Reply to Referee 2 (Prof Yuji Aruka)

This paper is quite an interesting paper on the recent Japanese economy by the time series segmentation study on the 36 Nikkei Japanese industry indices. This paper was written before the natural and earthquake disaster occurred in March 2011. However, the result and the prediction will not be largely corrected because the disaster did not have any fatal damage to the Japanese industrial/financial structure, though it must be true that the economy was forced to be relatively digressed due to the disaster. This paper forecasts a slugged recovery until early 2012. This will not necessarily be inadmissible. It was true that the investors over the world either on the stock prices or the currency exchanges were relatively calm during the ongoing disaster. Furthermore, the reconstruction needs after destructions urgently are required in Japan. So these needs will a fortiori be conducive for the economic recovery. The future of the economy may be still much influenced by the external fluctuations outside Japan.

The study reported in this paper was completed in November 2010, using data up till June 2010. If possible, we would like to look at the time series just before and just after the March 2011 disaster, to see if what kind of psychological changes are reflected in the stock market. Even without looking into the data, we agree with Prof Aruka that the industrial and finance structures were not damaged by the earthquake and ensuing nuclear crisis, in part because of the tremendous optimism displayed by the Japanese people.

This paper smartly depicts some adjustment/coordination processes from the internationally external shocks. It will be stimulating in a sense that the analysis verifies the internal transitions among the empirically clustered industrial groups during the targeted recent 14 years according to the statistically derived segmentations of the whole period of this analysis. As a result of using the high frequency data, the interactive process of the concerned period of analysis can hierarchically be visible. The segmentation analysis on the time series employed in this analysis is rather classical, but the derived conclusions do intuitively not seem counterfactual. The conclusion derived here seems us considerably persuasive.

This paper used three independent parameters as the analytical benchmarks: (i) the starting time, (ii) the duration, and (iii) the Jensen-Shannon divergence value at the start of the corresponding segment. "The starting times allow us to roughly map out the progress of volatility shocks, whereas the duration and Jensen-Shannon divergence tell us how strongly the shock impact the different industries in the Japanese economy." The Jensen-Shannon divergence justifies the necessity of the segmentation analysis. In the Jensen-Shannon divergence calculated for the period of 1997 to 2008, it may be apparent for us that the segmented spectra is much better interpreted on the concerned periods than the single spectrum. See Figure 1. Thus the method employed in the paper could be expected a better understanding.

As pointed out by Prof Aruka, time series segmentation in economics dates back to the pioneering work by Goldfeld and Quandt in 1973, and thus can be considered a classical method. Since then, many economists and econometricians have extended the method to find more sensitive statistics for identifying change points or structural breaks. In this literature, Bai and Perron stand out in terms of their contributions. In these previous studies, the object is to identify a small number of change points within the low-frequency econometric time series data, whereas in our study, we attempt to identify a large number of change points within high-frequency financial time series data.

More importantly, our study focused on segmenting a cross section of time series, instead of a single time series. The econometrics community started looking into this problem in the last couple of years (see references in the first paragraph of Section 2.4), and they frequently assume that the change points occur at the same time point across the entire cross section of time series. This is reasonable for low-frequency econometric time series, if we assume that structural changes are abrupt. For high-frequency time series, this assumption is hard to justify, and so we do not segment all 36 Nikkei 500 industry index time series simultaneously (which can be done by using cross section likelihoods in our definition for the Jensen-Shannon divergence, instead of likelihoods for individual time series). By segmenting the time series independently, and thereafter comparing the cross section of time series segments, we discover the statistically significant comovements with no assumptions beyond what was necessary to do the segmentation. Because the change points associated a given comovement are different when we discover them this way, we find that we can also determine the time scale over which the comovement (recovery or crisis) occurs.

On the segmentation of the given time series adopted from the 36 Nikkei Japanese industry indices, the authors have employed the Gaussian parameters instead of the Levy distribution, because their interest are limited to a longer period which case is inclined to be favorable to the Gaussian process. They also justify it in view of the calculation cost. This choice may be not bad, indeed. The recursive segmentation is applied by a conservative method to impose the cutoff point Δ_0 = 10 when the new optimized segment boundaries are not found. They then achieved the segmentations for the 36 Nikkei 500 industry indices.

For interested readers, let us compare the computational costs of Gaussian time series segmentation and Levy time series segmentation. The typical computational time to completely segment a 14-year time series modeled as a collection of Gaussian processes is about 30 minutes. In contrast, we took about three weeks of computational time to segment a two-year time series with the same half-hourly data frequency if we assume it is generated by a collection of Levy processes. It is therefore infeasible to Levy time series segment the 14-year time series of the 36 Nikkei 500 industry indices. In any case, when we compare the Levy time series segments to the Gaussian time series segments over the same two-year period, the most significant change points coincide. This gives us confidence that the faster Gaussian time series segmentation will find the most significant change points, even if the financial time series is better described by a collection of Levy processes.

In order to summarize the whole segmentations into our conventional phases: (i) growth; (ii) contraction; (iii) correction; (iv) crash, they used the hierarchical agglomerative clustering of the segments to make out the dendrogram. Hence the dendrogram can skillfully establish the four phases of the standard macroeconomic phases. The fourth cluster is verified to be statistically more robust than the five or six cluster, while the third cluster is coarser. See Figure 4. Here it is noted that the dendrogram at the level 4 (4 clusters) contains the two different choices of 6 clusters. The same story thus underlies the two different configurations. By the authors, the different configurations reflect the tinted color differences of the same photograph. So the start dates among the industries may be different in the same phase of the economy. See Tables 5 to 6.

In response to these remarks, we would like to point out once again that the two different choices of 6 clusters are not equally robust statistically. Ideally, if we split a collection of *n* objects into *m* clusters, we would like these *m* clusters to be equally robust. That is, the *m* clusters are stable over the same range of thresholds. In practice, depending on how the objects are related to each other, this is rarely possible. There will almost always be a cluster which is stable over a larger range of thresholds, and another cluster that is stable over a smaller range of thresholds. This is so whether we employ a uniform global threshold, or *m* local thresholds. From Figure 4, we see that when we choose a uniform global threshold of $\Delta = 32$, we obtained one dark blue/blue cluster, one green cluster, one yellow cluster, one orange cluster, and two red clusters. Of these, the dark blue/blue cluster is stable when Δ is increased, but breaks into two clusters, dark blue, and blue, when Δ is decreased slightly. Similarly, the two red clusters are stable when Δ is decreased, but merges when Δ is increased slightly. In contrast, when we select local thresholds to select one dark blue cluster, one blue cluster, one green cluster, one yellow cluster, one orange cluster, and one red cluster, we can vary these local thresholds over fairly large ranges without changing the classification. Therefore, this latter classification into six clusters using local thresholds can be considered more robust than the earlier classification into six clusters using a global threshold.

The authors have fully inspected the configurations of temporal distributions over 36 industries during 1996 to 2010. Those inspections can be useful to confirm the five macroeconomic periods generally accepted by the public, which unfolded between 1996 and 2010: (i) the Asian Financial Crisis (1997 to 1999); (ii) the Technology Bubble Crisis (2000 to 2002); (iii) economic growth (2003 to 2006);(iv) Subprime Crisis (2007 to 2008); and (v) Lehman Brothers Crisis (2008 to2010). As for these periods, the authors calculated the cross correlations between the 36 Nikkei 500 Japanese industry indices. Furthermore, in order to detect the internal transitions of each period, they then compared them against cross correlations over the entire 14-year period by employing the MST diameters. See Table 7.

Following the econophysics tradition initiated by Mantegna and coworkers, the authors utilize the MST, i.e., the Minimum Spanning Trees, derived from the cross-correlation matrices. "The diameter of a MST is the largest number of links that one has to traverse to get from one node to another." This measurement gives an important index to monitor the transition between the different phases. In this study, according to the authors, the cluster may be regarded as more open (closer), if the degree of a MST is bigger (smaller). By this definition, it follows that the financial contagion (positive sentiments) spreads faster on a closer MST (on a more open MST). This criterion may give us one of the useful judgments for the progress of the economy.

In this study, the matrices, from which the MST's are derived, are given in three ways by employing the time series of the entire duration, but also two-year intervals, and every individual particular segments. The normalized cross-correlation matrix (the zero-lag cross-correlation between Japanese industries *i* and *j*) has a next component:

$$C_{ij} = \sqrt{\frac{\frac{1}{T-1}\sum_{t=1}^{T}\left(x_{it}-\bar{x}\right)\sum_{t=1}^{T}\left(x_{jt}-\bar{x}\right)}{\frac{1}{T-1}\sum_{t=1}^{T}\left(x_{it}-\bar{x}\right)^2 \frac{1}{T-1}\sum_{t=1}^{T}\left(x_{jt}-\bar{x}\right)^2}}$$

In order to span a MST, firstly, we compute the distance d_{ij} for any two financial series pair (*i*, *j*):

$0 \le d_{ij} = 2(1 - C_{ij}) \le 2$

Then we must look for the minimum of it, then the next minimum, and so on. "We repeat this process with pairs (i_k, j_k) with increasingly larger distances (d_{ik}, d_{jk}) . If no cycles are formed after drawing a link between i_k and j_k, the link is accepted. Otherwise, it is rejected. The whole process stops when all time series are incorporated into the spanning graph."

In summary, at first, the authors calculated the cross correlations between the 36 Nikkei 500 Japanese industry indices over these five macroeconomic periods over the entire 14-year period. Next, they then calculated the MST diameters. Thus the authors discovered from the MST formations on each time period that the Chemicals and Electric Machinery industries are consistently the hubs through all the five macroeconomic periods. This is a wonderful scientific finding, which may instruct the industrial policy for the macroeconomic reconstruction. It is to be noted that such an actually acute point could never be obtained, as long as we cling to a traditionally monolithic analysis. Faced with the time series segmentation study developed here, the traditional macroeconomics will lose its significance.

When we looked at the MSTs of the US market in our previous paper (Zhang et al., 2011), we saw clear transitions between a compact topology during growth and an open topology during crisis, as far back as 2001. Assuming that publicly listed companies know how to tap the positive energy in a growing market to expand their businesses, either by direct sales of reserve stocks, or by securing bank loans offering their stocks as collateral, or by using their stock market position to attract further private investments, we interpret the compact MST topology as statistical evidence for one sector of the US economy taking advantage of growth in another sector to accelerate its own growth. Conversely, during crisis times, we assume companies source out more suppliers to avoid over-reliance on any single supplier, and also actively explore alternative customer bases to avoid over-reliance on any single group of customers. Assuming that stock market investors are well informed, this supply chain management strategy naturally weakens preferential correlations between the stock prices of companies that used to leverage on each other to grow. We therefore interpret the open MST topology as statistical evidence for sectors of the US economy strategically reducing any preferential connections, even though market-level correlations are high during crisis.

When we try to apply this picture to understand our results in this paper, we find that the MST does not always open up during crisis, and does not close in during growth. In fact, the MST during the Asian Financial Crisis is very compact, although it did open up during the Technology Bubble Crisis. However, in the ensuing growth period, the MST did not close in. Finally, comparing the MSTs of the Subprime/Lehman Brothers Crises and the Asian Financial Crisis, we find that the recent crisis MSTs are more open than the crisis MST from more than 10 years ago. We consider this anecdotal evidence for the Japanese market learning the 'proper' macroeconomic response to a financial crisis only recently. It remains to be seen whether Japan has learned the 'proper' macroeconomic response to accelerate growth, because correlations between industries as visualized through the MST do not appear to be optimal during the 2004–2005 growth period.

More interestingly, in addition, the authors interestingly observed the phenomena on the flight to quality in the Japanese stock market. In the period of turmoil, some investors change their portfolio of the so-called growth stocks to exchange with the stocks of good performing, i.e., quality stocks. For instance, the authors divided the Subprime and

Lehman Brothers Crises into 8 corresponding segments. In view of the MSTs, they identified the hubs in each stage and also the transition between these stages. In particular, it is quite interesting to learn the transition from Subprime5 to Subprime6.

In Subprime5, one part of the MST became elongated, whereas the other part become very compact and centered around Machinery (NELI), which became the only hub in the MST (see Fig. 11(e)). By comparison, we see that the cluster of industries around Electric Machinery (NELI) very quickly dissipated and reformed around Chemicals (NCHE), Machinery (NMAC), and surprisingly, Foods (NFOD), which is another industry perceived to be a quality industry. That is to say, the authors observed the both directions to growth and quality industries from Subprime5 to Subprime6. Consequently, they judged that there were two fairly distinct groups of optimistic and pessimistic investors.

Thus the authors reached their interesting conclusion on the recent Japanese economy: "Going through all eight MSTs, we see that *the growth hubs are fairly robust and persistent*, and recovers rapidly after short disappearances. The quality hubs, on the other hand, do not survive for very long. We take this to mean that Japanese investors are actually more optimistic than most people give them credit for."

In a new challenge like M. Aoki and H. Yoshikawa (2007): Reconstructuring Macroeconomics, a macroeconomic state must correspond to a multiple states of microeconomic structures. This kind of microeconomic multiplicity may generate a potentially new evolution at the macro level by recombining the existing industries. This paper statistically uses this fact. It has just provided us with a fully empirical foundation for the new macroeconomic evolutionary theory. Hence the reviewer highly welcomes this study in order to reconstruct macroeconomics. Naturally, the Japanese policy makers and analysts are recommended to have an analytical thought in line with this study. Investors also are recommended to learn some analytical points shown in this paper. It must be true that the intelligence of investors can also contribute to a smooth evolution.

We are happy to learn of the ideas of Aoki and Yoshikawa from Prof Aruka, and note the convergence between their ideas and ideas developed elsewhere. The idea that a large number of microscopic variables can combine to give a small number macroscopic variables with radically different dynamics is the basis behind the prevailing understanding of complex systems and complexity. The economy is clearly a complex system with emergent properties at various levels, and we are of the opinion that economics can benefit from theories, tools and methods coming out of the growing interdisciplinary community working on complex systems.

Also, in the statistical physics and applied mathematics literature, there have also been works on understanding the effects of coarse graining (replacing a collection of variables by their aggregate average) and statistical fluctuations. We now understand that many seemingly unrelated microscopic models yield the same macroscopic model under coarse graining. These universal macroscopic models are called Ginzburg-Landau theories in statistical physics, and normal form equations in applied mathematics. In addition, the roles of statistical fluctuations in determining which macroscopic models we end up in starting from a given microscopic model are also partially understood. We are of the opinion that if we look at macroeconomics from this perspective, there is actually no need to give up on general equilibrium theories, because these corresponds to Ginzburg-Landau theories of the aggregates. Instead, we believe the way to go forward profitably is look more closely at statistical fluctuations, and how they drive economies from one fixed point to another. In our paper, we really are only scratching at the surface of an enormous problem of how statistical fluctuations and their changes shape the trajectories of economies.

Finally, the study must demonstrate to be a useful analytical standard for analyzing a new macroeconomic fluctuation. But a macroeconomic fluctuation may also depend on a time series of technological innovations. Sometimes the innovation may forcefully derive the economy. The reviewer suggests the authors to consider the cross effects between the financial time series and the technological time series.

We thank Prof Aruka for the suggestion to look also at the technological time series. We will need to talk to economists on how to acquire this form of data. If any reader would offer to point us in the correction direction, we would very much appreciate it.