delta \log L$	—	1091.3	1007.8	6.3

Our base model does not include quarter dummies and in Table B4 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 B5: 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	−0.0134
age	−0.0121	−0.0121	−0.0116
long run 45–55	−0.1427	−0.1448	$-0.1386^{\dagger}$
long run 55+	−0.5041	−0.5067	−0.5344
short run	−0.0514	−0.0443	−0.0560
$\hat{\psi}$	$3.74^{\dagger}$	$3.52^{\dagger}$	$3.41^{\ddagger}$
$\Delta \log L$	—	735.2	

In Table B5, 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 B5, column 3 shows, the results are amazingly similar, demonstrating that the precise method of approximating the location is not so important.

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.

C Importance Ordering and Premia for Earthquake Risk

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.

Table C6: Influences of each component for each type and city, real prices. Interquartile range between brackets.

	intercept	ward(+)	ward(-)	macro	property(+)	property(-)	long-run risk	short-run risk
Type								
land & building	0.3254	0.0497	-0.0285	0.6533	0.0600	-0.0354	-0.0187	-0.0014
	(0.0320)	(0.0162)	(0.0104)	(0.0436)	(0.0273)	(0.0224)	(0.0085)	(0.0025)
land only	0.3152	0.0514	-0.0284	0.6637	0.0360	-0.0191	-0.0180	-0.0000
	(0.0394)	(0.0159)	(0.0095)	(0.0404)	(0.0296)	(0.0087)	(0.0085)	(0.0024)
condo	0.2936	0.0613	-0.0302	0.6911	0.0403	-0.0334	-0.0196	-0.0018
	(0.0217)	(0.0226)	(0.0108)	(0.0490)	(0.0216)	(0.0219)	(0.0076)	(0.0031)
City								
Tokyo	0.3119	0.0582	-0.0253	0.6598	0.0452	-0.0285	-0.0186	-0.0026
	(0.0316)	(0.0177)	(0.0079)	(0.0382)	(0.0270)	(0.0194)	(0.0076)	(0.0015)
Osaka	0.3095	0.0457	-0.0464	0.6899	0.0491	-0.0323	-0.0228	-0.0000
	(0.0354)	(0.0306)	(0.0102)	(0.0479)	(0.0302)	(0.0247)	(0.0050)	(0.0000)
Nagoya	0.3035	0.0498	-0.0269	0.6808	0.0518	-0.0313	-0.0280	-0.0000
	(0.0266)	(0.0152)	(0.0103)	(0.0437)	(0.0332)	(0.0208)	(0.0087)	(0.0000)
Fukuoka	0.2519	0.0535	-0.0324	0.7044	0.0487	-0.0363	-0.0061	-0.0000
	(0.0292)	(0.0243)	(0.0074)	(0.0590)	(0.0352)	(0.0240)	(0.0021)	(0.0000)
Sapporo	0.2442	0.0534	-0.0318	0.7145	0.0532	-0.0362	-0.0033	-0.0000
	(0.0308)	(0.0116)	(0.0040)	(0.0556)	(0.0375)	(0.0255)	(0.0032)	(0.0000)

We write the prediction based on our original model, labeled eq. (4) in ILMY, as

$$\hat{\hat{y}}_{it}^{(k)} = \hat{\alpha}_0^{(k)} + \hat{\alpha}_{c(i)} + \hat{\gamma}_{q(t)} + x_{i.}^{(k)\prime}\hat{\beta}_1 + x_{.t}^{(k)\prime}\hat{\beta}_2 + \bar{x}_{it}^{(k)\prime}\hat{\beta}_3 + r_{it}(\hat{\psi})'\hat{\beta}_4. \tag{C.19}$$

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 (C.19) symbolically as

$$\log(\text{price}) = \text{intercept} + W_+ - |W_-| + M + P_+ - |P_-| - |R_{lr}| - |R_{sr}|, \tag{C.20}$$

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 $x_{i\cdot}^{(k)\prime}\hat{\beta}_1$, respectively; 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 $\bar{x}_{it}^{(k)\prime}\hat{\beta}_3$; 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).

The ordering of importance is then characterized by what is arguably the most direct and natural way of characterizing the importance of the explanatory variables: based on their impact on the property prices. That is, we compute the influences as the impact, i.e., relative contribution, of the subsets of explanatory variables to the (log) property prices. Some subsets have a positive impact, some have a negative impact, and they sum to unity.

Table C6 presents the median of the 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 67%. The intercepts for type and city are also important, contributing around 31%. Location matters as well, as the two subsets of ward attractiveness regressors take up around 5% and −3% of the influence, while the two sets of individual property characteristics add up to another 5% and −3%. This leaves around −2% for long-run and (distorted) 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-seventh of the influence of long-run risk. The joint influence of long-run and distorted short-run earthquake risk, on average −2.0% of log property prices, translates in monetary terms into a marginal effect of around −7 million Japanese yen per property, slightly more than the average annual income of a middle-income Japanese household in the period 2006/Q2 to 2015/Q3 that we analyze (Source: e-Stat Portal Site of Official Statistics Japan).

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 C6 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 the two sets of individual property characteristics and the intercepts for type and city are relevant in explaining dispersion in property prices (2.6%, 1.8% and 3.1% on average, respectively), still surpassed by macroeconomic variables (4.4%), and quite closely followed by the two sets of ward characteristics (1.8% and 1.0%) and risk variables (1.1%). 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 −0.26% overall but −0.40% in 2011/Q2. Large earthquakes have an important short-run effect in Tokyo; see also below. 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, whether 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 C7: 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 C7 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 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 0.9%, 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.

D Brief Literature Review

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 for women the change in risk preference is inconclusive.

E Spatial Windows for the ETAS Model

In Section 2.1 of ILMY we estimate the ETAS model for each of the five cities that we consider, and we use the estimated intensities to generate 90-days probabilities of an earthquake exceeding a magnitude threshold of 5.5, for each city. We take the spatial windows somewhat larger than the city of interest, as seismic activity just outside the city may also impact the risk perception. Furthermore, this also helps to reduce the bias of the ETAS parameters stemming from seismicity originating outside of the extent of the spatial window.

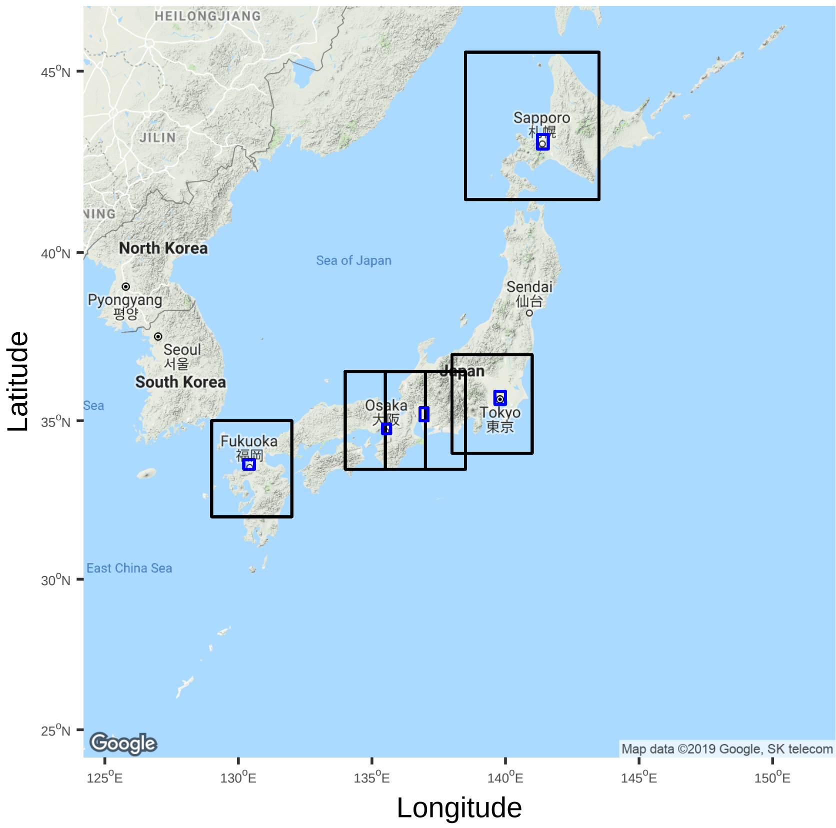

Figure E1: Spatial windows encapsulating each city

This is illustrated in Figure E1, which does not appear in the paper.

Additional 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.

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

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