The Lahore Journal of Economics
26: 1 (Summer 2021): pp. 1–30

Estimation and Forecasting of Industrial Production Index

Muhammad Ejaz and Javed Iqbal*

Abstract

It is essential that policymakers consider cyclical changes in output. Monthly industrial production is one of the most important and commonly used macroeconomic indicators for this purpose. However, monthly estimates of industrial production are not available for Pakistan. Instead, policymakers rely on a large-scale manufacturing (LSM) index that accounts for only 10 percent of GDP. Another limitation of this index is that it accounts primarily for private sector industry, leaving out the direct public sector presence in industrial production. Economic policymakers rely heavily on the LSM index to gauge economic activity in Pakistan. In this study, we compute a new industrial production index (IPI) that extends to the whole industrial sector in Pakistan, incorporating additional information that the LSM index misses. Post-estimation, we build seven econometric models reflecting conditions in the real, financial, and external sectors to estimate year-on-year changes in the new IPI. Our results show that the root mean square error of the ARDL model reflecting financial conditions is lowest of the models tested, which included AR, VAR, and BVAR, across all horizons.

Keywords: Economic indicator, industry studies, econometric forecasting, Pakistan.

JEL Classification: L600, C80, C530.

1. Introduction

Industrial production is a key variable of interest in short-term economic policy analysis. This is true despite the services sector accounting for the largest share of GDP. The industrial sector is important in explaining aggregate fluctuations because certain service activities are closely linked to industrial ones (Bruno & Lupi, 2004). It is for this reason that economists continue to consider industrial production a leading indicator of economic activity (see Banerjee, Marcellino & Masten, 2005).

* The authors are, respectively, joint director of the Monetary Policy Department at the State Bank of Pakistan, and joint director of the Research Department at the State Bank of Pakistan.

Industrial output is closely related to the level of income (Hettige, Lucas & Wheeler, 1992). Bean (1946) has shown that differences in per capita income are explained by the stages[1] and patterns[2] of industrialization across countries. Accordingly, industrial production is given due weight in policy analyses by central banks around the world. The Board of Governors of the Federal Reserve System accords great importance to industrial production in policy analysis, using it as an economic indicator to gauge movements in production and highlight structural developments in the economy.[3] Similarly, researchers and policymakers at the European Central Bank track industrial production closely to help them understand short-term changes in the European Union's real economic activity (Bodo, Golinelli & Parigi, 2000).

Industry accounts for 23 percent of the total GDP at producer prices in Pakistan (Table 1). Unfortunately, high-frequency data on industrial production in Pakistan is not available and it does not help that Pakistan's GDP is officially available only annually. Although the Pakistan Bureau of Statistics (PBS) once provided an industrial production series, available since 1999, it was discontinued in 2012. Arby (2008), Hanif, Iqbal and Malik (2013), and Tahir, Ahmed and Ahmed (2018) have made notable attempts to estimate the national accounts on a quarterly basis, but their estimates do not provide a series for industrial production as a separate component.

Table 1: Composition of industry in 2017/18 (P)

Constant basic prices 2005/06	PKR million	As % of total industry	As % of manufacturing
Industry	2,951,336		
M&Q	341,934	11.6	
Manufacturing	1,680,161	56.9	
LSM	1,338,010		79.6
SSM	232,413		13.8
Slaughtering	109,738		6.5
Energy	219,463	7.4	
Construction	349,778	11.9	
GDP + tax - subsidies	13,100,711		
LSM as % of GDP at PP	10.2%		
Industry as % of GDP at PP	23%		

Source: Pakistan Bureau of Statistics.

[1] This refers to the proportion of a country's working population engaged in primary occupations—agriculture, forestry, and fishing.

[2] By pattern of industrialization, Bean (1946) means the relative importance of secondary occupations (manufacturing, mining, and construction) and tertiary occupations (trade and services).

[3] https://fred.stlouisfed.org/series/INDPRO

Although data on industrial production tends to be high-frequency, containing cyclical information that is important for economic policy decisions (Bodo et al., 2000), there has been little effort to measure industrial production in Pakistan on a monthly basis. This underscores the need for monthly estimates of industrial production, especially when monetary policy in Pakistan is now conducted in alternate months of every fiscal year.

In the absence of complete information on industrial production, policymakers in Pakistan are compelled to make do with proxies such as large-scale manufacturing (LSM), which accounts for 10 percent of the overall industrial production (Table 1). More importantly, the PBS publishes the LSM index with a considerable lag.[4] Hussain, Hyder and Rehman (2018) attempt to address these data delays by providing early estimates of LSM, but these are not without limitations. They are estimated using statistical methods and are thus not completely reliable. Moreover, the issue of coverage remains salient because a substantial portion of information (about 13 percent) on industrial production is not accounted for due to the unavailability of an appropriate series. By relying only on a subcomponent (LSM) of the overall valued added by industry, policymakers are essentially missing crucial information to the tune of PKR1.6 trillion of GDP (Table 1) that might influence cyclical changes in industrial production.

Another shortcoming of LSM is that it represents primarily private sector industry, whereas in Pakistan, the public sector has a large presence in industrial production. For example, electricity and gas distribution account for nearly 7.4 percent of total industry. Realizing this need, we have developed a new industrial production index (IPI) for Pakistan, which addresses the aforementioned issues.

The contributions of this study are twofold. First, we fill a gap in the literature by providing an index of industrial production that not only enhances the coverage that LSM lacks but is also based on a higher frequency than most estimates of national accounts. Second, we attempt to forecast the proposed index using a range of econometric models in search of a true model. To this end, we use several models to forecast the proposed IPI. These models take into account conditions in the real, external, and financial sectors. We believe the forecasts generated by these models may serve as useful benchmarks for policymakers and researchers alike. We have

[4] As of finalizing this article on 15 December 2020, the last available value of LSM from the PBS was for the month of October 2020.

evaluated the forecasting performance of our models to show that they can be useful in obtaining forecasts.

The rest of the article is organized as follows. Section 2 provides a literature review. Section 3 describes our estimation methodology for the proposed IPI for all sectors and subsectors of industry. Section 4 analyzes and tests the newly estimated IPI. Section 5 discusses the data required to forecast the index. Section 6 specifies the models used to forecast the IPI. Forecast evaluations are discussed in Section 7, while Section 8 provides concluding remarks.

2. Literature Review

It is important to note prior efforts to make national income accounts available more frequently. These are helpful to policymakers for analyses of cyclical change in economic activity. A major contribution in this respect is that of Arby (2008) who has compiled the national accounts of Pakistan and provides a thorough account of the literature. Earlier studies, including Bengaliwala (1995) and Kemal and Arby (2004), are also notable for their efforts to enhance the frequency of national accounts. These studies generate national income accounts using 1980/81 as the base year. Similarly, Arby and Batool (2007) have quarterized the overall gross fixed capital formation. More recently, Hanif et al. (2013) provide estimates for quarterly private and public sector gross fixed capital formation separately, in addition to estimates of quarterly overall gross fixed capital formation.

While these studies have helped make the national accounts available at a higher frequency, their focus has been on estimating the aggregate national accounts and not industrial production. We have already established the need to track industrial production separately. To reiterate, it is critical to understand short-term changes in economic activity. This is best done by looking at industrial production in an economy. It is also the practice in the world of practical policymaking. This study fills this gap by providing monthly estimates of industrial production for the period 1990M7:2018M6.

3. Estimation of Monthly IPI

Industry accounts for 23 percent of total GDP at producer prices released provisionally by the PBS (see Table 1). It is subdivided into four major sectors. Within industry, manufacturing accounts for 56.9 percent, followed by mining and quarrying (M&Q) at 11.6 percent. Electricity generation and distribution and gas distribution takes up 7.4 percent, while construction constitutes 11.9 percent of the total value added by industry. Manufacturing within industry is further divided into three subsectors: LSM, small-scale manufacturing (SSM) and slaughtering, which account for 79.6, 13.8 and 6.5 percent of activity, respectively.

In the following subsections, we define a formal setup for estimating the IPI for each subsector within industry. In doing so, we rely on the methodology proposed by Arby (2008), which was subsequently used by Hanif et al. (2013) to estimate the quarterly GDP of Pakistan, using a production approach.

The monthly IPI is the sum of the monthly value added by each subsector, as given by equation (1):

$$IPI_m = \left[\frac{M\&Q_{m,y} + LSM_{m,y} + SSM_{m,y} + VAS_{m,y} + VAC_{m,y} + VAE_{m,y}}{\frac{1}{12}\sum_{m=1}^{12}(M\&Q_{m,2005-06} + LSM_{m,2005-06} + SSM_{m,2005-06} + VAS_{m,2005-06} + VAC_{m,2005-06} + VAE_{m,2005-06})}\right] * 100 \quad (1)$$

In equation (1), M&Q denotes mining and quarrying, LSM refers to large-scale manufacturing, SSM denotes small-scale manufacturing, VAS and VAC refer to the value added by the slaughtering and cement sectors, respectively, VAE is the value added by the energy sector, m denotes the value added by that sector in a given month, and y indicates the yearly value added by a sector. The methods employed to obtain monthly values for each sector are discussed below.

3.1. *Mining and Quarrying*

The M&Q subsector consists of 33 minerals, for which the PBS provides production figures in its monthly statistical bulletin. For each mineral, we add its respective production value over the 12 months of each fiscal year. We divide the resulting figure by the yearly M&Q estimate given by Hanif et al. (2013) to obtain the subsector's monthly weight in industry, and then weight the yearly figure for M&Q available from the PBS. Formally, this is expressed as follows:

$$M\&Q_{m,y} = \left[w_{m,y} * M\&Q_y\right] \quad (2)$$

where $w_{m,y}$ is the weighted average of the share of each item in production, calculated as follows:

$$W_{m,y} = \sum_{i=1}^{33}\left[\frac{P_{i,m,y}}{\sum_{m=1}^{12} P_{i,y}} * SGVA_i\right] \quad (3)$$

In the expression above, $SGVA_i$ is the share of gross value added by commodity *i* in base year 2005/06.

3.2. Manufacturing

Manufacturing is divided into three sectors: LSM accounts for 79.6 percent of total manufacturing, SSM for about 13.8 percent, and slaughtering for 6.5 percent. The methodology for estimating the value of each subsector is discussed below.

3.2.1. Large-Scale Manufacturing

The monthly value of LSM is calculated by estimating the monthly weights of LSM and multiplying by annual LSM production. The weights are determined by using the monthly and annual LSM indices available in the monthly statistical bulletin published by the PBS. We divide the monthly LSM index by its respective annual LSM index to obtain the weight and then multiply that weight by annual LSM production as expressed in equation (4) below:

$$LSM_{m,y} = \left[\frac{LSMI_{m,y}}{\sum_{m=1}^{12} LSMI_{m,y}}\right] * LSM\ (Rs\ in\ millions)_y \tag{4}$$

where $LSM_{m,y}$ is the value added by LSM in month m and fiscal year y and $LSM_{m,y}$ is the monthly index of LSM for month m in year y.

3.2.2. Small-Scale Manufacturing

The value of SSM is obtained in the same manner as for LSM, using the following expression:

$$SSM_{m,y} = \left[\frac{LSMI_{m,y}}{\sum_{m=1}^{12} LSM_{m,y}}\right] * SSM\ (Rs\ in\ millions)_y \tag{5}$$

We use the same weights as for LSM mainly because no separate SSM index is available in the monthly statistical bulletin published by the PBS. More importantly, the SSM is likely to exhibit similar behavior since developments in manufacturing will affect both LSM and SSM downstream.

3.3. Slaughtering

We cannot use the aforementioned methods to estimate monthly production for the slaughtering since no such index is available.

Accordingly, we rely on the quarterly weights estimated by the PBS (2002) and linearly decompose these into monthly weights. The respective quarterly weights are 0.18 for Q1, 0.25 for Q2, 0.35 for Q3, and 0.22 for Q4. We implement these by dividing the weights by 3 to obtain monthly weights and multiplying them by the annual value added under slaughtering (available from the PBS) to obtain the monthly value added, using the following expression:

$$S_{m,y} = w_m * SSM\ (Rs\ in\ millions)_y \tag{6}$$

where w_m is the weight for each month of the fiscal year. For m = 1, 2, 3, the respective weight is 0.06. For m = 4, 5, 6, the weight applied is 0.083. For m = 7, 8, 9, the weight used is 0.117 and for m = 10, 11, 12, the weight used is 0.073.

3.4. Construction

Cement production is a major component of the construction sector. We use cement production C with a three-month lag to estimate the total value added by construction, using the following expression:

$$VAC_{m,y} = \left[\frac{PoC_{m,z,y}}{\sum_{m=1}^{12} PoC_{m,z,y}}\right] * VAC\ (Rs\ in\ millions)_y \tag{7}$$

where $VAC_{m,y}$ is the value added by the construction sector in month m of fiscal year y. $PoC_{m,z,y}$ denotes the production of cement in month m of year z. Year z follows year y with a three-month lag (April to March). VAC_y denotes the value added by the construction sector for year y (available from the PBS).

3.5. Energy

Energy accounts for 7.4 percent of total industrial production and comprises electricity generation and distribution and gas distribution. The PBS provides monthly data for total electricity generation and gas distribution, but the annual value added of these components is provided as a combined value. Since these units of energy are different, they cannot be summed up to calculate monthly weights. This anomaly is dealt with by using the production values of gas and electricity expressed as a common unit, the TOE, available in the Energy Yearbook for the period 1999–2017.

For each year, the weights for gas and electricity are estimated using their production values for the period 1999–2017 and averaging these over the same period. This results in weights of 0.704 for electricity and 0.296 for gas. Next, the monthly shares of gas and electricity are calculated for each month of the fiscal year for the sample period and multiplied by their respective shares in the Energy Yearbook to arrive at a common share for each fiscal year in the sample. These weights are then used to estimate the monthly value added by energy in the industrial sector, using equation (8) below:

$$W_{m,y} = \left[\frac{PoE_{m,y}}{\sum_{m=1}^{12} PoE_{m,y}} * SoE_y\right] + \left[\frac{PoG_{m,y}}{\sum_{m=1}^{12} PoG_{m,y}} * SoG_y\right] \quad (8)$$

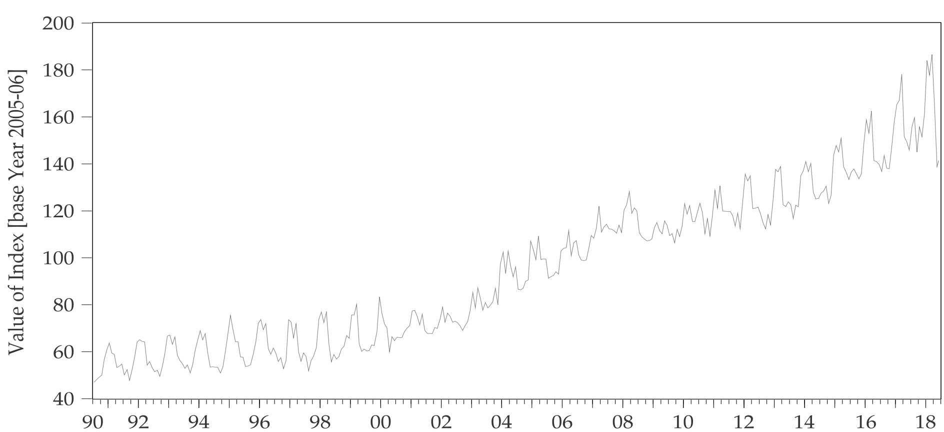
where $w_{m,y}$ is the weighted average of the share of electricity and gas. SoE_y and SoG_y are the respective shares of electricity and gas in total energy production in year y. These weights are used to arrive at an energy index using the expression below:

$$E\&G_{m,y} = \left[w_{m,y} * E\&G_y\right] \quad (9)$$

4. Analysis and Testing of Estimated IPI

Putting together all the subcomponents of equation (1) results in the IPI plotted in Figure 1. We also carry out various statistical and validation tests to lend credence to the calculated index.

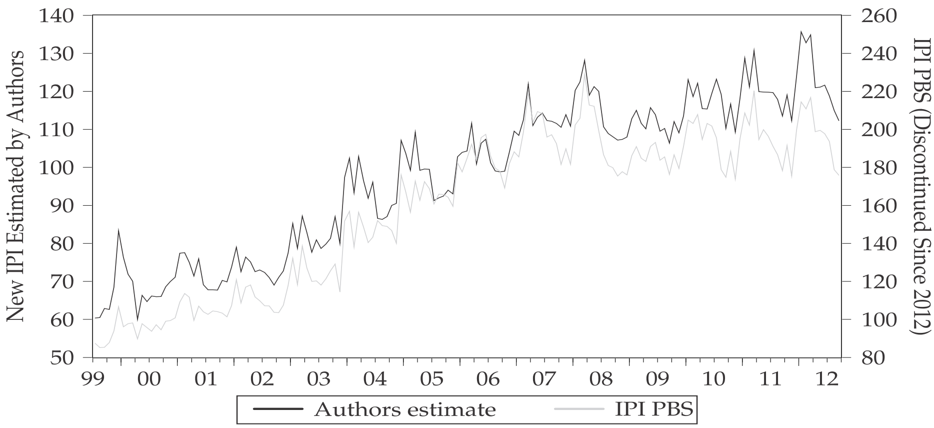
Figure 1: New Industrial Production Index Estimated by Authors



The sum of the monthly estimates of value added by industry adds up to the sectoral value added by industry available in the national income accounts. This can be verified using the actual series provided in the Appendix. Next, we plot our estimated IPI against the IPI officially published earlier by the PBS. We convert the old PBS series for base year

2005/06 and plot it against our calculated IPI. The juxtaposed series are given in Figure 2. Both seem to be highly correlated at a value of 0.97 (Table 2). The results in the table also show the Pearson correlation coefficient for testing for zero-correlation. The associated p-value shows that the strong correlation of 0.97 is statistically significant at 5 percent.

Figure 2: Comparison of PBS IPI and computed IPI



Note: The PBS's IPI series is available for the period 1999M7 to 2012M12, but has been discontinued since then. Therefore, the comparison is plausible only for this period.

We test for persistence by estimating the coefficients of both series for the common sample period, using autoregressive moving-average (ARMA) (1,1) models. The results in Table 2 show that the magnitude of the two coefficients is nearly the same and statistically significant, thus indicating strong persistence. We also regress the year-on-year (YOY) changes in the PBS series on our new IPI, using least squares. The coefficient is a strong 0.95 and statistically significant, with the model explaining 57 percent of the variation.

Table 2: Statistical tests performed on computed IPI

	IPI estimated by authors	IPI estimated by PBS	Sample period
ρ1	0.975*		1999M7: 2012M9
P-value	0.000		
Test of persistence			
AR (1)	0.984	0.986	1999M7: 2012M9
MA (1)	-0.251	-0.273	1999M7: 2012M9
R2	0.929	0.940	
Durbin-Watson	2.032	2.038	1999M7: 2012M9

Note: * = statistical significance at 5 percent, 1 Pearson correlation coefficient.
Source: Authors' estimates.

Overall, the computed series is strongly correlated with the old PBS series in the common sample and their AR (1) coefficients are high and statistically significant, indicating persistence. The YOY changes in the new series also explain the YOY changes in the old PBS IPI to a great degree.

5. Modeling and Forecasting of IPI

In this section, we identify determinants of industrial production in the empirical literature, which were used to develop specifications for various single- and multiple-equation econometric models.

5.1. Determinants of Industrial Production

Hettige et al. (1992) find that toxic intensity, environmental regulation, and trade policy (quotas and tariffs) negatively affect industry. Marchetti and Parigi (2000) use indicators of industrial activity, such as electricity consumption, to forecast industrial production.[5] Bruno and Lupi (2004) use the quantity of goods transported by railway and an index based on the opinions of entrepreneurs about production prospects to forecast industrial production. Jiranyakul (2006) studies the impact of international oil prices, the real exchange rate, and real money supply on Thailand's industrial production. Dutta and Ahmed (2004) show that there exists a unique long-run relationship between the aggregate growth function of industrial value added and its major determinants of the real capital stock, the labor force, real exports, the import tariff collection rate, and the secondary school enrolment ratio.

Tabak and Feitosa (2010) investigate whether the term spread can help predict the path followed by industrial production. Their results suggest that the yield spread contains relevant information for explaining industrial production. Enu, Hagan and Attah-Obeng (2013) identify real petroleum prices and the real exchange rate as negatively affecting industrial production, while imports of goods and services and government spending are positive factors influencing industrial production. Mohsen, Chua and Sab (2015) argue that industrial output is positively related to capital, manufactured exports, population, and agricultural output, but negatively related to oil prices. Jain, Nair and Jain (2015) find that foreign direct investment (FDI), imports and exports are important determinants of the industrial components of India's GDP.

[5] A theoretical rationalization of the use of electricity consumption data along with convincing empirical evidence of its reliability in forecasting industrial production has been established by Bodo and Signorini (1987) and Bodo, Cividini and Signorini (1991).

In the case of Pakistan, the literature on industrial production is limited. Perhaps the most relevant study is by Ajmair and Hussain (2017), who attempt to explain industrial production in an autoregressive distributed lag (ARDL) setup with external debt shocks, FDI, GDP, exports, inflation, and personal remittances. They find that trade and personal remittances are positive determinants of industrial production in Pakistan. Hussain et al. (2018) use a range of macroeconomic variables to forecast LSM.[6] However, they take into account only LSM as a proxy for industrial production (a fact they specifically mention) rather than the actual IPI, which was unavailable. Most of their variables have already been mentioned above. Others include total tax collection, credit, LSM, the wholesale price index (WPI), and interest spreads. Variables that could be important determinants—but for which no empirical support is found— include global GDP and unemployment.

In this study, it is not possible to consider all variables due to the issue of degrees of freedom. Also, population, labor force participation or unemployment, and secondary school enrolment are essentially various proxies for the main variable—labor supply. In such a situation, we have picked only those variables that best explain the model or are based on behavioral relationships that are most suitable in the case of Pakistan. Table 3 defines and classifies the variables used and gives their data sources. The classification is used to specify different models in later sections.

Table 3: List of variables and sample period

Variable	Definition and unit	Code	Sample	Source
Real sector				
Industrial production	Index estimated by authors	IPI	1990M7:2018M6	Authors' estimates
Electricity consumption	Electricity consumption index for industry	ECI	1990M7:2018M6	Authors' calculation using PBS data
Income	Paid-up capital, retaining earnings, preference shares of all listed nonfinancial corporates (PKR million)	CAPEX	1990M7:2018M6	SBP

[6] LSM is used as a proxy for GDP in Pakistan for both policy and research purposes. It is expected to be closely related to industrial production. Ex ante, we expect the IPI to be strongly correlated with LSM and therefore some of the variables considered by the authors may also be important determinants of this study.

Variable	Definition and unit	Code	Sample	Source
Tax	Tax provision/expenses (PKR million)	TAX	1990M7:2018M6	SBP
Labor	Labor force participation rate (%)	LFP	1990M7:2018M6	Authors' estimates
Price	Wholesale price index	P	1991M7:2018M6	SBP
External sector				
Exchange rate	Monthly average of USD/PKR exchange rate (USD/PKR)	ER	1991M7:2018M6	SBP
Foreign investment	Net figure from BOP financial account (USD million)	FDI	1997M7:2018M6	SBP
Int'l oil price	Average of Brent, Dubai and WTI (USD/barrel)	OP	1997M7:2018M6	Bloomberg
World output	US industrial production index: Total index, index 2012 = 100, monthly, not seasonally adjusted	WO	1990M7:2018M6	FRED
Exports	External trade, goods, value of exports FOB (USD million)	X	1990M7:2018M6	IMF
Financial/monetary sector				
Money supply	Monetary aggregate M2 (PKR million)	M2	1990M7:2018M6	SBP
Interest rate	Monthly average of offer side of 12-month tenor (PKR million)	KIBOR	1990M7:2018M6	SBP
Credit	Total finance = current + noncurrent liabilities (adjusted) on balance sheets of all listed nonfinancial firms on KSE (PKR million)	CR	1990M7:2018M6	SBP

5.2. *Data*

Before we specify models and make forecasts, it is important to discuss issues pertaining to the data used. In time-series econometrics, longer series with higher frequencies are crucial to better forecasting. With this in mind, we have exhausted all sources and avenues to ensure that the maximum data for our sample period (1990M7:2018M6) was available to ensure the accuracy and reliability of the forecasts generated by the models. However, as with any developing country, reliable data in its desired form is not easily available for Pakistan. Accordingly, we have tried to ensure that data for all variables used in the estimation was obtained from reliable sources. Where data was not available, proxies were used to estimate the series, based on ancillary information.

5.2.1. External Sector

We take monthly data on exports and imports of goods from the International Financial Statistics (IFS) database maintained by the International Monetary Fund. For exports, we take the value of goods exported (in USD million) free on board (FOB); for imports, we take the value of goods imported (in USD million) inclusive of cost, insurance and freight (CIF). Note that we exclude the import and export of services since our purpose is to explain industrial production. The IFS data was available only until 2017M9 and thus the remaining data until 2018M6 has been taken from monthly statistics on external trade published by the PBS.[7] For FDI, we consider net flows instead of solely inflows because industrial production could be affected by both inward and outward investment flows. While the IFS database provides this data for Pakistan on a quarterly basis, the State Bank of Pakistan (SBP) makes FDI data available for as far back as 1997M7. Therefore, we rely on the latter instead.

Although world GDP would be an appropriate indicator of world output, it is not available on a monthly basis. Therefore, we use the US IPI as a proxy for world output, this index being readily available for the period 1990M7:2018M6. For commodity prices, we take the average US dollar per barrel spot prices of the Brent, Dubai Fateh and West Texas Intermediate indices available from the World Bank. Data on the USD/PKR market exchange rate and nominal effective exchange rate was readily available from the IFS database for the entire sample period.

Since the first 11 months will be lost, as all the variables represent YOY growth, and data for FDI is available from 1997M7, the sample for this sector is restricted to 1998M7:2018M6. This comes to 240 months, of which 228 have been used for estimation and the remaining 12 months for forecasting.

---

[7] http://www.pbs.gov.pk/trade-detail?page=2

5.2.2. Real Sector

For electricity consumption, the most relevant indicator is the percentage of electricity consumed by industry. The share of industry in total electricity consumption is available from Haver/NEPRA for the period 2004–17. These shares have been interpolated constantly to arrive at monthly shares. For the years prior to 2004, we simply average the annual shares for 2004–07 and apply these retrospectively to calculate the monthly shares. These shares are applied to monthly data on the electricity consumed by industry (in million kWh) available in the PBS's monthly statistical bulletin. Using 2005/06 as the base year, we calculate the index of industrial electricity consumption for the period 1990M7:2018M6.

For monthly investment data, we had two options. One was to use the quarterly shares used by Hanif and Malik (2013) and interpolate them constantly to arrive at monthly investment on the right-hand side of the GDP. This data is available for 2000M7:2018M6. However, a more suitable proxy for industrial investment is the capital expenditure (CAPEX) incurred by corporates. We estimate the CAPEX for the industrial sector using annual data on retained earnings, paid-up capital and preference shares (available from the SBP),[8] while using the monthly shares of industry estimated in Section 3 above to yield monthly CAPEX data.

For a proxy for labor (L), we rely on annual data on labor force participation.[9] To convert this variable to monthly data, we use the annual share of industry and divide it equally over 12 months of the fiscal year to yield monthly shares. These weights are then applied to the labor force participation rate to arrive at monthly labor force participation (LFP) for industry. For prices, the relevant indicator is the WPI, available from the SBP for the period 1991M7:2018M6 with a single base year.

For income, the best proxy is corporate income, for which we use the net profit before taxation (in PKR million) of nonfinancial corporates, available from the SBP's monthly and annual statistical bulletins. Since this data is available on an annual basis, we use the same methodology as for the IPI to arrive at monthly estimates. We estimate the shares of manufacturing, construction and energy (gas and electricity) and use these to generate the weighted monthly shares of each sector. We then apply these shares to the

---

[8] See https://www.sbp.org.pk/publications/index2.asp for the monthly statistical bulletin and five-yearly handbook of statistics on the Pakistani economy.

[9] Available from Haver. The source cited is the PBS and Ministry of Planning, Development and Reforms.

net profit before tax of nonfinancial firms to arrive at their monthly corporate income. The data on monthly taxes paid by industry is estimated similarly, using the total tax provision/expense available from the same source.

As with the external sector, the first 11 months are lost since we are interested in YOY growth. Further, the WPI is available from 1991M7, and thus the sample for this sector is restricted to 1992M7:2018M6. This gives us 312 months of data, of which 300 months' data is used for estimation and the remaining 12 months for forecasting.

5.2.3. Financial Sector

The most relevant indicator of the price of corporate credit is the offer side of the 12-month KIBOR, which banks usually use as a reference rate to price credit to the private sector. However, this data is available only from 2004M4. Alternatively, we could use the auction rate of 12-month T-bills, for which the data goes back to 1998M7. However, there are 108 instances in which bids were rejected and thus the data for this indicator is not available.

There are two sources of data on credit to industry. The first is the SBP, which provides monthly data on private sector credit, available from 2006M6. The SBP also provides annual balance sheet data for nonfinancial firms listed on the KSE in its annual as well as monthly statistical bulletins. We rely on this source for two reasons. First, it gives data on the demand side of credit, which is preferable. Second, a longer time-series is available. We estimate the total finance received by industry by aggregating the current and noncurrent liabilities of nonfinancial firms, adjusted for employees' benefit obligations. To convert this into monthly data, we use the methodology described in Section 3. Accordingly, the data on finance available to industry is for 1990M7:2017M6. Data on the money supply (M2) is available from the SBP for the period 1990M7:2018M6.

As with the other two sectors, the first 11 months will be lost since we are interested in YOY growth. Further, the KIBOR is available from 2004M7 and thus the sample for this sector is restricted to 2005M7:2017M6. Of the total 157 observations, we use 145 for estimation and the remaining 12 to obtain forecasts. A list of final variables used, along with their data sources, is given in Table 3.

6. Methodology for Forecasting IPI

This section discusses the specifications of single- and multiple-equation econometric models capable of forecasting the IPI. These models are specified based on the behavioral relationships between the determinants of industrial production identified in Section 5.1.

6.1. ARIMA Model

We estimate an autoregressive integrated moving-average (ARIMA) model using equation (10) below. The lags of the autoregressive term and moving average are finalized based on the Akaike information criterion (AIC). The maximum lag structure is set at 12 each for the autoregressive and moving-average terms. These are then selected using a generalized-to-specific approach based on the lowest AIC. The final estimated model is as follows:

$$IPI_t = \alpha + \sum_{i=1}^{p} \beta_i IPI_{t-i} + \sum_{j=1}^{q} \alpha_j \varepsilon_{t-j} + \varepsilon_t \tag{10}$$

6.2. ARDL Models

Since the variables considered in this study can be integrated either of order 0 or 1, an ARDL model may be appropriate for explaining the IPI. We estimate three models reflective of real, external, and financial conditions as affecting industrial production.

In the real sector, growth in the IPI could be explained by a distributed lag component of a set of explanatory variables in the following setup:

$$GIPI_t = c_1 + \alpha_i \sum_{i=1}^{n} GIPI_{t-i} + \phi_i \sum_{i=1}^{n} EC_{t-i} + \beta_j \sum_{j=1}^{n} CAPEX_{t-j} - \gamma_k \sum_{k=1}^{n} P_{t-k} + \delta_l \sum_{l=1}^{n} LFP_{t-l} - \varphi_l \sum_{l=1}^{n} TAX_{t-l} + \varepsilon_{1t} \tag{11}$$

where GIPI is the growth in the IPI, EC is electricity consumption, CAPEX is capital expenditure, P is the WPI, LFP is labor force participation, and TAX is taxation.

Similarly, IPI growth could also be affected by conditions in the external sector. For this, we run a second ARDL model of the form expressed in equation (12):

$$GIPI_t = c_1 + \alpha_i \sum_{i=1}^{n} GIPI_{t-i} + \rho_i \sum_{i=1}^{n} X_{t-i} + \tau_k \sum_{k=1}^{n} ER_{t-k} + \theta_l \sum_{l=1}^{n} OP_{t-l} + \vartheta_m \sum_{m=1}^{n} FDI_{t-m} + \omega_o \sum_{o=1}^{n} WO_{t-o} + \varepsilon_{2t} \quad (12)$$

where ER is the exchange rate, OP is oil price and WO is world output.

Lastly, growth in the IPI could also be affected by monetary conditions. For this, we estimate the model in equation (13) below:

$$GIPI_t = c_1 - \alpha_i \sum_{i=1}^{n} KIBOR_{t-i} + \beta_j \sum_{j=1}^{n} M2_{t-j} + \gamma_k \sum_{k=1}^{n} CR_{t-k} + \varepsilon_{3t} \quad (13)$$

where M2 is the money supply, KIBOR is the Karachi interbank offer rate and CR is credit.

In the models above, the maximum lag structure has been set at 12 for both dependent and independent variables. The final lags are then selected using a general-to-specific approach based on the lowest AIC.

6.3. Vector Autoregressive Model

Sims (1980) criticized the large-scale macroeconomic models of the time because of the strong restrictions they imposed. The problem with these models was that they were highly specified with strong assumptions concerning the dynamic nature of the relationship between macroeconomic variables. He argued that the models were largely inconsistent with the notion that economic agents take the effect of today's choices on tomorrow's utility into account, which later became known as the Sims critique. The critique explained that, in a world with rational, forward-looking agents, no variable could be deemed exogenous. Sims proposed vector autoregressive (VAR) models as an alternative, which allowed one to model macroeconomic data without imposing strong restrictions. Since then, VAR models have become the mainstay of modern applied macroeconomics and it makes sense to use the VAR setup to forecast the IPI.

The VAR specifications we use are not completely devoid of the theory. They have been developed using the empirical literature already discussed. We estimate three kinds of VAR models, specified below.

6.3.1. Industrial Production and Financial Conditions (VAR 1)

Industries are capital-intensive and thus financial factors will be important. Conducive financial conditions, in the form of availability of credit for both working and fixed investment at reasonable interest rates, could be precursors to industrial production. Any change in the money supply (MS) will lead to changes in credit supply (CR), which will eventually affect market interest rates (KIBOR). A VAR model specifying such financial conditions is given below:

$$GMS_{t+1} = E_t[GMS_{t+1}] + \varepsilon_{t+1}^{GMS}$$

$$GCR_{t+1} = E_t[GCR_{t+1}] + \alpha_1\varepsilon_{t+1}^{GMS} + \varepsilon_{t+1}^{GCR}$$

$$KIBOR_{t+1} = E_t[KIBOR_{t+1}] + \alpha_2\varepsilon_{t+1}^{GMS} + \alpha_3\varepsilon_{t+1}^{GCR} + \varepsilon_{t+1}^{KIBOR}$$

$$GIPI_{t+1} = E_t[GIPI_{t+1}] + \alpha_4\varepsilon_{t+1}^{GMS} + \alpha_5\varepsilon_{t+1}^{GCR} + \alpha_6\varepsilon_{t+1}^{KIBOR} + \varepsilon_{t+1}^{GIPI}$$

where E_t is the conditional expectation operator and the α terms are the impulse response coefficients. A VAR model of the form above gives us the following recursive structural VAR system:

$$Y_{T+1} = AY_T + B\varepsilon_{t+1} \tag{14}$$

where $Y = (GMS, GCR, KIBOR)$, $\varepsilon = (\varepsilon^{GMS}, \varepsilon^{GCR}, \varepsilon^{KIBOR}, \varepsilon^{GIPI})$ and $B = \begin{bmatrix} 1 & 0 & 0 & 0 \\ \alpha_1 & 1 & 0 & 0 \\ \alpha_2 & \alpha_3 & 1 & 0 \\ \alpha_4 & \alpha_5 & \alpha_6 & 1 \end{bmatrix}$

6.3.2. Industrial Production and External Sector Conditions (VAR 2)

Conditions in the external sector are also expected to affect industrial production. Changes in global oil price (GOP) will affect the growth in global output (GWO), creating investment opportunities (FDI), which in turn will affect exports (X). This will have an effect on the equilibrium exchange rate (ER), which will eventually determine industrial production (IPI). Such a VAR model is specified as follows:

$$GOP_{t+1} = E_t[GOP_{t+1}] + \varepsilon_{t+1}^{GOP}$$

$$GWO_{t+1} = E_t[GWO_{t+1}] + \alpha_1\varepsilon_{t+1}^{GOP} + \varepsilon_{t+1}^{GWO}$$

$$GFDI_{t+1} = E_t[GFDI_{t+1}] + \alpha_2\varepsilon_{t+1}^{GOP} + \alpha_3\varepsilon_{t+1}^{GWO} + \varepsilon_{t+1}^{GFDI}$$

$$GX_{t+1} = E_t[GX_{t+1}] + \alpha_4\varepsilon_{t+1}^{GOP} + \alpha_5\varepsilon_{t+1}^{GWO} + \alpha_6\varepsilon_{t+1}^{GFDI} + \varepsilon_{t+1}^{GX}$$

$$GER_{t+1} = E_t[GER_{t+1}] + \alpha_7\varepsilon_{t+1}^{GOP} + \alpha_8\varepsilon_{t+1}^{GWO} + \alpha_9\varepsilon_{t+1}^{GFDI} + \alpha_{10}\varepsilon_{t+1}^{GX} + \varepsilon_{t+1}^{GER}$$

$$GIPI_{t+1} = E_t[IPI_{t+1}] + \alpha_{11}\varepsilon_{t+1}^{GOP} + \alpha_{12}\varepsilon_{t+1}^{GWO} + \alpha_{13}\varepsilon_{t+1}^{GFDI} + \alpha_{14}\varepsilon_{t+1}^{GX} + \alpha_{15}\varepsilon_{t+1}^{GER} + \varepsilon_{t+1}^{IPI}$$

where E_t is the conditional expectation operator and the α terms are the impulse response coefficients. A VAR model of this form gives us the following recursive structural VAR system:

$$Y_{T+1} = AY_T + B\varepsilon_{t+1} \tag{15}$$

where $Y = (GOP, GWO, GFDI, GM, ER, GIPI)$, $\varepsilon = (\varepsilon^{OP}, \varepsilon^{WO}, \varepsilon^{FDI}, \varepsilon^{X}, \varepsilon^{ER})$ and

$$B = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ \alpha_1 & 1 & 1 & 0 & 0 & 0 \\ \alpha_2 & \alpha_3 & 1 & 0 & 0 & 0 \\ \alpha_4 & \alpha_5 & \alpha_6 & 1 & 0 & 0 \\ \alpha_7 & \alpha_8 & \alpha_9 & \alpha_{10} & 1 & 0 \\ \alpha_{11} & \alpha_{12} & \alpha_{13} & \alpha_{14} & \alpha_{15} & 1 \end{bmatrix}$$

6.3.3. Industrial Production and Real Sector Conditions (VAR 3)

This specification takes into account conditions in the economy's real sector. Variables that are expected to affect industrial production include the tax rate (TX), growth in electricity consumption (GEC), investment growth (GCAPEX), and growth in the WPI (GP). The model is specified as follows:

$$TX_{t+1} = E_t[TX_{t+1}] + \varepsilon_{t+1}^{TX}$$

$$GEC_{t+1} = E_t[GEC_{t+1}] + \alpha_1\varepsilon_{t+1}^{TX} + \varepsilon_{t+1}^{GEC}$$

$$GCAPEX_{t+1} = E_t[GCAPEX_{t+1}] + \alpha_2\varepsilon_{t+1}^{TX} + \alpha_3\varepsilon_{t+1}^{GEC} + \varepsilon_{t+1}^{GCAPEX}$$

$$GP_{t+1} = E_t[P_{t+1}] + \alpha_4\varepsilon_{t+1}^{TX} + \alpha_5\varepsilon_{t+1}^{GEC} + \alpha_6\varepsilon_{t+1}^{GCAPEX} + \varepsilon_{t+1}^{GP}$$

$$GIPI_{t+1} = E_t[GIPI_{t+1}] + \alpha_7\varepsilon_{t+1}^{TX} + \alpha_8\varepsilon_{t+1}^{GEC} + \alpha_9\varepsilon_{t+1}^{GCAPEX} + \alpha_{10}\varepsilon_{t+1}^{GP} + \varepsilon_{t+1}^{GIPI}$$

where E_t is the conditional expectation operator and the α terms are the impulse response coefficients. A VAR model of the form above gives us the following recursive structural VAR system:

$$Y_{T+1} = AY_T + B\varepsilon_{t+1} \tag{16}$$

where $Y = (TX, GEC, GINC, GINV, GP, GIPI)$, $\varepsilon = (\varepsilon^{TX}, \varepsilon^{GEC}, \varepsilon^{GCAPEX}, \varepsilon^{GP}, \varepsilon^{GIPI})$, and

$$B = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ \alpha_1 & 1 & 0 & 0 & 0 \\ \alpha_2 & \alpha_3 & 1 & 0 & 0 \\ \alpha_4 & \alpha_5 & \alpha_6 & 1 & 0 \\ \alpha_7 & \alpha_8 & \alpha_9 & \alpha_{10} & 1 \end{bmatrix}$$

6.3.4. Bayesian Vector Autoregressive Model

Using the standard VAR model specified in equations (14), (15) and (16), we run its Bayesian vector autoregressive (BVAR) variant for the real, external, and financial sectors. Some discussion on the BVAR model is warranted here. The difference between a standard VAR and BVAR model is that the latter's parameters are treated as random variables and assigned prior probabilities. BVAR methods are used frequently to deal with the problem of over-parameterization typical of standard VAR models, since Bayesian priors provide a logical and consistent method of imposing parameter restrictions.[10]

The type of prior used in estimation is very important. For our purposes, we have used the ubiquitous Litterman/Minnesota prior, which specifies hyper-parameters using four scalars $\mu_1, \lambda_1, \lambda_2$, and λ_3. λ_1 denotes the overall tightness of the variance (of the first lag) and controls the relative importance of sample and prior information. If λ_1 is small, prior information dominates the sample information. λ_2 represents the relative tightness of the variance of other variables. Setting $\lambda_2 = 0$ implies that the VAR is collapsed to a vector of univariate models. $\lambda_3 > 0$ represents the relative tightness of the variance of the lags.

7. Model Estimation and Forecasting

The first step before estimating any model is a critical overview of the data itself. We have already discussed measurement issues pertaining to the sample in Section 5.2. Table 4 provides basic descriptive statistics for the sample used.

---

[10] See Litterman (1986); Doan, Litterman and Sims (1984); Sims and Zha (1998).

Table 4: Descriptive statistics for sample data, 1990M7–2018M6

Var	Definition and unit	Mean	Median	Max.	Min.	SD	N
CR	Liabilities of the firm (PKR million)	111,395	63,458	314,667	9,565	89,085	336
ECI	Electricity consumption index	71	68	141	31	23	336
ER	Exchange rate (PKR/USD)	64	60	119	22	27	336
FDI	Net FDI flows from BOP (USD million)	160	111	1,263	-54	174	252
CAPEX	Capital expenditures (PKR million)	32,488	25,619	79,654	2,096	24,586	336
P	Wholesale price index	105	72	234	25	68	324
M2	Money supply (PKR million)	4,285,433	2,489,854	15,763,268	340,652	4,165,971	336
KIBOR	Offer side of 12M (%)	10	10	16	3	3	171
IPI	Industrial production index	95	92	187	47	33	336
TAX	Tax provision/expenses (PKR million)	5,061	3,555	16,338	231	4,725	336
OP	Average of DF, WTI, Brent (index)	48	35	133	10	32	336
X	Goods exported FOB (USD million)	1,219	1,113	2,613	336	571	336
WO	US industrial production (index)	91	95	109	62	13	336
L	Labor force participation rate (%)	0.52	0.53	0.58	0.46	0.04	300

Source: Authors' estimates.

It is also essential to discuss the time-series properties of the sample used in this study. Table 5 provides the results of the standard augmented Dickey-Fuller (ADF) unit root tests. The test is carried out on YOY growth in dependent and regressors with a maximum of 13 lags without including the trend or intercept. Lag selection is based on the Schwarz information criterion (SIC). The test is carried out only for the periods of estimation. The results show that all variables are stationary either at a 5 percent or 10 percent level of significance.

Table 5: Unit root test results

G = YOY	ADF test statistics	P-value	Sample period
GIPI	-9.21*	0.000	1992M7:2017M6
GKIBOR	-5.26*	0.000	2005M7:2017M6
GCAPEX	-2.72*	0.032	1992M7:2017M6
GECI	-8.31*	0.000	1992M7:2017M6
GCR	-2.79**	0.062	2005M7:2017M6
GFDI	-11.27*	0.000	1998M7:2017M6
GX	-4.04*	0.001	1998M7:2017M6
GP	-1.93**	0.052	1992M7:2017M6
GOP	-3.92*	0.002	1998M7:2017M6
GWO	-4.09*	0.001	1998M7:2017M6
GM2	-2.96*	0.040	2005M7:2017M6
GTAX	-4.78*	0.000	1992M7:2017M6
GER	-2.57**	0.100	1998M7:2017M6
GL	-3.61*	0.006	1992M7:2017M6

Note: Null = variable has a unit root. Test statistic: ADF. Lag length: SIC. * = significant at 5 percent, ** = significant at 10 percent.
Source: Authors' estimates.

Table 6 provides diagnostics for all ten estimated models. For the autoregressive (AR) model, the lag length is finalized using the automatic ARIMA model in a general-to-specific-approach. The system evaluates 169 models and finalizes the AR model ($p = 9$, $q = 11$). The AIC of the final model is a modest 5.78 and the model explains about 55 percent of the variation. We test for autocorrelation up to nine selected lags and cannot reject the null hypothesis that there is no autocorrelation up to nine lags. Note that we have only reported the Q-stat at the ninth lag.

For the ARDL models, the lags of dynamic regressors are finalized using the AIC with a maximum lag of 12 months. The final lags are selected based on the AIC using a general-to-specific approach. The Breusch-Pagan-Godfrey F-statistics show that the errors of the estimated models are all homoskedastic, except for Model 2. Similarly, we test for no autocorrelation up to the final lag, using the Breusch-Godfrey serial correlation LM test. The results are acceptable for all models except Model 3. Of these, the model reflecting financial conditions appears to best explain growth in the IPI as well as the variation in growth in the IPI.

Table 6: Time-series properties of estimated single and multiple equation time-series models

#	Model	Regressors	R	Q-stat[d]	AIC	F-stat[a]	F-stat[b]	χ^2	LRE[c]	Sample
1	AR (9, 11)	GIPI, MA	0.55	0.39	5.78	-	-	-	-	1991M7: 2017M6
2	ARDL (5, 0, 1, 1, 0, 0)	GIPI, ECI, CAPEX, LFP, TAX, P	0.41	-	5.94	2.20*	0.68	-	-	1993M7: 2017M6
3	ARDL (12, 2, 0, 0, 0, 3)	GIPI, X, ER, OP, FDI, WO	0.51	-	5.93	1.43	3.08*	-	-	1998M7: 2017M6
4	ARDL (12, 5, 1, 5)	GIPI, KBIOR, M2, CR	0.66	-	5.19	0.99	0.58	-	-	2006M4: 2017M6
5	VAR (2, 2)	GECI GCAPEX, GP, GTAX^	0.36	-	-	-	-	282.18	17.57	1993M7: 2017M6
6	VAR (2, 2, 2)	GX, GFDI, GER, GWO^, GOP^	0.34	-	-	-	-	251.97	14.98	1998M7: 2017M6
7	VAR (2, 2, 2)	GCR, GM2, GKBIOR	0.49	-	-	-	-	239.18	28.03*	2005M6: 2017M6
8	BVAR	GECI GCAPEX, GP, GTAX^	0.34	-	-	-	-	304.65	87.57*	1993M7: 2017M6
9	BVAR	GX, GFDI, GER, GWO^, GOP^	0.32	-	-	-	-	248.23	75.48*	1998M7: 2017M6
10	BVAR	GCR, GM2, GKBIOR	0.45	-	-	-	-	286.28	64.31*	2005M6: 2017M6

Note: * = significant at 5 percent.
a Breusch-Pagan-Godfrey F-statistics with null: errors are homoskedastic.
b Breusch-Godfrey serial corr. LM test F**-statistics with null: no serial correlation up to lag *h*.
c LRE-state with null: no serial correlation at lag *h*.
d No autocorrelation up to order *k*.
χ^2 is the test of VAR residual heteroskedasticity (with no cross-terms).
^ treated as exogenous in the model setup.
Source: Authors' estimates.

The diagnostics are not as strong for the VAR models. However, the critical condition that all variables should be stationary is met comfortably, as evident from the results in Table 5. As for the multivariate normality of errors and serial correlation, the results are less than perfect. For the standard VAR specification in Models 5 and 6, the LRE-stat shows that there is no autocorrelation up to the final lag used. However, this condition is not met in the BVAR case. Similarly, the condition of heteroskedasticity appears to have been violated in both the standard and Bayesian VAR models. This may be due partly to specification issues.

7.1. Evaluation of Models

To evaluate the models, we perform out-of-sample forecasts by leaving out 12 months of the latest data, from 2017M7 to 2018M6, for all ten models and then calculating their deviations from the actual IPI values for that period. The root mean squared errors (RMSEs) obtained from the ten models estimated for three different horizons (h = 3, h = 6 and h = 12) are provided in Table 7. Generally, the ARDL models outperform the others in terms of a low RMSE. This is true across all three horizons.

Table 7: RMSEs of estimated models

No.	Model	Specification	h = 3	h = 6	h = 12
1	AR (9,11)	GIPI, MA	6.12	5.00	6.05
2	ARDL (5,0,1,1,0,0)	GIPI, ECI, CAPEX, LFP, TAX, P	5.72	5.35	5.54
3	ARDL (12,2,0,0,0,3)	GIPI, GX, GER, GOP, GFDI, GWO	4.84	4.33	5.52
4	ARDL (12,5,1,5)	GIPI, KBIOR, M2, CR	3.92	3.94	4.19
5	VAR (2,2)	GECI GCAPEX, GP, GTAX^	7.24	6.41	6.45
6	VAR (2,2,2)	GX, GFDI, GER, GWO^, GOP^	7.71	6.82	6.67
7	VAR (2,2,2)	GCR, GM2, GKBIOR	6.61	5.90	6.17
8	BVAR-real	GECI GCAPEX, GP, GTAX^	7.47	6.62	6.59
9	BVAR-external	GX, GFDI, GER, GWO^, GOP^	7.96	7.08	6.84
10	BVAR-financial	GCR, GM2, GKBIOR	6.76	6.01	6.24

Note: * treated as exogenous in the model setup.
Source: Authors' estimates.

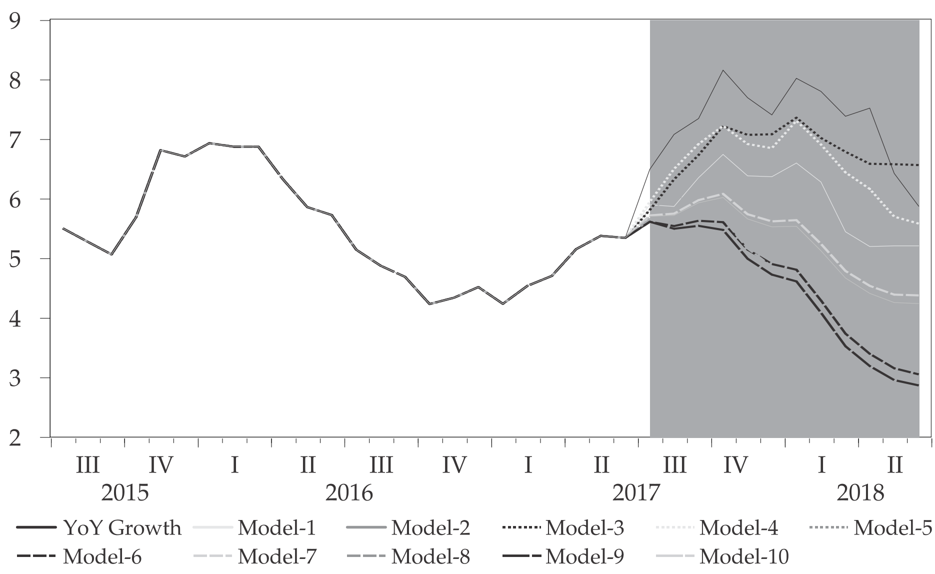
Among the ARDL models, the results show that Model 4 outperforms all other models, including the random walk model, across all horizons.[11] It takes into account financial conditions in an ARDL setup. The growth in the IPI is explained by the distributed lag components of the interest rate, monetary growth, and credit growth. The model also explains about 66 percent of the variation (see Table 6). This model includes lags of the estimated IPI, KIBOR, money supply, and credit. The coefficient of lagged growth of the IPI is 0.44 and statistically significant. Similarly, the coefficient of the KIBOR is also statistically significant but with a negative sign, indicating that a higher cost of credit is bad for industries, which tend to incur large financial costs. The coefficient of money supply is negative,

[11] In the empirical literature, it is not uncommon for single-equation models to beat BVAR model as far as forecasting accuracy is concerned. Hanif and Malik (2015) show that their ARDL model performs best among a suite of single- and multiple-equation models. Czudaj (2011) also shows that his ARDL-based forecasts of inflation obtained from the Philips curve are better than those from AR models.

which means that a higher money supply leads to inflation, thus increasing the cost of production, which causes production to be scaled back.

The coefficient of credit supply is negative, but statistically insignificant. There are several reasons why this model best explains the model. First, we do not provide a path for the independent variables used in the system: the dependent variable was estimated for data points up until 2018M6, whereas the independent variables have no actual path. This obviously reduces the error, which other systems could not have done, especially multiple setups that use a dynamic path for independent variables determined within the system. Intuitively, industrial production is usually buoyant in stable macroeconomic and financial conditions, which works through the expectations channel (Montes & Bastos, 2013). It is also worth mentioning that, as the horizon increases, the forecast accuracy of the models deteriorates generally. This is acceptable since uncertainty increases in the distant future. The combined paths of all the forecasts can be seen in Figure 3.

Figure 3: Comparison of 12-month moving average of YOY percentage growth of actual and forecast IPI (various models)



8. Conclusion

In the absence of an index to track industrial production, policymakers in Pakistan make do with LSM, which accounts for only 10 percent of the total value added to GDP even though industry overall accounts for 23 percent. We have tried to estimate an IPI that covers the overall value added by industry to the national GDP.

The index is the sum of monthly value additions in the following subsectors: M&Q, LSM, SSM, slaughtering, construction and energy. M&Q consists of 33 minerals, for which we obtain production figures from the monthly statistical bulletin published by the PBS. For each mineral, we add the respective production value over the 12 months of each fiscal year. We divide the resulting figure by the yearly estimate for M&Q given by Hanif et al. (2013) to obtain its monthly weight in industry, and then weight the yearly figure for M&Q.

The monthly value of LSM is estimated by computing its monthly weights and multiplying these by annual LSM production. The weights are determined using the monthly and annual LSM indices available in the monthly statistical bulletin published by the PBS. The monthly value added for SSM is obtained similarly. For the slaughtering sector, we obtain its value added using the quarterly weights estimated by the PBS and decomposing these linearly into monthly weights. The respective quarterly weights are 0.18 for Q1, 0.25 for Q2, 0.35 for Q3, and 0.22 for Q4. We implement these by dividing the weights by 3 to obtain monthly weights and multiplying them by the annual value added under slaughtering to obtain its monthly value added.

We use cement production as a proxy for the construction sector's value addition (VAC) since it is a major component in construction. We obtain monthly shares using actual cement production data in thousand metric tons. The annual VAC is decomposed into weighted monthly value additions. For the value added by the energy sector (VAE), we estimate a common unit for electricity generation and gas distribution, the TOE, from the Energy Yearbook and then construct weights to arrive at monthly value additions.

Our proposed index closely follows the old IPI disseminated earlier by the PBS and subsequently discontinued in 2012. The correlation between the two is more than 90 percent. Our calculated IPI passes a battery of statistical and validation tests, thus supporting the methodology used.

References

Ajmair, M., & Hussain, K. (2017). Determinants of industrial sector growth in Pakistan. *Journal of Poverty, Investment and Development*, *34*, 15–23.

Arby, M.F. (2008). *Some issues in the national income accounts of Pakistan (rebasing, quarterly and provincial accounts and growth accounting)* [Unpublished doctoral dissertation]. Pakistan Institute of Development Economics.

Arby, M.F., & Batool, I. (2007). *Estimating quarterly gross fixed capital formation* (Working Paper No. 17). State Bank of Pakistan, Research Department.

Banerjee, A., Marcellino, M., & Masten, I. (2005). Leading indicators for Euro-area inflation and GDP growth. *Oxford Bulletin of Economics and Statistics*, *67*(1), 785–813.

Bean, L.H. (1946). International industrialization and per capita income. In M. Gilbert, D. Brady & S. Kuznets (Eds.), *Studies in income and wealth* (vol. 8, 119–143). National Bureau of Economic Research.

Bengaliwala, K. (1995). *Temporal and regional decomposition of national accounts of Pakistan* [Unpublished doctoral dissertation]. University of Karachi.

Bodo, G., Cividini, A., & Signorini, L.F. (1991). Forecasting the Italian industrial production index in real time. *Journal of Forecasting*, *10*(3), 285–299.

Bodo, G., Golinelli, R., & Parigi, G. (2000). Forecasting industrial production in the Euro area. *Empirical Economics*, *25*(4), 541–561.

Bodo, G., & Signorini, L.F. (1987). Short-term forecasting of the industrial production index. *International Journal of Forecasting*, *3*(2), 245–259.

Bruno, G., & Lupi, C. (2004). Forecasting industrial production and the early detection of turning points. *Empirical Economics*, *29*(3), 647–671.

Czudaj, R. (2011). Improving Phillips curve based inflation forecasts: A monetary approach for the Euro area. *Banks and Bank Systems*, *6*(2), 5–14.

Doan, T., Litterman, R., & Sims, C. (1984). Forecasting and conditional projection using realistic prior distributions. *Econometric Reviews*, *3*(1), 1–100.

Dutta, D., & Ahmed, N. (2004). Trade liberalization and industrial growth in Pakistan: A cointegration analysis. *Applied Economics*, *36*(13), 1421–1429.

Enu, P., Hagan, E., & Attah-Obeng, P. (2013). Impact of macroeconomic factors on industrial production in Ghana. *European Scientific Journal*, *9*(28), 238–254.

Hanif, M.N., Iqbal, J., & Malik, M.J. (2013). Quarterisation of national income accounts of Pakistan. *SBP Research Bulletin*, *9*(1), 1–61.

Hanif, M.N., & Malik, M.J. (2015). Evaluating the performance of inflation forecasting models of Pakistan. *SBP Research Bulletin*, *11*(1), 43–78.

Hettige, H., Lucas, R.E.B., & Wheeler, D. (1992). The toxic intensity of industrial production: Global patterns, trends, and trade policy. *American Economic Review*, *82*(2), 478–481.

Hussain, F., Hyder, S.K., & Rehman, M. (2018). *Nowcasting LSM growth in Pakistan* (Working Paper No. 98). State Bank of Pakistan.

Jain, D., Nair, K. S., & Jain, V. (2015). Factors affecting GDP (manufacturing, services, industry): An Indian perspective. *Annual Research Journal of Symbiosis Centre for Management Studies, Pune*, *3*, 38–56.

Jiranyakul, K. (2006). The impact of international oil prices on industrial production: The case of Thailand. *NIDA Economic Review*, *1*(2), 35–42.

Kemal, A.R., & Arby, M.F. (2004). *Quarterisation of annual GDP of Pakistan* (Statistical Paper No. 5). Pakistan Institute of Development Economics.

Litterman, R.B. (1986). Forecasting with Bayesian vector autoregressions—five years of experience. *Journal of Business and Economic Statistics*, *4*(1), 25–38.

Marchetti, D.J., & Parigi, G. (2000). Energy consumption, survey data and the prediction of industrial production in Italy: A comparison and combination of different models. *Journal of Forecasting*, *19*(5), 419–440.

Mohsen, A.S., Chua, S.Y., & Sab, C.N.C. (2015). Determinants of industrial output in Syria. *Journal of Economic Structures*, *4*(1), Article 19.

Montes, G.C., & Bastos, J.C.A. (2013). Economic policies, macroeconomic environment and entrepreneurs' expectations. *Journal of Economic Studies*, *40*(3).

Pakistan Bureau of Statistics. (2002). *In-house study on slaughtering*. National Accounts Division.

Sims, C.A. (1980). Macroeconomics and reality. *Econometrica*, *48*(1), 1–48.

Sims, C.A., & Zha, T. (1998). Bayesian methods for dynamic multivariate models. *International Economic Review*, *39*(4), 949–968.

Tabak, B.M., & Feitosa, M.A. (2010). Forecasting industrial production in Brazil: Evidence from a wavelet approach. *Expert Systems with Applications*, *37*(9), 6345–6351.

Tahir, A., Ahmed, J., & Ahmed, W. (2018). *Robust quarterization of GDP and determination of business cycle dates for IGC partner countries* (Working Paper No. 97). State Bank of Pakistan.

Appendix

IPI estimated by authors [base year = 2005/06]

FY	July	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun
1990	47.06	48.16	49.18	50.02	56.68	60.70	-	-	-	-	-	-
1991	50.15	52.36	47.81	52.35	57.73	64.07	63.67	59.46	58.87	53.31	53.82	54.80
1992	51.43	52.14	49.49	53.72	59.39	66.69	65.10	64.39	64.15	54.32	55.79	53.29
1993	52.86	54.39	51.01	54.46	60.60	64.93	67.06	63.10	66.32	58.46	56.30	54.88
1994	53.43	53.32	50.95	53.56	60.29	67.84	68.96	65.03	67.71	59.51	53.46	53.60
1995	53.66	53.87	54.40	59.00	64.03	72.09	75.43	69.41	64.21	64.24	57.79	57.67
1996	59.39	55.82	57.44	52.82	56.19	73.60	73.71	69.33	71.99	61.29	58.92	61.58
1997	57.97	51.88	56.19	58.07	61.81	73.83	72.71	65.86	71.87	59.90	55.98	59.49
1998	56.89	57.92	61.07	62.24	66.89	65.76	76.96	72.45	76.82	63.44	55.65	58.77
1999	60.36	60.49	62.88	62.60	68.50	83.31	75.67	75.65	80.15	63.35	60.10	61.06
2000	66.18	66.00	66.03	68.58	69.96	71.13	76.32	71.89	70.03	60.01	66.35	64.70
2001	67.78	67.77	67.71	70.29	69.89	73.74	77.43	77.54	75.01	71.43	75.95	69.12
2002	72.36	71.03	69.04	71.06	72.75	77.49	78.96	72.56	76.43	75.14	72.57	72.99
2003	78.66	79.79	81.33	87.00	80.01	97.37	85.20	78.70	87.15	82.86	77.70	80.94
2004	86.59	86.31	87.09	90.00	90.50	107.09	102.42	93.33	102.85	96.49	91.87	96.09
2005	91.25	91.98	92.45	94.02	93.04	102.82	103.67	99.24	109.20	99.23	99.56	99.50
2006	101.22	99.03	98.80	99.00	103.94	109.51	103.94	104.31	111.54	100.91	106.35	107.39
2007	112.30	112.15	111.53	110.53	113.87	110.83	108.40	112.47	121.92	111.01	113.18	114.28
2008	110.68	108.89	108.08	107.18	107.34	107.98	120.28	122.46	128.14	118.97	121.23	120.00
2009	109.50	110.24	106.37	112.05	109.07	113.51	112.78	115.01	111.60	110.16	115.73	113.84
2010	123.18	119.19	110.26	116.63	109.32	118.12	123.06	118.64	122.14	115.43	115.38	119.42
2011	119.69	117.89	113.51	119.03	112.32	123.97	128.70	121.18	130.71	119.89	119.82	119.76
2012	118.78	114.96	112.29	118.41	114.08	124.37	135.64	132.76	134.84	120.96	121.12	121.62
2013	122.55	116.75	122.37	121.74	135.05	136.95	137.58	136.66	138.89	122.67	121.88	123.84
2014	127.52	128.40	130.55	123.09	126.58	143.70	141.04	136.67	140.17	127.95	125.08	125.34
2015	136.61	137.83	135.83	133.70	135.75	148.96	147.77	145.12	150.89	138.70	136.39	133.32
2016	136.77	143.56	138.20	137.93	147.28	157.63	158.74	153.07	162.48	141.31	140.99	139.71
2017	155.80	159.64	145.06	155.71	151.60	161.36	165.16	167.00	178.24	151.58	149.45	145.86
2018	-	-	-	-	-	-	184.01	177.84	186.55	165.12	138.78	142.60