A Quasi-Bayesian Analysis of Structural Breaks:
China's Output and Productivity Series

Xiao-Ming Li[*]

Department of Commerce, Massey University—Albany, New Zealand

Abstract

A quasi-Bayesian model selection approach is employed to detect the number and dates of structural changes in China's GDP and labour productivity data. It is shown that the predictive likelihood information criterion is valid only among models with well-behaved residuals.

Key words: structural change; predictive likelihood; GDP; labour productivity; China

JEL classification: C22; E30

1. Introduction

The literature on structural change in macroeconomic time series has been growing rapidly over the past decade. From a methodological point of view, such studies may be classified into two categories: one that adopts a hypothesis testing approach, and the other that takes a model selection approach (see Maddala and Kim, 1998, p. 406). Regarding hypothesis testing, there are two broad avenues to view. Tests of the unit root null hypothesis with or without structural change against the trend-stationarity alternative that allows for one or more breaks in the trend function fall into the first. Contributions in this area include Perron (1989), Inwood and Stengos (1991), Zivot and Andrews (1992), Raj and Slottje (1994), Perron (1997), Nunes et al. (1997), Lumsdaine and Papell (1997), Kanas (1998), Ben-David et al. (1999), and Newbold et al. (2001). The second line involves testing for alternative hypotheses against each other with respect to the number of breakpoints in the series, disregarding whether or not a unit root is present. Several testing techniques have recently been developed by Vogelsang (1997), Bai and Perron (1998), and Bai (1999) and applied in Ben-David and Papell (1998) and Ben-David and Papell (2000).

The problem of choosing between difference-stationary models and trend-stationary models that account for structural change and between models with

different numbers of breaks can also be tackled by a model selection approach. In fact, it should properly be so, as Maddala and Kim (1998) argue that ultimately determination of the number of breaks is best viewed as a model selection problem. However, the structural-change literature with model selection does not seem to be as vast as with hypothesis testing. Kim and Maddala (1991) propose a Bayesian approach which uses the BIC criterion for selecting the best model among the models with different numbers of breaks. A recent paper by Wang and Zivot (2000) also employs similar methods to select the most appropriate model from the data. While the above two studies may be categorised as following a Bayesian approach, Kashiwagi (1991) suggests what Maddala and Kim (1996a) term a "quasi-Bayesian" approach that uses the predictive likelihood in analysing the same problem of detecting structural changes. Bianchi (1995) extends this quasi-Bayesian approach to the case of autoregressive errors, and Bianchi and Kan (1996) apply the extended method in a study of the trend behaviour of the US GNP series. For differences between the Bayesian and quasi-Bayesian procedures, see Maddala and Kim (1996b).

More work needs to be done to compare the above two independently developed procedures for selecting the best model among alternatives with different numbers of breaks (Maddala and Kim, 1996b). This paper is part of the empirical work for that purpose. It reports the results of applying the quasi-Bayesian approach to detecting the structural breaks in China's GDP and labour productivity data, which may be used later for comparisons with the results of applying the Bayesian approach. Section 2 briefly describes the approach employed, Section 3 presents empirical results, and Section 4 concludes.

2. Predictive Likelihood

Consider a time series model with multiple breaks as follows:

$$y_t = a + bt + \sum_{i=1}^{n} a_i DU_{it} + \sum_{i=1}^{n} b_i DT_{it} + e_t . \tag{1}$$

In this model, y_t is a time series under investigation, t represents time trend, n is the number of breakpoints, e_t denotes the error term, and a, b, a_i, and b_i are model parameters to be estimated. The break dummy variables, DU_{it} and DT_{it}, take the following values:

$$DU_{it} = \begin{cases} 1 & \text{if } t \geq TB_i \\ 0 & \text{otherwise} \end{cases}$$

$$DT_{it} = \begin{cases} t - TB_i + 1 & \text{if } t \geq TB_i \\ 0 & \text{otherwise} \end{cases}$$

with the ranges of n break dates being:

$$1 < TB_1 < \ldots < TB_i < \ldots < TB_n < T.$$

Here T denotes the effective sample size and $TB_1, \ldots, TB_n$ are n break dates. We as-

sume that the errors e_t follow an AR(p) process with the order p to be determined. The breakpoints (including their number and timing) are to be searched for with the aid of the predictive likelihood information criterion as briefly described below.

If $p = 0$, the errors are serially independent and so the predictive likelihood criterion suggested by Kashiwagi (1991) is relevant, and its value, *PL*, is computed as:

$$PL = -\frac{T}{2} - \frac{T}{2}\ln(2\pi\sigma^2) - \frac{T(k+1)}{T-k-2}, \qquad (2)$$

where σ^2 is the variance of e_t and k is defined as $k \equiv 2(n+1) - m_1$. Here m_1 is the number of the periods that have only one observation.

If $p > 0$, the errors e_t become an autoregressive process of order p:

$$e_t = \mu_1 e_{t-1} + \mu_2 e_{t-2} + ... + \mu_p e_{t-p} + \varepsilon_t , \ e_t \sim N(0, \sigma_\varepsilon^2 \Omega_p), \ \varepsilon_t \sim N(0, \sigma_\varepsilon^2), \qquad (3)$$

where μ_1, …, μ_p are parameters and ε_t represents a normally, independently and identically distributed (NIID) error term. In this case, a heuristic criterion proposed by Bianchi and Kan (1996) may be employed. It is expressed as:

$$PL = -\frac{T}{2} - \frac{T}{2}\ln(2\pi\sigma_\varepsilon^2) - \ln\left(\left|\Omega_P\right|\right) - \frac{T(k+1)}{T-k-2}, \qquad (4)$$

with $k \equiv 2(n+1) - m_1 + p$. The idea behind adding p to the expression for k is to penalise for the number of autoregressive terms in the errors.

Equation (1) is estimated with the ordinary least square (OLS) method when p is assumed to be 0 and with the maximum likelihood method when p is assumed to be greater than 0. In the former case, use (2) to calculate *PL*. In the latter case, simply deducting from the estimated value of the log-likelihood function the bias correction term $T(k+1)/(T-k-2)$ will give *PL*.

3. Empirical Results

We apply the above-outlined methods to China's output (*GDP*) and labour productivity (*PRO*) annual data (in log form) obtained from the *China Statistical Yearbooks* for 1998 and 2000 and from *http://www.stats.gov.cn/tjgb/ndtjgb/qgndtjgb/ t20020331_15395.htm*: "The 2000 Statistical Communiqué of National Economy and Social Development, February 28, 2001" (in Chinese). The sample spans from 1952 to 2000, giving us a total of 49 observations.

Tables 1 to 4 display the models with the number of breaks associated with the highest *PL* values under four different assumptions about the error terms. They also report diagnostic statistics about the behaviour of the error terms. In particular, both the Ljung-Box *Q* and the Breusch-Godfrey (*B-G*) statistics are used to check for the presence of serial correlation. Since the former is valid only for a pure autoregressive process while the latter is also valid in the case where such variables as trends and dummies appear as regressors, we mainly rely on the latter for inference.

If only the predictive likelihood criterion is used, one would certainly pick out the models in Table 1 and conclude that GDP has 10 structural breaks and labour productivity has 9. However, although the *Q* statistics suggest that the assumption of independent errors cannot be rejected at the 5% level, the *B-G* statistics indicate a decisive rejection at a lower than 1% level for the two series. In addition, the productivity data suffer the problem of heteroscedasticity. Thus, the models that assume e_t to be an NIID error are not valid, and the two numbers of breakpoints suggested by the highest *PL* values may well be overestimated.

Table 1. The Selected Model When the Errors Are Assumed to Be AR(0)

Series	n	PL	Q(10)	B-G(10)	Normality	ARCH(3)
GDP	10	91.860	16.358 [0.090]	28.668 [0.001]	0.233 [0.890]	5.817 [0.121]
Productivity	9	99.283	18.248 [0.051]	26.655 [0.003]	1.396 [0.497]	8.631 [0.035]

Notes: For GDP, the 10 break dates are 1958, 1960, 1962, 1967, 1969, 1970, 1976, 1984, 1989, and 1993. For productivity the 9 break dates are 1961, 1963, 1966, 1969, 1971, 1976, 1984, 1990, and 1996. Numbers in parentheses are the lag lengths, and numbers in brackets are the p-values.

Table 2. The Selected Model When the Errors Are Assumed to Be AR(1)

Series	n	PL	Q(10)	B-G(10)	Normality	ARCH(3)
GDP	7	85.243	35.961 [0.000]	29.757 [0.001]	0.497 [0.780]	5.253 [0.154]
Productivity	11	99.199	23.762 [0.008]	64.470 [0.000]	0.321 [0.851]	2.496 [0.476]

Notes: For GDP, the 7 break dates are 1958, 1960, 1962, 1966, 1969, 1971, and 1976. For productivity the 11 break dates are 1954, 1960, 1962, 1966, 1969, 1971, 1976, 1981, 1984, 1990, and 1996. Numbers in parentheses are the lag lengths, and numbers in brackets are the p-values.

Table 3. The Selected Model When the Errors Are Assumed to Be AR(2)

Series	n	PL	Q(10)	B-G(10)	Normality	ARCH(3)
GDP	5	90.681	11.017 [0.356]	0.219 [0.640]	3.351 [0.187]	5.644 [0.130]
Productivity	4	91.963	13.106 [0.218]	0.021 [0.884]	0.632 [0.729]	0.935 [0.432]

Notes: For GDP, the 5 break dates are 1957, 1961, 1966, 1975, and 1977. For productivity the 4 break dates are 1956, 1961, 1976, and 1990. Numbers in parentheses are the lag lengths, and numbers in brackets are the p-values.

Table 4. The Selected Model When the Errors Are Assumed to Be AR(3)

Series	n	PL	Q(10)	B-G(10)	Normality	ARCH(3)
GDP	7	90.029	29.779 [0.001]	28.077 [0.002]	0.144 [0.931]	4.262 [0.234]
Productivity	4	84.310	19.738 [0.032]	18.346 [0.031]	0.148 [0.929]	3.810 [0.282]

Notes: For GDP, the 7 break dates are 1956, 1958, 1960, 1968, 1971, 1974, and 1976. For productivity the 4 break dates are 1957, 1961, 1976, and 1990. Numbers in parentheses are the lag lengths, and numbers in brackets are the p-values.

The models with AR(1) and AR(3) specifications for e_t do not qualify for a clean bill of health either. Tables 2 and 4 show that both the *Q* and *B-G* statistics signal serial correlation in ε_t, although normality and homoscedasticity are present. The only valid models seem to be those given in Table 3: incorporating the AR(2)

assumption for e_t now makes ε_t a NIID process, as suggested by all the reported diagnostic test statistics. Accordingly, 90.681 and 91.963 are the only valid highest *PL* values for the GDP and productivity data respectively, and the associated numbers of breaks are 5 for the former and 4 for the latter. Table 5 presents more details about the *PL* values and break dates for the models with AR(2) error terms, and the two columns with bold numbers represent the two best models among them for the two series respectively.

Table 5. The PL Values and Break Dates When the Residuals Are Assumed to Be AR(2)

GDP	$n = 0$	$n = 1$	$n = 2$	$n = 3$	$n = 4$	$n = 5$	$n = 6$
PL	54.234	67.767	80.099	88.880	89.892	**90.681**	90.372
*TB*1		1961	1961	1957	1957	**1957**	1957
*TB*2			1976	1961	1961	**1961**	1961
*TB*3				1976	1966	**1966**	1968
*TB*4					1976	**1975**	1970
*TB*5						**1977**	1975
*TB*6							1980
Productivity	$n = 0$	$n = 1$	$n = 2$	$n = 3$	$n = 4$	$n = 5$	$n = 6$
PL	55.154	67.447	77.411	86.156	**91.963**	91.290	90.564
*TB*1		1961	1961	1961	**1956**	1956	1956
*TB*2			1976	1976	**1961**	1961	1961
*TB*3				1990	**1976**	1973	1966
*TB*4					**1990**	1976	1969
*TB*5						1990	1976
*TB*6							1990

It is interesting to compare our results with those in Bianchi and Kan (1996). They show that the predictive likelihood approach is robust to different assumptions about the error terms e_t, as the number of structural breaks turns out to be the same ($n = 2$), and the parameter estimates for different error specifications are not very different. However, this conclusion does not apply in our study here, since we find that the number and the dates of breaks are sensitive to different error processes assumed. To further verify whether this is indeed the case, we conducted a small Monte Carlo study. More specifically, we generated a random variable y_t using equation (1) but assuming 2 breaks (i.e., DU_{1t}, DU_{2t}, DT_{1t} and DT_{2t}) with their dates and the values of all coefficients being given. The error term e_t was generated following an AR(1) process with the autoregressive coefficient predetermined. When the (correct) AR(1) specification is used for the error structure, the computed highest *PL* value equals 74.375 which corresponds to the two true breakpoints. However, if the (incorrect) AR(0) process was assumed for the error term, we obtain a higher *PL* value of 78.260 associated with 4 breaks instead of 2; that is, we end up with an over-inflated number of breakpoints. This suggests that the validity of the *PL* model selection approach is conditional on the correct specification of the error structure, at least in this application. Merely accepting the highest *PL* value among all models with different structural changes and varying error processes could therefore result in overestimating or underestimating the number of breakpoints and possibly in misidentifying their locations.

The best models we have selected, i.e., the models that have the highest *PL* values and well-behaved residuals ε_t, yield the following maximum likelihood estimates:

$$\ln GDP_t = \underset{(85.411)}{4.340} + \underset{(10.705)}{0.156t} - \underset{(-3.212)}{0.106\,DU_{57}} - \underset{(-7.629)}{0.232\,DU_{61}} + \underset{(2.256)}{0.061\,DU_{66}} +$$

$$\underset{(4.294)}{0.146\,DU_{75}} - \underset{(-2.776)}{0.067\,DU_{77}} - \underset{(-3.615)}{0.096\,DT_{57}} - \underset{(-0.807)}{0.018\,DT_{61}} +$$

$$\underset{(2.403)}{0.030\,DT_{66}} - \underset{(-5.019)}{0.117\,DT_{75}} + \underset{(6.555)}{0.140\,DT_{77}} + e_t \tag{5}$$

$$e_t = \underset{(17.053)}{1.034\,e_{t-1}} - \underset{(-14.932)}{0.905\,e_{t-2}} + \varepsilon_t$$

and

$$\ln PRO_t = \underset{(98.06)}{5.658} + \underset{(4.428)}{0.086t} - \underset{(3.919)}{0.132\,DU_{56}} - \underset{(-9.480)}{0.259\,DU_{61}} - \underset{(-4.653)}{0.107\,DU_{76}} +$$

$$\underset{(-5.685)}{0.145\,DU_{90}} - \underset{(-2.852)}{0.075\,DU_{56}} - \underset{(2.773)}{0.031\,DT_{61}} - \underset{(7.796)}{0.023\,DT_{76}} +$$

$$\underset{(5.125)}{0.021\,DT_{90}} + e_t \tag{6}$$

$$e_t = \underset{(10.966)}{0.908\,e_{t-1}} - \underset{(-9.834)}{0.815\,e_{t-2}} + \varepsilon_t.$$

Based on equations (5) and (6), Figures 1 and 2 depict the long-run growth paths of China's GDP and labour productivity. The growth rates of GDP during the six break periods can be calculated using the point estimates of the slopes given in (5). They are: 15.6% (1952–1956), 6.0% (1957–1960), 6.0% (1961–1965; note that the coefficient of DT_{61} in (5) is insignificant), 9.0% (1966–1974), −2.7% (1975–1976), and 11.3% (1977–2000). By the same calculation method and using the relevant information provided in equation (6), we obtain the growth rates of labour productivity for the five break periods as follows: 8.6% (1952–1955), 1.1% (1956–1960), 4.2% (1961–1975), 6.5% (1976–1989), and 8.6% (1990–2000).

Figure 1. The Long-Run Growth Path of Logarithmic Output
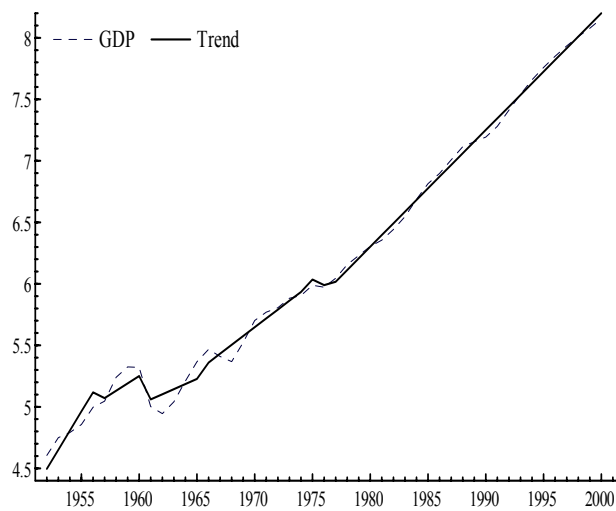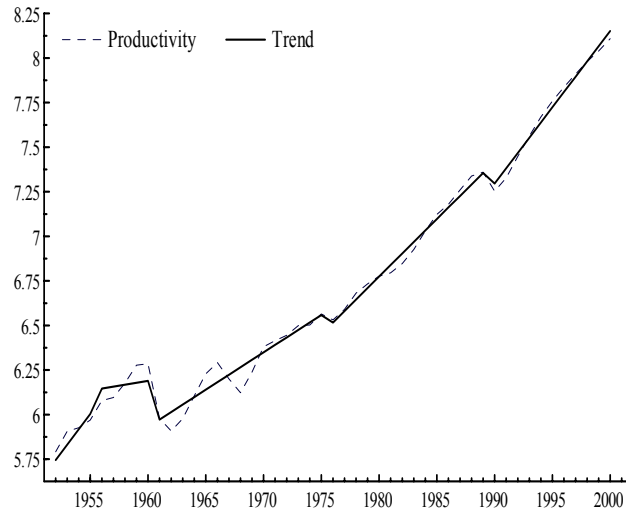
Figure 2. The Long-Run Growth Path of Logarithmic Labour Productivity

These changes in long-run growth rates can all be related to particular historical events, but there are three notably surprising results to point out here. First, the Cultural Revolution (1966–1976) did not have a negative impact on labour productivity growth but had a positive growth effect on GDP, contrary to what people normally believe. Second, it is since the end of the Cultural Revolution in 1976, and not since the inauguration of economic reforms in 1979, that both GDP and labour productivity have experienced growth takeoff. That is, economic reforms did not impact the long-run growth paths of these two series, again a seemingly counter-intuitive result. Third, 1990 is the first year of implementing a drastic program of macroeconomic austerity and is also the immediately subsequent year of the 1989 political crisis (the "June 4th event"). However, how these unfavourable shocks should cause a rise in the long-run growth rate of labour productivity is a mystery. Clearly, the currently adopted approach is unable to solve these puzzles.

4. Conclusions

In this paper, a quasi-Bayesian model selection approach developed by Kashwagi (1991) and extended by Bianchi and Kan (1996) is adopted to detect the number and timing of structural changes in China's GDP and labour productivity data. We have shown that one cannot mechanically or blindly use the highest value of the predictive likelihood information criterion to pick out the best model among models with and without well-behaved residuals: the criterion is valid only when the white-noise assumption about the error terms is satisfied. In our case, only within the class of models with AR(2) error terms can we compare the *PL* values and use validly the highest one to determine the number and dates of breakpoints.

The empirical results suggest 5 breakpoints for GDP in 1957, 1961, 1966, 1975, and 1977 and 4 breakpoints for labour productivity in 1956, 1961, 1976, and 1990. These breakpoints divide the sample into different growth periods. While most of the estimated growth results are expected, some changed/unchanged long-run growth rates do not seem to agree with common sense. Accordingly, further research to adopt other model selection approaches such as the Bayesian one is needed to confirm these new findings or check their validity.

References

Bai, J. and P. Perron, (1998), "Estimating and Testing Linear Models with Multiple Structural Changes," *Econometrica*, 66, 47-78.

Bai, J., (1999), "Likelihood Ratio Tests for Multiple Structural Changes," *Journal of Econometrics*, 91, 299-323.

Ben-David, D. and D. H. Papell, (1998), "Slowdowns and Meltdowns: Postwar Growth Evidence from 74 Countries," *Review of Economics and Statistics*, 80, 561-571.

Ben-David, D. and D. H. Papell, (2000), "Some Evidence on the Continuity of the Growth Process among the G7 Countries," *Economic Inquiry*, 38, 320-330.

Ben-David, D., R. L. Lumsdaine, and D. H. Papell, (1999), "Unit Roots, Post-war Slowdowns and Long-run Growth: Evidence from Two Structural Breaks," *NBER Working Papers*, No. 6397.

Bianchi, M., (1995), "Time Series Modelling in the Presence of Structural Change," Ph.D. Dissertation, Universite Catholique de Louvain.

Bianchi, M. and K. Kan, (1996), "Segmented Trend Modelling of the US GNP Series," *Applied Economics*, 28, 531-536.

Inwood, K. and T. Stengos, (1991), "Discontinuities in Canadian Economics Growth, 1879-1985," *Explorations in Economic History*, 28, 274-286.

Kanas, A., (1998), "Testing for a Unit Root in ERM Exchange Rates in the Presence of Structural Breaks: Evidence from the Bootstrap," *Applied Economics Letters*, 5, 407-410.

Kashwagi, N., (1991), "Bayesian Detection of Structural Changes," *Annals of the Institute of Statistical Mathematics*, 43, 77-93.

Kim, I. M. and G. S. Maddala, (1991), "Multiple Structural Breaks and Unit Roots in Exchange Rates," Paper presented at the Econometric Society Meeting at New Orleans, Dec. 1991.

Lumsdaine, R. L. and D. H. Papell, (1997), "Multiple Trend Breaks and the Unit-Root Hypothesis," *Review of Economics and Statistics*, 79, 212-218.

Maddala, G. S. and I. M. Kim (1996a), "Bayesian Detection of Structural Changes," in *Bayesian Analysis in Statistics and Econometrics*, D. A. Berry, K. M. Chaloner, J. K. Geweke, G. S. Maddala, and I. M. Kim eds., New York: Wiley, 359-370.

Maddala, G. S. and I. M. Kim, (1996b), "Structural Change and Unit Roots," *Journal of Statistical Planning and Inference*, 49, 73-103.

Maddala, G. S. and I. M. Kim, (1998), *Unit Roots, Cointegration, and Structural Change*, Cambridge, Cambridge University Press.

Newbold, P., S. Leybourne, and M. E. Wohar, (2001), "Trend-Stationarity, Difference-Stationarity, or Neither: Further Diagnostic Tests with an Application to U.S. Real GNP, 1875-1993," *Journal of Economics and Business*, 53, 85-102.

Nunes, L., C. P. Newbold, and C. M. Kuan, (1997), "Testing for Unit Roots with Breaks: Evidence on the Great Crash and the Unit Root Hypothesis Reconsidered," *Oxford Bulletin of Economics and Statistics*, 59, 435-448.

Perron, P., (1989), "The Great Crash, the Oil Price Shock, and the Unit Root Hypothesis," *Econometrica*, 57, 1361-1401.

Perron, P., (1997), "Further Evidence on Breaking Trend Functions in Macroeconomic Variables," *Journal of Econometrics*, 80, 355-385.

Raj, B. and D. J. Slottje, (1994), "The Trend Behaviour of Alternative Income Inequality Measures in the United States from 1947-1990 and the Structural Break," *Journal of Business & Economic Statistics*, 12, 479-487.

Vogelsang, T. J., (1997), "Wald-Type Tests for Detecting Breaks in the Trend Function of a Dynamic Time Series," *Econometric Theory*, 13, 818-849.

Wang, J. and E. Zivot, (2000), "A Bayesian Time Series Model of Multiple Structural Changes in Level, Trend, and Variance," *Journal of Business and Economic Statistics*, 18, 374-386.

Zivot, E. and D. W. K. Andrews, (1992), "Further Evidence on the Great Crash, the Oil-Price Shock, and the Unit-Root Hypothesis," *Journal of Business and Economic Statistics*, 10, 251-270.