Finance and Economics Discussion Series
Divisions of Research & Statistics and Monetary Affairs
Federal Reserve Board, Washington, D.C.

Nowcasting Indonesia

Matteo Luciani, Madhavi Pundit, Arief Ramayandi, and Giovanni Veronese

2015-100

Please cite this paper as:
Luciani, Matteo, Madhavi Pundit, Arief Ramayandi, and Giovanni Veronese (2015). "Nowcasting Indonesia," Finance and Economics Discussion Series 2015-100. Washington: Board of Governors of the Federal Reserve System, http://dx.doi.org/10.17016/FEDS.2015.100.

Nowcasting Indonesia*

Matteo Luciani
Federal Reserve Board
matteo.luciani@frb.gov

Madhavi Pundit
Asian Development Bank
mpundit@adb.org

Arief Ramayandi
Asian Development Bank
aramyandi@adb.org

Giovanni Veronese
Banca d'Italia
giovanni.veronese@bancaditalia.it

September 2015

Abstract

We produce predictions of the current state of the Indonesian economy by estimating a dynamic factor model on a dataset of eleven indicators (also followed closely by market operators) over the time period 2002 to 2014. Besides the standard difficulties associated with constructing timely indicators of current economic conditions, Indonesia presents additional challenges typical to emerging market economies where data are often scant and unreliable. By means of a pseudo-real-time forecasting exercise we show that our model outperforms univariate benchmarks, and it does comparably with predictions of market operators. Finally, we show that when quality of data is low, a careful selection of indicators is crucial for better forecast performance.

JEL codes: C32, C53, E37, O53

Keywords: Nowcasting, Dynamic Factor Models, Emerging Market Economies

---

* We would like to thank Dennis Sorino for excellent research assistance and participants of Economics and Research and Regional Cooperation Department, ADB seminar series for valuable comments. The project was prepared under RDTA8951: Macroeconomic modeling for improved economic assessment. This paper was written while Matteo Luciani was *chargé de recherches* F.R.S.-F.N.R.S., and gratefully acknowledges their financial support. Of course, any errors are our responsibility.

1 Introduction

It is well known that macroeconomic data are released with a substantial delay. Additionally, in emerging market economies, low frequency data, i.e. annual national accounts, rely on a smaller array of surveys and indicators than in advanced economies, and provide a partial picture of the economy. However, complete and up to date information on the current state of the economy is crucial for policy makers, market participants and public institutions. Indeed, agents periodically update their forecasts, and monitoring economic conditions in real-time helps them to assess whether the forecasts are on track or need to be revised. Similarly, the process of policymaking often requires long term projections of the economy that heavily rely on accurate initial conditions and forecasts. Therefore, constructing timely "predictions" of current economic conditions, namely *nowcasts*, is of fundamental importance for decision making.

A lot of information is contained in economic indicators that are available on a quarterly, monthly, weekly and even daily basis, and in principle it is possible to use this information to build "predictions" of the current state of the economy. However, high frequency data in emerging market economies are often scant, noisy, released with a lag, and can have missing information. This complicates the difficult task of real-time monitoring and decision making, particularly in an environment where growth volatility is typically high, and where there is considerable uncertainty surrounding trend growth as it may undergo changes due to rapid catching-up phases or persistent slowdowns. In other words, in addition to standard problems, constructing timely indicators on current economic conditions for emerging market economies presents some extra challenges.

In this paper we focus on Indonesia, the largest economy in Southeast Asia which is rapidly gaining influence in the world economy. With a number of high frequency data indicators available and yet facing problems that commonly plague emerging economy datasets, Indonesia provides an interesting training case for developing a nowcasting framework that can be applied to monitor other similar economies in the region.

Two main issues emerge with regard to monitoring in real-time: how many and which indicators to select, and what econometric model to use to extract information from the data. In this paper, we produce "predictions" of the current state of the Indonesian economy by estimating a dynamic factor model on a dataset of eleven indicators (also followed closely by market operators) over the time period 2002 to 2014. Our choice of the model is based on the fact that it is parsimonious and is able to cope with missing data and mixed frequency indicators; and can potentially be estimated on a large number of variables. Further, since the seminal paper of Giannone et al. (2008) this model has become a standard tool for monitoring economic activity, as it has proved to be successful in nowcasting several economies, including emerging ones such as China (Giannone et al., 2014) and Brazil (Bragoli et al., 2014).

The rest of the paper proceeds as follows: Section 2 presents Indonesia's GDP data, and discusses the problems of having several GDP series with different base years, and no official seasonally adjusted data. Section 3 discusses our nowcasting procedures. This section is divided in two parts: in the first part we describe the process of choosing a set of indicators that contains useful information on economic activity, and in the second we present the application of a dynamic factor model to Indonesia's data.

The evaluation of our model is presented in Section 4. Several results emerge. First, incorporating high frequency data in a rigorous framework leads to an improvement in the forecast accuracy of Indonesia's economy compared to simple univariate benchmarks. Second, too many variables are not always optimal for the purpose of monitoring as they can be noisy or uninformative (see also Bańbura et al., 2013; Luciani, 2014b), particularly so when the target variable, namely Indonesia's GDP growth, has limited number of observations. A careful selection of meaningful variables improves the forecast performance. Third, our model does well in predicting quarterly GDP growth when compared to private forecasters such as Bloomberg, and also does well in predicting annual GDP growth when compared to institutional forecasts of the International Monetary Fund and the Asian Development Bank.

Finally, Section 5 concludes.

2 Indonesia's GDP data: Patterns and issues

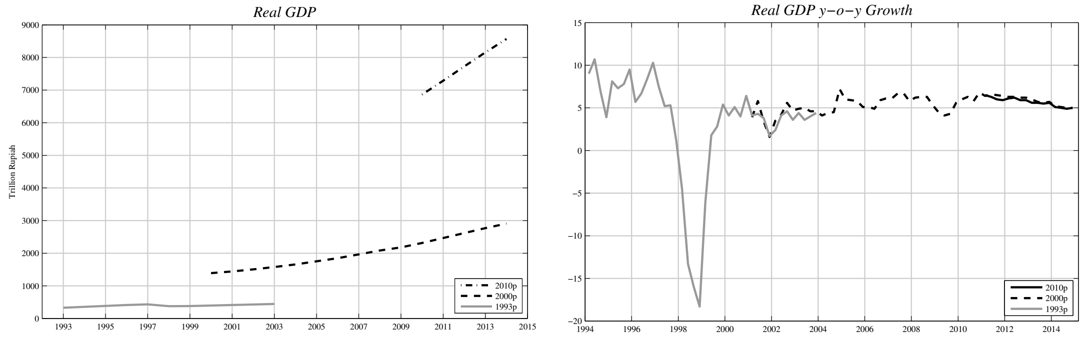
In this Section we present Indonesia's GDP data, and discuss a number of issues in the data that need to be carefully tackled even before starting any monitoring process. The first is that there is no single long series available from official statistical sources. The left hand side chart in Figure 1 plots the level of GDP at constant prices in trillion rupiah from 1993 to 2014, where the three lines refer to GDP across different base years, 1993, 2000 and 2010 respectively. The aggregation methodology was common between the 1993 and 2000 base years, but changed from SNA 1993 to SNA 2008 for the 2010 base year.[1] Base changes are common for GDP data as they can incorporate changes in the economy's structural composition. However the strikingly different slopes among the lines despite a few years of overlaps between series suggests that substantial revisions in data releases affect not only the level of GDP but also its growth rate.[2] This raises questions on the composition of the aggregate series and methodologies used in the construction, which is exacerbated by a lack of publicly available information on procedures used.

The right plot in Figure 1 shows the year-on-year (y-o-y) growth rate of quarterly

[1] A detailed explanation of the System of National Accounts (SNA) can be found in http://unstats.un.org/unsd/nationalaccount/sna.asp.

[2] In fact even nominal GDP data for the overlapping time periods are not comparable between the different bases for GDP series. Data are available in Table VII.1 at http://www.bi.go.id/en/statistik/seki/terkini/riil/Contents/Default.aspx.

Figure 1 Gross Domestic Product series for Indonesia



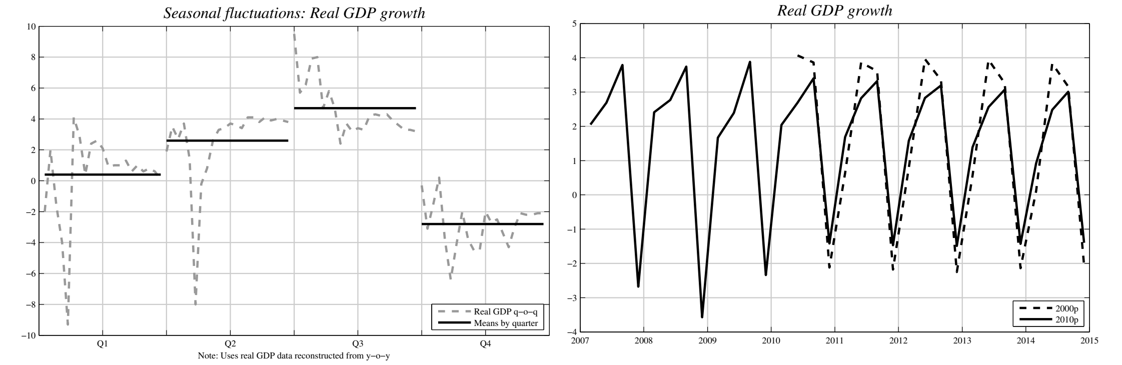
Notes: The figure on the left plots annual GDP at constant prices in levels starting in 1993, and the three lines represent the different base years. There was also an accounting change from SNA 1993 to SNA 2008. The figure on the right plots the year-on-year growth rate of quarterly GDP at constant prices, again corresponding to the three base years.

GDP. As we can see, GDP growth is characterized by large fluctuations due to different crises hitting the economy. In particular, the Asian financial crisis in the late 1990s stands out as a unique episode for Indonesia's growth path with GDP falling by more than 15% in 1998. Then, since 2002 the growth rate stabilized somewhat at a yearly rate of around 5%, with low volatility. As a consequence of this pattern, we exclude the Asian financial crisis years from our sample and use the series only from 2002. This is unavoidable because the evolution of GDP growth over the financial crisis would dominate our estimates if not excluded. It would capture features of the data (i.e. co-movements) that are potentially spurious and misleading for the nowcasting exercise since the relationships that held during the financial crisis could be different from those that hold during expansions and more moderate recessions.

The second big issue with Indonesia's GDP data is that the growth series exhibits a marked seasonal pattern (Figure 2), and yet, there is no seasonally adjusted GDP series available from official sources. So this leaves us with the problem of having to deal with seasonality in the data, particularly when trying to combine series with different base years to obtain coherent and long time series. Indeed, as shown in the right-hand side panel in Figure 2 which plots the quarter-on-quarter (q-o-q) growth rates, the seasonality of GDP data with 2010 base year exhibits a clear departure from the seasonal pattern in the series with 2000 base year. This adds to the complication of splicing the different series together. Suppose we start with the latest available GDP series which is of 2010 base, and extend it backwards using the q-o-q growth rates of the previous base series. As seen in the left hand side panel of Figure 3, it results in inconsistent seasonal patterns within the spliced series, i.e., between the actual data and the extended data. On the contrary, the reconstructed series based on y-o-y growth rates does not seem to have the same defect.

In order to deal with the two issues highlighted, we construct a long time series for GDP by using y-o-y growth rates as shown in the right hand side panel of Figure 3. Furthermore, in our analysis going forward we continue to use y-o-y growth rates. We are well aware that using y-o-y growth makes the series smoother as it effectively tackles the issue of seasonality in the data, but it also lags quarter-on-quarter

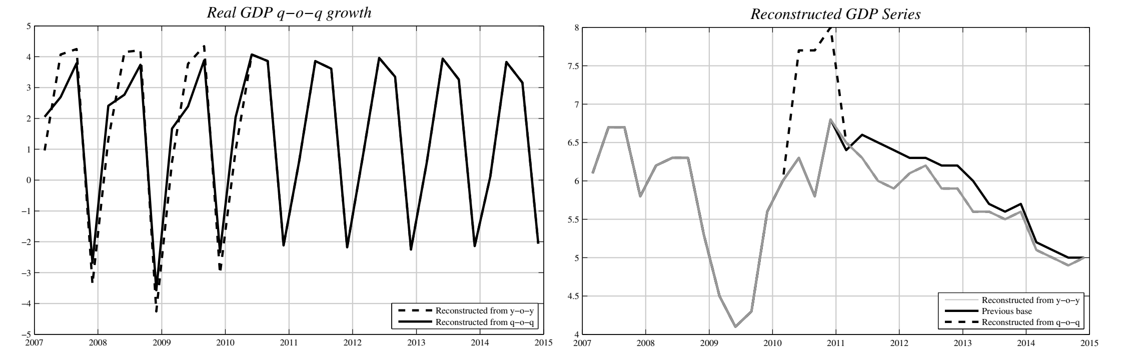
Figure 2 Quarter-on-quarter growth of real GDP



Notes: The figure on the left side shows the seasonal fluctuations in q-o-q GDP growth, where real GDP data is reconstructed using y-o-y growth rates. The figure on the right side plots q-o-q growth rates of real GDP series available, based on two different accounting methodologies, SNA 1993 and SNA 2008 and two different base years, 2000 and 2010.

growth. However, the alternative of using some standard procedure to seasonally adjust the series brings about non-trivial problems: first, any technique to eliminate seasonal fluctuations in the data would introduce an estimation bias associated with the specific procedure being utilized; second, it would not be possible to compare our "predictions" with any credible benchmark. Conversely, using y-o-y growth provides us with a comparable platform with many forecasters such as IMF, ADB and others who look at annual growth rates.

Figure 3 Reconstructing GDP series using y-o-y growth



Notes: Since the latest available data starts in 2010, we have to reconstruct a long time series and the left side graph displays the growth patterns using q-o-q and y-o-y growth rates. The figure on the right side plots the reconstructed series using different growth rates.

3 Nowcasting

Typically GDP data provide the most comprehensive picture of the economy, by aggregating activity of different sectors. Unfortunately, the data come with long delays and are not available on a high frequency basis. In most cases, it is published quarterly and in some developing countries, even annually.

For Indonesia, GDP data is reported quarterly with a delay of about five weeks. This means that the growth rate of the economy in the first quarter that ends in

March is not known until the second week of May. While this is a reasonable delay compared to most emerging and some advanced economies, it is still insufficient for the purpose of real-time monitoring as it is not possible to make an assessment of the strength of economic activity for almost five months into the year.

To assess high frequency movements in economic activity in emerging market economies, analysts typically use the industrial production index as a proxy for output as it is available on a monthly basis.[3] The use of this index relies on the assumption that movements in the industrial sector are a good approximation of aggregate economic activity. While the assumption may hold well for some economies, in others, the cyclical component of the index is not found to be sufficiently synchronized with GDP (Fulop and Gyomai, 2012). A close examination on the relation between Indonesia's GDP growth and its y-o-y growth rate of industrial production shows that the two are only weakly linked (Figure 4). The correlation coefficient between the two series is just 0.35, which suggests that industrial production in Indonesia may not be a reliable proxy for GDP.

Figure 4 Gross Domestic Product & Industrial Production

Notes: The black line is y-o-y GDP growth rate, while the grey line is y-o-y growth rate of Industrial Production Index. The data for Industrial Production were converted to quarterly frequency by averaging monthly observations.

To gauge the current state of the Indonesian economy in real-time, we need to construct a prediction of GDP growth before the official data is released. This means that at each point in time we want to predict not only the current and next quarter estimates for GDP growth (henceforth *nowcast* and *forecasts*), but also, wherever the official data has not been published yet, the past quarter GDP growth (*backcast*).

Ideally, if the true data sources and compilation methods for GDP were known, we could simply attempt to reverse engineer the process performed quarterly by the Indonesian statistical office, but at a higher frequency. Unfortunately though, as discussed in Section 2, very little information on these methods is available from the

[3] See, for example, Maćkowiak (2007) and Raghavan and Dungey (2015) for applications to a set of emerging economies, and Kasri and Kassim (2009) and Kubo (2009) for Indonesia specifically.

statistical agency. This forces us (1) to build an information set that is informative for describing the GDP growth process, possibly by including variables which may be outside the scope of the statistical office, but that may still contain useful leading information, and (2) to choose an econometric model to build our prediction.

In the next two subsections we address these issues. In particular, in Section 3.1 we explain how we construct the database, while in Section 3.2 we introduce Dynamic Factor models (DFMs). We choose to use a DFM since it is a parsimonious model that is able to cope with missing data, mixed frequency, and potentially can be estimated on a large number of variables. Furthermore, this model proved to be very successful in nowcasting several economies, including emerging ones.[4]

3.1 What variables to select?

There is a wide set of potentially useful monthly and quarterly series which could help to extract information on the state of the Indonesian economy. However, the limited number of time series observations available for our target, quarterly year-on-year GDP growth (see Section 2), constrains our choice of both the variables and the model. In principle, Dynamic Factor models are consistently estimated when the number of variables is diverging to infinity, and in practice they are usually estimated on relatively large datasets. However, when using a small number of time series observations, if the sample size is severely limited, then including too many variables is likely to introduce a lot of estimation uncertainty, ultimately worsening the prediction performance. This is a particular source of concern when dealing with Indonesian data, as only few series display a marked comovement with the GDP quarterly dynamics and hence the risk is to introduce excessive noise in the model estimation.[5] The questions then are: on the basis of which criteria should we select variables? And, how many variables should we select?

A possible strategy to draw from a large pool of variables could be to rely on a purely mechanical statistical selection procedure. For example, Bai and Ng (2008) suggest selecting with the LARS algorithm only those variables that are really informative for forecasting the target variable, while Camacho and Perez-Quiros (2010) suggest first selecting a core group of variables, and then evaluating if other possible predictors are useful.

An alternative strategy, pioneered by Bańbura et al. (2013) and followed by Luciani and Ricci (2014), Giannone et al. (2014), and Bragoli et al. (2014), is to exploit

[4] A non exhaustive list of countries and papers is: the US (Bańbura et al., 2011, 2013; Giannone et al., 2008), the Euro Area (Angelini et al., 2011; Bańbura and Rünstler, 2011), Germany (Marcellino and Schumacher, 2010), France (Barhoumi et al., 2010), Ireland (D'Agostino et al., 2012), Norway (Aastveit and Trovik, 2012; Luciani and Ricci, 2014), China (Giannone et al., 2014), Brazil (Bragoli et al., 2014), New Zealand (Matheson, 2010), the Global Economy (Matheson, 2013), and Latin America (Liu et al., 2012)

[5] Furthermore, the literature on nowcasting with large-dimensional DFMs has reached the conclusion that, unless one need to monitor the data flow with many data releases, there is no need of a large database when forecasting with Dynamic Factor Model as long as the variables on which the model is estimated are appropriately selected (see Bańbura et al., 2013; Luciani, 2014a,b).

the "revealed preferences" of professional forecasters who follow the Indonesian economy on the Bloomberg platform. These analysts subscribe to the Bloomberg news alert for specific data releases of the variables that they monitor, and use them to form their expectations on current and future fundamentals of Indonesia. Since Bloomberg constantly ranks the analysts' demand for these alerts by constructing a relevance index for each macroeconomic indicator, we can select variables based on this relevance index.[6]

We adopt the latter approach, since the automatic selection approach risks leading to an unstable choice of variables in a real-time scenario.[7] This instability would not only be difficult to justify from an economic standpoint, it would also complicate the interpretation of the forecasts' revisions. Moreover, we also tried the automatic selection approach and the performance of our model is worse in this case than when we used the revealed preference approach (see the Appendix).

It turns out that for Indonesia only a relatively small number of macroeconomic series are tracked in real-time by the markets (see Table 1). On the one hand, there are indicators describing macroeconomic developments (e.g. the GDP itself, car sales, exports, imports and manufacturing PMI). On the other hand, given their direct impact on the foreign exchange and fixed income markets, analysts also monitor indicators that directly describe the monetary policy stance. These are the central bank reference interest rate as well as key monetary aggregates.

Starting from the set of indicators in Table 1 followed by business analysts we constructed our database as follows:

1) We excluded all those indicators that either had too few observations, or we did not manage to retrieve. This is the case of PMI for which data are available only starting from June 2012, and of Danareksa Consumer Confidence and Motorcycle Sales for which we were not able to retrieve data.[8]
2) We screened each of the remaining indicators in order to understand whether they are followed by analysts because they convey information on the state of the real economy, or they are directly related to the stance of the Central Bank and its balance sheet. Therefore, since Bank of Indonesia has an inflation target, we discarded CPI, and furthermore, we also removed Foreign Reserves, and Net Foreign Assets as they are mainly related to the foreign exchange policy.
3) We then excluded those variables that are the sum of other variables in the database or are too similar to other series. So we kept Imports and Exports, but we excluded Current Account; and we kept M1, but discarded M2.

[6] The implicit assumption here is that since (also) based on their expectations on future fundamentals analysts allocate their investments, they (better) know what are the relevant series to monitor in order to form appropriate expectations on GDP growth.

[7] As shown by De Mol et al. (2008) since there is a lot of comovement among macroeconomic data, the set of indicators selected with statistical criteria is extremely unstable.

[8] As the index compiled by Danareksa was not available to us, we experimented with the household consumer confidence index compiled by the Bank of Indonesia. The latter however displays trending pattern which appears difficult to reconcile with the state of the economy, and we hence discarded it.

Table 1 Bloomberg Calendar: follow the market revealed preference

Variable	**Reference period**	**Release date**	**Freq**	**Rel.**
Bank Indonesia Reference Rate	17-Mar	Mar-17	D	95
GDP YoY	4Q	Feb-2	Q	64
PMI Mfg Markit	Mar	Apr-4	M	82
CPI YoY	Jan	Mar-6	M	86
Foreign Reserves	Dec	Mar-3	M	86
Trade Balance	Feb	Mar-15	M	23
CPI NSA MoM	Feb	Mar-6	M	27
CPI Core YoY	Feb	Mar-6	M	55
Exports YoY	Feb	Mar-15	M	27
GDP QoQ	4Q	Feb-2	Q	59
Consumer Confidence Index	Feb	Mar-4	M	64
Local Auto Sales	Feb	Mar-16	M	50
Motorcycle Sales	Feb	Mar-16	M	32
Net Foreign Assets IDR	Feb	Mar-28	M	45
Imports YoY	Feb	Mar-15	M	50
Money Supply: M2 YoY	Feb	Mar-28	M	32
Danareksa Consumer Confidence	Feb	Mar-5	M	36
Money Supply M1 YoY	Jan	Mar-28	M	32
BoP Current Account Balance	4Q	Mar-15	Q	14

Notes: From left to right: Variable reports the name of the variable; Reference period reports the period to which the data that will be released refers to, while Release Date reports when the data will be released. For example on February 2, 2015, the statistical office released the data for GDP Q4 2014. "Freq." reports at which frequency the variable is published. "Rel." is the relevance index in Bloomberg which counts the number of subscribers to the news alert alerting the release of the variable.

4) Finally, we use our "expert judgement" and added a few indicators that we think provide some extra information about the Indonesian economy. To this end, to capture information regarding the increasing role of construction activity for the Indonesian economy we include domestic cement consumption. To account for spillovers from the foreign sector into the domestic economy we included a variable from the very timely Markit PMI manufacturing survey. In particular we included the aggregate for emerging economies, which is dominated by developments in China as well as countries in the Asian region. Finally, we also included a sectoral breakdown of the imports series to better capture the possibly different lead/lag characteristics of each of these series with GDP growth.

By following this strategy, we end up with a data set of ten macroeconomic indicators plus GDP (see Table 2). While GDP and Business Tendency Index are quarterly series, the remaining are at monthly frequency. The column "Delay" reports the publication delay expressed in number of days in our stylized calendar, and as we can see there are substantial differences between series in terms of their publication delay. For example, the PMI for developing economies is published just four days after the reference month, while data on imports are released a month after the reference month.[9]

[9] For the policy rate we adopted the assumption that it is observed the first day of the month following the reference month. For example, the policy rate for January it is observed on February 1. Of course this is an approximation because we know what the policy rate is everyday in January. In principle, we could have accounted for daily observations in the interest rate since DFMs allow us to do so (Modugno, 2014). However, Bańbura et al. (2013) have shown that including data at

Table 2 Data Description and Data Treatment

Variable	**Freq.**	**Source**	**Start**	**Delay**	**Trans.**
Central Bank policy rate	M	Bank Indonesia	Jan-93	1	
PMI developing economies	M	JP Morgan	Apr-04	4	
Cement, domestic consumption	M	Statistics Indonesia	Jan-94	10	yoy
Exports	M	Statistics Indonesia	Jan-93	15	yoy
Car sales	M	PT Astra	Jan-93	16	yoy
Imports: Consumption Goods	M	Statistics Indonesia	Mar-01	34	yoy
Imports: Capital Goods	M	Statistics Indonesia	Mar-01	34	yoy
Imports: Raw materials	M	Statistics Indonesia	Mar-01	34	yoy
Gross Domestic Product	Q	Statistics Indonesia	Q1 1993	36	yoy
Business Tendency Index	Q	Bank Indonesia	Q2 2000	38	
M2	M	Bank Indonesia	Jan-93	59	yoy

Notes: From left to right: Variable reports the name of the variable; Freq. specifies whether a variable is monthly (M) or quarterly (Q); Source reports the original source of the data; Start specifies since when a variable is available; Delay reports the release delay expressed in number of days in our stylized calendar; and, finally, Trans specifies whether a variable has been transformed to year-on-year growth rates or it is considered in levels. Data for "PMI developing economies" are produced by JP Morgan and were downloaded from Thomson Reuters, Datastream. Data for "Car sales" are produced by PT Astra data and were downloaded from CEIC Data/ISI Emerging Markets.

3.2 Dynamic Factor models

Factor models are based on the idea that macroeconomic fluctuations are the result of few macroeconomic shocks, which affect *the whole* economy, and a number of sectoral/regional shocks that affect *a part of* the economy. Therefore, each variable in the dataset can be decomposed into a common part and an idiosyncratic part, where the common part is assumed to be characterized by a small number of common factors ($\mathbf{f}_t$) which are time series processes meant to capture the comovement in the data, *i.e.* the business cycle.

Formally, let x_{it} be the i-th stationary variable observed at month t, then

$$x_{it} = \boldsymbol{\lambda}_i \mathbf{f}_t + \xi_{it} \qquad i = 1, \dots, n \tag{1}$$

where $\mathbf{f}_t$ is an $r \times 1$ vector (with $r \ll n$) containing the common factors, and ξ_{it} is the i-th idiosyncratic component. The vector of common factors evolve over time as a VAR(p) process driven by the common shocks $\mathbf{u}_t \sim \mathcal{N}(0, \mathbf{I}_r)$, while each idiosyncratic component follows an independent AR(1) model driven by the idiosyncratic shocks e_{it}:

$$\mathbf{f}_t = \sum_{s=1}^{p} \mathbf{A}_s \mathbf{f}_{t-s} + \mathbf{u}_t \tag{2}$$

$$\xi_{it} = \rho_i \xi_{it-1} + e_{it}. \tag{3}$$

Equations (1)-(3) define the Dynamic Factor model used in this paper.[10] The model can be estimated by Principal Components (Stock and Watson, 2002a; Bai, 2003),

the daily frequency is not particularly useful for nowcasting GDP, so we adopted the convention that the interest rate is monthly and is observed on the first day after the reference month.

[10] This is the model studied in Doz et al. (2011, 2012), which is a special case of the model studied in Forni et al. (2009). In this model, the common shocks and the idiosyncratic shocks are assumed to be uncorrelated at all leads and lags, while the idiosyncratic shocks are allowed to

by using the Kalman Filter (Doz et al., 2011), or by maximum likelihood techniques through the EM algorithm (Doz et al., 2012). In this paper we will use maximum likelihood, and in particular we will use the EM algorithm proposed by Bańbura and Modugno (2014) which can handle both mixed frequencies and missing data.

In the next section we will use the model to produce real-time predictions of Indonesian GDP growth. DFM proved very successful in real-time forecasting, and when used for this task they work as follows: suppose that we are at day d, and that at date d it is available a given vintage of data: $\mathbf{X}^d$. Further suppose that on the basis of $\mathbf{X}^d$ we have constructed our prediction: $\hat{x}_{it}^d = \widehat{\boldsymbol{\lambda}}_i \hat{\mathbf{f}}_t^d + \hat{\mathbf{e}}_t$. Now, suppose that at day $d+1$ a new data is released (*eg.* Exports). Based on this new piece of information we can check if our stand about the business cycle is still correct or if we need to revised it, which is what the DFM does automatically. More specifically, at day $d+1$ we have now a new vintage of data: $\mathbf{X}^{d+1}$. Given this new vintage we can update our estimate of the factors, $\hat{\mathbf{f}}_t^{d+1}$, and hence update our prediction: $\hat{x}_{it}^{d+1} = \widehat{\boldsymbol{\lambda}}_i \hat{\mathbf{f}}_t^{d+1} + \hat{\mathbf{e}}_t$

4 Empirics

4.1 The forecasting exercise

To evaluate the performance of our model, we perform a *pseudo* real-time out-of-sample exercise. Predictions of Indonesian GDP growth are produced according to a recursive scheme, where the first sample starts in June 2002 and ends in December 2007, while the last sample starts in June 2002 and ends in December 2014. The model is estimated at the beginning of each quarter using only information available as of the first day of the quarter, and then the parameters are held fixed until the next quarter.

To perform our exercise we construct real-time vintages by replicating the pattern of data availability implied by the stylized calendar (Table 2), and every time new data are released, we update the prediction based only on information actually available at that time. We call this exercise *pseudo* real-time since we are not able to track the full set of data revisions, an issue that we will discuss further later.

For the estimation, we include two factors ($r = 2$) and two lags ($p = 2$) in the VAR model governing the evolution over time of the factors. The choice of including two factors deserves a comment. First the literature on factor models has shown that for forecasting it suffices to include a small number of factors (eg. Stock and Watson, 2002b; Forni et al., 2003). Furthermore, recent literature on small-medium Dynamic Factor models (Bańbura et al., 2013; Luciani and Ricci, 2014; Giannone et al., 2014; Bragoli et al., 2014) often include one factor only. Therefore, a natural

be cross-sectionally correlated, albeit by a limited amount (approximate factor structure). For a more comprehensive treatment of the DFM we refer the reader to the aforementioned references and to the survey by Luciani (2014b).

choice would be to follow the literature and to set $r = 1$. However, this literature estimates models for q-o-q growth rates, while we are estimating a model for y-o-y growth rates, and if the model for q-o-q growth rates has one factor, then the corresponding model for y-o-y growth rates has four factors.[11] Hence, we should set $r = 4$, and, indeed, by looking at the eigenvalues of the covariance matrix we can see clearly three/four diverging eigenvalues. However, among these three/four eigenvalues the first two clearly dominate suggesting that the other two carry mainly noise, which motivates our choice to set $r = 2$.[12]

4.2 Comparison against Statistical Benchmark

To judge the performance of our model and to evaluate the information contained in our dataset we start by comparing our model with three benchmark models.

Our first benchmark is the *naive* forecast, obtained from the random walk model on GDP growth: $y_t^Q = y_{t-1}^Q + \varepsilon_t$. The second is a forecast from an autoregressive model of order two on GDP growth: $y_t^Q = \rho_1 y_{t-1}^Q + \rho_2 y_{t-2}^Q + \varepsilon_t$. Given the high persistence in our target series introduced by the y-o-y transformation, these univariate benchmarks are inherently tough competitors to match in our real-time exercise.

Our last benchmark is a bridge model (Parigi and Schlitzer, 1995). Bridge models predict GDP growth by using its own past plus one or more monthly indicators.[13] Formally, let y_t be the y-o-y GDP growth observed quarterly, i.e., at month $t = 3, 6, 9, \dots$ and let x_t be a monthly variable, then the Bridge model is defined as follows:

$$y_t = \mu + \alpha y_{t-1} + \beta \tilde{x}_t + \varepsilon_t \tag{4}$$

where $\tilde{x}_t = \sum_{j=1}^{3} \frac{1}{3} x_{t-j}$ is the monthly indicator aggregated at the quarterly frequency by a simple average. In this paper, we estimated equation (4) by OLS, and when we have a missing observations in x_t we filled it by using an AR model. Furthermore, the predictions from the bridge model are obtained by first estimating a model for each monthly indicator in the database (except for the PMI series), and then by averaging the prediction.[14]

Table 3 shows the Root Mean Squared Error (RMSE) at the end of each month of the DFM, an AR(2) model, a Random Walk, and the Bridge model. The Table is

---

[11] Let X_t be a non-stationary variable in log-levels, and let $x_t^y = X_t - X_{t-4}$ be the y-o-y growth rates and $x_t^q = X_t - X_{t-1}$ be the q-o-q growth rates, so that $x_t^y = x_t^q + x_{t-1}^q + x_{t-2}^q + x_{t-3}^q$. Then, if the true model is $x_t^q = \lambda f_t + e_t$, we have $x_t^y = \lambda(1 + L + L^2 + L^3) f_t + (1 + L + L^2 + L^3) e_t$, which can be rewritten as $x_t^y = \lambda \mathbf{F}_t + (1 + L + L^2 + L^3) e_t$ where $\mathbf{F}_t$ is a 4×1 singular vector.

[12] The first eigenvalue account for 70% of the total variance, the second for 20%, the third for 5%, and the fourth for 3%. In the appendix we show robustness results when the model is estimated by setting either $r = 1$ or $r = 4$.

[13] As pointed out by Baffigi et al. (2004), differently from DFMs, bridge models are not concerned with particular assumption underlying the DGP of the data, but rather, the inclusion of specific explanatory indicators is based on the simple statistical fact that they embody timely updated information about the target GDP growth series.

[14] PMI developing countries was excluded because there are too few observations for this indicator and the prediction is volatile.

divided in three parts: the first part, labelled as "Forecast", reports the RMSE of the prediction of the next quarter; the second, labelled as "Nowcast", reports the RMSE of the prediction of the current quarter; finally, the last section, labelled as "Backcast", reports the MSE of the prediction of the previous quarter.

As we can see from the fact that the RMSE in Table 3 are decreasing with each month, the DFM is able to correctly revise its GDP prediction as more data becomes available. Furthermore, compared to the univariate benchmarks the RMSE of the DFM is consistently lower, up to a maximum reduction at the end of the first month after the reference month of 39% compared to the Autoregressive model, and of 15% compared to the Bridge model. This is an important finding since it tells us that there is valuable additional information in the Indonesian high frequency data that can be used to predict GDP growth.

Table 3 Root Mean Squared Error: End of month

	Month	DFM	AR	RW	Bridge
Forecast	1	0.595	0.703	0.847	0.627
	2	0.525	0.619	0.661	0.567
	3	0.467	0.619	0.661	0.536
Nowcast	1	0.441	0.609	0.666	0.494
	2	0.342	0.455	0.430	0.391
	3	0.298	0.455	0.430	0.361
Backcast	1	0.279	0.456	0.459	0.331

Notes: This table reports Root Mean Squared Error (RMSE) at the end of each month for the Dynamic Factor model (DFM), an AR(2) model, a Random Walk (RW), and the Bridge model. The upper panel labelled as "Forecast", reports the RMSE of the prediction of the next quarter; the mid panel, labelled as "Nowcast", reports the RMSE of the prediction of the current quarter; the bottom panel labelled as "Backcast", reports the MSE of the prediction of the previous quarter.

In Table 4 we investigate which data release carries more information for the prediction with the DFM. Identifying such variables is particularly important since it can help policymakers understand what series to track while monitoring the Indonesian economy. More precisely, Table 4 shows the RMSE associated with each data release. We can see that some variables are particularly relevant for correctly updating the prediction of y-o-y GDP growth. Among these, the most important one is Exports, which accounts for the largest reduction in RMSE when the data is released. Other relevant ones are GDP of the previous quarter, Imports and Cement. Notice also that upon the release of some variables the "average" forecasting performance reported in Table 4 appears to deteriorate. Thus it would be tempting to drop these variables in order to "improve" the overall forecasting performance. However, each of them may also improve the estimation of the model by exploiting the commonality in the data, and hence make our forecasts more robust to one-off changes in a particular variable.

We conclude this section with a caveat that will apply to most of the empirical exercise. As we argued in Section 2 we had to use data only from 2002 onwards, and this has limited us in two ways. First, as discussed in Section 3, we restrict the number of variables to include in the model. Second, since we are able to produce predictions for just 28 quarters, we are averaging over only 28 prediction errors to produce the RMSE.

Table 4 Root Mean Squared Error: Data Flow

	Day	Release	Forecast	Nowcast	Backcast
Month 1	1	Policy rate	0.624	0.485	0.315
	3	Imports*	0.606	0.455	0.297
	4	PMI	0.612	0.460	0.300
	10	Cement	0.611	0.456	0.296
	15	Exports	0.595	0.442	0.280
	16	Car sales	0.594	0.440	0.277
	28	M2	0.595	0.441	0.279
Month 2	1	Policy rate	0.594	0.443	0.278
	3	Imports*	0.592	0.458	0.285
	4	PMI	0.598	0.467	0.285
	5	GDP	0.589	0.436	
	7	BTI**	0.588	0.437	
	10	Cement	0.582	0.426	
	15	Exports	0.524	0.348	
	16	Car sales	0.525	0.342	
	28	M2	0.525	0.342	
Month 3	1	Policy rate	0.518	0.344	
	3	Imports*	0.509	0.327	
	4	PMI	0.514	0.337	
	10	Cement	0.509	0.333	
	15	Exports	0.459	0.298	
	16	Car sales	0.464	0.295	
	28	M2	0.467	0.298	

Notes: This table reports Root Mean Squared Error (RMSE) in correspondence of each data releases. Column "Forecast", reports the RMSE of the prediction of the next quarter; column "Nowcast", reports the RMSE of the prediction of the current quarter; column "Backcast", reports the MSE of the prediction of the previous quarter. *In this day 3 different series are released: Imports: Consumption Goods, Imports: Capital Goods, and Imports: Raw materials but results are grouped in one variable. ** BTI stands for Business Tendency Index.

4.3 Comparison against Market Benchmark

Another way to evaluate the forecasting performance of our model, is to compare the prediction obtained with the DFM with the prediction of market operators. In Figure 5 we compare our predictions with those of the Bloomberg Survey (BS). The BS consists of the median GDP prediction provided independently by a number of specialists a few days before GDP is released.[15] Therefore, since the Bloomberg Survey is released few days before previous quarter GDP, according to our terminology the BS prediction is a *backcast*, and we will compare it with our last prediction before GDP is released.

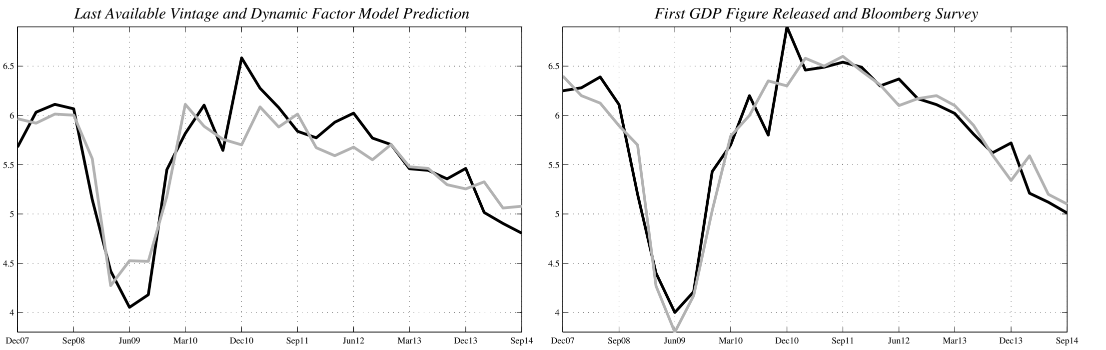
When comparing our prediction with Bloomberg we have to be careful with respect to data revisions. It is well known that GDP data are revised often, and as alluded in Section 4.1 we are not able to track data revisions. The literature on factor models has shown that these models are robust to data revisions (Giannone et al., 2008) as revision errors are by nature idiosyncratic and do not affect factor estimation. However the case of Indonesia appears rather exceptional since the statistical office has recently revised the GDP series substantially. The black line of the left plot in Figure 5 represents the last vintage available of the GDP series, while the black

[15] Notice that both the number of specialists that provide the prediction as well as the survey release day vary from quarter to quarter.

line in the right plot of Figure 5 represents the first release of GDP growth by the statistical office. Clearly, these two series are quite different, and in particular from 2011 onwards the statistical office has systematically revised down its GDP growth estimates. This fact is crucial when we attempt to make comparisons with truly historical forecasts. In particular, the Bloomberg survey is targeting the first release of year-on-year quarterly GDP growth (right plot), while the DFM we have estimated is designed to target the final release (left plot).

As we can see from Figure 5, the Bloomberg Survey is tracking the first release well, with a RMSE of 0.249. Similarly, we can also see that our prediction is good, and indeed our RMSE is 0.286 which is just 15% worse than that of the Bloomberg Survey. What is more striking though, is the magnitude of the revision error of the statistical office. Indeed, if we think of the first release as an estimate of the final release, we can than compute a RMSE which in this case is 0.334, higher than what we obtain with the DFM (albeit using only final figures). This result suggests that the process of GDP revisions is not entirely random, as it can be "improved upon" by exploiting the information available in our relatively small dataset. More generally, it suggests the difficulty in interpreting the reliability of official national accounts estimates.

Figure 5 Prediction of Previous Quarter GDP



Notes: The black line of the left plot represents the last vintage available of the GDP series, while the grey line is the prediction obtained with the DFM the last day before GDP is released. The black line in the right plot represents the first figure released by the statistical office of GDP growth, while the grey line is the median prediction from the Bloomberg Survey.

4.4 Comparison against Institutional benchmarks

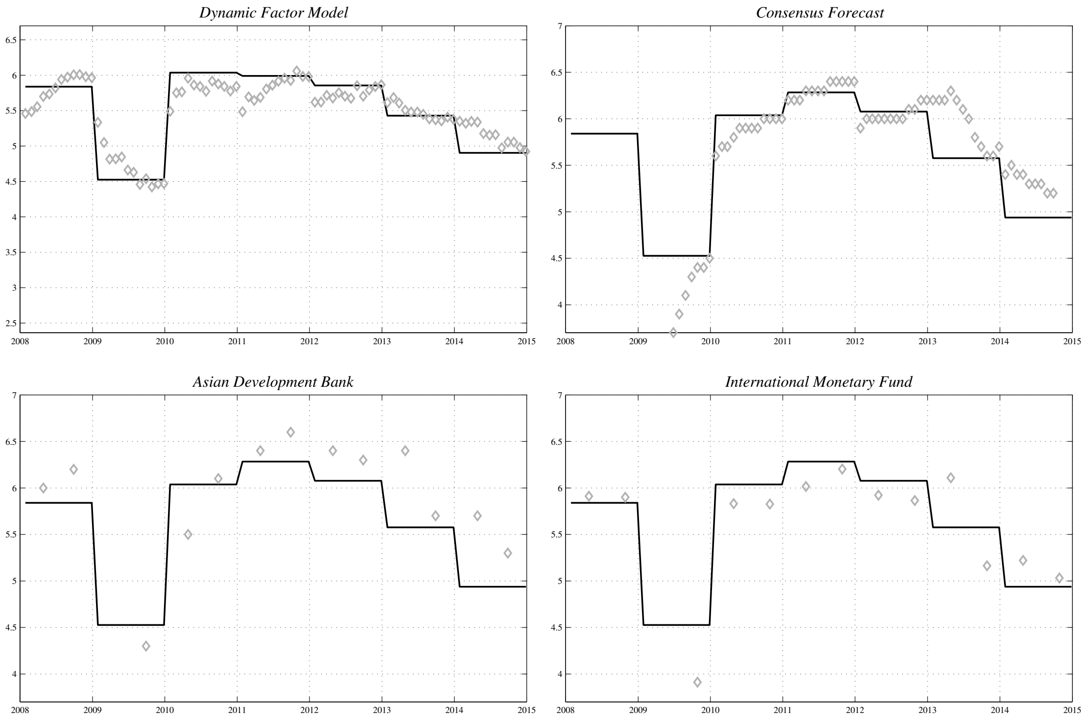
Finally, by following Luciani and Ricci (2014) we can construct predictions for the current annual growth rate and compare them with those published by policy institutions.[16] In details, we compare our predictions with those published by the

[16] Let $X_q^y = 100\times\log(GDP_q^y)$ be GDP of the q-th quarter of year y, and let $Z^y = 100\times\log(GDP^y)$ be GDP of year y. Then, by definition $x_q^y = X_q^y - X_q^{y-1}$ is the y-o-y growth rate, while $z^y = Z^y - Z^{y-1}$ is the annual growth rate. Following Mariano and Murasawa (2003), we make use of the approximation $Z^y \approx (X_1^y + X_2^y + X_3^y + X_4^y)/4$, which allow us to write the annual growth rate as a function of y-o-y growth rates: $z^y = Z^y - Z^{y-1} \approx (X_1^y + X_2^y + X_3^y + X_4^y)/4 - (X_1^{y-1} + X_2^{y-1} + X_3^{y-1} + X_4^{y-1})/4 = (x_4^y + x_3^y + x_2^y + x_1^y)/4$.

Asian Development Bank (ADB) in the Asian Development Outlook, those published by the International Monetary Fund (IMF) in the World Economic Outlook, and the prediction by Consensus Forecast (CF).[17] The ADB publishes its prediction of current annual GDP growth twice a year, approximately in April and in late September; also the IMF publishes twice a year its prediction but these are released on April and October. Predictions by CF are available each month.

The north west (NW) panel in Figure 6 shows annual GDP growth together with the prediction of current annual GDP growth obtained at the end of each month with the DFM. The other panels of Figure 6 show predictions from ADB, IMF, and CF together with annual GDP growth. Note however, that the annual GDP growth reported in the NW panel is different from that reported in the other panels. In the NW panel we are reporting annual growth computed on the basis of the last vintage of available data (reconstructed series using base year 2010 as described in Section 2), while the other panels report annual growth computed on the basis of the last vintage of the old GDP series (2000 basis). We do the latter because the ADB, the IMF, and CF predicted the series with the old base in real-time and not the new base.

Figure 6 Prediction of Annual GDP Growth Rate



Notes: The black line in the north-west plot is annual GDP growth, while the grey diamonds are the prediction obtained at the end of each month with the DFM. In all the other panels the black line is annual GDP growth computed by using the last vintage of the old GDP series (2000 basis), while the grey diamonds are the prediction of ADB, IMF, and CF. Data for CF are from Consensus Economics Inc.

From Figure 6 and the RMSE values in Table 5, we can see that the DFM is

[17] Consensus Economics Inc. forecasts comprise quantitative predictions of private sector forecasters. Each month survey participants are asked for their forecasts of a range of macroeconomic and financial variables for the major economies.

predicting annual GDP growth quite well, and in particular it correctly revises its prediction as more data becomes available during the calendar year. Furthermore, the prediction of the DFM is comparable to that of CF, and slightly superior to that of the ADB and the IMF. Of course here two caveats apply: the first is that using annual data we have only 7 observations, and the second is that our exercise is *pseudo* real-time, and therefore we have a better information set than the one available in real-time to forecasters.

Table 5 Root Mean Squared Error: Annual GDP Growth

Month	DFM	CF	ADB	IMF
January	0.484	0.408		
February	0.352	0.408		
March	0.295	0.577		
April	0.248	0.722	0.603	0.824
May	0.195	0.655		
June	0.152	0.431		
July	0.167	0.348		
August	0.080	0.237		
September	0.124	0.158	0.262	
October	0.128	0.080		0.306
November	0.118	0.097		
December	0.093	0.096		

Notes: This table reports Root Mean Squared Error (RMSE) of the prediction of annual GDP growth at the end of each month. The RMSE of the DFM is computed with reference to the last vintage available for the GDP series, while the RMSE for ADB, CF, and IMF, is computed with reference to the last vintage of the old GDP series (2000 basis). Data for CF are from Consensus Economics Inc.

5 Conclusions

In this paper we have applied state of the art techniques for nowcasting Indonesia's GDP growth. Our approach is based on a Dynamic Factor model, to efficiently exploit monthly and quarterly variables and to properly account for the sequence of macroeconomic data releases.

We find that relying on market "revealed preferences" for certain indicators on the Indonesian economy is an effective guide to choosing what variables to include in our information set. To this end we have relied on the Bloomberg platform, which tracks the relevance of each series for its subscribers, and also on our "expertise judgment" by including a few indicators that we think provide extra information on the Indonesian economy. Based on this, despite using a relatively narrow set of variables, when focusing on the year-on-year growth rate, the Dynamic Factor model nowcast error falls by 35% compared to the benchmark AR, and by almost 40% for the backcast. Lacking a full time series of GDP revisions as well as for the information set used in the Dynamic Factor model we cannot assess how well our model predictions perform compared to those of experts' forecast surveyed by Bloomberg. Still, our "pseudo-real-time" forecasting performance is comparable to the one achieved in a truly "real-time" setting by the median Bloomberg survey.

Furthermore, since our model can be used to forecast further ahead, we compute also calendar year annual growth rates. This exercise allows us to compare the tracking of our Dynamic Factor model forecasts on a smoother and medium-term indicator of Indonesia's growth, arguably more important for policy decisions than the (potentially erratic) quarterly growth rate. In this case our model compares well with the forecasts produced by the average of private sector expectations (Consensus Forecasts) as well as by the IMF-World Economic Outlook.

Finally, our exploration into Indonesia's data sheds light onto a lack of valuable high frequency statistics on economic growth, as well as on deficiencies in the statistical framework underlying national accounts.

References

Aastveit, K. A. and T. Trovik (2012). Nowcasting norwegian GDP: the role of asset prices in a small open economy. *Empirical Economics 42*(1), 95–119.

Angelini, E., G. Camba-Mendez, D. Giannone, L. Reichlin, and G. Rünstler (2011). Short-term forecasts of euro area gdp growth. *The Econometrics Journal 14*(1), C25–C44.

Baffigi, A., R. Golinelli, and G. Parigi (2004). Bridge models to forecast the euro area GDP. *International Journal of Forecasting 20*(3), 447–460.

Bai, J. (2003). Inferential theory for factor models of large dimensions. *Econometrica 71*, 135–171.

Bai, J. and S. Ng (2008). Forecasting economic time series using targeted predictors. *Journal of Econometrics 146*(2), 304–317.

Bańbura, M., D. Giannone, M. Modugno, and L. Reichlin (2013). Now-casting and the real-time data-flow. In G. Elliott and A. Timmermann (Eds.), *Handbook of Economic Forecasting*, Volume 2, pp. 195–237. Amsterdam: Elsevier-North Holland.

Bańbura, M., D. Giannone, and L. Reichlin (2011). Nowcasting. In M. P. Clements and D. F. Hendry (Eds.), *Oxford Handbook on Economic Forecasting*. New York: Oxford University Press.

Bańbura, M. and M. Modugno (2014). Maximum likelihood estimation of factor models on data sets with arbitrary pattern of missing data. *Journal of Applied Econometrics 29*(1), 133–160.

Bańbura, M. and G. Rünstler (2011). A look into the factor model black box: Publication lags and the role of hard and soft data in forecasting GDP. *International Journal of Forecasting 27*(2), 333–346.

Barhoumi, K., O. Darné, and L. Ferrara (2010). Are disaggregate data useful for factor analysis in forecasting French GDP? *Journal of Forecasting 29*(1-2), 132–144.

Bragoli, D., L. Metelli, and M. Modugno (2014). The importance of updating: Evidence from a brazilian nowcasting model. Finance and Economics Discussion Series 2014-94, Federal Reserve Board.

Camacho, M. and G. Perez-Quiros (2010). Introducing the euro-sting: Short-term indicator of euro area growth. *Journal of Applied Econometrics 25*(4), 663–694.

D'Agostino, A., K. McQuinn, and D. O'Brien (2012). Now-casting Irish GDP. *OECD Journal: Journal of Business Cycle Measurement and Analysis 2012*(2), 21–31.

De Mol, C., D. Giannone, and L. Reichlin (2008). Forecasting using a large number of predictors: Is bayesian shrinkage a valid alternative to principal components? *Journal of Econometrics 146*, 318–328.

Doz, C., D. Giannone, and L. Reichlin (2011). A two-step estimator for large approximate dynamic factor models based on kalman filtering. *Journal of Econometrics 164*(1), 188–205.

Doz, C., D. Giannone, and L. Reichlin (2012). A quasi maximum likelihood approach for large approximate dynamic factor models. *Review of Economics and Statistics 94*(4), 1014–1024.

Forni, M., D. Giannone, M. Lippi, and L. Reichlin (2009). Opening the Black Box: Structural Factor Models versus Structural VARs. *Econometric Theory 25*(5), 1319–1347.

Forni, M., M. Hallin, M. Lippi, and L. Reichlin (2003). Do financial variables help forecasting inflation and real activity in the Euro Area? *Journal of Monetary Economics 50*(6), 1243–1255.

Fulop, G. and G. Gyomai (2012). Transition of the OECD CLI system to a GDP-based business cycle target.

Giannone, D., S. Miranda Agrippino, M. Modugno, and L. Reichlin (2014). Nowcasting China. mimeo.

Giannone, D., L. Reichlin, and D. Small (2008). Nowcasting: The real-time informational content of macroeconomic data. *Journal of Monetary Economics 55*(4), 665–676.

Kasri, R. and S. H. Kassim (2009). Empirical determinants of saving in the Islamic banks: Evidence from Indonesia. *JKAU: Islamic Econ. 22*(2), 181–201.

Kubo, A. (2009). Monetary targeting and inflation: Evidence from Indonesia's post-crisis experience. *Economics Bulletin 29*(3), 1805–1813.

Liu, P., T. Matheson, and R. Romeu (2012). Real-time forecasts of economic activity for Latin American economies. *Economic Modelling 29*(4), 1090–1098.

Luciani, M. (2014a). Forecasting with approximate dynamic factor models: the role of non-pervasive shocks. *International Journal of Forecasting 30*(1), 20–29.

Luciani, M. (2014b). Large-dimensional dynamic factor models in real-time: A survey. Université libre de Bruxelles.

Luciani, M. and L. Ricci (2014). Nowcasting Norway. *International Journal of Central Banking 10*, 215–248.

Maćkowiak, B. (2007). External shocks, US monetary policy and macroeconomic fluctuations in emerging markets. *Journal of Monetary Economics 54*(8), 2512–2520.

Marcellino, M. and C. Schumacher (2010). Factor MIDAS for nowcasting and forecasting with ragged-edge data: A model comparison for German GDP. *Oxford Bulletin of Economics and Statistics 72*(4), 518–550.

Mariano, R. S. and Y. Murasawa (2003). A new coincident index of business cycles based on monthly and quarterly series. *Journal of Applied Econometrics 18*(4), 427–443.

Matheson, T. D. (2010). An analysis of the informational content of New Zealand data releases: The importance of business opinion surveys. *Economic Modelling 27*(1), 304–314.

Matheson, T. D. (2013). New indicators for tracking growth in real time. *Journal of Business Cycle Measurement and Analysis 2013*(1), 51–71.

Modugno, M. (2014). Now-casting inflation using high frequency data. *International Journal of Forecasting 29*(4), 664–675.

Parigi, G. and G. Schlitzer (1995). Quarterly forecasts of the italian business cycle by means of monthly economic indicators. *Journal of Forecasting 14*(2), 117–141.

Raghavan, M. and M. Dungey (2015). Should ASEAN-5 monetary policy-makers act pre-emptively against stock market bubbles? *Applied Economics 47*(11), 1086–1105.

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

Stock, J. H. and M. W. Watson (2002b). Macroeconomic forecasting using diffusion indexes. *Journal of Business and Economic Statistics 20*(2), 147–162.

A Appendix

A.1 Robustness

In this appendix we show robustness checks with respect to the number of factors and to the composition of the dataset.

As we discuss in Section 3.1 we constructed the database by selecting variables from the set of indicators that market analysts are monitoring. In Table 6 we show RMSE for a DFM parameterized as described in Section 4.1 but estimated on different datasets. In particular we considered the option of using all the indicators in Table 1 (Bloomberg Selection), and the option of selecting indicators automatically with the LARS algorithm (Automatic Selection) as in Bai and Ng (2008). As we can see our database clearly delivers the best performance, though at least in forecasting the performance of the DFM estimated over the indicators followed by Bloomberg is comparable. It is particularly disappointing the performance of the DFM when the indicators are selected with LARS. We believe that this is a consequence of having too few observations, and some missing values here and there in the time series of our indicators.

Then, in Section 4.1 we motivated our choice of including two factors in the model, but in doing so we explained that two possible meaningful options were to estimate a model with one factor, or a model with four factors. In Table 6 we also show the RMSE for a DFM estimated by including different number of factors, and as we can see the choice of including two factors proved to be optimal.

Table 6 Root Mean Squared Error: Different model specifications and datasets

	Month	Benchmark Model	One Factor	Four Factors	Automatic Selection	Bloomberg Selection
Forecast	1	0.595	0.614	0.716	0.921	0.588
	2	0.525	0.549	0.666	0.810	0.526
	3	0.467	0.517	0.564	0.751	0.533
Nowcast	1	0.441	0.499	0.499	0.726	0.542
	2	0.342	0.389	0.408	0.554	0.423
	3	0.298	0.362	0.383	0.507	0.433
Backcast	1	0.279	0.331	0.333	0.489	0.408

Notes: This Table reports Root Mean Squared Errors for the DFM estimated under different configurations or over database in which the selection process is different than the one explained in Section 3.1. The upper panel labelled as "Forecast", reports the RMSE of the prediction of the next quarter; the mid panel, labelled as "Nowcast", reports the RMSE of the prediction of the current quarter; the bottom panel labelled as "Backcast", reports the MSE of the prediction of the previous quarter.