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

This text serves as an appendix to the paper "Earthquake Risk Embedded in Property Prices: Evidence from Five Japanese Cities." For context, notation and definitions, see the paper. In the first section we provide some technical results for the multivariate three-error components model and in the second section we provide an analysis of the robustness of our estimation results.

A 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

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

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 (23), 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

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

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.

Table 1: 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 1 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 1 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 2.

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 'indus-

Table 2: 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

trial'), 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 3. In the first column we

Table 3: 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
$\hat{\psi}$	$1.9^{\ddagger}$	$4.04^{\dagger}$	$4.11^{\dagger}$

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 4: 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 4 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.

In Table 5, column 2 we see that the key parameters are also hardly affected, although

Table 5: 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	

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 5, 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.