Bank i Kredyt 52(5), 2021, 437-456

BANK I KREDYT
BANK & CREDIT

Should we recalculate the level of spillover effects if the alternative GDP measures for China are correct?

Anna Sznajderska*

Submitted: 27 April 2020. Accepted: 6 May 2021.

Abstract

Official statistics show very stable growth rates of production and prices in recent years in China. These statistics are widely criticized. Alternative measures of China's GDP suggest that China's growth rates are exaggerated. This indicates that the slowdown of the Chinese economy is even bigger than usually assumed. In this paper, we try to answer the question whether alternative data on Chinese GDP affect the level of spillover effects of the Chinese economy. To this end, we estimate two alternative GVAR models and compare the obtained results. The usage of the alternative GDP series with a lower growth rate than the official GDP growth rate appears to weaken the spillover effect.

Keywords: alternative GDP measures, spillover effects, GVAR model

JEL: C32, E01, E32, F10, O50

* Warsaw School of Economics; e-mail: asznajd@sgh.waw.pl.

1. Introduction

Reliable statistical data are necessary to analyse and evaluate the functioning of each economy. The Chinese economy is a particularly interesting case. China is one of the most important economies in the world. China started its reform and opening in 1979 and since then, China's share in global GDP increased from about 2.3% in 1980 to 18.75 in 2018 (World Economic Outlook Database). China is a major importer and producer of many raw materials. In addition, since China joined the World Trade Organization in July 2001, there has been a dynamic development of trade from and to China. The Chinese economy began to grow rapidly and many trade barriers were removed, which allowed the increase in the sales of consumer and investment goods to China.

The Chinese economy is changing from an export and investment driven economy to a consumption-based economy. The transformation of the economy leads to a slower but more stable rate of economic growth. The slowdown of the Chinese economy and the related spillover effects are widely discussed (see, for example, Cashin, Mohaddes, Raissi 2016; Dieppe et al. 2018; Inoue, Kaya, Ohshige 2015; Osorio, Unsal 2013). According to our knowledge, however, there is no study that looks at the differences in the size of China's spillover effect depending on different measures of China's GDP.

Official statistics on the level of GDP in China raise many doubts. First of all, these data are very smooth and stable, meaning that their values rarely change. Over the past twenty quarters (2015Q1 – 2019Q4), China's GDP growth rate has varied from 6.1% to 7.0%. It is difficult to find such stability in GDP growth rate in other countries. Secondly, the local governments are rewarded for meeting growth and investment targets that cause an overestimation of China's national accounts. The National Bureau of Statistics (NBS henceforth) in China is aware of this fact and tries to adjust the aggregate statistics. Some economists claim that this adjustment is sufficient. In Section 2 we present the arguments of official data supporters.

Many economic studies indicate that Chinese GDP is more volatile and pro-cyclical than official figures show. This is indicated by lower level data (GDP components), such as steel or coal production, iron ore imports, semiconductor prices or car sales volume. Also, the data smoothing is not reflected in data on the service sector only (whose role in China is constantly growing). Therefore, the level of GDP in China is estimated in a number of alternative ways (using, for example, energy consumption, luminosity, railroad freight, passenger travel, volume of bank loans, construction indicators or level of national taxes), that we discuss in Section 3.

We use the GVAR model to estimate the level of spillover effects of a negative output shock in China for the chosen countries. An interesting research problem is to check whether the results change when we use the modified series of the Chinese GDP. We estimate two GVAR models, the first one uses official GDP data and the second applies alternative GDP data delivered by Chen et al. (2019). We discuss and compare the obtained results in Section 4. Our estimates indicate that the magnitude of spillover effects is smaller for the model that uses alternative GDP measures than for the model that makes recourse to the official data.

2. China's official statistics

Many economists are skeptical about China's official GDP figures. Xiong (2018), for example, presents a theoretical framework in which competition between local governments increases both GDP and investment in China. Lyu et al. (2018) provide evidence that the regional growth target can be achieved by fabricating data. On the other hand, there are also economists who claim that China's official data are reliable. This group includes, for example, Michael Owyang (Vice President of Federal Reserve Bank of St. Louis), Hannah Shell, John Fernald, Israel Malkin and Mark Spiegel. Below we present some of their arguments.

Owyang and Shell (2017) indicate that one of the sources of problems with China's GDP statistics is the transformation from a centrally planned economy to a market economy. The NBS was created to report agricultural production and production data in state-owned enterprises to check if production targets were met. In the market economy, the scope of the collected data is different, and it takes time to understand and implement the new system. From 1993, China began to apply the United Nations' System of National Accounts, which is based on the concept of value added. This concept was completely new to many Chinese statisticians and administrative staff. In addition, the authors point out that measurement errors in an economy as large and complex as China is inevitable.

Moreover, the NBS statistical office is aware of the overstatement of the results by local authorities and is trying to make proper adjustments, e.g. in 2015, it reported a GDP value of 10.4 trillion dollars, which is 7% less than the sum of local GDP levels. In 1994, the country introduced census surveys to bypass statistical departments and check the quality of the data. In 1998, a reform was introduced to allow a structural break in the GDP time series to get rid of earlier revaluations.

It is worth noting that China has a complicated system of exercising power. Local authorities play a key role in the country's economic development. Local authorities are responsible for around 70% of the country's fiscal expenditure, including the development of economic institutions, infrastructure, opening new markets, construction of roads, highways and airports. Despite this, local authorities are designated by the central authorities and not elected by the local electorate. Bai, Hsieh and Song (2019) indicate that Chinese local governments have enormous political power and administrative capacity, providing so-called "special offers" for privileged private companies.

The overstatement of GDP figures is largely due to the assessment system of the provincial authorities on which their further career path depends. The level of GDP and the level of industrial production for a given province are indicators based on which the activity of local authorities is assessed. Moreover, as a consequence of the evaluation system, local statistical offices operate under strong pressure from provincial authorities.

Chen et al. (2019) show the scale of the NBS' corrections (see Figures 1 and 2 in Chen et al. 2019). It turns out that until 2003 official statistics were understated in relation to local data, while after 2003 this trend reversed. Since 2006, the gap between local estimates and official data has been 5% of the aggregate GDP. It turns out that the gap between estimates was mainly due to differences in the industrial sector. These results are confirmed by Holz (2014) and Ma et al. (2014). Slight discrepancies were noted in the consumption and inventory data. Data on net exports were overstated by official statistics, while data on the level of investment (gross fixed capital formation) were understated. Chen et al. (2019) prove that the scale of adjustment is insufficient (cf. Section 3).

The ambiguities associated with the calculation of the Chinese GDP, and perhaps above all the lack of clear rules for its calculation raise concerns about the independence of the NBS. According to Owyang and Shell (2017), Ben Bernanke and Peter Olson emphasized that the lack of transparency from the NBS in China is the result of low statistical reliability of the data, and not the political dependence of the NBS. The NBS provides data that is less volatile than in other countries, which can lead to the illusory impression that the data is unreliable or manipulated. However, the smoothness of the data is rather related to the technical aspect of their calculation than political manipulation.

In addition, the World Bank provides the Statistical Capacity Indicator that reflects the country's ability to produce and disseminate high-quality aggregate statistical data. The indicator is based on 25 criteria in the following areas: methodology, data sources, and periodicity and timeliness. The overall Statistical Capacity score is calculated as a simple average of all three area scores on a scale of 0–100. Statistical capacity score for China increased from 64 in 2004 to 80 in 2016, but it decreased to 78 in 2018.

Some economists believe that the scale of adjustment on the NBS side is insufficient and official data on the Chinese GDP are overstated despite corrections (at least recently). Some economists are uncertain about the quality of official data on the Chinese GDP and are interested in doing a robustness check. Both of these groups measure China's GDP using alternative methods. In the next section we discuss a number of such methods.

3. Alternative GDP measures for the Chinese economy

We can divide the papers on China's GDP into 5 categories. The first group consists of papers that claim that China's official data are generally accurate (Holz 2014; Perkins, Rawski 2008). The second group claims that China's official data is not accurate, but without specifying the direction of bias (WikiLeaks publication of the premier, Li Keqiang, in 2007; Fernald, Malkin, Spiegel 2013). The third group state that China's official data has an upward bias (Rawski 2001; Maddison, Wu 2006; Chen et al. 2019). The fourth group are papers that claim that China's official data has a downward bias (Clark, Pinkovskiy, Sala-i-Martin 2017). Whereas the last group shows time-varying bias (Nakamura, Steinsson, Liu 2016).

Below we present some of these papers and the chosen alternative GDP measures that were used for the Chinese economy. One of the alternative GDP measure is the Li Keqiang index, which is the index used by the current Chinese prime minister. The prime minister used this indicator when he was secretary of the Communist Party of Liaoning province. According to a State Department note that leaked through WikiLeaks, Li Keqiang told the US ambassador in 2007 that GDP data for Liaoning Province are unreliable and he himself uses simple arithmetic average of the growth rates of electricity production, railroad freight, and bank loans. He admitted also that the official statistics are man-made and for reference only.

Fernald, Malkin and Spiegel (2013) have developed an index based on the three characteristics recommended by Li Keqiang and have made this data available online. The authors estimated the regression model of the calculated Li Keqiang index on Chinese GDP in 2000–2009, and then using the estimated coefficients, created a GDP forecast for 2009–2012. It turns out that changes in the official GDP series after 2009 were in line with changes in the Li Keqiang index.

Another alternative measure of GDP is the change in energy consumption. In China the most important sector of the economy is industry, in which energy consumption is very high. The level of energy consumed is correlated with the production level and is independent of the possible manipulation of the Chinese authorities. The level of energy consumed is also a variable that Chinese statisticians in the centrally planned economy could easily measure.

Already at the end of the 90's Rawski (2001) shows that from 1997 to 2000 the level of Chinese GDP increased by 24.7 pp and the level of energy consumed dropped by 12.8 pp. This means a 30% decrease in energy consumption in these years, which is unlikely for an industrial economy. The author argues that the data provided demonstrate an overestimation of GDP data for China. He compares data from the Chinese economy with data from other countries. It turns out that double-digit GDP growth was always associated with double-digit energy consumption growth.

It is worth noting that using energy consumption as a measure of GDP may cause some problems. For example, the decrease in energy consumption with a simultaneous increase in GDP may be caused by an increase in the efficiency of its consumption, a change from an economy based on industry to an economy based on services or a change from production to energy consumption.

Nakamura, Steinsson and Liu (2016), on the other hand, using Engle curves, construct new consumption growth and inflation statistics for China for 1995–2011. The Engle curves are based on the empirical finding that as households become richer, a larger fraction of total expenditures are spent on luxuries and a smaller fraction are spent on necessities. The authors state that official statistics present a smoothed version of reality.

There are several private sector research companies that have developed their own measures of the Chinese GDP growth based on a wide range of indicators, including freight volume, passenger travel, electricity production, construction indicators, purchasing managers indexes, or financial indicators such as money supply and stock market index. These measures usually indicate an overestimation of the Chinese GDP during periods of economic downturn. Owyang and Shell (2017) call these indicators black boxes because we don't know exactly how they are calculated. Moreover, they indicate that different measures often lead to very different time series. As an example, the authors give the third quarter of 2015, when the Lombard Street Research measure indicated a GDP growth of 2.9%, the Bloomberg model showed a 6.6% growth, while the official estimate was 6.9%.

The indicators described above are not able to show whether the level of Chinese GDP has been overestimated for a long period. They also do not include the growing sector of services and agricultural production.

The situation is different with the next alternative method that uses satellite data that measures the intensity of man-made night lights (luminosity). Unlike human-made economic data, this data is resistant to falsification or misreporting. These data are used, for instance, by Henderson, Storeygard and Weil (2012) or Clark, Pinkovskiy and Sala-i-Martin (2017).

Henderson, Storeygard and Weil (2012) use data from the satellites of the United States Air Force Defense Meteorological Satellite Program to calculate alternative GDP measures for a panel of 188 countries between 1992 and 2008. These satellites orbit the Earth 14 times a day since 1970. Their estimates differ from official data by a maximum of 3 pp annually. The authors argue that data on luminosity reflect the level of GDP well, because the consumption of all goods in the evening requires light. To verify this claim, they confirmed that the diversity of illuminated pixels in individual countries is positively correlated with income.

Clark, Pinkovskiy and Sala-i-Martin (2017) quite surprisingly provide evidence that official China's GDP measures may be underestimated. The authors exploit the same data on luminosity as Henderson, Storeygard and Weil (2012), but instead of using nighttime lights as a component of measure of economic activity, they use it as an auxiliary variable to help uncover the correlation structure between the measures they use in their index.

Lastly, we present the work of Chen et al. (2019), which is particularly important for our study. Chen et al. (2019) correct the value of the Chinese GDP in two ways. First, they adjust the level of domestic GDP using the difference in value added reported by the NBS and the increased rate in VAT revenues reported by the State Administration of Taxation (SAT) for sectors where VAT is the main type of taxation. Chen et al. (2019) indicate that GDP estimates by the NBS were consistent with SAT data up to 2007/2008. After this period, in the years 2008 to 2016, the GDP growth rate was 1.7 pp lower than the rate reported by the NBS. In addition, the aggregate investment and savings rate was 7 pp lower than official statistics in 2016. Maddison and Wu (2006), to compare with, show revisions of GDP growth of 7.85% a year 1978–2003, compared with the official rate of 9.59%. It is important to note that Chen et al. (2019) assume that China's official GDP data before 2008 are generally accurate.

Secondly, the authors estimate the relationship between local levels of GDP and a group of economic factors before 2008. Then, they use the calculated coefficients to estimate levels of local GDP after 2008 based on the level of selected economic factors. Selected factors include satellite night lights, national taxes, electricity consumption, rail freight flow, export and import levels.

Chen et al. (2019) come to three basic conclusions. First of all, nominal GDP growth in China after 2008, and in particular after 2013, is lower than the growth reported in official statistics. Secondly, the authors point out that the true savings rate declined by around 10 pp from 2008 to 2016, with two--thirds of the decline showing up in the external surplus and one-third in the investment rate. Official statistics show a decrease of just 3% over these years. Thirdly, the authors' calculations suggest that the investment rate declined by around 3% of GDP between 2008 and 2016, whereas official statistics show that the investment rate increased during this period. Additionally, it is worth noting that Chen et al. (2019) emphasize the weak position of the NBS compared to the strong position of local leaders in the Chinese political system, which results in the inability to improve official statistics. It is worth noting that the problem is even greater because all attempts to start a discussion on the correctness of the official GDP data in China are treated as an attack on the powers of the Communist Party.

4. Does the adjustment of China's GDP data change the size of the spillover effects?

4.1. Methodology – GVAR model

The global vector autoregression model (GVAR) is the model that enables the analysis of the whole world economy. It consists of a few equations for each country, which usually describe the level of economic activity, the level of prices, the level of the exchange rate, the level of the interest rate, and also the level of other economic variables chosen by the researcher. Each equation consists of both domestic and foreign variables. The foreign variables are calculated as the weighted average of domestic variables and the weights usually depend on the level of trade between the economies.

The following vector autoregression model with exogenous variables (VARX(p_i, q_i)) is estimated for each country in the sample ($i = 1,\ldots N; N$ is the number of countries), where p_i and q_i are lag lengths selected by the Schwarz information criterion (SBC).

$$x_{it} = \alpha_{i0} + \alpha_{i1}t + \Phi_{i1}x_{i,t-1} + \ldots + \Phi_{ip_i}x_{i,t-p_i} + \Lambda_{i0}x^*_{it} + \ldots + \Lambda_{iq_i}x^*_{i,t-q_i} + \Psi_{i0}\omega_t + \ldots + \Psi_{iq_i}\omega_{t-q_i} + u_{it} \tag{1}$$

where:

x_{it} is a vector of domestic variables,
x^*_{it} is a vector of foreign variables,
ω_t is a vector of variables used in the dominant unit.

Pesaran and Chudik (2013) define the dominant unit as the unit that influences the rest of the variables in the model both directly and indirectly, and its effects do not vanish even as the dimension of the model (N) tends to infinity. The dominant unit acts as a dynamic factor in the regressions of the non-dominant units. One may model as dominant unit a particular economy (for example the United States) or a particular global variable (for example oil prices).

Importantly, the foreign variables are calculated as weighted averages of domestic variables:

$$x^*_{it} = \sum_{j=0}^{N} \rho_{ij} x_{jt}, \text{ where } \rho_{ii} = 0$$

ρ_{ij} are the weights calculated on the basis of trade flows $\left(\sum_{j=0}^{N} \rho_{ij} = 1\right)$ from IMF DOTS,
k_i is a number of domestic variables for the i^{th} economy,
k^*_{it} is a number of foreign variables for the i^{th} economy.

There are two improvements that we have considered here. The first one is to use trade-in-value--added data, which incorporate also services from the OECD database. The second one is to use financial flows data as in Backé, Feldkircher and Slacík (2013) from BIS. But in both cases, we ended up with too many missing data.

The dominant unit is estimated as:

$$\omega_t = \mu_0 + \mu_1 t + \kappa_1 \omega_{t-1} + \ldots + \kappa_{\tilde{p}} \omega_{t-\tilde{p}} + \lambda_1 \tilde{x}_{t-1} + \ldots + \lambda_{\tilde{q}} \tilde{x}_{t-\tilde{q}} + \eta_t \tag{2}$$

where $\tilde{x}_t$ are feedback variables, constructed as a weighted average of variables included in the models for non-dominant units, with weights based on PPP-GDP (see Smith, Galesi 2014, p. 153 for details).

In the first step the models for individual countries are estimated in the following VECMX(p_i, q_i) error correction form:

$$\Delta x_{it} = c_{i0} + \sum_{j=1}^{r_i} \gamma_{ij} ECT_{ij,t-1} + \sum_{p=1}^{p_i} \tilde{\phi}_{ip} \Delta x_{i,t-p} + \sum_{q=0}^{q_i} \tilde{\Lambda}_{iq} \Delta x^*_{i,t-q} + \sum_{q=0}^{q_i} \tilde{\Psi}_{iq} \Delta \omega_{t-q} + u_{it}$$

where *ECT* denote error correction terms and r_t is the number of cointegrating relations.

Most of the variables are nonstationary with stationary first differences; thus, the usage of VECM models allows us to capture long-run relationships that exist among the domestic and the country specific foreign variables.

In the second step, the corresponding VARX models (Eq. 1) are recovered from the estimated VECMX models. Then the individual country models are stacked into one model:

$$G_0 x_t = a_0 + a_1 t + G_1 x_{t-1} + \ldots + G_p x_{t-p} + \Psi_0 \omega_t + \ldots + \Psi_q \omega_{t-q} + u_t$$

$$G_0 = \begin{pmatrix} A_{00} W_0 \\ A_{10} W_1 \\ \ldots \\ A_{N0} W_N \end{pmatrix}, G_j = \begin{pmatrix} A_{0j} W_0 \\ A_{1j} W_1 \\ \ldots \\ A_{Nj} W_N \end{pmatrix}, a_0 = \begin{pmatrix} a_{00} \\ a_{10} \\ \ldots \\ a_{N0} \end{pmatrix}, a_1 = \begin{pmatrix} a_{01} \\ a_{11} \\ \ldots \\ a_{N1} \end{pmatrix}, u_t = \begin{pmatrix} u_{0t} \\ u_{1t} \\ \ldots \\ u_{Nt} \end{pmatrix}$$

where $A_{i0} = \left(I_{k_i}, -\Lambda_{i0}\right), A_{ij} = \left(\Phi_{ij}, \Lambda_{ij}\right),\ j = 1, \ldots p,\ p = \max(p_i), q = \max(q_i)$. W_i is a $\left(k_i + k_i^*\right) \times k\ \left(k = \sum_{i=0}^{N} k_i\right)$ link matrix for the i^{th} country. It consists of two blocks: the first one is the selection matrix for domestic variables and the second one is the matrix defined by trade weights ρ_{ij} for foreign variables.

$$\begin{pmatrix} x_{it} \\ x_{it}^* \end{pmatrix} = W_i x_t$$

When solving the GVAR model equations (2) and (3) are written as one equation, which is solved recursively. The equations simplify significantly when we estimate a model without a dominant unit. However, this is not the case used in this paper, hence we do not give further complicated formulas that are shown in Smith and Galesi (2014).

In the next step we calculate the generalised impulse response functions (GIRFs) that were developed in Koop, Pesaran, and Potter (1996) and Pesaran and Shin (1998), and we calculate 90% bootstrap confidence bands. We cannot use the standard IRFs that assume orthogonal shocks, because of the large number of variables in the GVAR model. GIRFs do not depend on the ordering of the variables. GIRFs show how changes in the Chinese GDP affect the other variables in the GVAR over time regardless of the source of the change. We do not know whether such a shock stems from a shift in the demand or supply of output in China or in other countries. The shock is equal to one standard deviation of the error in China's GDP equation.

4.2. Data

The following five domestic variables (x_{it}) are used in our GVAR model for each country: real GDP, CPI, stock price index, REER, and short-term interest rate. Two foreign variables (x_{it}^*) are used: real GDP and short-term interest rate. Furthermore, one global variable (ω_t) is used, i.e. the oil prices. The model uses a time-varying matrix of trade flow. The dominant unit for oil prices includes two feedback variables (real GDP and short-term interest rate). The same dataset was used in Sznajderska and Kapuściński (2020).

The sample consists of 59 economies, where the 19 euro area countries are grouped into one economy. We use quarterly observations. The time span of data is from 1995Q1 to 2017Q4.

In the Appendix, in Table 3 we present data sources for each economy. The primary data sources are IMF IFS for the GDP, prices and the interest rate, BIS for REER, and MSCI for the stock index. Real GDP, CPI, stock market indexes, and REER are 2010 indices (2010 = 100). The GDP and CPI data are seasonally adjusted and they are in logarithms.

Oil prices were taken from FRED, as an average from Brent, West Texas Intermediate and Dubai Fateh indexes. We used the World Economic Outlook database of the IMF for the construction of country-specific PPP-GDP weights. We took the average between 1995 and 2017 from the annual data on purchasing power parity measured in billions of international dollars.

The matrixes of trade flows were constructed on the basis of International Monetary Fund statistics, namely the Direction of Trade Statistics (DOTS). These are annual data, so they allowed us the construction of the matrix of trade flows for each year separately and subsequent estimation of the model with time-varying link matrixes.

As an alternative China's GDP measure, we decided to use data calculated by Chen et al. (2019). These data are publicly available (https://www.brookings.edu/bpea-articles/a-forensic-examination-of-chinas-national-accounts/). In addition, these are the most recent GDP estimates for China. Chen et al. (2019) are quite skeptical about the quality of official data, and their correction is certainly significant. Figure 1 shows two time series for real GDP growth rate in China, the first one is the growth rate based on official data, and the second one the growth rate calculated by Chen et al. (2019) based on VAT receipts.

The data used in the model ends in 2017Q4, while the estimates of Chen et al. (2019) end in 2016. However, Chen et al. (2019) estimate that the GDP growth rate in China is overestimated by an average of 2% per year. Thus, we decided to lower the GDP growth rate in China in 2017, i.e. the missing year, also by 2%.

In the next section we present the results of estimation of two GVAR models – the first one uses official data, and the second one uses adjusted data. Our aim is to check how the change in China's GDP series affects the size of the spillover effect.

4.3. The results of estimating two GVAR models differing in the GDP measure used

We obtained a stable GVAR model, meaning convergent persistent profiles,[1] eigenvalues lying in the unit circle, and non-explosive GIRFs. To arrive with a stable GVAR, we reduced the number of cointegrating relations for Argentina from 5 to 1 relation. Argentina is the only country for which the program has determined the maximum possible number of cointegrating vectors (and reported it as a full rank).

In the first step, we estimate the GVAR model with the official GDP series for China. In the second step, allowing model parameters to change (e.g. the lag length or the number of cointegrating relationships), we estimate GVAR with an alternative GDP series, calculated on the basis of the work of Chen et al. (2019). The data provided by Chen et al. (2019) are in annual frequency, therefore

[1] Persistent profiles show the time profiles of the effects of either variable or system specific shocks on the cointegration relations in the GVAR model. The value of persistent profiles is unity on impact and, if the vector under investigation is indeed a cointegration vector, it should tend to zero as the time horizon tends to infinity.

we converted them to quarterly frequency. We inserted the annual observation into the last period of the quarterly data and then we performed linear interpolation on the other values.

Table 2 shows the list of countries included in the model with lag lengths (p_i and q_i) and the number of cointegrating relations (r_i) for both model 1 (official GDP data) and model 2 (alternative GDP data from Chen et al. (2019)). The only difference is the number of cointegrating relations for the Czech Republic.

In order to compare the results from both models, we decided to scale the impulse response functions so that China's GDP shocks are normalized to -1% on impact. In the case of model 1, the GDP shock was equal -0.376% on impact, and in the case of model 2, the GDP shock was equal -0.519% on impact.

Table 1 presents some of the results. In the first column there are GIRFs for model 1 with official data and in the second column for model 2 with alternative data. In the first row the scaled China's GDP shock is presented. The shock is equal to 1% on impact and is persistent, meaning that it does not return to zero, but for the next 20 quarters it is equal to about -1%. The shock refers to a sudden unexpected change in China's GDP, but the source of this change is unknown. The shock is transmitted to other economies mainly through its impact on foreign GDP.

In the second row we present the average GDP reaction for all economies included in the model. For both models (1 and 2), the reaction is not statistically significant. The reaction is slightly stronger for the model that uses the official data than for the model that uses alternative data.

In the third row we present the average GDP reaction for China's neighbouring countries (South--East-Asia and Oceania). It is worth noting that the reaction is half as weak for the model that uses alternative data (minimum equal to -0.009% in the 4th quarter) than for the model that uses official data (minimum equal to -0.25% in the 3rd quarter). In both cases the reaction is not statistically significant.

In the following two rows we present the GDP reactions for Taiwan and Hong Kong. Hong Kong is China's financial center. Taiwan is one of China's largest trading partner that also aims to become Asian financial hub. Both regions are considered by the Chinese authorities as part of China's territory, although they have considerable autonomy. The response of GDP in these economies is statistically significant for both models. The model that uses official data indicates a twice as strong response of economic activity than the model that uses alternative data. In the case of model 1 that uses official data, the strongest GDP reaction is equal to -0.47% in the 7th quarter for Taiwan and -0.30% in the 3rd quarter for Hong Kong.

In the next two rows, we present the GDP reactions for Singapore and Thailand. In the case of model 1 that uses official data, the reactions are on the border of statistical significance. The maximum reaction of a GDP is equal to -0.33% in the 3rd quarter for Thailand and -0.40% in the 3rd quarter for Singapore. In the case of model 2 that uses alternative data, the reaction of the GDP for both economies is statistically insignificant and much weaker. The maximum reaction of the GDP is equal to -0.07% in the 3rd quarter for Thailand and -0.10% in the 5th quarter for Singapore.

The last, eighth row presents the results for the United States. In the case of both model 1 and model 2, the GDP reaction for the United States is statistically significant. The reaction is weaker in model 2 than in model 1. In the case of model 1 that uses official data, the maximum reaction is equal -0.22% in the 8th quarter, and in the case of model 2 that uses alternative data, the maximum reaction is equal -0.15% in the 5th quarter.

The results show that including alternative GDP measures for China in the GVAR model significantly affects the results. The model that uses alternative China's GDP measures generates weaker impulse response functions for the majority of economies and leads to lower level of spillover effects. It means that if the official GDP statistics in China are overestimated, then the levels of the spillover effect reported in the literature are probably overestimated.

5. Conclusions

The study presents a novel approach to measuring the spillover effects from China. We look at the differences in the size of China's spillover effect depending on the measure of China's GDP.

The quality of the official Chinese GDP data is disputable. It is believed that the Chinese statistical office (National Bureau of Statistics) should show greater transparency in the calculation of its statistics and the data collection process, which would certainly increase their credibility.

Either because of concerns about the quality of data or because of the need for a robustness check, economists use alternative GDP measures for China. In this article we describe some of these measures, between them the Li Keqiang index, a method based on changes in the level of energy consumption, a method based on satellite data that measures the intensity of artificial lights at night, a method based on the rate of increase in VAT, and models based on many different factors that are measurable for the Chinese economy, but difficult to manipulate by the authorities.

In the empirical section we compare two GVAR models, the first model uses standard official GDP data for China and the second one applies alternative GDP data for China from Chen et al. (2019). Chen et al. (2019) provide revised estimates of China's national GDP using data on value-added taxes. The paper, published by Brookings Papers in Economic Activity, provides one of the most recent estimates, calculated in a clear and transparent way for the reader.

The results show that including alternative GDP measures for China in the GVAR model affects the results significantly. The model that uses alternative China's GDP measures generates weaker impulse response functions for the majority of economies and leads to weaker spillover effects. The reaction is at least half as weak for the Asian countries and for the United States. In the case of Singapore and Thailand, it becomes statistically insignificant for the model that uses alternative data.

Our preliminary results show that the level of China spillover effects should be recalculated if the alternative GDP measures for China are correct. We think that the topic is worth studying further, using different data on China's GDP and different models measuring the level of spillover effects.

References

Backé P., Feldkircher M., Slacík T. (2013), Economic spillovers from the euro area to the CESEE region via the financial channel: a GVAR approach, *Focus on European Economic Integration*, Q4/13, 50–64.

Bai Ch., Hsieh Ch., Song Z. (2019), Special deals with Chinese characteristics, *NBER Macroeconomics Annual*, 34, 341–379.

Cashin P., Mohaddes K., Raissi M. (2016), *China's slowdown and global financial market volatility; is world growth losing out?*, IMF Working Papers, 16/63.

Chen W., Chen X., Hsieh Ch., Song Z. (2019), *A forensic examination of China's national accounts*, NBER Working Papers, Spring, Brookings Papers in Economic Activity.

Chudik A., Pesaran M. H. (2013), Econometric analysis of high dimensional VARs featuring a dominant unit, *Econometric Reviews*, 32(5–6), 592–649.

Clark H., Pinkovskiy M., Sala-i-Martin X. (2017), *China's GDP growth may be understated*, NBER Working Paper, 23323.

Dieppe A., Gilhooly R., Han J., Korhonen I., Lodge D. (2018), *The transition of China to sustainable growth – implications for the global economy and the euro area*, Occasional Paper Series, 206, European Central Bank.

Fernald J., Malkin I., Spiegel M. (2013), *On the reliability of Chinese output figures*, FRBSF Economic Letter, 8, 25 March.

Henderson J.V., Storeygard A., Weil D.N. (2012), Measuring economic growth from outer space, *American Economic Review*, 102(2), 994–1028.

Holz C.A. (2014), The quality of China's GDP statistics, *China Economic Review*, 30, 309–338.

Inoue T., Kaya D., Ohshige H. (2015), *The impact of China's slowdown on the Asia Pacific region: an application of the GVAR model*, Policy Research Working Paper Series, 7442, The World Bank.

Koop G., Pesaran M.H., Potter S.M. (1996), Impulse response analysis in nonlinear multivariate models, *Journal of Econometrics*, 74(1) 119–147.

Lyu Ch., Wang K., Zhang F., Zhang X. (2018), GDP management to meet or beat growth targets, *Journal of Accounting and Economics*, 66(1), 318–338.

Ma B., Song G., Zhang L., Sonnenfeld D.A. (2014), Explaining sectoral discrepancies between national and provincial statistics in China, *China Economic Review*, 30, 353–369.

Maddison A., Wu H. (2006), *China's Economic Performance: How Fast has GDP Grown? How Big is it Compared with the USA?*, University of Queensland.

Nakamura E., Steinsson J., Liu M. (2016), Are Chinese growth and inflation too smooth? Evidence from Engel curves, *American Economic Journal: Macroeconomics*, 8(3), 113–144.

Osorio C., Unsal D.F. (2013), Inflation dynamics in Asia: causes, changes, and spillovers from China, *Journal of Asian Economics*, 24(C), 26–40.

Owyang M.T., Shell H.G. (2017), China's economic data: an accurate reflection, or just smoke and mirrors?, *The Regional Economist*, 25(2), 7–12.

Perkins D.H., Rawski T.G. (2008), Forecasting China's economic growth to 2025, in: L. Brandt, T.G. Rawski (eds.), *China's Great Economic Transformation*, Cambridge University Press.

Pesaran M.H., Shin Y. (1998), Generalized impulse response analysis in linear multivariate models, *Economics Letters*, 58(1), 17–29.

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

Sznajderska A., Kapuściński M. (2020), Macroeconomic spillover effects of the Chinese economy, *Review of International Economics*, 28(4), 992–1019.

Xiong W. (2018), *The Mandarin model of growth*, NBER Working Paper, 25296.

Acknowledgements

This work was supported by the National Science Centre under Grant No. 2016/21/D/HS4/02798.

Appendix

Figure 1
Real GDP growth in China

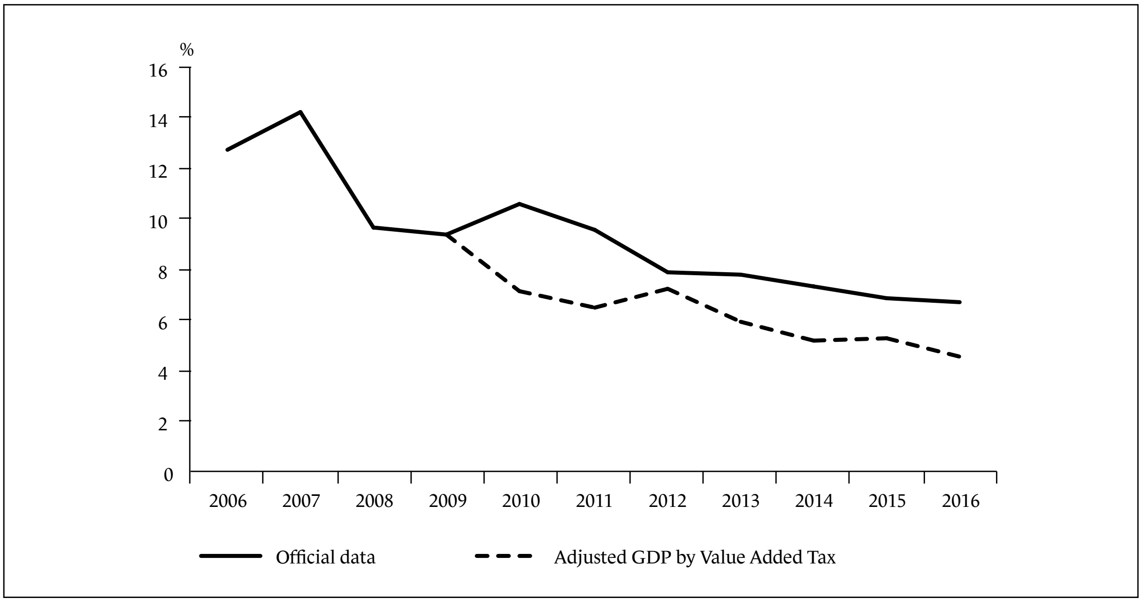


Table 1
Comparison of GIRFs from model 1 (with official data) and from model 2 (with alternative data from Chen et al. 2019)

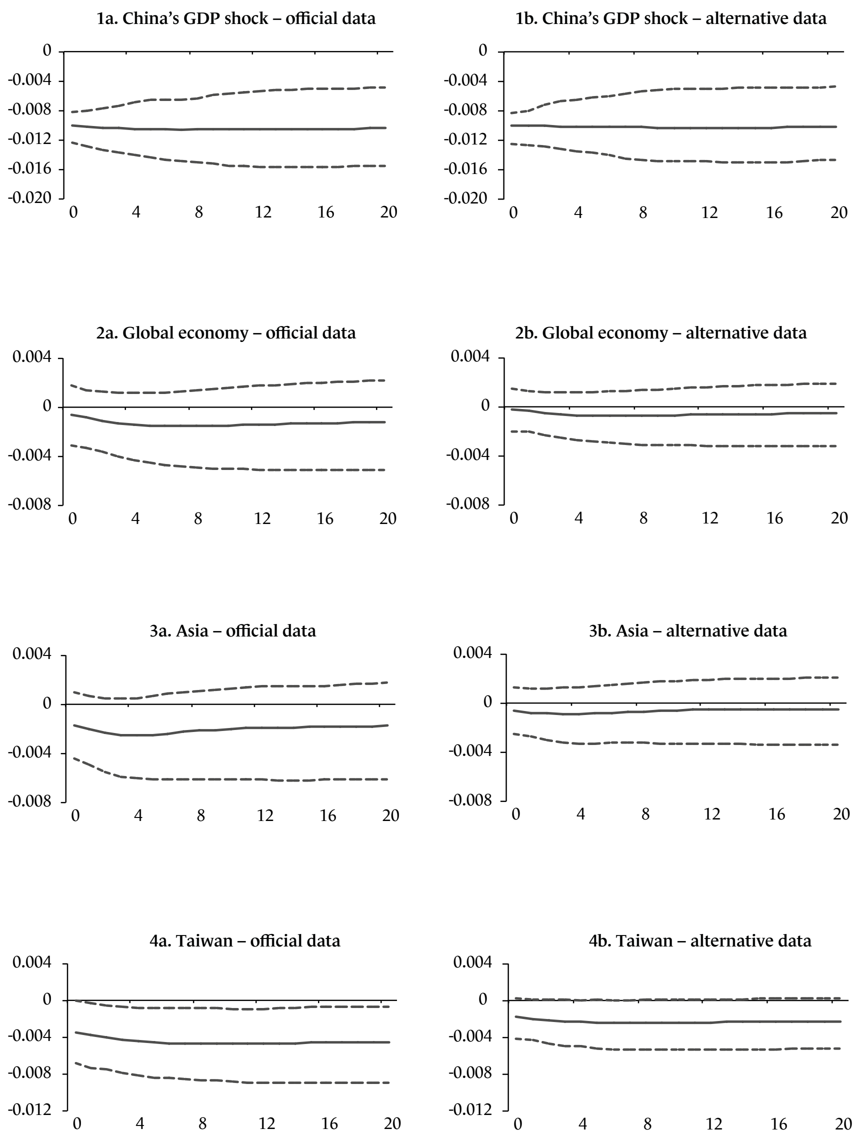
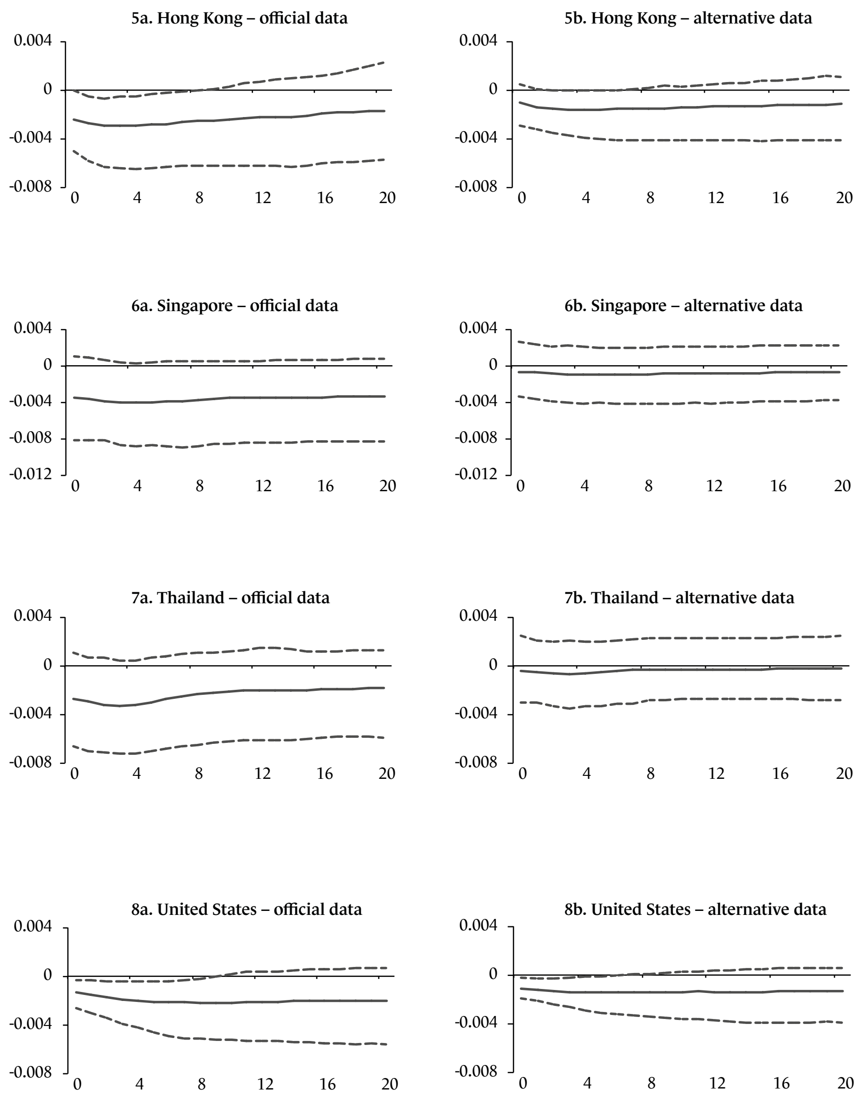


Table 1, cont'd



Note: dashed lines show 90% bootstrap confidence bands. The quarters after the shock are on the horizontal axis. Percentage deviations from the base level are on the vertical axis.

Table 2
Lag length (p_i and q_i) and the number of cointegrating relations (r_i) for model 1 (with official data) and model 2 (with alternative data from Chen et al. 2019)

	Model 1			Model 2		
	p	q	r	p	q	r
Algeria	2	1	3	2	1	3
Argentina	1	1	1	1	1	1
Australia	2	1	2	2	1	2
Brazil	1	1	2	1	1	2
Canada	1	2	2	1	2	2
Chile	1	1	3	1	1	3
China	1	1	2	1	1	2
Taiwan	1	1	2	1	1	2
Colombia	1	1	3	1	1	3
Croatia	1	1	3	1	1	3
Czech Republic	1	1	2	2	1	3
Denmark	1	1	2	1	1	2
Euro area	1	1	3	1	1	3
Hong Kong SAR	1	1	3	1	1	3
Hungary	1	1	2	1	1	2
Iceland	1	1	3	1	1	3
India	1	1	2	1	1	2
Indonesia	2	1	2	2	1	2
Israel	1	1	2	1	1	2
Japan	1	1	1	1	1	1
Korea	1	1	4	1	1	4
Malaysia	1	1	2	1	1	2
Mexico	1	1	2	1	1	2
New Zealand	1	1	1	1	1	1
Norway	1	1	1	1	1	1
Peru	1	1	3	1	1	3
Philippines	1	1	3	1	1	3
Poland	1	1	4	1	1	4
Romania	1	2	2	1	2	2
Russia	1	1	4	1	1	4
Saudi Arabia	1	1	1	1	1	1
Singapore	1	1	1	1	1	1
South Africa	1	1	3	1	1	3
Sweden	1	1	2	1	1	2
Switzerland	1	1	4	1	1	4
Thailand	1	1	3	1	1	3
United Kingdom	1	1	1	1	1	1
United States	1	1	2	1	1	2

Table 3
Data sources

	Country	Real GDP	Price level	Stock price index	REER	Short term interest rate
1	Algeria	IMF WEO	IMF IFS	–	BIS	IMF IFS
2	Argentina	IMF IFS + IMF WEO	national	MSCI + IMF IFS	BIS	IMF IFS
3	Australia	OECD	IMF IFS	MSCI	BIS	IMF IFS
4	Brazil	IMF IFS + IMF WEO	IMF IFS	MSCI	BIS	IMF IFS
6	Canada	IMF IFS + OECD	IMF IFS	MSCI	BIS	IMF IFS + OECD
7	Chile	OECD + national	IMF IFS	MSCI	BIS	IMF IFS + OECD
8	China	national	IMF IFS	MSCI	BIS	national*
9	Taiwan	national	national	MSCI	BIS	national
10	Colombia	IMF IFS + OECD	IMF IFS	MSCI	BIS	IMF IFS
11	Croatia	IMF IFS + IMF WEO	IMF IFS	–	BIS	IMF IFS + national
12	Czech Republic	IMF IFS + IMF WEO	IMF IFS	MSCI + OECD	BIS	IMF IFS
13	Denmark	IMF IFS	IMF IFS	MSCI	BIS	IMF IFS
14	Euro area	IMF IFS	IMF IFS + OECD	MSCI	BIS	IMF IFS + OECD
15	Hong Kong SAR	IMF IFS	IMF IFS	MSCI	BIS	IMF IFS
16	Hungary	IMF IFS	IMF IFS	MSCI	BIS	OECD
17	Iceland	IMF IFS + IMF WEO	IMF IFS	OECD + IMF IFS	BIS	IMF IFS
18	India	OECD + IMF WEO	IMF IFS	MSCI	BIS	IMF IFS + national
19	Indonesia	IMF IFS + OECD	IMF IFS	MSCI	BIS	IMF IFS + OECD
20	Israel	IMF IFS	IMF IFS	MSCI	BIS	OECD
21	Japan	IMF IFS + OECD	IMF IFS	MSCI	BIS	IMF IFS + OECD
22	Korea	IMF IFS	IMF IFS	MSCI	BIS	IMF IFS

Table 3, cont'd

	Country	Real GDP	Price level	Stock price index	REER	Short term interest rate
23	Malaysia	IMF IFS + national	IMF IFS	MSCI	BIS	IMF IFS
24	Mexico	IMF IFS + OECD	IMF IFS	MSCI + OECD	BIS	IMF IFS
25	New Zealand	IMF IFS	IMF IFS	MSCI	BIS	OECD
26	Norway	IMF IFS	IMF IFS	MSCI	BIS	OECD
27	Peru	IMF IFS + national	IMF IFS	MSCI + national	BIS	IMF IFS + national
28	Philippines	IMF IFS + OECD	IMF IFS	MSCI	BIS	IMF IFS
29	Poland	IMF IFS + OECD	IMF IFS	MSCI	BIS	IMF IFS
30	Romania	IMF IFS	IMF IFS	–	BIS	IMF IFS
31	Russia	IMF IFS	IMF IFS	MSCI + national	BIS	IMF IFS + OECD
32	Saudi Arabia	IMF IFS + IMF WEO	IMF IFS	–	BIS	IMF IFS + national
33	Singapore	IMF IFS + national	IMF IFS	MSCI	BIS	IMF IFS
34	South Africa	IMF IFS + OECD	IMF IFS	MSCI	BIS	IMF IFS
35	Sweden	IMF IFS	IMF IFS	MSCI	BIS	IMF IFS + OECD
36	Switzerland	IMF IFS	IMF IFS	MSCI	BIS	IMF IFS + OECD
37	Thailand	IMF IFS	IMF IFS	MSCI	BIS	IMF IFS
39	United Kingdom	IMF IFS	IMF IFS	MSCI	BIS	IMF IFS
40	United States	IMF IFS	IMF IFS	MSCI	BIS	IMF IFS + OECD

* We decided to use the Lending Rate of Financial Institutions, 1 year or less from Macrobond for China.

Czy powinniśmy ponownie obliczyć wielkość efektów *spillover*, jeśli alternatywne miary PKB dla Chin są poprawne?

Streszczenie

Oficjalne statystyki dotyczące PKB i poziomu cen w Chinach są w ostatnich latach bardzo stabilne. Statystyki te są jednak często krytykowane. Powstało wiele alternatywnych miar PKB dla Chin, z których większość wskazuje, że oficjalne statystyki są w ostatnich latach zawyżane. Oznacza to, że spowolnienie gospodarcze w Chinach jest większe niż oficjalnie szacowane. W niniejszym artykule staramy się odpowiedzieć na pytanie, czy przyjęcie alternatywnej miary PKB dla Chin wpływa na wielkość efektów *spillover* z chińskiej gospodarki. W szczególności badamy wpływ negatywnego szoku popytowego w Chinach na poziom aktywności gospodarczej w innych gospodarkach. Główna teza artykułu brzmi: jeśli alternatywne miary PKB w Chinach są poprawne, to należy ponownie policzyć wielkości efektów *spillover*.

W tym celu szacujemy globalne modele wektorowej autoregresji (modele GVAR). Modele te pozwalają na jednoczesne modelowanie globalnej gospodarki, składającej się w naszym przypadku z 59 krajów. Można je zatem wykorzystać do policzenia wielkości efektu *spillover*. Efekt ten mierzony jest jako reakcja poziomu realnego PKB w danym kraju na negatywny szok popytowy w Chinach.

Zakres czasowy analizy to okres od 1. kwartału 1995 do 4. kwartału 2017 r. Dla każdego kraju szacujemy pięć równań: dla realnego PKB, poziomu cen CPI, głównego indeksu giełdowego, realnego efektywnego kursu walutowego i dla krótkookresowej stopy procentowej. Dodatkowo w równaniach występują zmienne zagraniczne dotyczące poziomu aktywności gospodarczej i stopy procentowej. Zmienne zagraniczne tworzone są na podstawie macierzy przepływów handlowych pomiędzy krajami.

W pracy opisano różne alternatywne miary PKB dla Chin, na przykład metody oparte na poziomie zużycia energii czy intensywności palonych w nocy świateł, które widoczne są na zdjęciach satelitarnych. Następnie zdecydowano, że najlepszą i najbardziej aktualną z dotychczas opisanych jest miara zaproponowana przez Chena i in. (2019). Autorzy wyliczyli poziom PKB w Chinach na podstawie wpływów z podatków VAT.

Wyniki estymacji pokazują, że jeśli alternatywna miara PKB w Chinach jest poprawna, to efekty *spillover* są niższe niż te, które uzyskuje się przy użyciu oficjalnych statystyk. Wyniki dla modelu wykorzystującego alternatywną miarę PKB wskazują na co najmniej dwukrotnie słabszą reakcję PKB dla krajów azjatyckich i dla Stanów Zjednoczonych niż w przypadku modelu wykorzystującego oficjalne statystyki. W przypadku Singapuru i Tajlandii reakcja PKB staje się nieistotna statystycznie dla modelu wykorzystującego miarę alternatywną.

Artykuł pokazuje wstępne szacunki i zwraca uwagę na niepodjęty dotychczas w literaturze temat. Badania tego rodzaju mogą być kontynuowane, w szczególności z wykorzystaniem innych alternatywnych miar PKB w Chinach, jak również innych modeli mierzących wielkość efektu *spillover*.

Słowa kluczowe: alternatywne miary PKB, efekty *spillover*, model GVAR