BOFIT Discussion Papers
1 • 2011

Aaron Mehrotra and Jenni Pääkkönen

Comparing China's GDP statistics with coincident indicators

SUOMEN PANKKI • FINLANDS BANK •
EUROJÄRJESTELMÄ
EUROSYSTEMET

Bank of Finland, BOFIT
Institute for Economies in Transition

BOFIT Discussion Papers
Editor-in-Chief Aaron Mehrotra

BOFIT Discussion Papers 1/2011
18.2.2011

Aaron Mehrotra and Jenni Pääkkönen: Comparing China's GDP statistics with coincident indicators

ISBN 978-952- 462-696-5
ISSN 1456-5889
(online)

Suomen Pankki
Helsinki 2011

Contents

Tiivistelmä ....

Abstract ....

1 Introduction 1

2 Methodology 3

3 Empirical evidence 6

4 Conclusions 16

References 23

Aaron Mehrotra and Jenni Pääkkönen

Comparing China's GDP statistics with coincident indicators

Tiivistelmä

Tässä tutkimuksessa käytetään faktorianalyysiä useiden makrotalouden indikaattoreiden sisältämän informaation tiivistämiseen siten, että saadut faktorit voidaan tulkita Kiinan talouden tilaa kuvaavina suhdanneindikaattoreina. Tutkimuksessa verrataan estimoitujen faktoreiden ja BKT:n dynamiikkaa sekä mallin tuottamia faktoreita muihin, jo olemassa oleviin Kiinan talouden suhdanneindikaattoreihin. Tuotetun indikaattorin ja Kiinan BKT-sarjan liikkeet vastaavat varsin hyvin toisiaan, ja sarjojen erot ovat vähäisiä. Silloin kun näiden sarjojen välillä esiintyy eroja, ne näyttävät johtuvan Kiinan talouskasvuun kohdistuvista šokeista, sillä sen paremmin BKT:n autoregressiivinen prosessi kuin suhdanneindikaattoritkaan eivät onnistu ennustamaan talouden kasvua näillä periodeilla.

Avainsanat: faktorimallit, pääkomponentti, BKT, Kiina

Comparing China's GDP Statistics with Coincident Indicators*

Aaron Mehrotra† Jenni Pääkkönen‡

February 7, 2011

Abstract

We use factor analysis to summarize information from various macroeconomic indicators, effectively producing coincident indicators for the Chinese economy. We compare the dynamics of the estimated factors with GDP, and compare our factors with other published indicators for the Chinese economy. The indicator data match the GDP dynamics well and discrepancies are very short. The periods of discrepancies seem to correspond to shocks affecting the growth process as neither autoregressive models for GDP itself nor various coincident indicators are able to forecast them satisfactorily.

Keywords: Factor models, principal component, GDP, China.

JEL Classification: C38, O4, P2.

---

*We thank Carsten Holz, Ali Kutan, Henri Nyberg, participants in the 2011 AEA-ASSA Meetings in Denver, ACES session on *Perspectives on China's Growth*, the 2011 Finnish Economic Association Annual Meeting in Oulu, and a Bank of Finland seminar, for their helpful comments and suggestions. All views expressed in the paper are personal and do not necessarily reflect those of the Bank of Finland or the Government Institute for Economic Research. Any errors are our own.

†Bank of Finland, Institute for Economies in Transition (BOFIT), PO Box 160, FI-00101 Helsinki, FINLAND. e-mail aaron.mehrotra@bof.fi.

‡Government Institute for Economic Research, PO Box 1279, FI-00101 Helsinki, FINLAND. e-mail jenni.paakkonen@vatt.fi.

1 Introduction

China's growing role in the world economy has prompted international observers and researchers to develop leading and coincident indicators for analyzing the current dynamics and prospects of the Asian powerhouse. The OECD, for example, has published a leading indicator for China since 2006. In May 2010, the US Conference Board released both a coincident and a leading indicator for China, providing backdata until 1986. Indeed, assessment of China's growth has become so critical that the simple announcement by the Conference Board in July 2010 that it was revising its April 2010 calculation for its leading indicator was enough to put international financial markets on edge.

At the same time, research has been conducted pointing to plausible limitations with Chinese data. This arises despite the fact that many of the concerns expressed are relevant for many emerging economies, and the National Bureau of Statistics (NBS) in China is aware of shortcomings with the data. Zheng (2001) notes that the quarterly national accounts rely heavily on estimates and excessive aggregation, and are particularly weak with respect to the transportation and real estate sectors and the price system. Xu (2004, 2008) elaborates on China's statistical practices and describes the discrepancies with GDP measurements between China's practice and the 1993 SNA guideline (although such discrepancies do not impact estimation of the size of GDP). He (2010) also reports discrepancies in China's GDP accounting, attributing them mainly to the revision methodology used after the 2004 census.

International observers have challenged the Chinese data from several perspectives (IMF, 2006).[1] The World Bank (1997) claimed growth may have been about one percentage point lower than official figures state during 1978-1995 due to underestimation of consumption and investment deflators.[2] Some au-

[1] In 2010, the overall score for China's data in the World Bank Statistical Capacity Indicator (SCI) was 58 on a scale 0-100, somewhat below the international average (65). The World Bank Statistical Capacity Indicator (SCI) consists of three assessment areas: methodology (where China's score is 50), data sources (score 40), and periodicity and timeliness (score 83). See World Bank (2011).

[2] Young (2003) suggests that systematic understatement of inflation by enterprises accounts for 2.5% growth per year in the nonagricultural economy during the first two decades of the reform period (1978–98). Henderson et al. (2009) compare the official output growth to

thors question the credibility of Chinese statistics (e.g. Rawski, 2001), while others point out discrepancies between provincial and national GDP figures. Holz (2008) presents the view that the economic census results of 2004 actually provide evidence in favor of the provincial numbers against the national aggregates.

This paper combines the interest in building coincident indicators to evaluate China's growth with an examination of how well they match the data on GDP. We use factor analysis to summarize information from various macroeconomic indicators, effectively producing a subset of coincident indicators for the Chinese economy. We compare the dynamics of the estimated factors with GDP, and compare our factors with other published indicators for the Chinese economy. The results suggest that our indicator data, summarized by principal components, closely match the GDP dynamics and that the discrepancies between GDP data and components are small. Moreover, the dynamics of our indicator are extremely close to those published by the US Conference Board and China's NBS (especially since 2001). The periods of discrepancies between GDP data and the coincident indicators seem to correspond to shocks affecting the growth process, as neither autoregressive models for GDP itself nor various coincident indicators are able to forecast GDP growth at these periods satisfactorily.

The contribution of the study is that our paper brings previously suggested episodes of statistical inconsistencies into an empirical test by means of factor analysis.[3] This differs from previous papers that have established alternative GDP series for China by means of growth accounting, or have pointed out data discrepancies without using econometric techniques to evaluate their exact timing or extent. Moreover, we are unaware of any other study that compares published coincident indicators for China with the aim of analyzing how well these capture the dynamics of the reported GDP. This is important; the different published indicators on China's development now attract considerable public attention, yet their relationship with China's GDP figures has not been

that suggested by satellite data on lights at night during the period 1992/1993 – 2005/2006. According to their results, the adjusted growth rate for China is somewhere between 7.0-9.0% p.a.

[3] There are a few studies that use principal component analysis in tracking the state of the economy (see e.g. Federal Reserve Bank of Chicago, 2001; Giannone et al., 2008).

evaluated. The benefit of our approach is that the relative importance of the different factors can be examined and the factors can be given an economic interpretation, advancing the understanding of the drivers of growth in China.

The outline of the paper is as follows. Section 2 presents the methodology of the study, and includes a discussion about the relevant data issues. This is followed by Section 3 presenting the estimation results. Section 4 closes the paper.

2 Methodology

We compare various indicators of growth with the officially reported GDP data. Using static factor analysis, we combine information from numerous production indicators by a principal components approach, and regress the reported GDP growth figures on the estimated factors.[4] The estimated factors provide a proxy for the dynamics of growth, which could be treated as a type of latent variable (see e.g. Aigner et al., 1984). By analyzing the fit of the model across time, we can detect periods when the reported growth figures are at odds with our indicators.

Since a complete coverage of economic production is hard to achieve even in advanced economies, the OECD (2002) attempted to provide a standard on measurement for the non-observed economy (NOE).[5] The principles for measuring the NOE are quite general and useful in selecting the indicators of economic activity. However, observed deviances between our measure of economic activity and China's GDP may stem from the fact that our measure better captures the non-observed economy not present in official statistics. A similar index is used by the Chicago Fed to measure "real-time" economic activity in the US providing support to our approach (Federal Reserve Bank of Chicago, 2001).[6]

[4] See Stock and Watson (2002), who construct forecasts by principal components from a large number of predictors. They show that as the number of predictors and time-series observations grow large, the forecasts are asymptotically efficient and consistent.

[5] According to the OECD, non-observed activities are those missing from the basic data used to compile the national accounts because they are underground, illegal, informal, household production for own final use, or due to deficiencies in the basic data collection system.

[6] The Chicago Fed National Activity Index is the first principal component comprising 85 variables representing four categories of the US data: i) production and income; ii) employ-

Given the data at hand, our approach is best described as a mixture of the production approach and income-based methods, although some demand-side data are used as control variables (OECD Handbook, Ch. 5). The production approach attempts to measure the industry-specific production from agriculture, construction, trade etc. using indicators such as fertilizers, cement and import statistics. Income-based methods use information on household income such as disposable income or net income. The exact variables used to estimate the static factor are specified in Appendix (A). There are 83 variables in total.[7]

In selecting the exact variables to be included in the factor model, we benefit from the analysis by Rawski (2001), who highlighted numerous inconsistencies between standard data during the slowdown of 1997-2000. We bring several of these suggested inconsistencies into a statistical test. We include indicators of energy production, noting Rawski's observation that while reported real GDP grew by close to 25% during 1997-2000, energy consumption dropped by almost 13%.[8] Similarly, the inclusion of the production figures of numerous industrial products is justified by the persistently large share of industry in China's GDP (roughly 50%) and the fact that the trends in production of these goods are sometimes at odds with aggregate industrial production figures. Among the industrial products, we include steel and cement. Again, Rawski makes the claim that the high reported growth in investment during 1997/1998 is inconsistent with the weaker growth in steel consumption and cement output.

In order to capture demand-side pressures, we include consumer price inflation, imports from Asia, and cargo at ports.[9] Income developments are taken into account by including the growth of disposable income per capita in the urban areas, together with the cash income per capita of rural households. Rawski

ment, unemployment and hours; iii) personal consumption and housing; and iv) sales orders and inventories. It seems to track the US business cycle surprisingly closely.

[7] The results by Bai and Ng (2002) suggest that the number of variables to construct the factor need not be extremely large for the principal components approach to yield precise estimates. Thus, the number of indicator variables in our analysis could be even smaller. However, we include all the available variables that could give reasonable information about the growth dynamics.

[8] The reason for using energy production, instead of consumption, is data availability.

[9] Imports from Asia and cargo handled at ports are included in the OECD Composite Leading Indicator for China. Our empirical analysis suggests that these data provide meaningful information to evaluate the coincident dynamics in China as well.

(2001) suggests that aggregate retail sales figures are at odds with household income figures as higher retail sales imply an increasing propensity to consume. In fact, the savings rate of Chinese households has been increasing over time. Income developments are also partly reflected in the measures of service (transportation) sector we employ such as the overall number of tourists and passengers on highways, waterways, railways, and in aviation. We also include profits of industrial enterprises as an indicator of overall profitability of the economy.

Principal component analysis aims at reducing the dimensionality of a dataset consisting of a large number of interrelated variables, while retaining as much of the variation present in the dataset as possible (Jolliffe, 2002). The principal components methodology applies the variance structure of the indicator variables and represents a solution method for a factor model. The presentation of the model here closely follows Johnson and Wichern (2002).

Consider an observable $(p\times 1)$ random vector X, with mean μ and covariance matrix Σ. According to a general factor model, X is linearly dependent on unobservable random variables $F_1, F_2, ..., F_m$, also called the common factors, and on p additional sources of variation, $(\varepsilon_1, \varepsilon_2, ..., \varepsilon_p)$, the specific factors. An orthogonal factor model can be written in matrix notation as $X = \mu + LF + \varepsilon$. Here the matrix L is a $(p \times m)$ matrix of factor loadings, with its components l_{ij} representing the loading of the ith variable on the jth factor. The idea is to capture the most variation in X with the principal components $C = LF$. Principal components (PCs) are ordered so that the first component retains the highest share of variation in the dataset compared to the other indicators. Ideally, the PCs have meaningful interpretations, but only in the limits of the accompanying economic theory or the context of a particular application.

To obtain the principal components solution, let us consider the sample correlation matrix R specified in terms of its spectral decomposition. Let R be a $(k\times k)$ positive definite matrix with the spectral decomposition $A = \sum\limits_{i=1}^{k} \lambda_i e_i e_i'$. Assume that the normalized eigenvectors are the columns of another matrix $P = [e_1, e_2, ..., e_k]$. Then it holds that $A = \sum\limits_{i=1}^{k} \lambda_i e_i e_i' = P\Lambda P'$, with $PP\prime = P\prime P = I$ and Λ is a diagonal matrix with $\lambda_1, \lambda_2, ..., \lambda_k$ on the diagonal. We specify the sample correlation matrix R in terms of its eigenvalue-eigenvector pairs $(\widehat{\lambda}_1, \widehat{e}_1), (\widehat{\lambda}_2, \widehat{e}_2), ..., (\widehat{\lambda}_p, \widehat{e}_p)$, where $\widehat{\lambda}_1 \geq \widehat{\lambda}_2 \geq ... \geq \widehat{\lambda}_p$. If $m \prec p$ is the number of common factors, the matrix of estimated factor loadings $\left\{\widetilde{l}_{i,j}\right\}$ is given by $\widetilde{L} = \left[\sqrt{\widehat{\lambda}_1}\widehat{e}_1 \vdots \sqrt{\widehat{\lambda}_2}\widehat{e}_2 \vdots \ldots \vdots \sqrt{\widehat{\lambda}_m}\widehat{e}_m\right]$.

As we want to combine information from a relatively large number of indicators into a small number of factors, we are interested in the share of sample variance contributed by the various factors. The proportion of the total sample variance explained by the jth factor is $\widehat{\lambda}_j/p$, with $(\sqrt{\widehat{\lambda}_j}\widehat{e}_j)'(\sqrt{\widehat{\lambda}_j}\widehat{e}_j) = \widehat{\lambda}_j$. While there is no consensus over how to choose the PCs for a regression analysis, it is obvious the PCs should retain enough of the variation of the original data. In our case, the question is simplified by our underlying aim to capture the dynamics of GDP. However, since the data are noisy and some indicators are perhaps poorly measured, some PCs with a high variance could, in fact, be poor explanatory variables of the underlying "latent" variable – GDP growth. Since we do not know a priori which variables are measured well, we include all the original 83 variables in our dataset to compute the principal components. Therefore, it is plausible that some of the first PCs are not closely related to actual GDP growth as they may capture noise or the underground economy, while some PCs with a low variance may have high explanatory power.[10]

After the principal components are determined, estimating a principal component regression is straightforward. A standard regression model is $Y = Z\beta + u$, where Y represents China's GDP growth rate, Z corresponds to the vector of independent explanatory variables, β is the vector of regression coefficients and u represent the i.i.d. errors – deterministic components are omitted for simplicity. We then use our estimated principal components as the explanatory variables in the Z vector.

3 Empirical evidence

Due to data availability limitations and the significant reforms that China underwent in early 1990s, our estimation sample spans 1997Q1 to 2009Q4.[11] This time period captures two episodes of challenging international crises for Chi-

[10] For more discussion see Jolliffe (2002), Chapter 8.

[11] According to Holz (2003), the World Bank accepted the official Chinese GDP data for its own publications in 1999. However, to include the Asian crisis, we begin our sample in 1997Q1.

nese policymakers: the Asian and the global crisis. Prior to the start of the sample period, China experienced overheating with inflation surging alongside demand pressures in 1993-1994. Thereafter, China's policymakers applied stringent macroeconomic policies that brought inflation back under control. China even experienced deflationary episodes in 1998-2000 and 2002. Some observers attribute the period of deflation to productivity increases and tariff cuts due to WTO membership (see IMF, 2003), but weak demand is likely to have contributed as well, given that the regional environment experienced a negative shock due to the Asian crisis in 1997-1998. Strong growth ensued in 2003, and the government used contractionary macroeconomic policy at that time. Nevertheless, economic overheating became a serious concern for officials in 2006. The global financial crisis hit China's economy through the collapse in international trade in late 2008, but growth rates remained relatively high in international comparison. The lowest year-on-year growth rate (6.2%) was recorded in the first quarter of 2009. A large fiscal stimulus package announced in late 2008 was instrumental in maintaining economic growth; its impact was seen most directly in infrastructure investment.

It is necessary to obtain stationary data to apply the principal components analysis. For this purpose, all non-negative series were transformed into logarithms. Year-on-year growth rates of the resulting series were next taken to be consistent with our dependent variable, the reported (year-on-year) GDP growth rate.[12]

The estimated first principal component, applying the sample correlation matrix, from our data sample explains 22% of total sample variance, while 8% is explained by the second component. We depict the first principal component in Figure 1, together with the dependent variable, the y-o-y growth in GDP. There appears to be a rather strong comovement between the two series. With the ten first principal components we are able to explain just under 70% of the total sample variance, as shown in Table 1.

[12] The official consumer price inflation series is already reported in year-on-year terms.

Component	Proportion, %	Cumulative proportion, %
$pc1$	22.39	22.39
$pc2$	8.29	30.68
$pc3$	7.42	38.10
$pc4$	7.04	45.13
$pc5$	5.32	50.45
$pc6$	4.55	55.01
$pc7$	4.09	59.10
$pc8$	3.80	62.90
$pc9$	3.26	66.16
$pc10$	3.06	69.22

Table 1: Principal component analysis

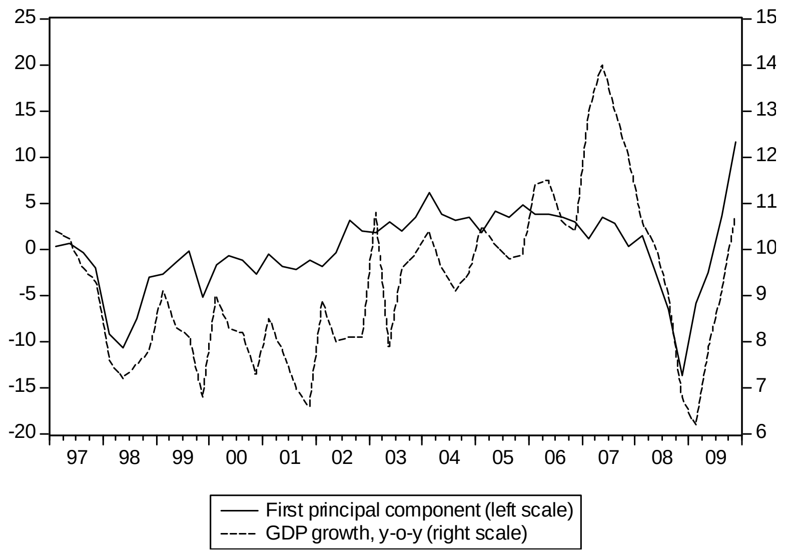


Figure 1. First principal component and GDP growth

As a starting point for our analysis, we regress the reported GDP growth rate on the ten first principal components, a constant and a linear trend, applying an OLS estimation (Model 1 in Table 2).[13] We then sequentially eliminate all

[13] Boivin and Ng (2006) find that the importance of the various factors depends on the exact macroeconomic time series to be explained; for some macro variables the estimated factors beyond the first three are quite important.

regressors with the lowest t-values, at each step re-estimating the model, until all coefficients satisfy the 10% significance threshold. In the resulting Model 2 in Table 2, only the components $pc1$, $pc4$ and $pc7$ maintain their statistical significance, together with the constant term and a linear trend. What is the composition of these three components in terms of the factor loadings?

The first principal component has high factor loadings on industrial indicators, specifically electricity production and production indicators for individual industrial goods. As the industrial sector corresponds to 50% of China's GDP from the production side, the close relationship between this component and GDP growth depicted in Figure 1 is not surprising. The fourth principal component has significant loadings on passenger numbers, i.e. indicators closely linked to the service sector, while the interpretation of the seventh component is somewhat less clear. However, given that the highest factor loadings are on household incomes, both rural and urban, and on some consumption goods such as garments, ventilators and coke, the latter used in heating and cooking, the seventh component could represent household income and consumption.[14] If the three principal components are interpreted as industrial production, service sector and household consumption/income, respectively, our model suggests these are the drivers of growth in China.

Given the possibility that some of the indicator series may be associated with GDP with a lag, we also consider a model including the first lags of the principal components $pc1$, $pc4$ and $pc7$ in the regression. However, a standard Wald test does not reject a hypothesis that all the first lags can be set jointly to zero (p-value 0.70).[15] We therefore continue with Model (2) of Table 2 as the benchmark. The residuals from this regression are depicted in Figure 2.[16]

[14] While coke is also used for electricity production, our interpretation here builds on the fact that each PC adds *new* information to the previous ones and the PCs are orthogonal. Since electricity production ranked high in the first component, the new information present in coke is probably more related to household than industry use. Also inflation, which affects household savings, has a high loading in the seventh component.

[15] The fit of the regression could be considerably improved by including a lagged dependent variable. However, this would run counter to the aim of using indicator variables to proxy growth dynamics in the economy.

[16] A standard ADF test on the residuals of Model 2 provides strong evidence that the residuals are stationary.

Variable	Coefficient estimate (1)	Coefficient estimate (2)
$pc1$	0.129 (0.038)	0.163 (0.033)
$pc2$	0.139 (0.111)	—
$pc3$	-0.047 (0.062)	—
$pc4$	0.268 (0.118)	0.190 (0.106)
$pc5$	-0.019 (0.061)	—
$pc6$	0.078 (0.056)	—
$pc7$	0.504 (0.136)	0.405 (0.108)
$pc8$	-0.128 (0.113)	—
$pc9$	0.011 (0.072)	—
$pc10$	-0.157 (0.101)	—
c	5.370 (1.432)	6.605 (0.865)
Adj. R-squared	0.556	0.550

Table 2: Estimation results. HAC (Newey-West) standard errors in parentheses. Trend not displayed.

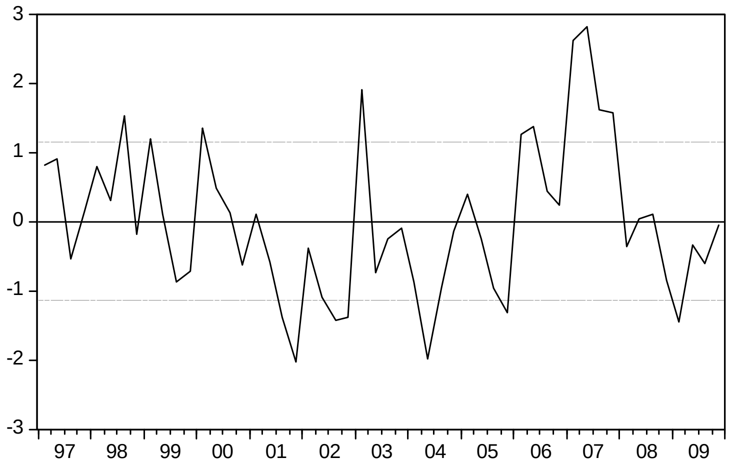


Figure 2. Residuals from benchmark regression

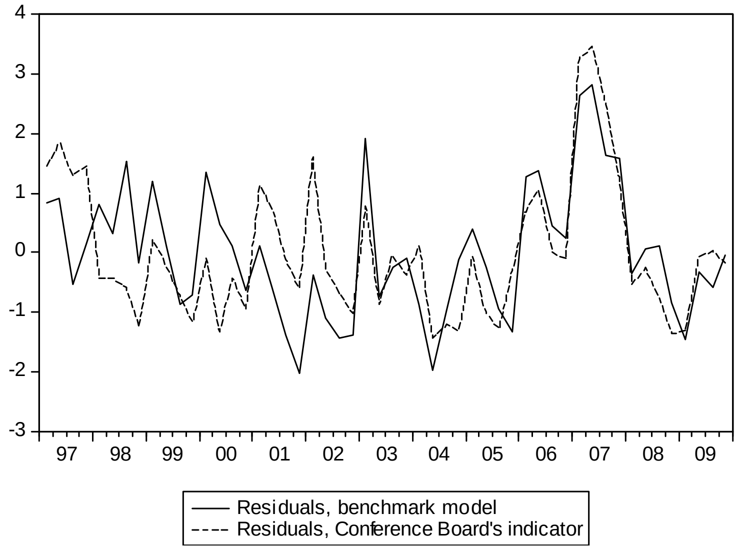
While for most of the sample the fit of the regression is perhaps surprisingly good, there are some periods with more significant discrepancies. The significant residuals appear during the Asian crisis, the period 2001-2003, the overheating of the economy in 2007 and the global crisis in end-2008. Rawski (2001) suggested that the slowdown in the late 1990s was more severe than reported in the GDP data. Our results similarly show discrepancies between the estimated factors and GDP data in 1998 during the Asian crisis.[17] For most of the late 1990s and through to early 2001, however, the estimated factor indicates no major discrepancies. Our results obtain some support from Chow (2006), who does not think it plausible that GDP growth rates in 1999-2001 would have been extremely low and yet increase dramatically in the following years. In contrast, when the GDP growth rate was in the range of 12%-14% in 2007, our analysis claims that this dynamic is not in line with the estimated factors. A similar finding is obtained for the first quarter of 2009, when the global financial crisis impacted China mainly through a decline in international trade.

An obvious check to evaluate the robustness of the above findings is to regress the announced GDP growth rates on another set of indicators. The Conference Board in 2010 started to publish both a coincident and leading indicator for China, providing backdata all the way to 1988.[18] The Coincident Economic Index is comprised of value added of industrial production, retail sales of consumer goods, electricity production, volume of passenger traffic and manufacturing employment. Similarly, the NBS reports a coincident index among its three macro-economic climate indices, providing data back to 1991. The coincident index is reported to "reflect the basic trend in the economy," and it is calculated using the following four data: industrial production, employment, investment, consumption and foreign trade, and social income, the latter including government tax revenue, enterprise profits and the income of residents (see e.g. NBS, 2010). We regress the announced GDP growth rate on the coincident indicators separately (together with a constant and trend) and compare

[17] Holz (2003) notes that there may have been substantial revisions to the NBS's energy data in 1997/1998. A re-estimation of the model starting from 1998 does not bring about statistically significant changes in the estimated parameters, however.

[18] When the Conference Board's coincident indicator was first published, the *Economist* (2010) noted that since China's economic present is almost as unclear as its future, the coincident indicator is of interest. This is, of course, the rationale for our factor analysis.

the residuals from these regressions with those using the estimated factor (our benchmark model). The residuals from these regressions are shown in Figure 3 (with the coefficient estimates reported in Appendix B).



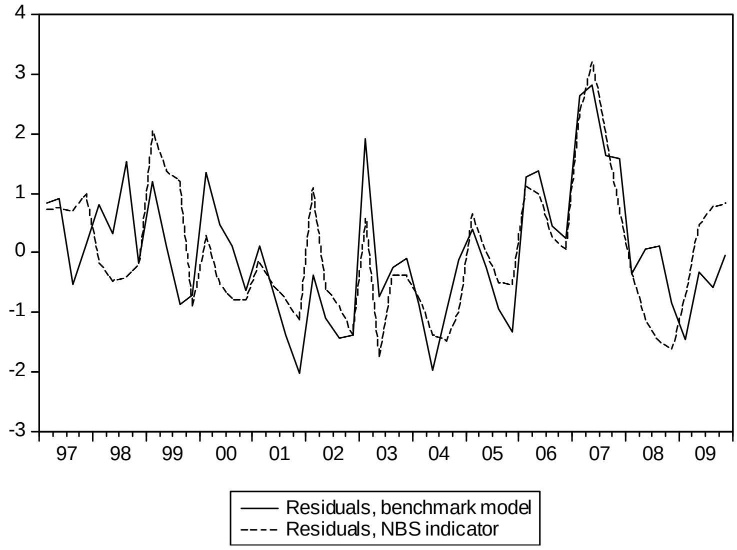


Figure 3. Residuals from benchmark model, together with regressions with Conference Board's coincident indicator (above) and NBS's coincident index (below)

Interestingly, the residuals are largely in line with our model, especially from 2002 onwards.[19] There are various possible explanations for this, not all mutually exclusive. It can be that both of the coincident indicators, and our estimated factors, are missing some important component of GDP, leading to the observed residuals. The output of the service sector is the most obvious candidate, although both our factors and the NBS's coincident indicator include corporate enterprise profits and residents' income that could partly reflect China's service sector developments. Another explanation is that there could be measurement errors or other quality problems with the GDP data as the residuals from the regressions are largely in line with another. However, we should emphasize that the number of outliers among the residuals is actually rather small, and for most of the sample the dynamics of the coincident indices are in line with the GDP data. Holz (2004) argues on the basis of Keidel's (2001) comparison of China's production and expenditure side GDP figures that China's real GDP data are not likely to be systematically biased, and are, in fact, rather reliable. One could read the outcome of our factor analysis as supporting this argument.

As the dynamics of the estimated residuals vary prior to 2002, we evaluate the performance of the coincident indicators in capturing the GDP dynamics from the start of the sample through 2001. When we estimate the regression models during the sample 1997Q1-2001Q4, we find that our first estimated principal component maintains its statistical significance during the period. While the NBS coincident indicator is also statistically significant, the Conference Board indicator is not. This suggests that our first estimated factor has important explanatory value during the earlier part of the sample, when the dynamics between the different indicators vary. Given that the fourth component had high factor loadings on indicators related to the service sector and the seventh may represent the incomes and consumption of the household sector, the insignificance of the fourth and seventh components in the shorter sample could reflect structural changes, including those in the growth model, of the Chinese economy over time.[20] We have also constructed a vector autoregressive model, specified

[19] In Appendix C, we also graphically compare the residuals from the Conference Board's model with that of the NBS.

[20] Similarly, recursive residuals suggest that the significance of the fourth and seventh components improves in the latter part of the overall sample. This could suggest that industrial

with four lags, with the variables included in Model (2) of Table 2. Analyzing the variance decomposition from this system, the importance of shocks to $pc1$ in driving GDP growth is consistently higher than those of $pc4$ or $pc7$, despite different tested variable orderings. This provides some further evidence about the importance of the first principal component.[21]

We next look deeper into the observed residuals for the entire sample. We take two time periods of large discrepancies, one with a negative (at 2001Q4) and one with a positive sign (2007Q1), and do a simple forecasting exercise around those time periods.[22] In particular, we wish to evaluate whether it is possible to obtain accurate forecasts around the time of the outliers by using (contemporaneous) values of the coincident indicators, or whether a simple autoregressive process for GDP growth would be more useful. The structure of the estimated models is shown in Table 3, and the forecasts from the models are shown in Figure 4.

	Benchmark factor model
Explanatory variables	$pc1_t, pc4_t, pc7_t, constant, trend$
Dependent variable	$GDPgrowth_t$
	AR(1)
Explanatory variables	$GDPgrowth_{t-1}, constant, trend$
Dependent variable	$GDPgrowth_t$
	Conference Board
Explanatory variables	$Coincident\ indicator_t$ (y-o-y), $constant, trend$
Dependent variable	$GDPgrowth_t$
	NBS
Explanatory variables	$Coincident\ index_t, constant, trend$
Dependent variable	$GDPgrowth_t$

Table 3: Models used for forecasting

production has only gradually led to an increase in household incomes and supported private consumption and the demand for services in the economy. These results are available from the authors upon request.

[21] These results are available from the authors upon request.

[22] Our approach here is related to "nowcasting," where monthly data releases are used to produce current-quarter forecasts of GDP growth (see e.g. Giannone et al., 2008).

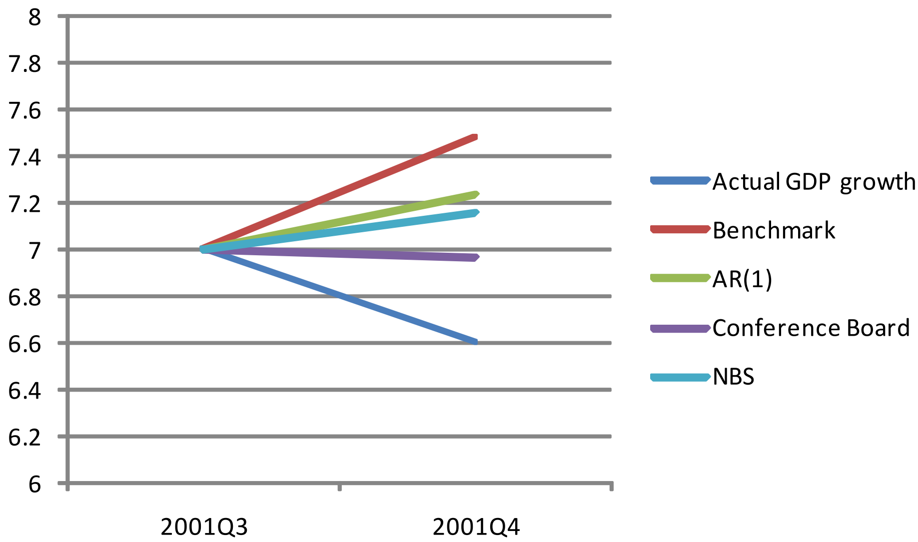
We estimate the models to 2001Q3 and 2006Q4, and evaluate the one-step ahead forecasts at these time periods, when there was a jump in the residual series for all of the three models (Benchmark factor model, Conference Board, and the NBS model). Not surprisingly, neither the benchmark factor models nor the coincident indices do a good job in forecasting at the times of outliers. More surprising perhaps is that the AR(1) model using only lagged GDP growth and the deterministic terms does not forecast well either.[23] This suggests that there could be a structural break in the process around that time, possibly due to an external shock hitting the economy.

We also look at the forecasts by international experts at these time points, in particular those published in the World Economic Outlook (WEO) of the International Monetary Fund (IMF).[24] China's growth prospects were slightly revised downwards in December 2001 due to the 9/11 terrorist attacks in the US, and China's main trading partners' prospects were subject to an even bigger revision. There was a major upward revision during April-September 2007 for China's GDP forecast for that year, by a total of 1.5 percentage points. This suggests that there could indeed be a break in the growth process at that time. However, as the *current* values of the coincident indicators also fail to capture the GDP dynamics in 2001Q4 and 2007Q1, there is a break in the process that is not explainable by the indicators used in the construction of the coincident indicators. It is also interesting to note that there was significant ex-post revision of the GDP data for both 2001 (up by 1.0 percentage points) and 2007 (up by 2.3 percentage points), which suggests that there may have been important data collection problems at that time as well.[25]

[23] There is a difference between the two forecasting periods in that while the actual observation is inside the 95% confidence intervals for all models in 2001Q4, it is outside the confidence intervals for all models in 2007Q1. We have also tested for higher order AR processes, but lags above the first one were not statistically significant. This finding is in line with Galbraith (2003).

[24] Pons (2000) analyzes the accuracy of IMF and OECD forecasts for the G7 countries. He does not find evidence of consistent over- or under-estimation in the forecasts.

[25] The magnitude of the revision is calculated by comparing the reported revised figures in 2010 with those originally published in China's Statistical Yearbooks in 2002 and 2008.



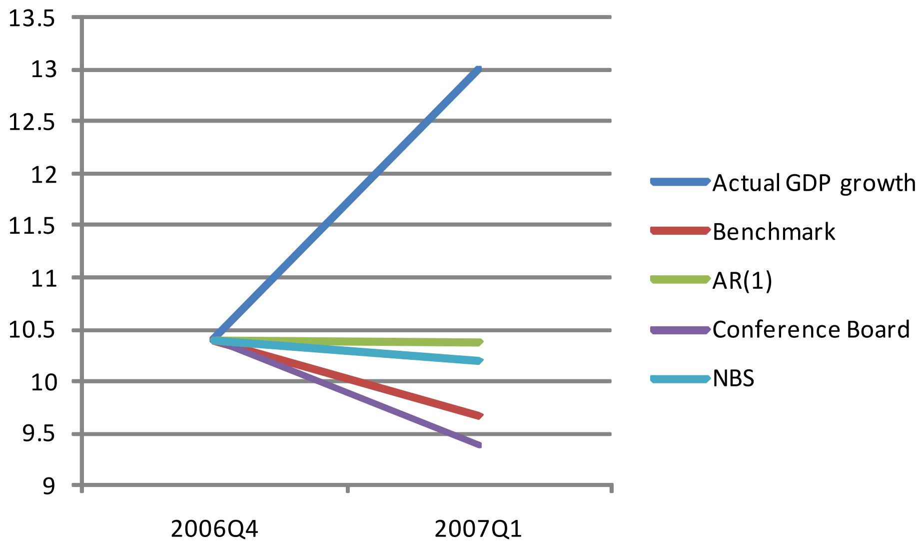


Figure 4. One-step ahead forecasts for 2001Q4 (top) and 2007Q1 (bottom).

4 Conclusion

The increased economic importance of China has prompted international observers and researchers to evaluate its economic developments with coincident and leading indicators. Indeed, the National Bureau of Statistics in China itself now publishes both a coincident and leading indicator. Both types of indicators usually combine a broad range of information from many variables, aggregating them by statistical techniques. At the same time, some researchers have expressed concerns about the quality, and even credibility, of China's GDP statistics. Then, other indicators for the macroeconomy could be meaningfully compared with the GDP statistics, in order to bring insight both to the current and possible future state of the economy.

We combine the previous considerations by aggregating information from various macroeconomic indicators for China by means of factor analysis, where we consider a type of coincident indicator and compare the dynamics of the estimated factors with the GDP statistics. In particular, we regress GDP growth rates on the estimated static factors and examine the residuals to find out when the GDP dynamics are possibly at odds with the estimated factors. Our sample runs from 1997 to 2009, covering both the Asian and the global crisis, with an episode of fast economic growth, but relatively subdued inflation, in between. We also compare our estimated factors with two other coincident indicators published for the Chinese economy, i.e. the ones produced by the US Conference Board and China's National Bureau of Statistics. Our factors also provide insights into how the drivers of growth in the Chinese economy may have changed over time.

We find that the dynamics of the GDP data match the estimated factors relatively well, and there are only very short periods (mostly individual quarters) with discrepancies between the series. Perhaps unsurprisingly, these discrepancies appear at turning points of business cycles, or times of international crises. Interestingly, the dynamics of the Conference Board's and the NBS coincident series are very similar to our estimated factors. Moreover, it is not possible to obtain reliable forecasts at the times of the discrepancies even with AR forecasts of the GDP growth rate itself, and there have been forecast revisions by international observers at these times as well. This suggests that the discrepancies occur at times of possible structural breaks in the series, and subsequent data revisions imply that there may have been data collection problems at these time periods. Nevertheless, we emphasize that during a major part of the sample the

GDP dynamics match well those of the coincident indicators.

One avenue of future research could be an evaluation of the various leading indicators for China and how well they forecast the future path of GDP. Such analyses would be especially relevant in ascertaining possible turning points in business cycles and would have important implications for economic policymaking.

A Data

Table 4 lists the variables used to estimate the static factor.

Industrial Production		
Air Conditioner	Household Refrigerator	Steel Products
Alternating Current Generator	HH. Washing Machines	Sulphur Acid
Autom.: Buses & Coaches	Iron Alloy	Synthetic Ammonia
Autom.: Cars	Kerosene	Synthetic Detergents
Autom.: Loading Vehicles	Lubricant Oil	Synthetic Feed Stuffs
Bicycles	Metal Cutting Machines	Synthetic Rubber
Camera	Micro Computer	Television Sets: Colour
Canned Food	Motor Cycles	Total Energy Production
Caustic Soda	Oily Gas Ventilator	Tractors
Cement	Paints	Woolen Goods
Chemical Fertilizer	Passenger Coaches	Woolen Yarn
Cemical Fibre: Artificial	Pharmac. and Medicine	Yarn
Chemical Fibre: Synthetic	Pig Iron	
Civil Steel Ships	Plastic Products (PP)	*Other Series*
Cloth: CF, Pure Chemical	PP: Membrane for Agric.	Cargo Handled at Ports
Cloth: Pure Cotton	Power Generated (PG)	Enterprise Deposits
Coke	PG: Thermal Power	Imports from Asia
Color Kinescope	PG Equipment	Inflation
Computer	Processed Crude Oil	M2
Concentrated Nitric Acid	Program Control Switchbr.	Number of tourists
Dairy Products	Pure Benzene	Disposable income *per capita*, urban
Diesel Oil	Rubber Tyre	Enterprise profits
Dye	Salt	Rural cash income
Ethylene	Semiconduct. Integr. Circuit	Passengers, aviation
Freight Wagons	Sewing Machines	Passengers, highway
Fuel Oil	Silk	Passengers, railway
Garments	Small Tractors	Passengers, waterway
Gasoline	Soda Ash	
Hi Fi	Steel	

Table 4: List of variables.

B Estimation results for other models

Model	Coefficient estimate
AR(1)	
GDPgrowth_{t-1}	0.708 (0.101)
Constant	2.022 (0.784)
Conference Board	
$\text{Coincident indicator}_t$ (y-o-y)	0.422 (0.075)
Constant	4.575 (0.812)
NBS	
$\text{Coincident index}_t$	0.469 (0.083)
Constant	−36.150 (7.895)

Table 5: Estimation results for competing models. HAC (Newey-West) standard errors in parentheses. Trend not displayed.

C Comparison of residuals

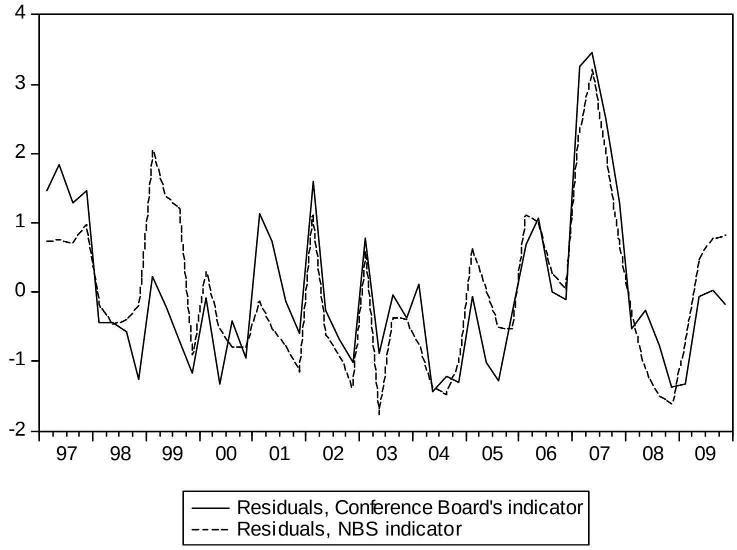


Figure. Residuals from regressions with Conference Board's coincident indicator and NBS's coincident index

References

[1] Aigner, D. J., Hsiao, C., Kapteyn, A., Wansbeek, T., 1984. Latent variable models in econometrics. In *Handbook of Econometrics*, Volume II. Eds. Z. Griliches and M.D. Intriligator. Elsevier Science Publishers BV.

[2] Bai, J., Ng, S., 2002. Determining the number of factors in approximate factor models. *Econometrica* 70(1), 191–221.

[3] Boivin, J., Ng, S., 2006. Are more data always better for factor analysis? *Journal of Econometrics* 132, 169-194.

[4] Chow, G., 2006. Are Chinese Official Statistics Reliable? CESifo Economic Studies 52(2), 396-414.

[5] Economist, 2010. Reading China's palm – A new leading indicator for China's economy. May 20th, 2010.

[6] Federal Reserve Bank of Chicago, 2001. Background on the Chicago Fed National Activity Index. www.chicagofed.org/digital_assets/publications/cfnai/background/cfnai_background.pdf

[7] Galbraith, J.W., 2003. Content horizons for univariate time-series forecasts. *International Journal of Forecasting*, 19, 43–55.

[8] Giannone, D., Reichlin, L., Small, D., 2008. Nowcasting: The real-time information content of macroeconomic data. *Journal of Monetary Economics* 55, 665–676.

[9] He, X., 2010. Noteworthy Discrepancies in China's GDP accounting. *China & World Economy* 18(4), 88-102.

[10] Henderson, J.V., Storeygard, A., Weil, D.N., 2009. Measuring economic growth from outer space. NBER Working Paper 15199.

[11] Holz. C. A., 2003. "Fast, Clear and Accurate:" How reliable are Chinese output and economic growth statistics? *The China Quarterly* 173,122-163.

[12] Holz, C. A., 2004. China's statistical system in transition: Challenges, data problems, and institutional innovations. *Review of Income and Wealth* 50(3), 381-409.

[13] Holz, C. A., 2008. China's 2004 economic census and 2006 benchmark revision of GDP statistics: More questions than answers? *The China Quarterly* 193,150-163.

[14] IMF, 2003. *Deflation: Determinants, risks and policy options – Findings of an interdepartmental task force.* International Monetary Fund, Washington, DC.

[15] IMF, 2006. *People's Republic of China: 2006 Article IV Consultation—Staff report; staff statement; and public information notice on the Executive Board discussion.* IMF country report No. 06/394

[16] Johnson, R. A., Wichern, D., 2002. *Applied multivariate statistical analysis.* 5th edition. New Jersey: Prentice Hall.

[17] Jolliffe, I. T., 2002. *Principal component analysis.* 2nd edition. Springer series in statistics.

[18] Keidel, A., 2001. China's GDP expenditure accounts. *China Economic Review* 12(4), 355-367.

[19] NBS, 2010. China Monthly Economic Indicators, 2010.8. National Bureau of Statistics: Beijing.

[20] OECD, 2002. *Measuring the non-observed economy. A Handbook.* OECD.

[21] Pons, J., 2000. The accuracy of IMF and OECD forecasts for G7 countries. *Journal of Forecasting* 19, 53–63.

[22] Rawski, T. G., 2001. What is happening to China's GDP statistics? *China Economic Review* 12, 347–354.

[23] Stock, J. H. and Watson, M. W., 2002. Forecasting using principal components from a large number of predictors. *Journal of the American Statistical Association* 97(460), 1167–1179.

[24] World Bank, 1997. China 2020: Development Challenges in the New Century. Washington, DC.

[25] World Bank, 2011. Bulletin Board of Statistical Capacity, http://go.worldbank.org/6EIB0ZUV00.

[26] Xu, X., 2004. China's gross domestic product estimation. *China Economic Review* 15, 302–322.

[27] Xu, X., 2008. Some differences in GDP measurement between China's practice and 1993 SNA guideline. *China Economic Review* 19, 480–488.

[28] Young, A., 2003. Gold into base metals: Productivity growth in the People's Republic of China during the reform period. *Journal of Political Economy*, 111(6), 1220–1261.

[29] Zheng, J., 2001. China's official statistics. Growing with full vitality. *China Economic Review* 12, 333–337.

Earlier BOFIT Discussion Papers

For a complete list of Discussion Papers published by BOFIT, see www.bof.fi/bofit

2010	No 1	Anatoly Peresetsky: Bank cost efficiency in Kazakhstan and Russia
	No 2	Laurent Weill: Do Islamic banks have greater market power?
	No 3	Zuzana Fungáčová, Laura Solanko and Laurent Weill: Market power in the Russian banking industry
	No 4	Allen N. Berger, Iftekhar Hasan and Mingming Zhou: The effects of focus versus diversification on bank performance: Evidence from Chinese banks
	No 5	William Pyle and Laura Solanko: The composition and interests of Russia's business lobbies: A test of Olson's "encompassing organization" hypothesis
	No 6 to	Yu-Fu Chen, Michael Funke and Nicole Glanemann: Off-the-record target zones: Theory with an application Hong Kong's currency board
	No 7	Vladimir Sokolov: Bi-currency versus single-currency targeting: Lessons from the Russian experience
	No 8	Alexei Karas, William Pyle and Koen Schoors: The effect of deposit insurance on market discipline: Evidence from a natural experiment on deposit flows
	No 9	Allen N. Berger, Iftekhar Hasan, Iikka Korhonen, Mingming Zhou: Does diversification increase or decrease bank risk and performance? Evidence on diversification and the risk-return tradeoff in banking
	No 10	Aaron Mehrotra and José R. Sánchez-Fung: China's monetary policy and the exchange rate
	No 11	Michael Funke and Hao Yu: The emergence and spatial distribution of Chinese seaport cities
	No 12	Alexey A. Ponomarenko and Sergey A. Vlasov: Russian fiscal policy during the financial crisis
	No 13	Aaron Mehrotra and Alexey A. Ponomarenko: Wealth effects and Russian money demand
	No 14	Asel Isakova: Currency substitution in the economies of Central Asia: How much does it cost?
	No 15	Eric Girardin and Konstantin A. Kholodilin: How helpful are spatial effects in forecasting the growth of Chinese provinces?
	No 16	Christophe J. Godlewski, Zuzana Fungáčová and Laurent Weill: Stock market reaction to debt financing arrangements in Russia
	No 17	Zuzana Fungáčová, Laurent Weill , Mingming Zhou: Bank capital, liquidity creation and deposit insurance
	No 18	Tuuli Koivu: Monetary policy, asset prices and consumption in China
	No 19	Michael Funke and Michael Paetz: What can an open-economy DSGE model tell us about Hong Kong's housing market?
	No 20	Pierre Pessarossi, Christophe J. Godlewski and Laurent Weill: Foreign bank lending and information asymmetries in China
2011	No 1	Aaron Mehrotra and Jenni Pääkkönen: Comparing China's GDP statistics with coincident indicators

Bank of Finland
BOFIT – Institute for Economies in Transition
PO Box 160
FIN-00101 Helsinki

☎ + 358 10 831 2268
bofit@bof.fi
http://www.bof.fi/bofit