ECONSTOR
Make Your Publications Visible.

A Service of
ZBW
Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

Fernald, John; Hsu, Eric; Spiegel, Mark M.

**Working Paper**

# Is China fudging its figures? Evidence from trading partner data

BOFIT Discussion Papers, No. 29/2015

**Provided in Cooperation with:**
Bank of Finland, Helsinki

*Suggested Citation:* Fernald, John; Hsu, Eric; Spiegel, Mark M. (2015) : Is China fudging its figures? Evidence from trading partner data, BOFIT Discussion Papers, No. 29/2015, ISBN 978-952-323-069-9, Bank of Finland, Institute for Economies in Transition (BOFIT), Helsinki, https://nbn-resolving.de/urn:nbn:fi:bof-201511031430

This Version is available at:
https://hdl.handle.net/10419/212839

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

WWW.ECONSTOR.EU

Mitglied der Leibniz-Gemeinschaft

BOFIT Discussion Papers
29 • 2015

John Fernald, Eric Hsu, and Mark M. Spiegel

# Is China fudging its figures? Evidence from trading partner data

Bank of Finland, BOFIT
Institute for Economies in Transition

BOFIT Discussion Papers
Editor-in-Chief Laura Solanko

BOFIT Discussion Papers 29/2015
15.10.2015

John Fernald, Eric Hsu and Mark M. Spiegel: Is China fudging its figures? Evidence from trading partner data

ISBN 978-952-323-069-9
ISSN 1456-5889 (online)

This paper can be downloaded without charge from http://www.bof.fi/bofit.

Suomen Pankki
Helsinki 2015

# Contents

Abstract ..... 4
1. Introduction ..... 5
2. Data and methodology ..... 8
3. Full sample results ..... 12
   3.1 Preliminary regression results ..... 12
   3.2 Structural breaks ..... 12
4. Results for 2008:Q1–2014:Q4 sub-sample ..... 13
5. Robustness checks ..... 15
   5.1 Does GDP have additional explanatory power? ..... 15
   5.2 Evidence from the early time period ..... 16
   5.3 Adding a 2nd principal component ..... 17
   5.4 Predicting official GDP ..... 18
   5.5 Overall in and out of sample performances ..... 18
6. Conclusion ..... 19
References ..... 22
Appendix Data sources ..... 23
Appendix Are imports a good measure of economic activity? ..... 26

John Fernald, Eric Hsu, and Mark M. Spiegel

# Is China fudging its figures?
# Evidence from trading partner data

## Abstract

How reliable are China's GDP and other data? We address this question by using trading-partner exports to China as an independent measure of its economic activity from 2000–2014. We find that the information content of Chinese GDP improves markedly after 2008. We also consider a number of plausible, non-GDP indicators of economic activity that have been identified as alternative Chinese output measures. We find that activity factors based on the first principal component of sets of indicators are substantially more informative than GDP alone. The index that best matches activity in-sample uses four indicators: electricity, rail freight, an index of raw materials supply, and retail sales. Adding GDP to this group only modestly improves in-sample performance. Moreover, out of sample, a single activity factor without GDP proves the most reliable measure of economic activity.

Keywords: China, GDP, principal components, structural break, forecasting.
JEL Classification: C53, C82, E20, F17.

**John Fernald**, Federal Reserve Bank of San Francisco. E-mail: john.fernald@sf.frb.org.

**Eric Hsu**, Federal Reserve Bank of San Francisco. E-mail: ehsu91@gmail.com.

**Mark M. Spiegel**, Federal Reserve Bank of San Francisco. E-mail: mark.spiegel@sf.frb.org.

We thank Israel Malkin, Andrew Tai, and Bing Wang for excellent research assistance. Eric Swanson and participants in the CUHK/BOFIT conference on China's Financial Intermediation provided helpful comments. The views expressed in this paper are those of the authors and do not necessarily reflect those of the Federal Reserve Bank of San Francisco or the Federal Reserve System.

# 1 Introduction

Observers of the Chinese economy have long questioned the accuracy of Chinese output figures.[1] In this paper, we assess the reliability of Chinese output figures by using trading-partner exports to China as an independent measure of its economic activity from 2000–2014. We find that the information content of Chinese GDP improves markedly after 2008. Nevertheless, even after 2008, simple activity factors—derived from the first principal component of sets of alternative indicators such as electricity or rail shipments—are more informative than GDP alone. And combining multiple indicators into a factor is more informative than using the indicators individually.

Under any circumstances, measuring Chinese GDP would be difficult. China's economy has grown rapidly and undergone extensive structural changes (e.g. Holz, 2008). Many observers further worry that output figures may be distorted, particularly by local and provincial officials in an effort to meet quotas handed down by the government. As a result, many analysts of Chinese economic activity rely instead on alternative, non-GDP indicators.[2]

Skepticism about the accuracy of Chinese data has been shared by prominent Chinese officials. For example, in 2007 current Premier Li Keqiang, was reported as saying that his province's government focused on "alternative indicators," rather than official GDP data (Wikileaks, 2007). Li mentioned three indicators: 1) electricity consumption; 2) the volume of rail cargo, which he suggests is fairly accurately measured because fees are charged for each unit of weight; and 3) the amount of loans disbursed, which may be more accurate because of regulatory oversight. By looking at these three figures, Li said he can measure with relative accuracy the speed of economic growth. Li reportedly said with a smile, "All other figures, especially GDP statistics, are 'for reference only.'"

The challenge in assessing the quality of reported Chinese output figures is to find an independent benchmark to compare with reported data. Henderson, et al (2012) use satellite data on light emissions to gauge growth in economic activity for a cross-section of countries, including China. China's reported GDP growth rate appears to be exceptionally high relative to its growth in observable light. Nakamura, et al (2014) use household consumption data to estimate Engel curves for China. They find that official aggregate consumption data are too smooth relative to what is implied by household spending patterns.

[1] See Sinclair (2012) for extensive references.
[2] For examples of informal press discussions, see Noble (2015), Sharma (2013), and Bradsher (2012).

In this paper, we use trading-partner-reported exports to China as an independent measure of Chinese economic activity. Specifically, we examine (inflation-adjusted) exports to China or Hong Kong as reported by its three major trading partners: the United States, the Euro area, and Japan. These data are not subject to manipulation or mismeasurement by Chinese authorities, but should be closely associated with economic activity in China. Specifically, since the data correspond to Chinese imports, they reflect both the use of intermediate inputs for production—an important aspect of China's economy—as well as finished goods imported for final consumption by Chinese residents. As the appendix describes, for economies with good statistical systems, imports comove very closely with GDP.

We compare movements in externally-reported exports to China to reported GDP, as well as to various combinations of "alternative indicators" of Chinese activity. If we find that movements in externally-reported exports to China are closely associated with movements in reported Chinese data, then we can conclude that these data are relatively reliable as measures of true Chinese output.

We begin by examining the first principal component of combinations of 10 widely cited and easily available (non-GDP) economic indicators produced by Chinese authorities. Our goal is to identify which indicators, singly or in combination, best explain China's externally-reported imports. Principal components estimation proves useful for yielding a parsimonious specification. Some of the individual indicators that we use might be subject to manipulation or systematic mismeasurement; but, if so, our tests would find that they are not related to our externally-reported Chinese-import data. Even if the indicators are informative, they might be noisy. By extracting an activity factor as the first principal component, we reduce the idiosyncratic noise in order to focus on the signal.

Our initial approach compares the information in a small set of potential activity indicators over the full sample of data. The set includes officially reported GDP, the first principal component of all 10 indicator variables, and the first principal component of the three Li indicators. GDP turns out to be only weakly related to externally-reported Chinese imports. The activity factors correspond much more closely to imports. Moreover, our principal component of all 10 indicators outperforms Li's set. In particular, although electricity and rail freight—two of the Li indicators—are strongly associated with imports, the lending indicator is much less important. Nevertheless, we find relatively little sensitivity to the exact group of included activity indicators in our comparisons of different groups of predictors.

The initial results do suggest that the accuracy of reported GDP—as well as the activity factors—has improved over time. Formal break tests confirm the existence of one or more structural breaks, with the most recent (and most substantive) one occurring at the onset of the global financial crisis. The improvements could reflect rising dissatisfaction by Chinese officials about the quality of their statistics. For example, Chinese officials have increasingly and openly discussed their concerns, as the Li quotation suggests. This dissatisfaction could only have increased with the onset of the global financial crisis, as successful implementation of the aggressive counter-cyclical measures adopted by the Chinese central government during the crisis required accurate assessments of prevailing economic conditions.

Given the break-test results, the remainder of our study concentrates on the period following this structural break in the first quarter of 2008. We begin with the set of ten alternative output indicators. We construct the first principal component of all 1023 possible permutations of these variables and relate them one-by-one to externally reported Chinese imports. This principal-component methodology allows us to focus on a parsimonious relationship and to identify a preferred (in sample) index of activity.

Using this methodology, we identify the ten "best-performing" sets of alternative indicators on the basis of fit. The activity factors from our top-performing sets explains most of the variation of imports within the sample. Our preferred set of indicators is electricity, rail freight, usage of raw materials, and retail spending. Individually, these indicators all have a statistically significant relationship with Chinese imports. In contrast, lending levels, one of the indicators highlighted by Premier Li, is not statistically related to imports.

Of course, our alternative indicators by construction focus on specific areas of the China economy. For example, a number, such as raw materials usage, are specifically related to manufacturing activity. As such, it is likely that the time series of Chinese imports does not follow those of our alternative indicators exactly. This raises the possibility that even after including our best alternative-indicator-based principal component reported GDP will still retain some independent explanatory power. After all, GDP is supposed to be the broadest measure of economic activity. We therefore add reported gross domestic product as a robustness check concerning the explanatory power of our indicator variables.

Our results show that, in sample, adding GDP marginally improves the statistical fit of most combinations of activity factors. Nevertheless, it does not markedly improve the fit of our best-performing indicator combinations, suggesting that, while there is additional

information, it is relatively modest. This still represents an improvement over the performance of GDP in the earlier period, where GDP had no additional information relative to our best-performing combinations of activity indicators.

We then conduct tests of the robustness of these results. We consider the use of two principal components from our ten potential indicator set. When added to our best-performing combinations, the second principal component is always statistically insignificant and adds little to the fit. This supports our single principal component specification.

We also examine the ability of our alternative indicators to fit reported GDP over the period since 2008. Most, but not all, of our indicators are closely related to reported GDP.

Finally, we compare in and out-of-sample performances of our alternative indicators and GDP. We truncate our sample in 2012Q4 and estimate performances of all permutations of alternative indicators between 2008Q1 and 20012Q4 and identify our best performing indicators. We then examine the performances of our indicators in predicting out-of-sample with and without GDP included over the final two years 2013Q1–2014Q4. We find that the inclusion of GDP worsens the performances of our best-performing alternative indicators.

Our emerging picture seems to be one where reported GDP is a more accurate depiction of Chinese output than it used to be. Nevertheless, the most reliable measure of activity is based on our best-performing alternative indicators alone.

The remainder of this paper is divided into six sections: Section 2 describes our data, including our independently verified, trade-based measure of China's imports, and discusses our methodology for identifying indicators. Section 3 presents our Bai-Perron evidence concerning structural breaks in the time series. Section 4 shows our main results for the most recent period. Section 5 conducts a number of robustness tests, including the examination of relative out-of-sample performances. Section 6 concludes.

# 2 Data and methodology

We use Chinese imports as reported by major trading partners as an independent (non-Chinese-source) indicator of Chinese economic activity. Specifically, we focus on exports to China or Hong Kong from the United States, the Euro area, and Japan. As the appendix shows, imports are a reliable indicator of economic activity for many countries. For example,

for the United States, the correlation of imports with GDP is about 0.8. Data are quarterly, and measured as year-over-year changes from the 2000:Q1 to 2014:Q4.

We obtained these data from original sources in these trading partner nations. We include exports to Hong Kong as well as to China, since many of the goods passing through that port are primarily destined for the Chinese mainland. Statistical authorities in, say, the United States have plausibly changed the degree to which they are able to track the ultimate destination over time—that is, a good that previously would have been recorded as an export to Hong Kong might now be recorded as an export to China. Using the combination of Hong Kong and China makes the data more comparable over time.[3] We convert all data to nominal U.S. dollars using market exchange rates and then convert to real values using a China-specific U.S. export deflator, as discussed in the appendix.[4]

These data on China's imports are not controlled in any manner by Chinese authorities. (Henceforth, when we refer to imports, it's always as reported by trading partners.) Trading-partner governments have no incentive to misrepresent their trade volumes with China. Of course, the rapid growth of trade with China could still cause some measurement challenges for these countries. However, these data still have the advantage of being measured at foreign ports. Moreover, while Chinese trade is growing as a share of total trade for these countries, overall trade is not growing nearly so fast. So tracking trade volumes, including those destined for or originating from China, is less challenging.

From Chinese-source data, we also identified 10 alternative (non-GDP) indicators on the basis of data availability. The 10 indicators were all available from the beginning of our sample (the fourth quarter of 2000), and were downloaded from CEIC Asia. To avoid the challenges of seasonal adjustment, we again look at all data in year-over-year terms. Many of the series are available monthly, but we convert all data to quarterly terms. Doing so facilitates comparisons with quarterly GDP data, smooths some high-frequency measurement error, and avoids problems with the timing of the Chinese New Year (which sometimes occurs in January, sometimes in February, and sometimes overlaps both). See the data appendix for further details.

How should we assess the informational content of these indicators? A misleading approach would be to simply regress China's imports on all 10 of the indicators. Because of

[3] Fernald, Edison, and Loungani (1999) argue that statistically as well as economically, it makes sense to combine Hong Kong with China.
[4] The U.S., Euro area, and Japan constitute about 35-40 percent of world exports to China, based on IMF DOTS statistics. We focus on these three partners, rather than the world, because they are likely to be more accurately reported, and also to have a less heterogeneous mix of products.

multicollinearity, few of these indicators are statistically significant when all are included, but such a regression would have a high $R^2$. Nevertheless, because of overfitting, using all 10 indicators would perform poorly out of sample relative to a more parsimonious specification.[5]

To minimize the risk of spurious fit, we instead use principal components. Doing so captures the key common information in the indicators — known as "activity factors" — in a parsimonious way. Principal components are defined by the property that all factors (or components) are orthogonal, with the first component explaining the maximum variation in the included data, the second one explaining the second most variation, and so forth.

Figure 1 shows full-sample (2000Q1–2014Q4) values of Chinese imports, GDP, and one possible activity factor. The factor is the first principal component of all 10 alternative indicators, so it is agnostic about which indicators have more informational content. All variables are in growth rates, normalized to have zero mean and unit standard deviation.

Clearly, the activity factor and imports are very highly correlated. For example, during the global financial crisis, both series drop about 3 standard deviations below their respective means. In the recovery, both series rise to above 2 standard deviations above their means. Thus, reassuringly, imports and the activity factor tell the same story about economic activity.

However, the relationship of reported GDP with either the activity factor or imports is less strong. The correlation is still positive and significant, but GDP rises more prior to the crisis than either imports or the activity factor, and falls less during the crisis.

The activity factor in Figure 1 is agnostic about which indicators have information about true economic activity and whether that information has changed over time. A key goal of the sections that follow is to identify which indicators (including GDP) are particularly informative. In this regard, note that the addition of an irrelevant data series, which is idiosyncratic in terms of China's imports, can reduce the explanatory power of the first principal component. The reason is that the first principal component will try to explain that idiosyncratic variation as well as the systematic variation that matters for imports.

[5] For example, we regressed the import data on all 10 indicators from the start of our sample until end-2012 and predicted out-of sample thereafter. For comparison, we also regressed the data on the first principal component of these indicators, as well as the first principal component of the three Li indicators. As expected, the regression with all 10 indicators individually had the lowest (best) RMSE in sample, 0.53 versus 0.70 and 0.75 for the first principal component of all 10 and the Li indicators respectively. However, the regression with all 10 indicators included had the highest (worst) RMSE out of sample: 0.70 versus 0.52 and 0.46 for the first principal component of all 10 and the Li indicators respectively. These results are available on request from the authors.

In light of these issues, we identify "best indicators" by constructing the first principal component of all possible subsets of these 10 variables, considering a total of 1023 combinations.[6] For example, 10 of the combinations have just a single indicator (each of the 10 variables); at the other extreme, one combination uses all 10 variables at the same time (our "all 10 indicators" factor plotted above). For each subset, we then regress growth in Chinese imports from the United States, the Euro area and Japan on the first principal component as well as real exchange rate values (which plausibly affect import levels independently of output).

Our baseline specification is thus

$$\Delta^4 m_t = c + \beta PC_{1t} + \gamma \Delta^4 RMB_t + \upsilon_t \,. \tag{1}$$

$\Delta^4 m_t$ is reported quarterly growth in real Chinese imports from (measured as real exports to China by) the United States, the euro area, and Japanr; $PC_{1t}$ is the contemporaneous value of the first principal component from the year-over-year growth in the chosen set of alternative indicators of Chinese economic activity; $\Delta^4 RMB_t$ is the four quarter change in the renminbi-dollar exchange rate; and $\upsilon_t$ is an error term. We estimate with ordinary least squares and show Newey-West standard errors that allow for heteroskedasticity and autocorrelation.

A concern is that we will choose a set of indicators that, by chance, work well in sample. For this reason, we also look at performances out of sample. We find that indicators that work well in sample tend to work well out of sample as well. This is not a surprise, since the method already takes an average of the informational content of a set of indicators.

[6] Other than the null set.

# 3 Full sample results

## 3.1 Preliminary regression results

We begin with full-sample results to illustrate our approach. Figure 2 shows fitted values of Chinese imports from estimating equation (1) from 2000:Q4 to 2014:Q4. We compare results with three potential explanatory variables. The first uses reported real GDP (GDP). The second uses the first principal component of all 10 non-GDP indicators (*ALL10).* The third uses the first principal component of the Li indicators *(LI).*

Visually, the two principal component indicators have similar fit. However, reported GDP fits much worse, consistent with the simple plot in Figure 1.[7] It follows that for the full sample, our principal component indicators outperform GDP in explaining Chinese imports. In particular, the *GDP* indicator initially under-predicted and then over-predicted Chinese imports during the crisis period relative to the principal component indices.

## 3.2 Structural breaks

As Figure 2 shows, there is a tighter fit for the more recent portion of our sample, particularly since 2008. Figure 3 shows this improved fit visually by estimating the same versions of equation (1) for 20-quarter rolling samples and then plotting the R-squared values. The figure confirms that there is indeed a marked increase in estimated R-squareds in 2008, after which they remain elevated for the remainder of our sample. Our regression results therefore suggest a structural break in the relationship around the time of the global financial crisis. This period also follows closely the comments by Li concerning the quality of official Chinese data.

To investigate this possibility formally, we conduct Bai and Perron (1998, 2003) tests for multiple structural breaks. This method searches for one or more breakpoints and evaluates the set of break dates that minimize the sum of squared residuals, either sequentially or in terms of re-optimizing when an additional break date is added.

Table 1 shows the break results. The regressions that identify a break find one in either 2007Q4 or 2008Q1. The regressions that identify a break find one in either 2007Q4

[7] In sample, the mean value of the root-mean-squared errors (RMSEs) for the three principal components is 8.1 while the RMSE for reported GDP is 10.6. Out of sample, the mean RMSE for the three principal component indicators is about 6.6, much lower than the value for GDP of 9.4. Out of sample, we would note that if we used the 10 indicators individually in a single regression, the RMSE is yet higher at 13.2. These preliminary estimates show the value of parsimony. All estimates are available from the authors on request.

or 2008Q1. We also observe a high incidence of statistically significant structural breaks in the third or fourth quarter of 2005.

We are interested in GDP as well as the alternative indicators. Hence, combining the formal tests with the visual evidence from Figure 3, we concentrate on the sample from 2008:Q1 through 2014Q4. We can be more confident that parameters are stable over this sample. When we later compare to earlier data, we combine subsamples despite often finding a structural break around 2005. The reason is that, with three sub-periods, the intermediate period is too short for reliable estimation.

# 4 Results for 2008:Q1–2014:Q4 sub-sample

Given the structural break results, we concentrate on the period 2008:Q1 through 2014Q4 to evaluate the quality of alternative activity indicators. Table 2 summarizes our estimation results. The estimated parameter values are not interesting per se, so we do not show them. We instead focus on (i) indicator names and sets; (ii) the statistical significance of the principal component; and (iii) fit as measured by root-mean-squared error (RMSE) and $R^2$.

The top of the table shows results for each of the 10 indicators individually, so the principal component approach is equivalent to using the indicator itself in the regression. The indicators are listed in order of their fit (highest $R^2$ /lowest RMSE). For comparison, we also include regression results using GDP as the indicator. China's real exchange, included as a control, is insignificant throughout, but consistently enters with its expected negative sign.

Eight of the ten individual indicators are statistically significant in explaining imports at a one percent confidence level. Over this period, rail freight performs best among the individual alternative indicators in explaining Chinese imports with an RMSE of 0.54. Electricity is a close second with an RMSE of 0.61.

Strikingly, over this period, the rail freight variable is the only one of the individual indicators that outperforms GDP in explaining Chinese imports. Reported GDP over this sub-period is statistically significant at the 1% level with an RMSE of 0.60. That is modestly superior to the performance of the second-place electricity variable. The other individual alternative indicators are not close. GDP also does markedly better over this turbulent sub-

sample than it did for the full sample discussed above. It is therefore clear that the explanatory power of GDP in fitting Chinese imports has increased during the latter portion of our sample.

That said, our best-performing combinations of the alternative indicators typically outperform GDP in explaining Chinese imports, sometimes substantially. The lower panel of the table shows results for our top ten indicator combinations, ranked on (in-sample) RMSE. The first principal component of the various sets enters statistically significantly in all specifications and the fit is markedly better than for any of the individual components alone. All ten of these sets of alternative indicators track Chinese imports well. Our best-fitting specification has an impressive R-squared value of 0.88. However, there is little sensitivity across these best-performing sets of indicators; even our tenth-best combination is comparable at 0.85.

The table also shows the indicators included in each of the top 10 performing combinations. The top combination turns out to include the indicators electricity, rail, raw materials production and retail sales. Sets 2 through 10 are very similar, only substituting one indicator or another. Rail appears in 9 of our top ten indicator combinations, while electricity appears in 7, and raw materials and retail sales appear in all 10. This result is confirmation that these four indicators provide valuable (and to some extent distinct) information about economic activity in China over the most recent sample.

These results already hint at the point that adding additional, lower-relevance variables do not necessarily improve explanatory power. To further illustrate this point, the set that includes all 10 indicators ranks 240$^{th}$ overall, with an RMSE of 0.48 that is only modestly better than the best individual indicators.

Finally, the table also includes the index based on the set of indicators publicized by Li Keqiang (the Li index). When it comes to explaining imports, the Li index is relatively poor. Indeed, it only marginally outperforms GDP over this period. The reason, of course, is that although it includes the relatively well-performing electricity and rail indicators, it also includes the largely irrelevant lending variable. In contrast, the Li indicator did rather well for our larger full sample above. These results suggest that the Li indicators are no longer as reliable measure of true economic activity in China.[8]

[8] The inferior performance of the Li indicator for explaining overall Chinese output may reflect that he was only trying to asset the economic performance of Liaoning Province. Its output bundle could differ systematically from the rest of China as a whole in a manner that favors the lending variable.

# 5 Robustness checks

## 5.1 Does GDP have additional explanatory power?

In Section 4, we found that activity factors based on a small number of indicators have exceptional explanatory power for imports. These factors substantially outperform GDP alone. Still, GDP is a broad measure of economic activity, whereas our top-performing indicators may disproportionately represent certain sectors of the Chinese economy. (For example, floor space added is likely to be closely correlated with construction volumes.) Hence, it seems a priori plausible that GDP may have additional explanatory power for Chinese imports. On the other hand, if Chinese GDP data are heavily manipulated or excessively noisy relative to true economic activity, then it might be dominated by the alternative activity factors.

We therefore add reported GDP to the regressions for Chinese imports, as a robustness check concerning their explanatory power. We again concentrate on the sample following the most recent structural break, i.e. 2008:Q1 through 2014:Q4.

Table 3 shows our results. The top panel shows that, with all 10 individual indicators, we find that GDP enters with a statistically significant positive coefficient.[9] Moreover, the fit markedly improves. For example, adding GDP to rail freight, our top-performing individual alternative indicator, increases the R-squared from 0.75 to 0.79.

Moreover, the addition of GDP reduces some of the explanatory power of our previously most informative alternative indicators. Retail and floor space, which were significant at a 1% confidence level on their own, are statistically insignificant with GDP added to the specification. Other indicators also exhibit a decline in significance. However, some of our other indicators gain in significance. In particular, two indicators that were insignificant on their own, air passengers and lending, are now significant at 5 and 1 percent confidence levels respectively. Hence, these indicators appear to contain information different from GDP. In contrast, retail and floor space are highly collinear with GDP based on their explanatory power in Table 2. Still, it is interesting that the regression prefers to load on GDP rather than these indicators.

Nevertheless, GDP adds much less information to our top combinations of alternative indicators in the bottom panel. GDP is statistically significant in only half of the 10 top

[9] The coefficient enters at a 1% confidence level for 8 of the 10 indicators, the exceptions being electricity and rail, with which GDP enters at 10% and 5% confidence levels respectively.

sets of indicators. With the top three sets, adding GDP does not change the fit of the regressions. In several other cases the fit improves modestly. Thus, while GDP has some additional informational content in explaining imports, even the in-sample contributions of GDP relative to our best-performing indicator combinations are quite modest.

## 5.2 Evidence from the early time period

We next consider evidence from the sample prior to our structural break. Although our Bai-Perron tests indicated more than one potential structural break date, we consolidate our data before the 2008 break to ensure that we have adequate data for estimation.

The top panel of Table 4 shows the earlier sub-sample results by individual indicator. The relative ranking of individual indicators changes somewhat. Those that did relatively well in the later sample tend to do well here, such as raw materials, rail, or electricity. But several are no longer significant, such as retail, air passengers, and property. Moreover, GDP fails to enter significantly on its own or in addition to eight of the 10 individual indicators.

The bottom panel shows that the best-performing combinations of alternative indicators perform much better than the individual indicators and continue to enter statistically significantly at a 1% confidence level. However, there are differences in the alternative indicators that are most prevalent in this group; for example, lending levels and reported exports from China to the rest of the world are more informative than they were in the later sample. GDP adds little if anything when added to our best-performing combinations, only entering with statistical significance in one of the ten best-performing combinations of alternative indicators.

Our examination of the earlier time period has two implications: First, GDP was not nearly as informative in the early portion of our sample as it was in the later sample period. GDP fails to enter significantly in most of our specifications for the pre-2008 period. This suggests that the contribution of GDP is stronger than it used to be. Second, the alternative indicators that best-predict Chinese imports changed somewhat over different samples. For example, our best-performing set of indicators from the most recent period, the combination including electricity, rail, raw materials, and retail, only placed 612$^{th}$ in the early period, with an R-squared of 0.13. We also see a poor performance by the Li indicators. This suggests caution in relying on a small set of alternative indicators to infer Chinese economic activity.

## 5.3 Adding a 2nd principal component

A priori, there need not be a single activity factor that describes China's imports. Hence, as a further robustness test, we add the second principal component from each combination of our ten potential indicator variables. Our specification with two principal components satisfies

$$\Delta^4 m_t = c + \beta PC_{1t} + \beta PC_{2t} + \gamma RMB_t + \upsilon_t \,. \qquad (2)$$

The variables are the same as in equation (1) with $PC_{2t}$ added as a second principal component.

Our results are shown in Table 5. We first take the ten best-performing combinations of alternative indicators from our single principal component exercise and add the best-performing second principal component. The second principal component does not enter significantly in a majority of our specifications. Moreover, the best combinations that include a second principal component only modestly reduce RMSE or increase R-squared relative to the single principal component results in Table 2.

However, it is possible that there may be other combinations of principal components that outperform using the initial set of ten best combinations obtained from single principal component specifications. To check if this is the case, we repeated our investigation for all possible combinations of two principal components.

Our results in the bottom panel still provide little compelling evidence in favor of the inclusion of a second principal component. The second principal component now comes in significantly more often than not. However, adding a second principal component again subtracts little from RMSE and adds little to R-squared. In other words, we achieve comparable fit through a variety of combinations in single principal component specifications. In particular, our best-performing alternative indicator set with two principal components is the same as it was in Table 1—with the same core set of indicators and nearly identical RMSE and R-squareds. Overall, then, there is little gain from adding a second principal component to our specification and for parsimony reasons we continue to concentrate on single principal component results.

## 5.4 Predicting official GDP

There may be independent interest in predicting official reported GDP figures, even if combinations of alternative indicators are the better estimators of actual Chinese economic activity. Table 6 examines how well alternative indicators explain reported GDP over the period 2008:Q1-2014:Q4. We use the same format as our base specification results from Table 2.

The top of the table shows results for the individual indicators. The relative rankings at explaining GDP are fairly similar to those explaining Chinese imports. Rail, property, and electricity still do best, although there are some modest changes in relative positions. The top seven indicators also significantly (at 5% level) explain GDP; lending is significant at a 6% confidence level. The bottom two (consumer index and air passengers) are insignificant.

We also list the top ten indicator combinations in explaining Chinese GDP. All ten combinations are significant at a 1% confidence level; all have an RMSE lower than, and an R-squared higher than, our top individual indicator rail freight. We do identify some surprising differences in the combinations that perform best in explaining GDP. Lending now plays a prominent role, entering in all of our ten top combinations. In contrast, the raw materials indicator is far less prevalent, only entering in four of our top ten specifications. Still, there are a lot of similarities, as electricity, rail freight, and retail sales all play prominent roles in fitting GDP data, as they did for our base import data specification.

## 5.5 Overall in and out of sample performances

Given the relatively stable performances of our alternative indicators in predicting imports, a question arises about how to use the indicators for prediction. We address this question by looking at which combinations of indicators work best out of sample. Given our relatively short sample available since the structural break we observe in 2008, we have a relatively small amount of data to use for parameter estimation prior to an out-of-sample period we can use for assessing predictive power.

We therefore choose our preferred indicators based on a combination of in and out-of-sample performances. The in-sample performances are estimated over longer samples. However, the in-sample rankings also favor less parsimonious specifications. We therefore also conduct out-of-sample exercises and choose our preferred indicator combination on the

basis of a combination of these performances. Our reported in-sample results use the sample 2008Q1–2012Q4, while our out-of-sample results are from the period 2013Q1–2014Q4.

We first examine the relative performances of individual indicators. We take all possible combinations of alternative indicators with five or fewer indicators and average all combinations of RMSE's among combinations of alternative indicators in which the indicator in question enters.

Our results are shown in the first part of Table 7. The results are largely in concert with our rankings of combinations. We again see rail freight, electricity, and property scoring particularly well, both in and out of sample. Reassuringly, we see much consistency across in-sample and out-of-sample performances. However, two exceptions are retail and raw materials. The former does far better in sample, whereas the latter is far better out of sample.

The bottom portion of Table 7 reports the best-performing combinations of indicators with and without GDP.[10] While there are some discrepancies between these and the best-performing combinations chosen above based solely on in-sample fit (Table 3), we again see rail freight, electricity, raw materials and retail prominently represented. Our best-fitting combination in-sample contains electricity, rail, air passengers, raw materials, and retail.

This combination enters as the fourth best out-of-sample among combinations of alternative indicators with GDP added. Adding GDP to this combination results in a modest improvement in fit within sample, reducing the RMSE from 0.37 to 0.32, but results in a substantive deterioration in out-of-sample performance: Average RMSE out of sample rises from 0.40 to 0.60. Indeed, across the board, the out-of-sample performances get substantively worse with reported GDP. These results suggest that for forecasting, one should use the overall best-performing combination of alternative indicators without adding GDP.

# 6 Conclusion

In this paper, we find that since 2008, reported Chinese GDP figures have been notably more reliable than earlier in capturing fluctuations in economic activity. We also identify activity factors based on a small set of alternative (non-GDP) indicators. Both GDP and these activity factors appear to have independent information, although the performances of our very-best-performing combinations of indicators are not improved by the addition of reported GDP.

[10] Real exchange rate changes are included, as always.

To reach these conclusions, our innovation is to use an independently verified indicator of economic activity in China: Imports as measured by trading partners. For an open economy like China, we expect imports to be closely related to economic activity, and trading-partner-reported exports should be an unbiased measure of China's imports, free from manipulation. We find that since 2008, reported GDP is significantly related to this measure of activity. However, one can still forecast best with a preferred set of the alternative indicators.

These results contrast with those we obtain for the period prior to 2008, when the performances of reported GDP was far inferior. Our results for this period support the widespread view—held by Chinese officials as well as others at that time—that there was little informational content in official Chinese GDP data. In short, our results suggest that official Chinese output data has improved and now contains information about Chinese economic activity. However, one still does best out-of-sample by using our preferred set of alternative indicators without GDP, as the inclusion of GDP deteriorates out-of-sample forecasting.

Our best-fitting set of alternative indicators used an index constructed from electricity use, rail freight volume, raw materials consumed, and retail sales. We found that principal component constructed from these indicators provide a good measure of economic activity.

We conclude with several caveats. First, imports are an imperfect measure of activity and may underweight certain activities, notably services and other non-tradable sectors. Still, imports are very highly correlated with our preferred activity factor. And that factor includes both relatively narrow indicators (like exports and, possibly, raw materials) and broader ones (such as electricity and new floor space constructed). Moreover, even if imports or the activity factor are imperfect, here is no reason to think they are necessarily worse than GDP alone.

Second, even for the pre-2008 period—when GDP is a poor measure of economic activity in China—we cannot say for sure whether GDP was manipulated, or merely limited in its coverage. If manipulation was rampant, we would expect it to be more prevalent during periods of exceptionally high or low economic activity, as data might be changed to more closely meet trend output goals. There appears to be some evidence of that here during the global financial crisis, but we cannot say whether the level and variability in GDP are accurate.

Finally, as China's economy and statistical system continue to evolve, indicators that do well historically might do less well going forward. For example, that rail and property do better after 2008Q1 than during the 2000Q1–2007Q4 period could reflect either changes in the composition of activity, changes in the quality of activity, or could be chance. Nevertheless, it is reassuring that our core set of indicators performs well across our two sample periods.

# References

Bai, J., Perron, P., (1998), "Estimating and testing linear models with multiple structural changes," Econometrica 66, 47–78.

Bai, J., Perron, P., (2003), "Computation and analysis of multiple structural change models," Journal of Applied Econometrics 18, 1–22.

Bradsher, Keith (2012). "China Data Mask Depth of Slowdown, Executives Say," New York Times, June 23. http://www.nytimes.com/2012/06/23/business/global/chinese-data-said-to-be-manipulated-understating-its-slowdown.html?pagewanted=all

Fernald, John, Israel Malkin, and Mark M. Spiegel, (2013), "On the Reliability of Chinese Output Figures," FRBSF Economic Letter, 2013–08, March 25.

Fernald, John & Edison, Hali, and Loungani, Prakash, (1999), "Was China the first domino? Assessing links between China and other Asian economies," Journal of International Money and Finance, Elsevier, vol. 18(4), pages 515–535, August.

Henderson, J. Vernon, Adam Storeygard, and David N. Weil, (2012), "Measuring Economic Growth from Outer Space," American Economic Review, 102(2), 994–1028.

Holz, Carsten A. (2008). "China's 2004 Economic Census and 2006 Benchmark Revision of GDP Statistics: More Questions than Answers?" The China Quarterly, March.

Holz, Carsten A. (2013). "Chinese Statistics: Output Data"

Holz, Carsten A. (2003). "'Fast, Clear and Accurate:' How Reliable Are Chinese Output and Economic Growth Statistics?" The China Quarterly, no. 173 (March 2003): 122–63.

Nakamura, Emi, Jon Steinsson, and Miao Liu, (2014), "Are Chinese Growth and Inflation Too Smooth?: Evidence from Engel Curves," mimeo, Columbia University, January.

Noble, Josh (2015). "Doubts rise over China's official GDP growth rate." *Financial Times*, September 16, 2015. http://www.ft.com/intl/cms/s/0/723a8d8e-5c53-11e5-9846-de406ccb37f2.html#axzz3lvULIfcE (accessed September 21, 2015).

Sharma, Ruchir (2013). "China's Illusory Growth Numbers." Wall Street Journal, October 31, 2013.

Sinclair, Tara (2012). "Characteristics and Implications of Chinese Macroeconomic Data Revisions." Manuscript, George Washington University. http://www.gwu.edu/~iiep/assets/docs/papers/Sinclair_IIEPWP2012-09.pdf

Wikileaks (2007). http://wikileaks.org/cable/2007/03/07BEIJING1760.html.

# Appendix Data sources

The chart below shows the raw data we used in the paper. All data were accessed in April 2015, mainly from CEIC Asia database.

| **Series** | **Description** | **Source** |
|---|---|---|
| Electricity | Electricity production, billions of kilowatt hours | National Bureau of Statistics (CEIC series 3662501) |
| Rail | Railway freight traffic, millions of tons | China Railway Corporation, National Railway Administration, (CEIC series 12915101) |
| Lending | Bank loans, billions of RMB | The People's Bank of China, (CEIC series 7029101) |
| Property | Real estate investment (Residential bldgs.), millions of RMB | National Bureau of Statistics, (CEIC series 3948701) |
| Air passengers | Air passenger traffic, millions of persons | Civil Aviation Administration of China (CEIC series 12916401) |
| Exports | Exports (FOB basis), millions of US dollars | General Administration of Customs (CEIC series 5823501) |
| Consumer Index | Consumer Expectation Index | National Bureau of Statistics (CEIC series 5198601) |
| Floor space | Floor space started, thousands of square meters | National Bureau of Statistics (CEIC series 3963901) |
| Raw materials | Index of raw materials supply, derived from a survey of managers from 5000 companies. Respondents are asked for views on adequacy of supplies of raw materials. | The People's Bank of China, (CEIC series 8003501) |
| Retail | Retail sales of consumer goods, billions of RMB | National Bureau of Statistics (CEIC series 5190001) |
| GDP | Real GDP index, available as 4-quarter growth rates | National Bureau of Statistics (CEIC series 1692001) |
| Exchange rates between Yen, Euro, USD, and RMB | | Bloomberg |
| Imports and exports between Japan and other countries | Thousands of Yen | Ministry of Finance (http://www.customs.go.jp/toukei/info/index_e.htm) |
| Imports and exports between European Union and other countries | Euros | Eurostat ("EU27 trade since 1988 by CN8" database) |
| Imports and exports between United States and other countries | | Census Bureau via Haver Analytics |

## Trade data (monthly & quarterly)

Trade Flows (millions of USD)

US_exp_China: US exports to China
US_imp_China: US imports from China
US_exp_HK: US exports to Hong Kong
US_imp_HK: US imports from Hong Kong
US_exp: Total US exports to China & Hong Kong
US_imp: Total US imports from China & Hong Kong
EU_exp_China: EU exports to China
EU_imp_China: EU imports from China
EU_exp_HK: EUexports to Hong Kong
EU_imp_HK: EU imports from Hong Kong
EU_exp: Total EU exports to China & Hong Kong
EU_imp: Total EU imports from China & Hong Kong
Japan_exp_China: Japan exports to China
Japan_imp_China: Japan imports from China
Japan_exp_HK: Japan exports to Hong Kong
Japan_imp_HK: Japan imports from Hong Kong
Japan_exp: Total Japan exports to China & Hong Kong
Japan_imp: Total Japan imports from China & Hong Kong
Trio_exp_China: Trio exports to China
Trio_imp_China: Trio imports from China
Trio_exp_HK: Trio exports to Hong Kong
Trio_imp_HK: Trio imports from Hong Kong
Trio_exp: Total Trio exports to China & Hong Kong
Trio_imp: Total Trio imports from China & Hong Kong

Note: Trio = US + EU + Japan

World_exp_China: World exports to China
World_imp_China: World imports from China
World_exp_HK: Worldexports to Hong Kong
World_imp_HK: World imports from Hong Kong
World_exp: Total World exports to China & Hong Kong
World_imp: Total World imports from China & Hong Kong

Sources

US trade data: Census Bureau via Haver
EU trade data: Eurostat ("EU27 trade since 1988 by CN8" database)
Japan trade data: Ministry of Finance (http://www.customs.go.jp/toukei/info/index_e.htm)
World trade data: Direction of Trade Statistics (IMF CD's)

## Price index for exports to China (monthly & quarterly)

We deflate trading-partner exports to China using a Chinese-specific deflator for U.S. exports. This deflator weights growth in overall U.S. agriculture and non-agriculture deflators by corresponding shares of U.S. exports to China and Hong Kong. Fernald, Malkiel, and Spiegel (2013) found that this “simple” deflator corresponds closely to a more sophisticated deflator for exports to China that uses detailed commodity-by-country data. (Those detailed data are available only after 2005:12.) The data sources are U.S. Census data on "Trade in Goods by NAICS-Commodity By Country," and "Export Price Indexes by NAICS." Data were accessed via Haver Analytics August 3, 2015.

*** The prefix "d12_" represents the 12-month percent change in that series
*** The prefix "d4_" represents the 4-quarter percent change in that series

## Adjustments made

Monthly proxy series converted to quarterly via summing over the quarter.

Missing observations around Chinese New Year:

- ElectricityConsumption missing January, filled in with half of February cumulative value
- Retail missing January and February, 1st quarter percent change filled in with March-to-March 4-quarter change
- Rail shipments: Series has a break in level in January 2004. We adjusted the series by splicing. (Done in code d02_input_proxy_data.do).
- Li: We use the adjusted rail data rather than the raw data.

## Trade

Data on China and Hong Kong summed and used for Chinese import/export numbers All trade series converted to USD using bilateral exchange rate data (Bloomberg, “USDEUR Curncy” and “USDJPY Curncy” series). Thus:

- EU trade in millions USD = (EU trade in Euros)/(1000000*(Euro-USD exchange rate)
- Japanese trade in millions USD = (Japanese trade in thousands Yen)/(1000*Yen-USD exchange rate)

# Appendix Are imports a good measure of economic activity?

Table A1 shows that China's raw correlation between imports and GDP is relatively low at only 17 percent. In contrast, the United States – a more closed economy, but one with a more reliable statistical system – has a much higher $R^2$ at 77 percent.

## Table A1 Relating Imports and Real GDP

Table 1: Predicting RGDP growth with import growth

| country | rsquared | GDP2000 | GDP2012 | exports2000 | exports2012 | imports2000 | imports2012 |
|---|---|---|---|---|---|---|---|
| United States | .77 | 45956 | 50865.93 | 10.63 | 13.52 | 14.33 | 16.89 |
| Malaysia | .68 | 15688 | 21897.32 | 119.81 | 87.14 | 100.6 | 75.28 |
| Singapore | .56 | 51491.19 | 74609.19 | 189.18 | 195.08 | 176.89 | 172.65 |
| Thailand | .51 | 8938.88 | 13736.22 | 66.78 | 74.98 | 58.14 | 73.85 |
| Korea, Rep. | .47 | 20756.78 | 31901.07 | 35.01 | 56.34 | 32.94 | 53.55 |
| Philippines | .38 | 4242.55 | 6004.78 | 51.37 | 30.8 | 53.36 | 33.99 |
| Indonesia | .21 | 5551.96 | 8855.01 | 40.98 | 24.29 | 30.46 | 25.86 |
| China | .17 | 3609.32 | 10756.46 | 23.33 | 27.32 | 20.92 | 24.5 |
| India | .14 | 2599.76 | 5050.11 | 12.77 | 24 | 13.66 | 30.74 |

Source: World Bank and author's calculations. Note that we use GDP per capita, PPP (constant 2011 international $). Imports and exports are reported as % GDP.

China is relatively low in both imports and exports as a share of GDP. Other countries at similar development and openness levels to China (e.g. Indonesia and India) exhibit roughly similar correlation figures. This raises the possibility that low levels of development, which are perhaps associated with output measurement errors rather than systematic data manipulation, may explain observed discrepancies for China.

Table 1 Structural break tests

| | Breaks (sequential) | Break dates (sequential) | Breaks (BIC) | Break dates (BIC) |
|---|---|---|---|---|
| GDP on Li | 2 | 2005q4<br>2008q1 | 4 | 2002q4<br>2005q4<br>2008q1<br>2011q4 |
| GDP on all components | 3 | 2005q3<br>2007q4<br>2010q1 | 3 | 2005q3<br>2007q4<br>2010q1 |
| Trio exports on GDP | 4 | 2003q2<br>2005q4<br>2008q1<br>2011q4 | 4 | 2003q2<br>2005q4<br>2008q1<br>2011q4 |
| Trio exports on Li | 0 | | 0 | |
| Trio exports on all components | 0 | | 0 | |

Note: Results for Bai-Perron structural break tests, with identified number of breaks and corresponding quarters found to be statistically significant at a 5% confidence level. Sequential method searches for a break date, then searches for a second, taking the date of the first as given. BIC method "reoptimizes" by searching for one, and then searching for two (potentially with neither identical to the date chosen for a single break date), etc.

Table 2 Explaining imports with principal components, 2008Q1–2014Q4

| Individual variables | Significance | RMSE | R-sq. |
|---|---|---|---|
| Retail | 0.00 | 0.81 | 0.44 |
| RawMat | 0.00 | 0.77 | 0.49 |
| FloorSp | 0.00 | 0.74 | 0.54 |
| Consumer | 0.01 | 0.91 | 0.29 |
| ChinaExp | 0.00 | 0.81 | 0.44 |
| AirPass | 0.66 | 1.07 | 0.02 |
| Property | 0.00 | 0.70 | 0.58 |
| Lending | 0.48 | 1.06 | 0.04 |
| Rail | 0.00 | 0.54 | 0.75 |
| Electricity | 0.00 | 0.62 | 0.68 |
| GDP | 0.00 | 0.61 | 0.69 |
| Combinations | | | |
| Li | 0.00 | 0.60 | 0.70 |
| All indicators | 0.00 | 0.48 | 0.80 |
| Best 10 combinations | | | |
| Electricity Rail RawMat Retail | 0.00 | 0.38 | 0.88 |
| Electricity Rail AirPass RawMat Retail | 0.00 | 0.38 | 0.88 |
| Electricity Rail Lending AirPass RawMat Retail | 0.00 | 0.38 | 0.88 |
| Electricity Rail Lending RawMat Retail | 0.00 | 0.40 | 0.86 |
| Electricity Rail Property AirPass RawMat Retail | 0.00 | 0.41 | 0.86 |
| Electricity Rail Property RawMat Retail | 0.00 | 0.41 | 0.86 |
| Rail Property AirPass RawMat Retail | 0.00 | 0.42 | 0.85 |
| Rail Property RawMat Retail | 0.00 | 0.42 | 0.85 |
| Electricity RawMat Retail | 0.00 | 0.42 | 0.85 |
| Rail RawMat Retail | 0.00 | 0.42 | 0.85 |

Notes: Reported p-values use Newey-West standard errors. All regressions include the real exchange rate as described in the text. The $R^2$ values are adjusted for degrees of freedom.

Table 3 Explaining imports with principal components and GDP, 2008–2013

| Individual variables | Tuple sig. | GDP sig. | RMSE | R-sq. |
|---|---|---|---|---|
| Retail | 0.21 | 0.00 | 0.60 | 0.70 |
| RawMat | 0.00 | 0.00 | 0.47 | 0.82 |
| FloorSp | 0.16 | 0.00 | 0.60 | 0.70 |
| Consumer | 0.04 | 0.00 | 0.53 | 0.77 |
| ChinaExp | 0.00 | 0.00 | 0.50 | 0.79 |
| AirPass | 0.05 | 0.00 | 0.57 | 0.74 |
| Property | 0.08 | 0.00 | 0.58 | 0.72 |
| Lending | 0.00 | 0.00 | 0.48 | 0.81 |
| Rail | 0.00 | 0.05 | 0.51 | 0.79 |
| Electricity | 0.06 | 0.07 | 0.56 | 0.74 |
| GDP | 0.00 | | 0.62 | 0.69 |
| Combinations | | | | |
| Li | 0.07 | 0.15 | 0.58 | 0.72 |
| All indicators | 0.00 | 0.06 | 0.46 | 0.83 |
| Best 10 combinations | | | | |
| Electricity Rail AirPass RawMat Retail | 0.00 | 0.45 | 0.38 | 0.88 |
| Electricity Rail RawMat Retail | 0.00 | 0.94 | 0.38 | 0.88 |
| Electricity Rail Lending AirPass RawMat Retail | 0.00 | 0.83 | 0.39 | 0.88 |
| Rail RawMat | 0.00 | 0.00 | 0.39 | 0.88 |
| Electricity Rail Lending RawMat Retail | 0.00 | 0.64 | 0.40 | 0.87 |
| Rail Lending Property AirPass RawMat Retail | 0.00 | 0.06 | 0.40 | 0.87 |
| Electricity Lending RawMat Retail | 0.00 | 0.02 | 0.41 | 0.86 |
| Electricity Rail Lending ChinaExp RawMat Retail | 0.00 | 0.03 | 0.41 | 0.86 |
| Rail RawMat Retail | 0.00 | 0.17 | 0.41 | 0.86 |
| Electricity Rail ChinaExp RawMat Retail | 0.00 | 0.06 | 0.41 | 0.86 |

Notes: Reported p-values use Newey-West standard errors. All regressions include the real exchange rate as described in the text. The $R^2$ values are adjusted for degrees of freedom.

Table 4 Explaining imports with principal components and GDP, 2000–2007

| Individual variables | Tuple sig. | GDP sig. | RMSE | R-sq. |
|---|---|---|---|---|
| Retail | 0.15 | 0.04 | 0.71 | 0.21 |
| RawMat | 0.00 | 0.46 | 0.61 | 0.42 |
| FloorSp | 0.29 | 0.17 | 0.73 | 0.17 |
| Consumer | 0.06 | 0.23 | 0.73 | 0.17 |
| ChinaExp | 0.00 | 0.70 | 0.65 | 0.34 |
| AirPass | 0.91 | 0.29 | 0.75 | 0.13 |
| Property | 0.46 | 0.25 | 0.74 | 0.15 |
| Lending | 0.10 | 0.54 | 0.67 | 0.31 |
| Rail | 0.00 | 0.05 | 0.56 | 0.51 |
| Electricity | 0.04 | 0.74 | 0.68 | 0.29 |
| GDP | 0.23 | | 0.75 | 0.13 |
| Combinations | | | | |
| Li | 0.64 | 0.37 | 0.75 | 0.13 |
| All indicators | 0.80 | 0.27 | 0.75 | 0.13 |
| Best 10 combinations | | | | |
| Electricity ChinaExp FloorSp RawMat | 0.00 | 0.57 | 0.52 | 0.58 |
| Electricity AirPass ChinaExp FloorSp RawMat | 0.00 | 0.57 | 0.52 | 0.58 |
| Electricity Lending ChinaExp FloorSp RawMat | 0.00 | 0.57 | 0.52 | 0.57 |
| ChinaExp FloorSp RawMat | 0.00 | 0.73 | 0.53 | 0.57 |
| ChinaExp FloorSp | 0.00 | 0.48 | 0.54 | 0.54 |
| Electricity FloorSp RawMat | 0.00 | 0.81 | 0.55 | 0.54 |
| Rail | 0.00 | 0.05 | 0.56 | 0.51 |
| Electricity ChinaExp FloorSp | 0.00 | 0.55 | 0.57 | 0.50 |
| Electricity AirPass ChinaExp Consumer FloorSp RawMat | 0.00 | 0.51 | 0.57 | 0.49 |
| Electricity Lending ChinaExp FloorSp | 0.00 | 0.77 | 0.57 | 0.49 |

Notes: Reported p-values use Newey-West standard errors. All regressions include the real exchange rate as described in the text. The $R^2$ values are adjusted for degrees of freedom.

Table 5 Explaining imports with 2 principal components and GDP, 2008–2014

| Individual variables | PC1 sig. | PC2 sig. | RMSE | R-sq. |
|---|---|---|---|---|
| Combinations | | | | |
| Li | 0.00 | 0.00 | 0.52 | 0.78 |
| All indicators | 0.00 | 0.67 | 0.49 | 0.81 |
| Best 10 combinations | | | | |
| Electricity Rail Lending RawMat Retail | 0.00 | 0.01 | 0.37 | 0.88 |
| Electricity Rail RawMat Retail | 0.00 | 0.17 | 0.38 | 0.88 |
| Electricity Rail AirPass RawMat Retail | 0.00 | 0.40 | 0.38 | 0.88 |
| Electricity Rail Lending AirPass RawMat Retail | 0.00 | 0.55 | 0.39 | 0.88 |
| Electricity Rail FloorSp RawMat Retail | 0.00 | 0.00 | 0.40 | 0.87 |
| Electricity Rail Lending FloorSp RawMat Retail | 0.00 | 0.00 | 0.40 | 0.87 |
| Electricity Rail Property RawMat Retail | 0.00 | 0.07 | 0.40 | 0.87 |
| Electricity Rail Property AirPass RawMat Retail | 0.00 | 0.02 | 0.40 | 0.87 |
| Electricity Rail Lending Property AirPass RawMat Retail | 0.00 | 0.00 | 0.41 | 0.86 |
| Electricity Rail Lending Property RawMat Retail | 0.00 | 0.03 | 0.41 | 0.86 |
| Best 10 combinations (without pc2) | | | | |
| Electricity Rail RawMat Retail | 0.00 | 0.17 | 0.38 | 0.88 |
| Electricity Rail AirPass RawMat Retail | 0.00 | 0.40 | 0.38 | 0.88 |
| Electricity Rail Lending AirPass RawMat Retail | 0.00 | 0.55 | 0.39 | 0.88 |
| Electricity Rail Lending RawMat Retail | 0.00 | 0.01 | 0.37 | 0.88 |
| Electricity Rail Property AirPass RawMat Retail | 0.00 | 0.02 | 0.40 | 0.87 |
| Electricity Rail Property RawMat Retail | 0.00 | 0.07 | 0.40 | 0.87 |
| Rail Property AirPass RawMat Retail | 0.00 | 0.50 | 0.42 | 0.85 |
| Rail Property RawMat Retail | 0.00 | 0.25 | 0.42 | 0.86 |
| Electricity RawMat Retail | 0.00 | 0.09 | 0.41 | 0.86 |
| Rail RawMat Retail | 0.00 | 0.68 | 0.43 | 0.85 |

Notes: Reported p-values use Newey-West standard errors. All regressions include the real exchange rate as described in the text. The $R^2$ values are adjusted for degrees of freedom. “Best 10 combinations” are the best 10 combinations in specifications that include second principal component while “Best 10 combinations (without pc2)” are the best 10 combinations that exclude the second principal component, i.e. the same 10 as those in our base specification.

Table 6 Predicting GDP from principal components 2008Q1–2014Q4

| Individual variables | Significance | RMSE | R-sq. |
|---|---|---|---|
| Retail | 0.00 | 0.53 | 0.48 |
| RawMat | 0.01 | 0.64 | 0.25 |
| FloorSp | 0.00 | 0.49 | 0.56 |
| Consumer | 0.15 | 0.71 | 0.06 |
| ChinaExp | 0.02 | 0.64 | 0.23 |
| AirPass | 0.66 | 0.73 | 0.02 |
| Property | 0.00 | 0.44 | 0.64 |
| Lending | 0.06 | 0.66 | 0.20 |
| Rail | 0.00 | 0.40 | 0.71 |
| Electricity | 0.00 | 0.44 | 0.63 |
| GDP | 0.00 | 0.00 | 1.00 |
| Combinations | | | |
| Li | 0.00 | 0.34 | 0.78 |
| All indicators | 0.00 | 0.40 | 0.70 |
| Best 10 combinations | | | |
| Electricity Lending Property Retail | 0.00 | 0.25 | 0.88 |
| Electricity Rail Lending Property Retail | 0.00 | 0.26 | 0.87 |
| Electricity Rail Lending Retail | 0.00 | 0.27 | 0.87 |
| Electricity Rail Lending Property AirPass Retail | 0.00 | 0.27 | 0.87 |
| Electricity Lending FloorSp RawMat Retail | 0.00 | 0.28 | 0.85 |
| Rail Lending Property AirPass Retail | 0.00 | 0.29 | 0.85 |
| Electricity Rail Lending Property AirPass FloorSp RawMat Retail | 0.00 | 0.29 | 0.85 |
| Electricity Rail Lending FloorSp RawMat Retail | 0.00 | 0.29 | 0.85 |
| Rail Lending Property AirPass FloorSp RawMat Retail | 0.00 | 0.29 | 0.85 |
| Electricity Lending Retail | 0.00 | 0.29 | 0.84 |

Notes: Reported p-values use Newey-West standard errors. All regressions include the real exchange rate as described in the text. The $R^2$ values are adjusted for degrees of freedom.

## Table 7 Overall individual and combination indicator performances

### 7a Individual indicators

| Individual indicators | Avg. IS RMSE | Avg. OS RMSE | Avg. |
|---|---|---|---|
| Electricity | 0.55 | 0.72 | 0.63 |
| FloorSp | 0.56 | 0.80 | 0.68 |
| Rail | 0.51 | 0.76 | 0.64 |
| Property | 0.54 | 0.76 | 0.65 |
| Retail | 0.54 | 0.80 | 0.67 |
| RawMat | 0.60 | 0.77 | 0.69 |
| AirPass | 0.64 | 0.86 | 0.75 |
| Lending | 0.65 | 0.88 | 0.77 |
| Consumer | 0.61 | 0.77 | 0.69 |
| ChinaExp | 0.58 | 0.78 | 0.68 |

Notes: or each combination of alternative indicators (with five or fewer indicators), we calculate the in-sample and out-of-sample RMSE from a regressing real imports on the first principal component of that indicator (and the real exchange rate). For each indicator shown, we average the RMSEs in sample (IS) and out of sample (OS) for all combinations including that indicator. See text for more details.

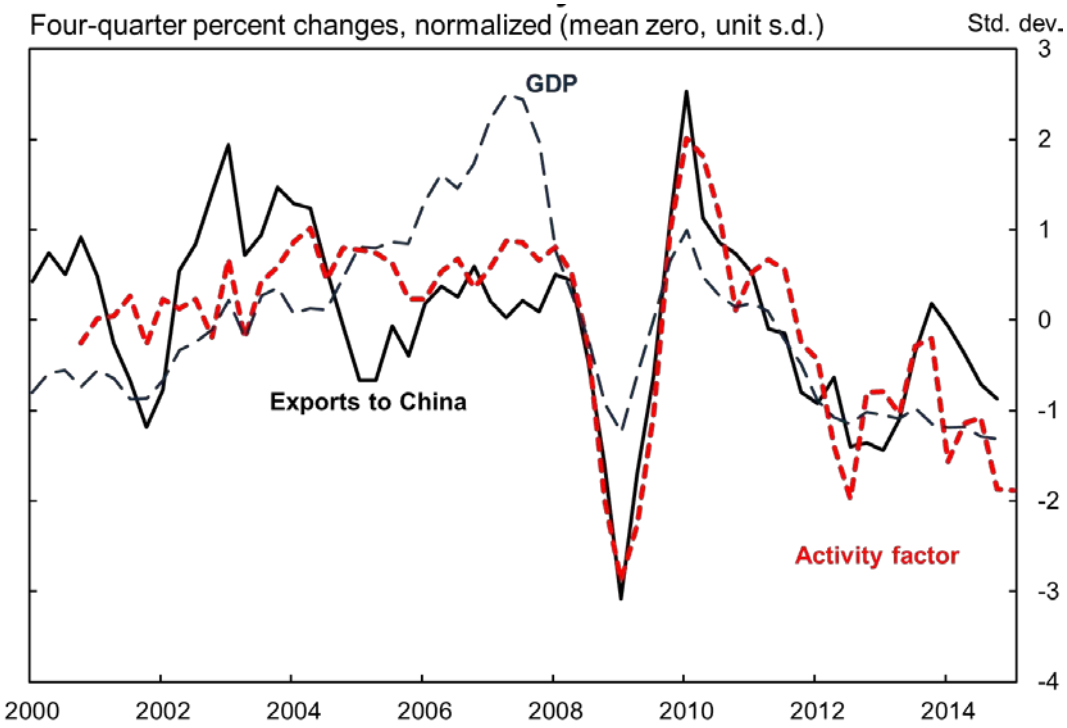
### 7b Combinations

| Indicator combinations without GDP | IS RMSE | OS RMSE | Avg. RMSE |
|---|---|---|---|
| Electricity Rail AirPass RawMat Retail | 0.38 | 0.43 | 0.40 |
| Electricity Rail Lending AirPass RawMat Retail | 0.36 | 0.49 | 0.43 |
| Electricity Rail RawMat Retail | 0.36 | 0.50 | 0.43 |
| Electricity Rail AirPass RawMat | 0.50 | 0.37 | 0.43 |
| Electricity Rail RawMat | 0.49 | 0.38 | 0.43 |
| Electricity RawMat Retail | 0.42 | 0.46 | 0.44 |
| Rail RawMat | 0.46 | 0.44 | 0.45 |
| Electricity Rail Lending Consumer RawMat Retail | 0.46 | 0.46 | 0.46 |
| Electricity Rail Consumer RawMat Retail | 0.46 | 0.46 | 0.46 |
| Electricity Rail ChinaExp RawMat Retail | 0.41 | 0.51 | 0.46 |

| Indicator combinations with GDP | IS RMSE | OS RMSE | Avg. RMSE |
|---|---|---|---|
| Rail RawMat | 0.36 | 0.61 | 0.49 |
| Rail Lending RawMat | 0.36 | 0.64 | 0.50 |
| Electricity Rail AirPass RawMat Retail | 0.34 | 0.69 | 0.51 |
| Electricity Lending RawMat Retail | 0.34 | 0.70 | 0.52 |
| Electricity Rail RawMat Retail | 0.34 | 0.71 | 0.52 |
| Electricity Rail Lending AirPass RawMat Retail | 0.34 | 0.71 | 0.53 |
| Rail Property AirPass RawMat | 0.38 | 0.69 | 0.53 |
| Electricity RawMat Retail | 0.35 | 0.72 | 0.54 |
| Electricity Rail Lending ChinaExp RawMat | 0.37 | 0.70 | 0.54 |
| Electricity Lending RawMat | 0.40 | 0.68 | 0.54 |

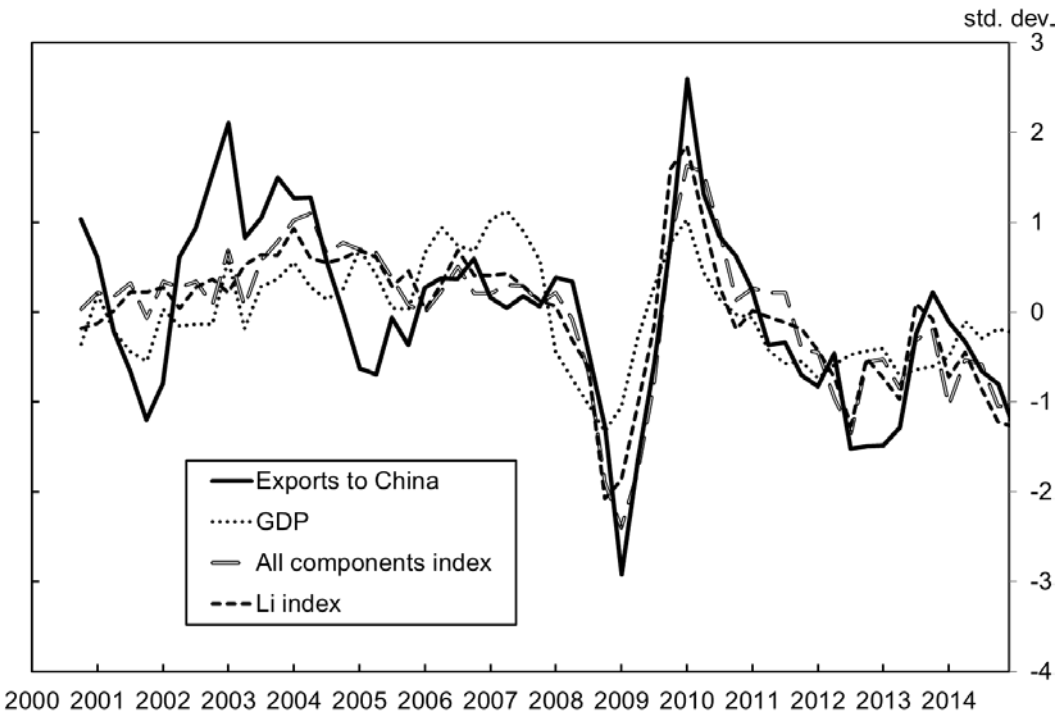
Note: Rankings of performances of individual and combinations of indicators in and out of sample. RMSE's are averages of all possible combinations with indicator included. In and out of sample rankings are based on these RMSE', with "overall" performance based on average of in and out of sample RMSE averages.

Figure 1 Indicators of economic activity in China



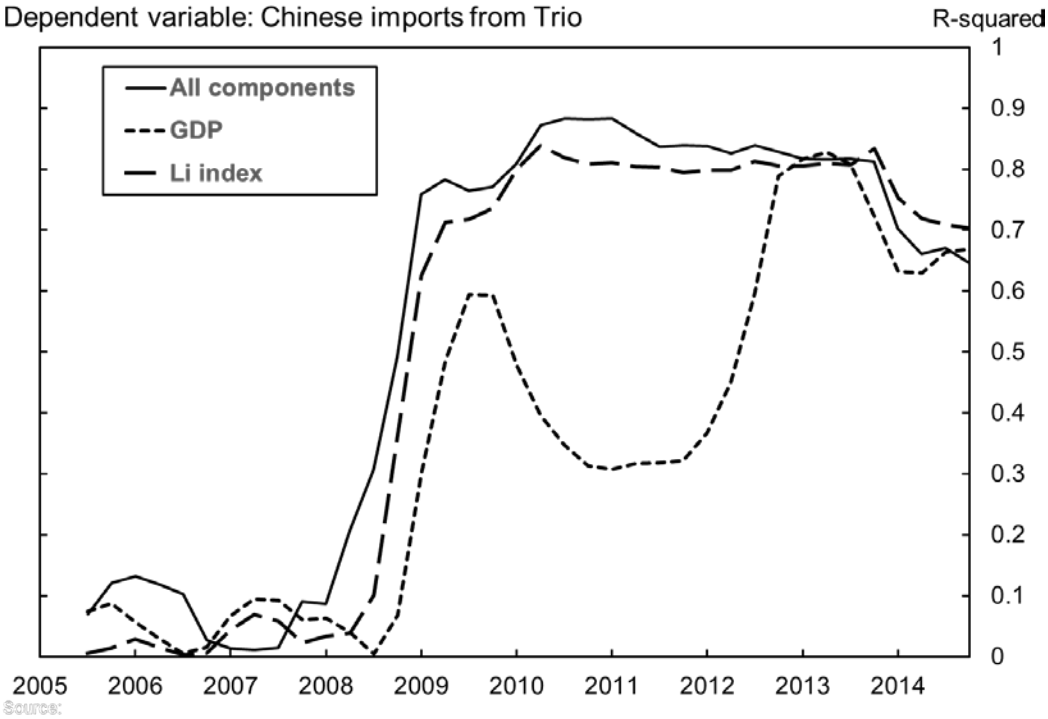

Note: Normalized GDP and "TRIO" exports to China 2000Q4–2014Q4. "All 10 indicators" series is normalized first principal component of all 10 activity indicators. See text for details.

Figure 2 Predicting Chinese imports from Trio




Note: Raw series of "TRIO" exports to China, with fitted series of GDP, all 10 activity indicators, and Li indicators. See text for details.

Figure 3 Quality of fit in rolling regressions




Note: R-squared values for rolling samples over 20-quarters, regressing Chinese imports on first principal components of all 10 indicators, GDP, and Li indicators, respectively. Chinese real exchange rate and constant term included in all specifications.

# BOFIT Discussion Papers

A series devoted to academic studies by BOFIT economists and guest researchers. The focus is on works relevant for economic policy and economic developments in transition / emerging economies.

2014

No 1 Vikas Kakkar and Isabel Yan: Determinants of real exchange rates: An empirical investigation
No 2 Iftekhar Hasan, Krzysztof Jackowicz, Oskar Kowalewski and Łukasz Kozłowski: Politically connected firms in Poland and their access to bank financing
No 3 Carsten A. Holz and Aaron Mehrotra: Wage and price dynamics in a large emerging economy: The case of China
No 4 Zuzana Fungáčová, Anna Kochanova and Laurent Weill: Does money buy credit? Firm-level evidence on bribery and bank debt
No 5 Jitka Poměnková, Jarko Fidrmuc and Iikka Korhonen: China and the World economy: Wavelet spectrum analysis of business cycles
No 6 Christopher A. Hartwell: The impact of institutional volatility on financial volatility in transition economies: a GARCH family approach
No 7 Christian Dreger, Tongsan Wang and Yanqun Zhang: Understanding Chinese consumption: The impact of hukou
No 8 John Bonin, Iftekhar Hasan and Paul Wachtel: Banking in transition countries
No 9 Chun-Yu Ho: Switching cost and deposit demand in China
No 10 Zuzana Fungáčová and Laurent Weill: Understanding financial inclusion in China
No 11 Anna Krupkina, Elena Deryugina and Alexey Ponomarenko: Estimating sustainable output growth in emerging market economies
No 12 Qing He, Chang Xue and Chenqi Zhu: Financial Development and patterns of industrial specialization: Regional evidence from China
No 13 Carsten A. Holz: Wage determination in China during the reform period
No 14 Thorsten Beck, Hans Degryse, Ralph De Haas and Neeltje van Horen: When arm's length is too far. Relationship banking over the business cycle
No 15 Boris Blagov and Michael Funke: The credibility of Hong Kong's currency board system: Looking through the prism of MS-VAR models with time-varying transition probabilities
No 16 Philip Molyneux, Hong Liu and Chunxia Jiang: Bank capital, adjustment and ownership: Evidence from China
No 17 Yin-Wong Cheung and Dagfinn Rime: The offshore renminbi exchange rate: Microstructure and links to the onshore market
No 18 Marko Melolinna: What is the role of Emerging Asia in global oil prices?
No 19 Yiwei Fang, Iftekhar Hasan and Lingxiang Li: Banking reform, risk-taking, and earnings quality – Evidence from transition countries
No 20 Yanrui Wu: Local government debt and economic growth in China
No 21 Christophe J. Godlewski, Rima Turk-Ariss and Laurent Weill: Do the type of sukuk and choice of shari'a scholar matter?
No 22 Elena Deryugina and Alexey Ponomarenko: A large Bayesian vector autoregression model for Russia
No 23 Yin-Wong Cheung, Menzie D. Chinn and Xingwang Qian: The structural behavior of China-US trade flows
No 24 Claudio Cozza, Roberta Rabellotti and Marco Sanfilippo: The impact of outward FDI on the performance of Chinese multinationals

2015

No 1 Qing He, Liping Lu and Steven Ongena: Who gains from credit granted between firms? Evidence from inter-corporate loan announcements Made in China
No 2 Ke Pang and Pierre L. Siklos: Macroeconomic consequences of the real-financial nexus: Imbalances and spillovers between China and the U.S.
No 3 V.V. Mironov, A.V. Petronevich: Discovering the signs of Dutch disease in Russia
No 4 Joshua Aizenman: The internationalization of the RMB, capital market openness, and financial reforms in China
No 5 Yu-Fu Chen, Michael Funke and Kunyu Tao: Financial market reform – A new driver for China's economic growth?
No 6 Jarko Fidrmuc and Iikka Korhonen: Meta-analysis of Chinese business cycle correlation
No 7 Jarko Fidrmuc, Zuzana Fungáčová and Laurent Weill: Does bank liquidity creation contribute to economic growth? Evidence from Russia
No 8 Elena Deryugina, Olga Kovalenko, Irina Pantina and Alexey Ponomarenko: Disentangling loan demand and supply shocks in Russia
No 9 Michael Funke, Petar Mihaylovski and Haibin Zhu: Monetary policy transmission in China: A DSGE model with parallel shadow banking and interest rate control
No 10 Riikka Nuutilainen: Contemporary monetary policy in China: A move towards price-based policy?
No 11 Iftekhar Hasan, Nada Kobeissi, Haizhi Wang and Mingming Zhou: Banking structure, marketization and small business development: Regional evidence from China
No 12 Linlin Niu, Xiu Xua and Ying Chen: An adaptive approach to forecasting three key macroeconomic variables for transitional China
No 13 Heiner Mikosch and Stefan Neuwirth: Real-time forecasting with a MIDAS VAR
No 14 Alexander Libman and Björn Vollan: Anti-Western conspiracy thinking and expectations of collusion: Evidence from Russia and China
No 15 Mikhail Stolbov: Causality between credit depth and economic growth: Evidence from 24 OECD countries
No 16 Kefei You: What drives China's outward FDI? A regional analysis
No 17 José R. Sánchez-Fung: Estimating the impact of monetary policy on inequality in China
No 18 Christian Hattendorff: Economic concentration and finance: Evidence from Russian regions
No 19 Alexey Porshakov, Elena Deryugina, Alexey Ponomarenko and Andrey Sinyakov: Nowcasting and short-term forecasting of Russian GDP with a dynamic factor model
No 20 Qing He, Iikka Korhonen, Junjie Guo and Fangge Liu: The geographic distribution of international currencies and RMB internationalization
No 21 Haining Wang, Zhiming Cheng and Russell Smyth: Does consuming more make you happier? Evidence from Chinese panel data
No 22 Mikhail Mamonov and Andrei Vernikov: Bank ownership and cost efficiency in Russia, revisited
No 23 Rui Mao and Yang Yao: Fixed exchange rate regimes, real undervaluation and economic growth
No 24 Elena Deryugina, Alexey Ponomarenko, Andrey Sinyakov and Constantine Sorokin: Evaluating underlying inflation measures for Russia
No 25 Christian Dreger, Jarko Fidrmuc, Konstantin Kholodilin and Dirk Ulbricht: The Ruble between the hammer and the anvil: Oil prices and economic sanctions
No 26 Zuzana Fungáčová, Riikka Nuutilainen and Laurent Weill: Reserve requirements and the bank lending channel in China
No 27 Simo Leppänen, Laura Solanko and Riitta Kosonen: Could climate change affect government expenditures? Early evidence from the Russian regions
No 28 Mariarosaria Comunale: Current account and REER misalignments in Central Eastern EU countries: an update using the macroeconomic balance approach
No 29 John Fernald, Eric Hsu and Mark M. Spiegel: Is China fudging its figures? Evidence from trading partner data

www.bof.fi/bofit

BOFIT Discussion Papers
http://www.bof.fi/bofit_en • email: bofit@bof.fi
ISSN 1456-4564 (print) // ISSN 1456-5889 (online)