On-line Supplement: Natural Disasters and Regional Industrial Production Efficiency: Evidence from Pre-war Japan

Appendix A. Methodology of Stochastic Frontier Analysis (SFA) (Section 5)

A-1. Methodology

In traditional non-frontier approaches to productivity analysis, all economic agents are assumed to be homogeneous units of production, and productivity growth takes place as a movement of the production frontier (Solow, 1957). A producer is technically efficient if he/she produces maximum output, for a given technology, from a given amount of inputs, and operates on the production efficiency frontier (Coelli et al. 2005). However, empirical studies demonstrate that in reality some production units are more efficient and operate on the technological frontier and are technically efficient, while others lag behind (Caves, 1989). In line with this, the SFA model is underpinned by the theoretical notion that production agents may behave sub-optimally and produce below the ideal "frontier" leading to technical inefficiency. This approach accounts for possible inefficient behavior by measuring inefficiency as the potential increase in the observed value of production against the maximum technically achievable value defined by the production frontier. Estimation of this frontier is based on the notion that a maximum achievable output exists, but is constrained by available inputs, and inefficiencies decrease production below the frontier. Technical inefficiency scores are thus calculated as the distance from current output to the frontier.

Statistically, the SFA is a parametric approach where the form of the production function is assumed to be known and allows other parameters of the production technology to be estimated. As such, it specifies a regression model characterized by a composite error term that can be decomposed into two parts. The first error component is assumed to follow a symmetric distribution and is a standard error term, while the second component captures technical inefficiency. Technical inefficiency scores are therefore free from distortion and statistical noise. The SFA also allows for the measurement of inefficiency and random shocks outside the control of the producer to affect output (Wadud, 2003; Coelli et al., 2005).

To apply the SFA model it is necessary to impose an a priori functional form and to specify distributional assumptions to separate the two components of the error term. We assume a general regional Cobb–Douglas production function as follows:

$$\log(OUTPUT_{it}) = \beta_0 + \beta_1 log(EMPLOYMENT_{it}) + \beta_2 \log(FACTORY_{it}) + \beta_3 \log(HORSEPOWER_{it}) + \mu_i + \lambda_t + V_{it} - IE_{it} \tag{1}$$

where *OUTPUT* is the value of output produced deflated to 1920 prices (yen), *EMPLOYMENT* is the number of persons employed in production, *FACTORY* is the number of factories, and *HORSEPOWER* is the power for production machines in factories measured by horsepower in prefecture *i* in year *t*; μ and λ are prefecture fixed effects and time dummies respectively, while *IE* is a nonnegative random variable accounting for technical inefficiency in the production function and *V* is the usual error term where both are independently distributed for all production units ($i=1, 2, \ldots, N$). Importantly, IE_{it} stands for time-varying technical inefficiency scores in prefecture *i* in year *t*. If IE_{it} is equal to zero, then prefecture *i* in year *t* is defined as being totally technically efficient and is at its maximum output level given the inputs used and technology available. If IE_{it} is greater than zero, then prefecture *i* in year *t* is defined as being technically inefficient. In essence, the *IE* measures the distance to the production possibilities frontier where a greater distance implies greater inefficiency, measured in terms of logged output. One may want to note that given our short time period the prefecture fixed effect is likely to take account of any location choices of industries and firms. Location choice can be assumed to be time invariant over our time period and accounted for by the prefecture fixed effect in the econometric estimation. In the 1920s, our sample period, interregional capital mobility was relatively low and productive firms were not very mobile.[1] To empirically show the low capital mobility of firms, Caruana-Galizia et al. (2021) measured the degree of geographical concentration on manufacturing within Japan from 1920 to 1960 and found that the Gini index of prefecture-level manufacturing output and employment had not changed in the 1920s. Arguably, most firms were single establishment without multiple plants and branch offices, except large stock companies and Zaibatsu firms. The location of firms highly adhered to the local transaction partners for a long time. Regarding disaster shocks, damaged firms did not choose relocation so much even after the big earthquake (Imaizumi, 2008)[2]. On the other hand, new heavy manufacturing industries arose in Tokyo, Osaka and major cities to access capital and financial markets. Thus, we can conclude that the location of firms and industries was likely to be invariant over time in the 1920s.[3]

A-2 Determinants of Efficiency

To estimate the impact of natural disasters through deaths and injuries on technical efficiency, we utilize the inefficiency scores obtained from the SFA model in Eq. (1) and run the following general regression equation:

$$-IE_{it} = \alpha_0 + \sum_{j=0}^{J} \alpha_{EQ,t-j} EQ_{i,t-j} + \sum_{j=0}^{J} \alpha_{CL,t-j} CLI_{i,t-j} + \rho_i + \tau_t + \varepsilon_{it} \qquad (2)$$

[1] Theoretically, Baldwin and Okubo (2006, 2014) constructed models where productive firms can choose profitable locations such as a big market with less risk of natural disasters. However, when relocation costs are large and/or transport costs are large, even productive firms are not able to choose their location.

[2] Imaizumi (2008) studied location patterns of machinery factories in Tokyo city before and after the Great Kanto Earthquake and found that damaged and survived firms in central Tokyo city tended to keep on their operation rather than the relocation to peripheral Tokyo area.

[3] In the 1930s, some heavy manufacturing firms initiated the construction of production plants in peripheral areas.

where EQ is the number of deaths per 1,000 people caused by earthquakes and CLI is the number of deaths per 1,000 people due to climate-based natural disasters (high tide, floods, and typhoons), and ρ and τ represent prefecture and year dummies. One should note that we use the negative value of IE as the dependent variable, so that the estimated coefficients can be more intuitively interpreted as impacts on technological efficiency, rather than inefficiency. As for specification (1) it is noteworthy that the prefecture fixed effect in (2) is likely to pick up any location choices due to prior knowledge of distributions of natural disasters across space, and we are hence only left with random realizations from such distributions.

Appendix B The Coverage of Manufacturing Census (Section 3-1)

Yuan et al. (2018) compared the data from the manufacturing census with macro-level manufacturing production, estimated by Shinohara (1972), and concluded that the manufacturing census data correspond well with the macro-level manufacturing production data. However, they pointed out that manufacturing production by small factories with less than 4 employees is not negligible, constituting around 10-15% of total manufacturing production, but is not covered by the census.

Appendix C Compound Event (Section 6.4)

Our base specification assumes that the impact of sequential events are independent of each other. Feasibly, however, having a damaging natural disaster in the previous period could affect its impact in the current period, commonly known as the impact of compound events (Zscheischler et al, 2020). To explore this possibility we created for each disaster type a dummy variable indicating whether there was a damaging occurrence in the previous period

(EQDUMMY and CLIDUMMY for earthquakes and climatic events, respectively), and interacted these with EQ and CLI in specification (2).

The results of this are shown in the final column for textiles and machinery in Tables 2 and 3 (in the main text), respectively. As can be seen, this produces some interesting findings. For machinery we find a negative contemporaneous impact of climatic disasters on efficiency only if there was also an event in the previous year. In terms of earthquakes allowing for a role of compound events produces dynamics on the efficiency effect where there was previously none. More specifically, while the contemporaneous boosting impact of an earthquake remains and is not affected if there was also an event the previous year, we now also find an (compound event) independent positive impact at t-2, i.e., two years after the disaster. Examining the interaction terms with EQDUMMY, however, shows that earthquakes reduce production efficiency if they were also preceded by an occurrence the prior year. This compounding effect is found consistently throughout the five lags employed in our analysis.

In terms of textiles including the interaction terms produces a negative independent effect of climatic disasters three years after the event. However, within the same time horizon a compound event enhances productivity, and thus explains the overall insignificant net impact at that time lag found in the non-interacted specification shown in the first column. For earthquakes the independent productivity enhancing effect continues up to at least five years after the event remains as found before. Examining the interaction terms with EQDUMMY, however, shows that if the earthquake was preceded by a damaging one in the previous year, this productivity boost is reduced. Comparing the coefficient to that of the independent impact suggests that this reduction is roughly between 10% and 60% of the independent boost.

On-line Appendix Tables and Figure

Table A1: Summary Statistics (Textiles)

Variable	Average	Std. Dev.	Min.	Max.
OUTPUT (mill. yen)	58.3	83.4	0.64023	59.5
EMPLOYMENT (person)	21022	26690	53	128361
Num FACTORY (number)	708.06	918.20	5	5564
HORSEPOWER (HP)	5790.70	16315.77	3	152446
EQ(Num death per 1000 by earthquake)	0.104	1.372	0	21.467
CLI(Num dead per 1000 by typhoon)	0.005	0.022	0	0.317
INEFFICIENCY (log of output)	17.096	0.842	13.743	18.989

Table A2: Summary Statistics (Machinery)

Variable	Average	Std. Dev.	Min.	Max.
OUTPUT (yen)	13.3	34.7	0.02227	212
EMPLOYMENT (person)	4810	11018	6	70194
Num FACTORY (number)	207.05	541.16	2	4705
HORSEPOWER (HP)	3291.1	12460.8	4	155066
EQ(Num death per 1000 by earthquake)	0.104	1.372	0	21.467
CLI(Num death per 1000 by typhoon)	0.005	0.022	0	0.317
INEFFICIENCY (log of output)	14.162	1.427	10.807	17.902

Table A3: Damage by the Great Kanto Earthquake

	Human damage		Physical damage	
Prefecture	**Number of dead and missing**	**Percentage to the population**	**Num of buildings completely burnt or destroyed**	**Percentage to the total buildings**
Total	104,619	0.89	464,909	20.4
Tokyo	70,497	1.75	328,646	39.8
Kanagawa	31,859	2.31	115,353	42.1
Chiba	1,420	0.11	42,945	24.5
Saitama	316	0.02	13,372	5.1
Yamanashi	20	0.00	562	0.5
Shizuoka	492	0.03	4,562	1.9
Ibaraki	15	0.00	157	0.1

Source: Tokyo City Government (1925), pp.160-163.

Figure A1: Map of Japan: Prefecture Names and Codes

1	Hokkaido	11	Saitama	21	Gifu	31	Tottori	41	Saga
2	Aomori	12	Chiba	22	Shizuoka	32	Shimane	42	Nagasaki
3	Iwate	13	Tokyo	23	Aichi	33	Okayama	43	Kumamoto
4	Miyagi	14	Kanagawa	24	Mie	34	Hiroshima	44	Oita
5	Akita	15	Niigata	25	Shiga	35	Yamaguchi	45	Miyazaki
6	Yamagata	16	Toyama	26	Kyoto	36	Tokushima	46	Kagoshima
7	Fukushima	17	Ishikawa	27	Osaka	37	Kagawa	47	Okinawa
8	Ibaraki	18	Fukui	28	Hyogo	38	Ehime		
9	Tochigi	19	Yamanashi	29	Nara	39	Kochi		
10	Gunma	20	Nagano	30	Wakayama	40	Fukuoka		

References

Baldwin, R., and Okubo, T. (2006). Heterogeneous firms, agglomeration and economic geography: spatial selection and sorting. *Journal of Economic Geography*, *6*(3), 323-346.

Baldwin, R., and Okubo, T. (2014). International trade, offshoring and heterogeneous firms. *Review of International Economics*, *22*(1), 59-72.

Caruana-Galizia, P., Okubo, T. and Wolf, N. (2021) "Underlying Sources of Growth First and Second Nature Geography", Ch.14 in The Cambridge Economic History of the Modern World, II, Broadberry, S. and Fukao, K. Cambridge University Press.

Caves, R.E., (1989). Mergers, takeovers, and economic efficiency: foresight vs. hindsight. *International Journal of Industrial Organization*, 7(1), 151-174.

Coelli, T.J., D.S.P. Rao, C.J. O'Donnell, and G.E. Battese. (2005). *An introduction to efficiency and productivity analysis*, 2nd ed. Springer.

Imaizumi, A. (2008) "Tokyo-hu Kikai Kanren Kogyo Shuseki niokeru Kanto Daishinsai no Eikyo" (The Impact of Great Kanto Earthquake on Clusters of Machinery Industries in Tokyo Prefecture), *Shakai Keizai-shigaku* 74-4 23-45.

Shinohara, M (1972) "Koukou gyou (Mining and Manufacturing)" Okawa, K., Shinohara, M., and Umemura M., Choki Keizai Toukei, Long-run Economic Statistics, Toyo Keizai Shinpo-sha.

Solow, R. (1957). "Technical change and the aggregate production function". *Review of Economics and Statistics*. 39 (3): 312–320

Tokyo City Government (1925) *Shinsai niyoru Nihon no Sonshitsu* (Damage to Japan by the Great Kanto Earthquake).

Wadud, M.A. (2003). Technical, allocative, and economic efficiency of farms in Bangladesh: A stochastic frontier and DEA approach. *Journal of Developing Areas* 37(1): 109–126.

Yuan, T, T. Settsu, J-P. Bassino and K. Fukao (2009) “Senzenki Nihon no Kennai Souseisan to Sangyo Kozo”, (Prefectural GDP and Industrial Structure in pre-war Japan) *Keizai Kenkyu* 60 (2)

Zscheischler, J., Van Den Hurk, B., Ward, P.J. and Westra, S., 2020. Multivariate extremes and compound events. In Climate extremes and their implications for impact and risk assessment, 59-76.