What do we gain from Seasonal Adjustment of the Indian Index of Industrial Production (IIP)?

No. 322
07-October-2020

Radhika Pandey, Amey Sapre and Pramod Sinha

National Institute of Public Finance and Policy
New Delhi

What do we gain from Seasonal Adjustment of the Indian Index of Industrial Production (IIP)?

Radhika Pandey, Amey Sapre and Pramod Sinha

Abstract

In this paper we conduct a seasonal adjustment (SA) of the 2011-12 base series of Index of Industrial Production (IIP). We use the x-13 ARIMA-SEATS iterative process and follow an indirect approach of first identifying seasonality at the product level and then recompile the manufacturing index with seasonally adjusted series. The SA process shows identifiable seasonality in 206/405 (almost 50%) items spread within broad NIC groups of food, beverages, textiles, leather & apparels. Seasonally adjusted levels also provide a smooth and low fluctuation series that can be used for extrapolation in the advance and provisional estimate stage of GDP estimates. However, the SA process reveals several data quality issues of inexplicable outliers, growth rates and changes in pattern of individual items. While seasonal adjustment has advantages, the process pre-supposes pristine data quality and given the trends shown by item level data, both the SA and actual IIP are inadequate in explaining the growth performance of the manufacturing sector.

Keywords: Seasonal Adjustment, X-13 ARIMA-SEATS, Index of Industrial Production, Fluctuations, India JEL C43, C50, P44

Acknowledgments

We are thankful to Pronab Sen and GC Manna for helpful comments and suggestions.

What do we gain from Seasonal Adjustment of the Indian Index of Industrial Production (IIP)?

Radhika Pandey, Amey Sapre and Pramod Sinha

June, 2020

Abstract

In this paper we conduct a seasonal adjustment (SA) of the 2011-12 base series of Index of Industrial Production (IIP). We use the x-13 arima-seats iterative process and follow an indirect approach of first identifying seasonality at the product level and then recompile the manufacturing index with seasonally adjusted series. The SA process shows identifiable seasonality in 206/405 (50%) items spread within broad NIC groups of food, beverages, textiles, leather & apparels. Seasonally adjusted levels also provide a smooth and low fluctuation series that can be used for extrapolation in the advance and provisional estimate stage of GDP estimates. However, the SA process reveals several data quality issues of inexplicable outliers, growth rates and changes in pattern of individual items. While seasonal adjustment has advantages, the process pre-supposes pristine data quality and given the trends shown by item level data, both the SA and actual IIP are inadequate in explaining the growth performance of the manufacturing sector.

Keywords

Seasonal Adjustment, X-13 arima-seats, Index of Industrial Production, Fluctuations, India

JEL

C43, C50, P44

Acknowledgments

We are thankful to Pronab Sen and GC Manna for helpful comments and suggestions.

Radhika Pandey, Amey Sapre and Pramod Sinha are at the National Institute of Public Finance and Policy (NIPFP), New Delhi. The views expressed in the paper are of the authors and not of the institute. |
Correspondence: `amey.sapre@nipfp.org.in`

Contents

Tables

Figures

All models are wrong, but some are useful.
– George Edward Pelham Box

1 Introduction

In this paper we conduct a seasonal adjustment of the 2011-12 base year series of the Index of Industrial Production (IIP). In India the IIP is an important high frequency indicator of industrial output and is also used as a critical input in estimation of value addition in the manufacturing sector. Similar to IIP, most macroeconomic series ought to be adjusted for seasonality before they are used in any empirical work or for policy inputs. However, in absence of seasonally adjusted series, the analysis is usually conducted using quarterly or annual growth rates. Year-on-Year growth captures changes relative to the previous year whereas the nature of high frequency indicators is to capture recent information such as month-on-month changes [See Bhattacharya et.al (2016) for a discussion on other macroeconomic series]. In this context, IIP assumes policy and research importance on two counts; (i) it serves as a lead indicator of industrial activity, especially for analyzing short term movements in production of manufactured items and (ii) it opens a variety of issues about measurement of industrial activity, growth performance of the economy and for extrapolation of value addition in the national accounts.

Despite the importance of IIP the index is marred with several practical problems. Changes in measurement units of items, unexpected movements of growth rates, lack of coverage and adequate representation of firms in the IIP frame, revisions and inconsistent trends with other measures of industrial activity, and host of other issues pose several challenges in effectively using the index for research and policy. [see Pandey *et. al* (2018), Rajakumar (2017), Sastry (2011), Manna (2010, 2015), Nagaraj (2002, 1999), Singhi (2000) and CSO (Undated) for a detailed discussion on related issues. However, for making an assessment of the growth performance of the manufacturing sector or understanding the trends of industrial output, addressing measurement issues alone does not solve all problems related to the index.

One of the areas that does not receive adequate attention is the impact of seasonality in high frequency indicators. Seasonal fluctuations are movements that recur with a similar intensity in a given period (such as months) each year, thus showing a clear pattern of peaks or troughs over a sufficiently long time period. Fluctuations in production play a crucial role in influencing levels and growth rates of IIP especially when the index is used as a means to extrapolate value addition estimates at the Advance and Provisional Estimate stage of GDP estimation.

Large swings in levels in either directions have at least three major implications; (i) incorrect analysis of rising or falling trends of levels and growth rates, (ii) imprecise growth rates due to a 'base' effect, and (iii) higher volatility in the aggregate which can lead to an imprecise estimation when used as an input. Also, without seasonal adjustment, one cannot distinguish between genuine seasonal fluctuations and the ‘noise’ of increase or decrease in industrial production. For instance, if a large number of items in the index show seasonal variation, then given the weight structure of the index, seasonal components would drive the index movements, which are distinct from the growth trend. In general, the impact of such movements cannot be assumed to be negligible. Conclusions on growth performance are usually drawn from the overall index, whereas ignoring seasonality can present an inaccurate picture, particularly on month-to-month variation.

The literature on seasonal adjustment of production (or industrial output) has made considerable progress in developing methods and policies on conducting automated seasonal adjustment of macro indicators. The EuroStat has established guidelines for seasonal adjustment within the European Statistical System (ESS) for harmonizing and maintaining comparability of infra-annual statistics, especially the Principal European Economic Indicators (PEEIs) (see ESS (2015) for details). Similarly the United States Bureau of Economic Analysis (US-BEA) and UN Industrial Development Organization (UNIDO) have outlined various procedures for conducting seasonal adjustments on a host of macroeconomic indicators. (see UNSD (2010)).

In the Indian context, the 2011 report on statistical audit of IIP had highlighted the following; (i) according to the IRIIP 2010 (UNSD (2010), countries should consider producing and disseminating seasonally adjusted series as an integral part of their long-term programme of quality enhancement of their industrial production statistics, (ii) seasonal adjustment must be performed only when there is clear statistical evidence and economic interpretation of the seasonal/calendar effects. The industry level (3 or 4 digit) of the estimates at which to apply seasonal adjustment will depend on the level at which IIP compilers are able to produce reliable seasonal adjustment estimates. It is generally recommended to perform seasonal and working day adjustments at the lowest level provided that the data have sufficient accuracy to enable reliable adjustment to be performed. (See Sastry (2011)). The audit report also recommended;

> (i) Seasonal adjustment of all India IIP should be taken up as an official responsibility by the CSO, (ii) Only the finally revised series of the All India IIP should be used for seasonal adjustment, (iii) Experimenting with various possible alternative methods and subjecting the seasonal adjustment results to validation using a wide range of quality measures, evolve an appropriate method of seasonal adjustment, (iv) Among others, the absence of residual seasonal and/or calendar effects as well as the stability of the seasonally adjusted pattern has to be carefully assessed, (v) The results of the experimental exercises may be released as a research study, (vi) Release the official seasonally adjusted series of the All India IIP along with details of the methodology used. See Sastry (2011) for details of the audit report.

In contrast, the 2014 report of the Working Group on compilation of the IIP stressed only on the need for conducting seasonal adjustment, without describing any method or procedure. The report highlighted two view points on seasonal adjustment; (i) seasonal adjustment is best left to the users’ and that the statistical agency (CSO) will not compile and publish such data and (ii) seasonal items ought to be identified and their confidence limit can be calculated using the method of moving averages.

From a data users' perspective, seasonal adjustment involves making a considerable number of empirical choices. Broadly, (i) choosing a level of disaggregation, i.e. whether to seasonally adjust broad sub-groups or individual items or the overall index, (ii) treatment of outliers, seasonal, calendar, effects and (iii) choice of models, etc. are driven primarily by user discretion or the problem at hand.

Seasonal adjustment is significantly affected by data quality. Typically, seasonal algorithms are applied assuming near-perfect data quality implying that empirical methods can identify genuine fluctuations and provide a reasonably fair description of seasonal factors. However, as empirical methods cannot account for poor data quality, an automated procedure overruns most data problems and often leads to inconclusive results. The difference is clearly visible when we choose to analyze seasonality at the item level, instead of the overall index.

Similarly, a major concern with data quality is revision in data. It is well known that seasonal adjustment methods are sensitive to outliers and revisions may alter data either by introducing new outliers or changing position of existing ones. Although there are guidelines on revisions, these limitations lead to difficulties in choosing an appropriate empirical model for adjusting outliers.

While seasonal adjustment has clear advantages, it also has few limitations (i) the adjustment process causes loss of information on account of data transformation, thus seasonally adjusted data *cannot* be considered as a replacement of actual data, (ii) conducting seasonal adjustment also involves subjectivity and judgment in the choice of models, level of aggregation, calendar events, etc. as they are based on the problem at hand and (iii) seasonal adjustment is based on *general guiding principles* and not on any 'best' or unique method.

1.1 Motivation: What do we do?

Seasonal adjustment is relatively simple if done at the aggregate level and extremely complicated if done at the item level (or lowest level of disaggregation). Detecting seasonal items opens up several possibilities of model choice, outlier effects, data quality problems and aggregation. Given quality of data, the index is a combination of seasonal and non-seasonal items, stable and moving seasonality, level shifts, inexplicable outliers, base effects, etc. Thus, item level seasonal adjustment makes it harder to distinguish between genuine seasonal movements, data problems and overall effect on growth rates.

Building on this premise, we outline one complete approach to construct a seasonally adjusted IIP (SA-IIP) from the item level. Our work builds on Mazumder & Chakraborty (2013) and Bhattacharya *et. al.* (2016) on seasonal adjustment of macro series. While Mazumder & Chakraborty (2013) applied seasonal adjustment using the X-12 ARIMA method on all item groups in the 2004-05 IIP series, Bhattacharya *et. al* (2016) applied the X-13 ARIMA-SEATS on the aggregate 2004-05 IIP series.

We deal with the following questions; (i) which items in the IIP show seasonality?, (ii) how does seasonality vary across items and months?, (iii) do seasonal components impact the overall IIP? and (iv) how does the seasonally adjusted IIP compare with the unadjusted index?.

While the process of seasonal adjustment is complex and involves user discretion, seasonally adjusted series must also make statistical sense and be useful for policy. Detecting seasonality relies almost entirely on a series of diagnostic tests and in-turn the statistical methods assume near perfect data quality. Dealing with item level data brings these two aspects upfront as outliers, inexplicable trends relate more to data quality than genuine movements in production. We follow the ESS (2009, 2015) and UNSD (2010) route that highlights the role of quality and meaningful procedures on seasonal adjustment as opposed to an automated approach for estimation.

The paper is divided in three parts. In the first part we discuss the composition, data and the issues related to the IIP before seasonal adjustment. In the second part, sections 4 outline the process of seasonal adjustment using the X-13 ARIMA procedure and we briefly discuss the mechanics of the automated process. In the third part, we deal with making sense of seasonality and discuss the relevance of SA-IIP for policy.

PART I

BEFORE SEASONAL ADJUSTMENT

2 Why seasonal adjustment?

Seasonal components consists of effects that are specific to calendar events and recur over time. In general, seasonal components are also taken to *recur with a similar intensity* and such components may have a specific direction of change as compared to trends in normal times. Typically, the magnitude of seasonality causes the overall aggregate to either sharply rise or fall in a consistent manner thereby leading to higher volatility in the values of the aggregate. However, over a sufficiently long period, the seasonal component may itself vary and not exhibit a fixed pattern. These are cases of varying seasonality which may alter the basic patterns of peaks and troughs over time. Broadly, seasonality arises from several calendar related events such as;

- Weather based factors: monsoon, winter or summer months
- Agricultural seasons: harvest or sowing season
- Administrative procedures: tax filings, financial year closure, working days, etc.
- Festivals: Diwali, Christmas, etc.
- Institutional: Annual budgets or Fiscal year ending
- Social and Cultural factors: Statutory holidays, etc.

Seasonal factors are also influenced by calendar related effects that may not be systematic on a month-to-month or year-to-year basis, such as;

- Trading Day effects: Stock market trading days in a month
- Moving (or shifting) Holiday effects
- Tourist arrivals

In addition to routine movements, irregular fluctuations may also occur due to a combination of unpredictable or unexpected factors. These are; unseasonal/ unpredictable changes in weather, natural disasters, civil unrests, strikes, lockouts, etc. leading to disruptions in man-days in production. However, in several cases, irregular fluctuations are indistinguishable from data problems that arise from lack of validation checks.

At the product level, several cases show large irregular fluctuations and are indistinguishable from source data problems. Such cases are treated as outliers (within the irregular component) that may or may not have any identifiable reason and are possibly data induced. Some examples are; (i) changes in source data (such as a reporting factory), (ii) changes in measurement or product classification, (iii) changes in taxation structure, (iv) lack of data validation. These abrupt changes distort the seasonal pattern and often lead to misleading interpretation of seasonality and overall trend of industrial production. We deal with several such cases in sections 4.3.1 and 5. These facts about the index make it necessary to isolate seasonal and irregular components so as to get an index that conveys 'news' about trends in industrial production.

The international literature in this domain has led to development of methods, best practices and general guidelines that most statistical agencies follow. We summarize some of the relevant portions as under before dealing with the IIP.

2.1 International guidelines on seasonal adjustment

The EuroStat (ESS 2009, 2015) and UNSD (2010) provide the general framework and guidelines for conducting seasonal adjustment of macro time series. Both EuroStat and UNSD stress on the need to identify and de-seasonalize an index like the IIP at the lowest possible aggregation level. Although no unique method for seasonal adjustment has been outlined, the guidelines deal extensively with use of automatic seasonal adjustment programs (X-12, X13, TRAMO, etc.) and country specific inputs that influence industrial production.

The ESS goes beyond mechanics of seasonal adjustment to implement a quality framework for seasonal adjustment for all EU member states that covers five dimensions of statistical quality as per the European Statistics Code of Practice; (i) relevance (ii) accuracy and reliability (iii) timeliness and punctuality (iv) coherence and compatibility and (v) accessibility and clarity. Specifically, the procedure mentions; "Accuracy and reliability are measurable quantitatively through statistical tests to assess whether the seasonally adjusted time series display suitable characteristics; the measures should not be limited by software choice. For example, if they are not embedded in the seasonal adjustment software, they should be defined elsewhere".

The ESS also describes a validation policy for seasonal adjustment whereby a wide range of measures such as descriptive, graphical and statistical test results are implemented to assess the validity of seasonal adjustment of important indicators. The policy highlights that "seasonally adjusted data must have a meaningful interpretation. As a consequence implausible data should not be validated even when statistical tests are successful." (See ESS (2015) for details).

Presently, in India the official statistical agency does not publish seasonally adjusted high frequency indicators either separately or before they are used as inputs in the national accounts. In other studies, Shastry (2011), Bhattacharya, et. al (2016) and Mazumdar & Chakraoborty (2013) deal with several aspects of seasonality in IIP but their findings are limited to the 2004-05 series and with every base year revision of IIP, the same exercise needs to be conducted. Similarly, each new base year presents new challenges as the composition of the index differs in terms of commodities and source data. In the following sections we describe the components of the 2011-12 IIP, its composition, followed by seasonal properties of its components.

3 IIP composition, data and need for seasonal adjustment

3.1 Composition and Data

The first part of the exercise is to compute the IIP using its individual items. The 2011-12 IIP series was introduced after a comprehensive revision of sources and updation in the item basket, (See CSO (2014)) for items and method.) The broad description of the index is in Table 1.

Table 1: Description of the Index of Industrial
Production 2011-12 Base Year series

Data	2011 (Apr) to 2019 (Nov)	Capital goods	67
Frequency	Monthly	Consumer Durables	86
Time Lag	42 days*	Consumer Non-Durables	100
Broad Index	3	Infra./Construction goods	29
Sub-Index	23	Intermediate goods	110
Items	407	Primary goods	15
Items (Manufacturing)	405	WPI Deflators	114
Items (Electricity)	1		
Items (Mining)			
Volume based items	298		
Value based items	109		

* for the current month, Source: Item level data available from Center for Monitoring Indian Economy, 2019

The index is computed as a weighted average of production relatives of all items and is a fixed-weight Laspeyres Index given by;

$$I = \frac{\sum_i R_i W_{i,0}}{\sum_i W_{i,0}} \quad \text{; where,} \tag{1}$$

- $W_{i,0}$ is the weight of the ith item in base year (0)
- R_i Production relative of item i and $R_i = P_{i,t}/P_{i,0}$
- $P_{i,t}$ is the value of production of the ith item in time period t
- $P_{i,0}$ is the value of production of the ith item in base year 0

The index also includes several items that are taken on value (i.e. in monetary) terms as opposed to physical volumes. These value items are first *deflated* by a representative WPI deflator and then aggregated to form the overall index. Using these two types of commodities, equation 1 can be reformulated as the sum of value and volume based items in the index. Specifically, let C_{vai} and C_{voi} denote value and volume based commodities, and w_{vai} and w_{voi} their respective weights at NIC five digit level in each industry category. By definition, summing over both types of commodities, the weighted average of value and volume based commodities gives the level of the index for the industry group at the two-digit level, i.e.;

$$I_i = \sum_i w_{vai}.C_{vai} + \sum_i w_{voi}.C_{voi} \tag{2}$$

where I_i represents industry group, and weights as per different compilation categories. The advantage of splitting the index into value and volume based commodities is the understand the role of the deflator (WPI in this case) that can also drive the trends of the index (See Pandey et. al. (2018) for a detailed discussion on this issue). Within the IIP, the primary interest is in the manufacturing sub-index that has a total of 405 items. Table 2 describes the volume and value based items and their corresponding weights in the index.

Compiling the index is based on a bottom-up approach. We select the closet matching WPI deflator for each compilation category so as to capture an approximate volume movement of the item. In some cases, weighted combination of more than one WPI deflator is used to deflate a value based item. The bottom-up approach can be visualized in a flow diagram.

Table 2: Number and weight of items in
Manufacturing index of IIP, 2011-12 series

NIC 2 Digit	Group	Volume N	Volume W	Value N	Value W	Total $\sum N$	Total $\sum W$
10	Food	53	5.304	–	–	53	5.304
11	Beverages	7	1.036	–	–	7	1.036
12	Tobacco	2	0.558	1	0.241	3	0.799
13	Textiles	15	3.164	3	0.128	18	3.292
14	Apparels	–	–	4	1.323	4	1.323
15	Leather	1	0.355	2	0.147	3	0.502
16	Wood	4	0.129	2	0.064	6	0.193
17	Paper	2	0.641	1	0.231	3	0.872
18	Media	2	0.237	4	0.443	6	0.68
19	Petroleum products	10	11.775	–	–	10	11.775
20	Chemicals	63	7.51	2	0.368	65	7.878
21	Pharmaceuticals	–	–	20	4.982	20	4.982
22	Rubber plastic	21	1.995	4	0.43	25	2.425
23	NM Minerals	12	3.795	3	0.293	15	4.088
24	Basic metals	28	12.807	–	–	28	12.807
25	Metal products	7	1.035	12	1.621	19	2.656
26	Computer electr.	12	0.974	6	0.598	18	1.572
27	Electrical equip.	28	2.142	5	0.86	33	3.002
28	Machinery equip.	15	1.988	22	2.782	37	4.77
29	Vehicles	2	1.339	5	3.519	7	4.858
30	Transport equip.	5	1.587	1	0.189	6	1.776
31	Furniture	–	–	3	0.131	3	0.131
32	Other manuf.	7	0.066	9	0.878	16	0.944
	Total	296	58.437	109	19.228	405	77.665

N & W respectively denote number & weight of items

The indirect approach, though lengthy and cumbersome has one clear advantage over the direct method; analyzing individual series helps in distinguishing between variation in data vs. genuine seasonal fluctuations. It is also important to highlight this aspect as official statistics use *averages of indexes*, and not *de-seasoned index growth rates* as inputs in building national account aggregates. Thus, benchmarking and extrapolation of macro aggregates is based on average of 12 month variation, which is different from average of 12 month of seasonally adjusted growth. These differences are brought out clearly when average year-on-year growth is compared with sum of monthly growth rates.

Bottom up approach to construction IIP and Seasonal IIP

Individual monthly raw data series [A]

↓	↓
Volume based	Value based
↓	↓
	Deflated series
	↓

↓

Computed Manufacturing IIP [B]

↓	↓
Volume based	Deflated Value based

↓	↓	↓	↓
Seasonal (SA)	Non-Seasonal (NSA)	Seasonal (SA)	Non-Seasonal (NSA)
↓	↓	↓	↓
SA series	NSA series	SA series	NSA series

↓

Computed Manufacturing SA IIP [C]

The first two steps of compilation are further broken up to visualize the steps involved in computation.

Individual monthly raw data series [A]

↓

Volume

↓

Yes	No
1. Compute average value of each series for the period Apr 2011-March 2012. Denote this value as 100 2. Convert each monthly value from Apr 2012 as; value * (100/ Avg. value)	1. Compute average value of each series for the period Apr 2011-March 2012. Denote this value as 100 2. Convert each monthly value from Apr 2012 as; value * (100/ Avg. value) 3. Identify a representative WPI deflator 4. The reduced form (Index value/100) is used to deflate the value obtained in step 2 4a. In case of more than one deflator for the item, the weights are split

↓

Values are multiplied by item weights and summed to form the index

Figure 1 and Table 3 describe the comparison of the actual and computed manufacturing index. The computed index matches the trend and the relevant moments (mean and standard deviation) of the actual index. This process lend credence to the deflators used for deflating value based items and the aggregation process. The next step is to analyze variation before passing each series into a seasonal adjustment procedure.

Figure 1: Comparison of actual and computed
Manufacturing IIP, 2011-12 series

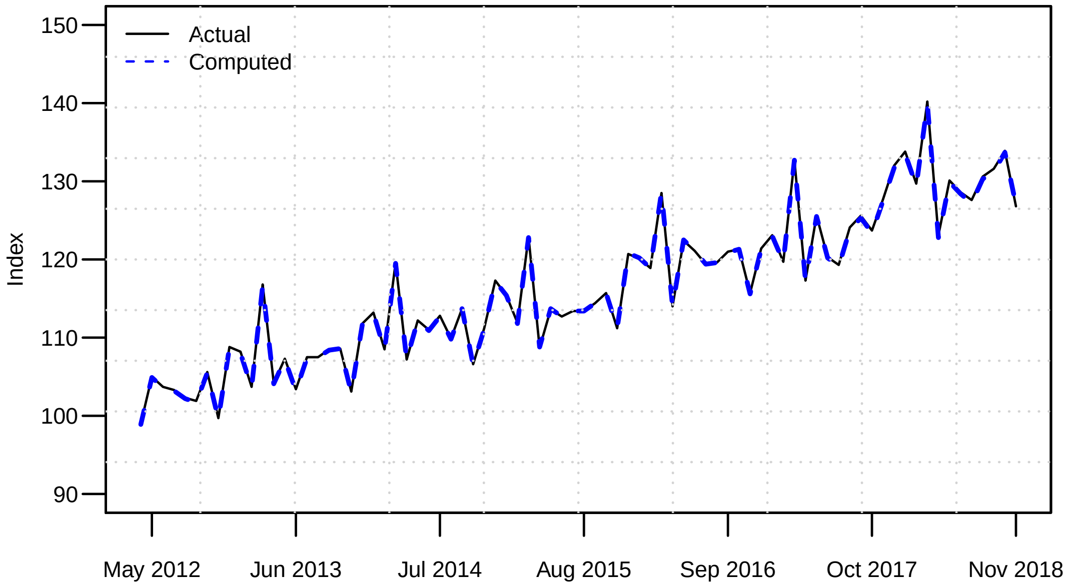


Table 3: Descriptive stats of actual & computed
Manufacturing IIP, 2011-12 series

Var.	N	Mean (μ)	SD (σ)	Min	Max
April 2012 – Nov. 2019					
IIP (Actual)	91	118.30	10.45	99.00	144.60
IIP (Computed)	91	118.20	10.39	98.90	144.30
April 2013 – Nov. 2019					
Growth rate IIP (Actual)	79	3.58	2.87	–4.26	10.37
Growth rate IIP (Computed)	79	3.57	2.87	–4.27	10.29

3.2 Variation in data vs. seasonal fluctuation

An important prerequisite for seasonal adjustment is to distinguish between variation in data and genuine seasonal fluctuations. This distinction is critical in case of IIP as individual commodities show wide to inexplicable month-on-month fluctuations that are different from routine peaks and troughs. Wide fluctuations distort averages and variance which results into large confidence intervals, and thus lead to unreliable estimates. Primarily, unexplained variation in data occurs due to; (i) changes in collection points of source data or substitution of reporting units, (ii) incorrect product description while data collection, (iii) imputation of missing values, and (iv) lack of consistency checks or data validation.

Three kinds of inexplicable variations are commonly observed; (i) one time level shift, (ii) systematic fall of peaks, (iii) one time change in pattern of seasonality, (iv) different pattern of past and recent seasonality. One of the confounding effects of such variation is on outliers. As earlier, most outliers may have an identifiable reason, but in case of IIP several series show inexplicable and abrupt changes. A list of outliers is presented in section 4.3.1.

Variation in data is largely understood by analyzing the pattern of levels and growth rates of individual items. Presently, the method followed by CSO (2014, Shastri (2011), Singhi (2000)) is to put thresholds for low and high variance and place confidence intervals around averages of individual series.

3.2.1 Existing method

CSO (2014, undated) studied the internal consistency of the 2004-05 IIP to identify the commodity groups that had a high impact on the growth of the manufacturing sector. They found food products, basic metals, machinery and motor vehicles to have a high impact on the growth of the manufacturing sector. Using annual data at a 5 digit NIC, they showed that the coefficient of variation in growth of around 324 commodities was within the range of 0–30%; 30–60% for 59 commodities and over 60 % for the remaining 14 commodities. Specifically, two conditions were followed, (i) Low weight items: $\mu \pm 3\sigma$ and for High weight items: $\mu \pm 2\sigma$, thus giving the tolerance range by the lower and upper bounds of the interval $[\mu - 2\sigma, \mu + 2\sigma]$. Further,

> (a) Current month's production data is checked to see whether it falls below the minimum production value in the past 12 months, or above the maximum production for the same period. In the case it lies outside this range, it is calculated (in percentage terms) how much they fall below (above) the minimum (maximum) production value and items showing huge variations are then highlighted and their production figures confirmed by the respective source agencies. Identification relies on;
>
> - $>$ 20% variation between current and previous months' production
> - $>$ 30% variation between current and corresponding month in the previous year
> - $>$ 20% variation between previous months' 1st RE and QE production figure
> - $>$ 10% variation between previous 3rd month's final revision and first revision

These measures are critical for understanding the month on month variation, outliers and eventually seasonality. However, for a longer time span, these measures prove to be inadequate for several reasons, (i) given a long time span (10 years), the mean and variance is significantly affected by outliers, thus leading to large confidence intervals, (ii) variation is data induced in cases of changes in data source and (iii) extreme changes remain in the series despite setting a variance threshold may not explain the nature of actual production. As an illustration, 4 tabulates the count of times for some series for which the monthly (point-on-point) and yearly growth exceeds the confidence range of 20% for P-o-P and 30% for Y-o-Y, within the limited time span of about 100 months.

Extreme variation in data points more towards data quality rather than changes in actual production. For instance, assuming a limited incremental capacity of a factory unit, the confidence range $[\mu \pm 3\sigma]$ is sufficiently large enough to capture variation in production. However, frequent breaches of the confidence regions by several series raises questions that neither relate to a seasonal pattern nor actual nature of production.

Table 4: Count of times monthly and year growth exceeds confidence range, May 2012 - Nov. 2019

	Series	P-o-P > 20%	Y-o-Y > 30%
1	Groundnut Oil	70	70
2	Anti-malarial drugs	60	69
3	Air Coolers	67	67
4	Human hair- articles thereof	67	67
5	Sunflower Oil	65	66
6	Creams and lotions	70	65
7	Cottonseed Oil	63	65
8	Digital indicator - all types	63	64
9	Solar power system etc.	43	63
10	Flat products of alloy Steel	34	63
11	Medicated shampoos	43	62
12	Moulding machine	66	60
13	Pistols and guns	63	60
14	Electric heaters	60	60
15	ACSR Conductors	22	60

Table 5 describes the maximum and minimum growth rates during the period April 2011 - Nov. 2019 of some of the series, arranged as per descending order of Y-o-Y growth rates.

Table 5: Max and Min of Year on Year and Point on Point growth rates in production, % (in descending order of Max Y-o-Y growth rate)

	Series	Y-o-Y (%)		P-o-P (%)	
		Max	Min	Max	Min
1	Air Coolers	80733.33	-99.99	19848.28	-97.12
2	Sunflower Oil	15684.51	-94.43	2317.08	-94.13
3	Other meats of crustacean/seafood	9351.43	-68.13	5564.29	-93.20
4	Electric heaters	7975.00	-100.00	3319.44	-94.81
5	Groundnut Oil	6356.75	-99.53	1668.97	-96.88
6	Solar power system/equipment	4219.00	-98.99	4298.18	-97.44
7	Castor seed oil	3811.20	-100.00	11082.91	-100.00
8	Scientific instruments/ etc.	2275.86	-56.72	170.67	-67.82
9	Coconut Oil	1990.35	-94.12	1218.47	-89.93
10	Machinery & equipment for defence	1924.73	-97.07	944.79	-97.99
11	Fragrances & Oil essentials	1166.67	-71.59	316.00	-72.53
12	Digestive enzymes and PPI drugs	1117.24	-30.56	352.94	-54.96
13	LCD/ LED monitor	1007.36	-68.35	2321.71	-90.93
14	Sugar	902.41	-91.05	2056.60	-92.73
15	Cottonseed Oil	800.22	-100.00	1001.62	-100.00

The period April 2011 to Nov. 2019 (103 months) Annual and monthly growth rate shows numerous cases of inexplicable maximum and minimum growth rates that do not show any pattern of seasonality. The seasonal adjustment procedure works to reduce the volatility in the series by lowering the standard deviation and shrinking the *range* of values. However, in several individual cases, seasonally adjusted growth rates also do not convey any meaningful trajectory. See Table 17 for a comparison of pre-and-post adjusted growth rates.

3.3 Monthly patterns: The bigger picture

To understand the seasonal pattern, one has to first take stock of the trends of several sub-indexes of the IIP. These trends give us the broader picture as to which type of commodities have a seasonal high and low value. We begin with computing the month averages by broad NIC group.

Table 6: Average of month wise index value by major groups, April 2011 - Nov. 2019

NIC		w	Winter		Spring		Summer		Monsoon		Autumn		Winter	
2D	Group		Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
10	Food	5.304	130.30	122.19	121.76	103.53	96.76	93.15	98.08	96.29	97.13	102.64	110.90	132.54
11	Beverages	1.036	103.47	104.36	126.56	126.69	140.24	117.36	96.43	92.80	95.31	96.88	91.74	96.14
12	Tobacco	0.799	119.56	119.56	121.27	94.61	116.65	116.48	106.94	107.24	110.58	109.24	111.04	117.40
13	Textiles	3.292	117.40	111.76	119.19	112.48	114.55	113.10	116.64	119.31	117.23	115.99	112.17	117.64
14	Apparels	1.323	142.29	140.29	162.80	130.48	133.01	128.56	126.23	121.94	121.48	119.19	112.33	140.41
15	Leather	0.502	124.63	124.69	128.79	124.86	128.26	123.43	124.08	113.51	114.26	109.29	109.03	121.51
16	Wood	0.193	91.13	96.69	106.67	91.15	94.93	102.69	100.24	102.09	105.26	93.11	89.33	102.44
17	Paper	0.872	111.09	104.31	115.01	107.39	110.70	107.53	111.19	111.18	108.15	110.19	106.03	111.59
18	Media	0.680	103.16	99.89	109.54	94.05	98.38	100.93	99.53	98.38	100.00	97.40	96.80	111.09
19	Petro products	11.775	120.09	110.20	122.56	108.84	116.71	114.59	116.71	116.54	111.06	119.35	115.10	119.73
20	Chemicals	7.878	114.44	107.23	119.19	106.95	113.23	111.76	115.80	115.68	114.93	118.11	108.59	114.17
21	Pharma.	4.982	161.41	160.66	168.36	141.00	160.56	151.20	158.33	161.44	162.76	154.29	152.80	163.81
22	Rubber, plastic	2.425	113.49	108.54	118.29	112.24	113.04	110.05	111.59	110.69	111.21	109.21	108.40	115.83
23	NM Minerals	4.088	116.70	110.61	123.63	113.39	114.70	112.68	108.85	103.40	105.39	109.25	101.90	112.37
24	Basic metals	12.807	130.34	125.11	137.74	123.01	132.89	128.71	129.34	128.75	125.65	127.74	121.29	126.73
25	Metal products	2.656	105.53	106.30	123.81	94.74	100.70	102.53	99.66	98.40	102.48	99.13	96.73	108.84
26	Computer electr.	1.572	128.34	123.80	135.69	122.08	134.28	128.56	136.50	139.39	152.08	144.21	113.26	115.83
27	Electrical equip.	3.002	110.40	114.19	126.16	103.76	114.13	112.06	106.61	110.90	126.64	126.46	122.47	124.74
28	Machinery equip.	4.770	104.81	113.06	141.61	97.08	105.05	109.74	103.09	107.08	117.35	106.44	104.17	116.79
29	Vehicles	4.858	106.81	107.01	116.06	106.04	107.78	103.24	106.98	104.43	106.14	104.63	102.17	99.56
30	Transport equip.	1.776	116.09	117.07	122.91	113.29	121.56	117.59	121.74	122.08	133.05	127.01	113.59	104.39
31	Furniture	0.131	153.16	160.46	188.24	143.51	157.55	162.69	159.19	160.05	165.96	156.41	153.11	170.97
32	Other manuf.	0.944	114.43	108.39	128.73	112.94	103.76	111.60	104.20	107.49	109.48	101.63	86.59	106.11

Cell values denote average of the index value of same month in each year

The average index level by month captures two essential features of the IIP, (i) commodity groups show a distinct variation in monthly average of levels, (ii) weights of seasonal items is sufficiently large to move the index. Table 6 highlights average values of the index by *seasons* for a sufficiently long period (April 2011 - Nov. 2019). Food items show a distinct pattern with a rise during winters and troughs during summers. Similarly, beverages show a reverse trend as the index tends to increase during months of spring and summer as compared to winters. Clothing (textile, ready-made apparels, leather) also show a pattern with higher production levels during winters as compared to summers.

Among other household consumables, electronics, computer peripherals, medicines also show a modest rise post monsoon as compared to the other months. Household consumables account for nearly 23% weight in the manufacturing index which is large enough to move the overall index in some systematic ways; (i) the index tends to be higher during months of winter and spring and shows a dip during summers and autumn, (ii) the rise in household consumables also coincide with festivals that are usually during autumn (October to November) and (iii) index values are amongst the highest during month of March, thus leading to a higher base (and lower growth rates) for the next month.

Since growth rates of sub-indexes are more useful for analyzing news about industrial production, we can extend the information in the Table 6 to tabulate the count of Point-on-Point (PoP) negative growth rates by each month.

Table 7: Count of negative growth rates by month, May 2012 - Nov. 2019

NIC	Winter		Spring		Summer		Monsoon		Autumn		Winter		
2D	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Sum
10	6	7	3	7	8	7	0	5	3	1	1	0	48
11	0	2	0	4	0	8	8	6	2	4	7	0	41
12	3	1	4	6	1	5	5	7	2	4	3	1	42
13	4	7	0	7	2	7	1	1	6	6	5	0	46
14	2	4	0	7	4	5	6	5	5	6	5	0	49
15	2	4	1	5	2	8	4	8	4	5	4	1	48
16	7	1	0	7	1	2	5	3	3	8	3	0	40
17	5	6	0	7	1	5	1	5	6	3	6	0	45
18	6	5	1	7	1	2	6	5	3	6	4	0	46
19	3	7	0	7	0	5	3	3	6	1	5	1	41
20	3	7	0	7	0	5	1	4	5	5	5	1	43
21	4	5	1	7	0	6	3	2	4	7	3	0	42
22	6	7	0	5	1	6	3	5	4	6	5	0	48
23	0	7	0	7	3	6	5	7	3	1	6	0	45
24	1	7	0	6	0	7	3	4	7	2	4	0	41
25	4	2	0	7	1	3	5	5	3	5	4	0	39
26	0	6	2	6	1	4	2	5	1	4	7	2	40
27	7	3	0	7	0	4	7	1	1	4	5	3	42
28	7	1	0	7	1	1	6	3	0	8	4	0	38
29	0	4	0	7	1	7	0	6	1	5	4	5	40
30	0	3	1	6	0	7	1	5	0	5	7	5	40
31	6	1	0	7	1	3	5	3	2	6	4	1	39
32	2	6	0	7	5	3	4	3	3	5	7	0	45
Sum	78	103	13	150	34	116	84	101	74	107	108	20	988

Table 7 highlights that the maximum number of negative growth rates (over previous month) are seen in April, followed by monsoon and autumn months (June, October, November). These patterns are consistent with the average of levels by month wherein we see higher production levels during March, May, September and December.

Broadly, the pattern also reveals some interesting cases; (i) seasonality is present across items and is visible in months of March, April, May and December and (ii) within items, a clear pattern is visible mostly for food items. Empirically, these patterns provide us the first clue on whether it would be sufficient to de-season at the sub-index level or identify individual commodities within each group. The choice depends primarily on an identification strategy which is elaborated in the following sections.

PART II

ABOUT SEASONAL ADJUSTMENT

4 The mechanics of seasonal adjustment

4.1 Direct vs. Indirect seasonal adjustment

A direct approach is to de-seasonalize the aggregate index, without taking into account the seasonality of its individual items. Thus, if we begin with an additive model, then an index A_t^D can be written as a decomposition of its components, i.e. $A_t^D = T_t + C_t + S_t + TD_t + FD_t + R_t$, where each component respectively denotes; *trend*, *cyclic*, *seasonal*, *trading day*, *festival day* and *residual or irregular* component. (See Gomez et. al (1997), Findley, et. al (1998), Astolfi, et. al (2001) Ladiray & Quenneville (2001) Scheiblecker (2014) for a discussion and decomposition.)

Alternatively, in the indirect adjustment method each individual item is checked for seasonality and the index is recreated using seasonally adjusted series. Thus, if we have (a_n) items in an index A_t, then $A_t^I = \sum_{n=1}^{N} \omega_n a_{n,t}$ where ω_n is the fixed weight structure of the individual (a_n) components or items. Post identification of seasonal and non-seasonal items, the index can be reconstructed as; $A_t^I = \sum_i w_{si}.C_{ai} + \sum_j w_{nsj}.C_{nsj}$, where (s, ns) denote seasonal and non-seasonal items. The key difference between the two methods is that in the indirect method the weight structure is assumed to be the same for the seasonal items. The results of both methods in general do not coincide but can lead to similar estimates under some conditions. For example,

- if the underlying components of the series are additive
- if the individual series do not have outliers, and
- if the overall index closely follows the patterns of seasonal items, both methods lead to a similar result
  Whereas, in the IIP, compared to the overall index and its individual series, we find;
- - several individual series have multiplicative components, hence are modeled using a log transformation
- - most series have outliers that distort the components, and
- - no clear pattern of visually identifiable seasonality exists for several series

As no clear patterns are visible for most sub-indexes for a substantial part of the calendar year, it is more appropriate to identify seasonality in individual items and reconstruct the index. In the next section we briefly discuss the mechanics of the automated process and identify two key components, viz. outliers and seasonal series.

4.2 Process: Automation, defaults and model choice

Before we proceed to estimation, it is essential to review the basic elements of seasonal adjustment and the X-13 ARIMA-SEATS method. We summarize the general approach to seasonal adjustment and briefly describe the workings of the X–13 ARIMA-SEATS package available in `R`, the selection procedure for seasonal candidates, choice of model and the type of outliers. The details of the theory and the algorithm are elaborated in the appendix. A detailed treatment of the theory and the package are available in US Bureau of the Cenus (2013). The X-13 AS is an automated iterative procedure that has capabilities of determining the most appropriate model, fitting and estimating various components of the time series and doing a series of diagnostic tests. The process is performed by an *AutoMDL* and *RegARIMA* algorithm that iterates in the following way;

- First, begin with a default AR $(p, d, q)(p, d, q)_s$ model of the type $(011)(011)_s$, where s stands for seasonal
- If time series effects such as trading day, holidays, etc. are specified, the algorithm checks for their statistical significance using the default model.
- A constant term is introduced in the AR model and residuals are checked using a t stat for its inclusion.
- Next, the algorithm checks for various types of outliers
- Iteratively, time series effects previously specified are checked again to determine whether they remain statistically significant after removing outliers
- The diagnostics are compared with the default model to check for improvements in residual statistics. If the model fails to show improvement, the default model is selected.
- The algorithm then determines the order of differencing of the time series based on an AR and MA process. Differencing is also an iterative process which first starts with an ARIMA process $(2, 0, 0)(1, 0, 0)_s$ and computes diagnostics for several ARMA variations $(11)(11)_s$ till AR parameters are close to unity.
- After determining the order of differencing, the remaining AR and MA parameters are determined based on the lowest BIC criterion. Iteratively, BIC2 is computed for all ARIMA models of the form $(3, d, 0)(P, D, Q)_s$, where d and D are the previously determined regular and seasonal orders of differencing. The program chooses the pair of values P and Q that minimize BIC2.
- The series is finally seasonally adjusted depending on the model choice and estimates of seasonal factors, stability and test statistics are obtained.

The user intervention in case of X-13 AS is in specifying the time series effects such as calendar, holidays, festivals, etc. Table 8 describes the time series effects that are typically included in the estimation process. Trading and working day effects are calendar related events that drive production levels primarily on account of man-days. The leap year adjusts the length of February to make it consistent with other months. Holidays and festivals are typically specific to a country but account for events like Thanksgiving, Christmas, etc. as they are commonly observed across countries. A brief list of type of time series effects is as follows;

Table 8: Types of seasonal effects time series components and outliers

Component Effects		Outlier Effects
1. Trend - Cycle	3. Calendar	Additive (AO)
2. Seasonal	3.1 Trading days	Level shift (LS)
2.1 Seasons	3.2 Working days	Transitory (TS)
2.2 Holidays	3.3 Leap year	
2.3 Festivals	4. Irregular	
	4.1 White noise	

In the Indian case, festivals like Dussera, Diwali can be added as they are important events particularly in the later part of year. The time series effects are part of pre-adjustment processes and are incorporated into the model as separate regressors. The mechanics are briefly summarized as under and are also discussed in the section on outliers.

4.3 Specification of the regression function

The regression function incorporates several dummy variable type regressors in the X-13 AS procedure. Some of the important specifications are as follows;

Effect	Specification
Fixed seasonal	$\mathrm{M}_{i,t} = \begin{cases} 1 & \text{if ith Month is January} \\ 0 & \text{otherwise} \\ -1 & \text{for December} \end{cases}$
Seasonal Outlier	$\mathrm{S}_{i,t_0} = \begin{cases} 0 & \text{for } t \geq t_0 \\ 1 & \text{for } t < t_0 \text{ for the same month} \\ -1 & -1/(s-1) \ \text{ otherwise} \end{cases}$
Outlier (seasonal)	at point S_0; $\mathrm{SO}_t^{t_0} = \begin{cases} 0 & \text{for } t \geq t_0 \\ 1 & \text{for } t < t_0 \text{ for the same month} \\ -1/(s-1) & \text{otherwise} \end{cases}$
Outlier (additive)	at point t_0; $\mathrm{AO}_t^{t_0} = \begin{cases} 1 & \text{for } t = t_0 \\ 0 & \text{for } t \neq t_0 \end{cases}$
Temporary Change	TC at point t_0; $\mathrm{TC}_t^{t_0} = \begin{cases} 0 & \text{for } t < t_0 \\ \alpha^{t-t_0} & \text{for } t \geq t_0 \end{cases}$ $(0 < \alpha < 1)$ where α is the decay rate back to the previous level
Level shift	at point t_0; $\mathrm{LS}_t^{t_0} = \begin{cases} -1 & \text{for } t < t_0 \\ 0 & \text{for } t \geq t_0 \end{cases}$
Temporary level shift	t_0 to t_1; $\mathrm{TL}_t^{t_0,t_1} = \begin{cases} 0 & \text{for } t < t_0 \\ 1 & \text{for } t_0 \geq t \geq t_1 \\ 0 & \text{for } t > t_1 \end{cases}$
Trend (constant)	$(1-B)^{-d}(1-B^s)^{-D}I(t \geq 1)$, where; $\mathrm{I(t \geq 1)} = \begin{cases} 1 & \text{for } t \geq 1 \\ 0 & \text{for } t < 1 \end{cases}$

4.3.1 Detection of outliers: Beyond data problems

X-13 AS deals with the following types of abrupt changes in the levels of a time series: Additive Outliers (AOs), Level Shifts (LS), Temporary Changes (TCs), Seasonal Outliers (SOs), Temporary Level Shifts (TLS) and ramps of various kinds. Theoretically, AOs are assumed to affect only *one observation*, whereas LS type increase or decrease *all observations* from a certain identifiable point by a constant value. *Temporary shifts* are detected when the level shifts abruptly but returns in an exponential way to its previous position. *Seasonal outliers* are detected as abrupt rise or fall in the level of the *seasonal pattern* which is compensated for in the other months. A *ramp type* flow is to detect a linear rise or fall in the level over a specified interval or between temporary level changes. Table 9 lists the total number of identified AO and LS type outliers in each individual series over the available data period and tabulates the *count by each broad group.*

Table 9: Count of outliers in levels in each individual series by NIC, Apr. 2011 - Nov. 2019

NIC	Num (LS)		Num (AO)		Sum		$\sum_i w$		$\sum$(LS)	$\sum$(AO)	Total	$\sum W$	
2D	Vol.	Val.	Vol.	Val.	Vol.	Val.	Vol.	Val.				n_S	
10	29	–	52	–	81	–	5.304	–	29	52	81	53	5.304
11	4	–	2	–	6	–	1.036	–	4	2	6	7	1.036
12	0	0	1	0	1	0	0.558	0.241	0	1	1	3	0.799
13	8	3	11	5	19	8	3.164	0.128	11	16	27	18	3.292
14	–	4	–	5	–	9	–	1.323	4	5	9	4	1.323
15	2	1	1	0	3	1	0.355	0.147	3	1	4	3	0.502
16	1	2	4	1	5	3	0.129	0.064	3	5	8	6	0.193
17	1	0	2	1	3	1	0.641	0.231	1	3	4	3	0.872
18	7	2	4	4	11	6	0.237	0.443	9	8	17	6	0.68
19	6	–	10	–	16		11.775	–	6	10	16	10	11.775
20	37	1	92	1	129	2	7.51	0.368	38	93	131	65	7.878
21	–	24	–	27	–	51	–	4.982	24	27	51	20	4.982
22	15	6	11	9	26	15	1.995	0.43	21	20	41	25	2.425
23	6	1	14	1	20	2	3.795	0.293	7	15	22	15	4.088
24	72	–	63	–	135	–	12.807	–	72	63	135	28	12.807
25	14	16	12	16	26	32	1.035	1.621	30	28	58	19	2.656
26	10	2	7	9	17	11	0.974	0.598	12	16	28	18	1.572
27	23	2	29	10	52	12	2.142	0.86	25	39	64	33	3.002
28	9	10	17	17	26	27	1.988	2.782	19	34	53	37	4.77
29	1	2	5	5	6	7	1.339	3.519	3	10	13	7	4.858
30	3	0	3	1	6	1	1.587	0.189	3	4	7	6	1.776
31	–	0	–	0	–	0	–	0.131	0	0	0	3	0.131
32	12	5	8	6	20	11	0.066	0.878	17	14	31	16	0.944
Sum	260	81	348	118	608	199	58.437	19.228	341	466	807	405	77.665

Qualitatively, AO type outliers capture inexplicable changes that are visible only once or are limited in number even in a longer time span. These could also be data induced as abrupt *one time* changes are unlikely to represent actual changes in production. LS type outliers could capture changes that happen due to changes in reporting factory, changes in product specification, etc. as they lead to changes in *all subsequent* observations. LS type outliers also separate two distinct patterns of seasonality as the underlying data may not be from the same source.

These type of outliers pose the basic challenge as most test statistics for identifiable and stable seasonality give inconclusive results. The pattern of outliers also influences month-on-month changes in two specific ways; (i) base effect and (ii) volatile growth rates. Table 10 tabulates the number of outliers (AO+LS) found in *all individual series* within each group.

Table 10: Count of outliers in levels by month by NIC group, Apr. 2011-Nov. 2019

NIC	n_S	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Sum	$\sum w$
10	53	5	4	4	8	6	8	11	13	8	6	4	4	81	5.304
11	7	1	0	0	1	1	0	0	2	0	0	1	0	6	1.036
12	3	0	0	0	1	0	0	0	0	0	0	0	0	1	0.799
13	18	1	1	3	3	5	1	3	5	2	1	1	1	27	3.292
14	4	0	2	1	1	1	0	0	1	2	1	0	0	9	1.323
15	3	0	1	0	0	1	0	0	0	0	0	1	1	4	0.502
16	6	0	1	0	2	0	2	0	0	0	2	0	1	8	0.193
17	3	0	0	1	1	0	1	0	0	0	0	1	0	4	0.872
18	6	1	1	2	2	2	1	1	1	2	1	2	1	17	0.68
19	10	2	0	1	1	2	0	0	3	0	6	1	0	16	11.775
20	65	8	3	13	18	20	7	9	12	9	20	5	7	131	7.878
21	20	1	3	6	5	4	4	5	2	2	5	8	6	51	4.982
22	25	1	1	4	8	3	5	4	2	4	4	0	5	41	2.425
23	15	1	1	2	5	1	1	1	0	4	3	1	2	22	4.088
24	28	6	8	19	28	12	8	11	9	4	10	7	13	135	12.807
25	19	7	4	5	8	4	1	5	6	3	5	2	8	58	2.656
26	18	2	0	1	8	2	2	2	4	2	1	1	3	28	1.572
27	33	4	3	7	5	2	7	6	6	6	5	7	6	64	3.002
28	37	2	4	5	8	1	3	4	3	5	7	5	6	53	4.77
29	7	1	0	2	2	1	1	1	2	0	0	2	1	13	4.858
30	6	0	0	3	1	0	1	1	0	0	0	0	1	7	1.776
31	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0.131
32	16	2	5	0	5	1	1	3	3	1	2	3	5	31	0.944
Sum	405	45	42	79	121	69	54	67	74	54	79	52	71	807	77.665

10: Food, 20: Chemicals, 21: Pharmaceuticals, 24: Basic metals, 25: Metal products,
27: Electrical equip., 28: Machinery equip.

Within the available time span, four groups (10, 20, 24 & 27) account for nearly 50% of the number of outliers. Outliers are also not expected to follow any set pattern, however a sizable number are clustered around the period March, April and May. Given the data pattern and the nature of outliers, the penultimate step is to identify seasonal and non-seasonal series.

5 Identification of seasonal components

Identifiable Seasonality is an iterative and stepwise process that is determined by a series of diagnostic tests. First, the seasonal component is defined as either (i) intra year variation that occurs repeatedly or constantly thus implying *stable* seasonality or (ii) is visible in an evolving fashion from year to year, thus implying *moving* seasonality.

The empirical approach is to determine both stable and moving seasonality in conjunction. The algorithm conducts two tests, namely, an F test for seasonality *assuming stability* and a Nonparametric Test for the seasonality assuming stability. The Kruskal-Wallis (KW) test is used in case of stable seasonality and identifiable seasonality is determined by combining the

F tests for stable and moving seasonality. For instance, let F_m and F_s denote the F tests for moving and stable seasonality. If the H_0 of *no stable seasonality is not rejected* at the 1% significance level, then the output of the algorithm is 'Identifiable seasonality not present' and the series is considered *nonseasonal*. If H_0 is rejected, then, combined test $(T_1 + T_2)/2$ is computed where $T_1 = 7/F_s$ and $T_2 = 3F_m/F_s$. Next, if the H_0 of *no moving seasonality* is rejected at the 5.0% significance level and the combined test statistic $T \geq 1$ and if the H_0 of identifiable seasonality not present is *not rejected* then the conclusion is '*Identifiable seasonality not present*'.

However, if the H_0 of identifiable seasonality not present *has not been accepted* but $T_1 \geq 1, T_2 \geq 1$ or the KW χ^2 test fails to reject at the 1% significance level, the algorithm returns the conclusion, 'Identifiable Seasonality *probably not present*'. Lastly, if the H_0 in the F_s and the KW χ^2 test are rejected and if none of the combined tests fail, then the conclusion is 'Identifiable seasonality *Present*'.

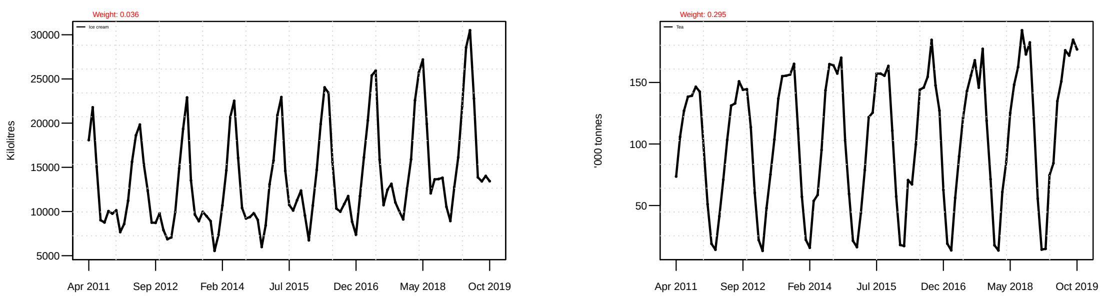
Based on these tests, one can obtain the number of series that show identifiable (and probable) seasonality for passing them through the adjustment algorithm. Statistically, the F tests does not always provide the most accurate results. Two problems have been highlighted in this regard. Findley et al. (1998) show that a spectrum, sliding-span and ARIMA diagnostics can be use as a better alternative to the F test as it tends to indicate false positive. Second, given a small sample size, log transformation and low autocorrelation, the F test tends to detect spurious seasonality. The spectral diagnostics are typically computed with 96 observations in default, but require a minimum 60 observations to perform the test.

5.1 Fluctuation in data: Identifiable and Non-identifiable patterns

Seasonal adjustment relies primarily on identifiable seasonality. Figure 2 presents some series that have a clear identifiable pattern over months and conclusively are candidates for seasonal adjustment.

Figure 2: Identifiable seasonal patterns

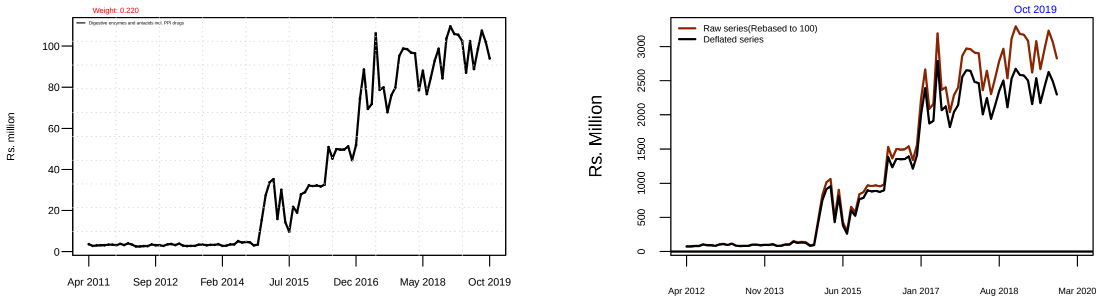
(a) Icecream (b) Tea



Trends in value based items also need to be analyzed carefully. The role of the deflator is critical in these cases as it can amplify the outliers. Figures 3 and 3 show the pre-and-post deflated case of one of the value based items.

Figure 3: Pattens of value based items (non-deflated)

(a) Digestive Enzymes (b) Digital Media



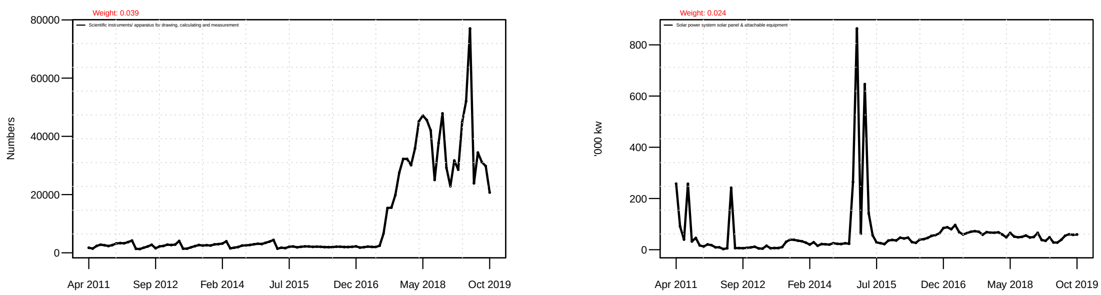
The post deflated series is similar in pattern but with has some noticeable differences. The deflated series is expected to show only level differences as the price series is not supposed to alter trends in production. In Figure 3, initially the pre and post deflated series almost overlap, thus suggesting a negligible effect of the deflator.

The effect of the deflator is pronounced in recent years where the nominal and 'real' values diverge. These effects pose problems for seasonal adjustments as the primary objective is to understand production levels and not trends in monetary values. Other than deflator issues, there are several cases where the patterns are neither identifiable nor explain the nature of production. These are cases of additive and level shift outliers and no visual seasonality.

5.1.1 Non-Identifiable patterns

Figure 4: Non-Identifiable patterns

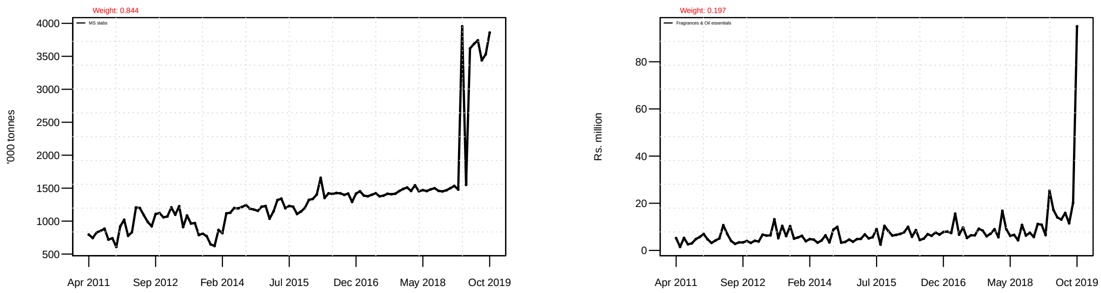
(a) Scientific instuments (b) Solar Power Equip.



Some of the illustrative cases suggest that changes in the trends could be data induced. If we assume that reporting factories are unchanging within a sampling frame then a large number of series show changes that seem unusual. The inexplicable patterns are also corroborated by unusually high maximum growth rates that are primarily on account of outliers. In short, there are three issues with such non-identifiable patterns;

Figure 5: Non-Identifiable patterns

(a) Mild Steel Slabs (b) Fragrance, Ess. Oils



- Non-identifiable patterns could depict an actual change in production, in which case there is no element of seasonality and the series only needs to be adjusted for outliers.
- Non-identifiable patterns show two or more distinct changes in trends, which could be result of changes in data sources. In such cases, there is no definite conclusion on seasonality.
- Non-identifiable patterns can also relate to past vs. recent seasonality. Detection requires a sufficiently long series ($n > 96$) and most non-identifiable patterns are too short for any meaningful trend analysis. Test statistics M9–M11 provide some insights into this aspect.

5.2 Candidates for seasonal adjustment

Based on a number of test diagnostics, Table 11 tabulates the count of series that have identifiable seasonal patterns. The number of series with multiplicative components were logarithmically transformed which totaled to 289/405, i.e. $\approx$ 71%. Large number of series with multiplicative component suggests that a direct seasonal adjustment of the index would not be appropriate (see Section 4.1 for general conditions for similarity between results of direct and indirect adjustment).

From the count of identifiable seasonality, 206/405 $\approx$ 51% of series have identifiable pattern, 156 of which are volume and 50 are value items. Broadly, seasonal items account for a substantial 44% weight in the overall index and therefore seasonality cannot be assumed to have a negligible effect on the movement of the index.

Table 11: Identification of seasonal items

NIC 2D	Transform	Iden. Seasonality	–	NIC 2D	Transform	Iden. Seasonality
10	43	33		22	20	15
Vol.	43	33		Vol.	16	13
Val.	-	-		Val.	4	2
11	7	6		23	10	6
Vol.	7	6		Vol.	8	5
Val.	-	-		Val.	2	1
12	1	0		24	23	10
Vol.	0	0		Vol.	23	10
Val.	1	0		Val.	-	-
13	12	6		25	16	12
Vol.	11	6		Vol.	4	5
Val.	1	0		Val.	12	7
14	3	2		26	12	9
Vol.	-	-		Vol.	7	7
Val.	3	2		Val.	5	2
15	3	2		27	22	16
Vol.	1	1		Vol.	18	14
Val.	2	1		Val.	4	2
16	3	2		28	29	27
Vol.	1	1		Vol.	10	13
Val.	2	1		Val.	19	14
17	1	1		29	4	4
Vol.	1	1		Vol.	1	2
Val.	0	0		Val.	3	2
18	6	4		30	4	3
Vol.	2	1		Vol.	3	2
Val.	4	3		Val.	1	1
19	6	6		31	2	2
Vol.	6	6		Vol.	-	-
Val.	-	-		Val.	2	2
20	37	27		32	9	9
Vol.	35	26		Vol.	4	4
Val.	2	1		Val.	5	5
21	16	4		Sum	289	206
Vol.	-	-				
Val.	16	4				

Within value items, $50/109 \approx 45\%$ show seasonality which are of interest. Value based items are mostly in NIC 25 & 28 (Metal Products, Machinery) that are both for industrial and final use. Food group has the maximum number of seasonal items, followed by Chemicals and Machinery Equipments. Together, these three groups account for nearly $87/206 \approx 42\%$ of the seasonal items.

Table 12: Non-seasonal and Seasonal items in IIP

NIC 2D	Non-Seasonal items						Seasonal items						Total (Mfg. IIP)					
	Value		Volume		Total		Value		Volume		Total		Value		Volume		$\sum_N$	$\sum_{Wt.}$
	N	Wt.	N	Wt.	N	Wt.	N	Wt.	N	Wt.	N	Wt.	N	Wt.	N	Wt.	N	Wt.
10	0	0.000	20	1.673	20	1.673	0	0.000	33	3.631	33	3.631	0	0.000	53	5.304	53	5.304
11	0	0.000	1	0.039	1	0.039	0	0.000	6	0.997	6	0.997	0	0.000	7	1.036	7	1.036
12	1	0.241	2	0.558	3	0.799	0	0.000	0	0.000	0	0.000	1	0.241	2	0.558	3	0.799
13	3	0.128	9	1.902	12	2.030	0	0.000	6	1.262	6	1.262	3	0.128	15	3.164	18	3.292
14	2	0.274	0	0.000	2	0.274	2	1.049	0	0.000	2	1.049	4	1.323	0	0.000	4	1.323
15	1	0.081	0	0.000	1	0.081	1	0.066	1	0.355	2	0.421	2	0.147	1	0.355	3	0.502
16	1	0.011	3	0.031	4	0.042	1	0.053	1	0.098	2	0.151	2	0.064	4	0.129	6	0.193
17	1	0.231	1	0.253	2	0.484	0	0.000	1	0.388	1	0.388	1	0.231	2	0.641	3	0.872
18	1	0.230	1	0.015	2	0.245	3	0.213	1	0.222	4	0.435	4	0.443	2	0.237	6	0.680
19	0	0.000	4	2.320	4	2.320	0	0.000	6	9.455	6	9.455	0	0.000	10	11.775	10	11.775
20	1	0.197	37	3.679	38	3.876	1	0.171	26	3.831	27	4.002	2	0.368	63	7.510	65	7.878
21	16	4.114	0	0.000	16	4.114	4	0.868	0	0.000	4	0.868	20	4.982	0	0.000	20	4.982
22	2	0.321	8	0.798	10	1.119	2	0.109	13	1.197	15	1.306	4	0.430	21	1.995	25	2.425
23	2	0.049	7	0.794	9	0.843	1	0.244	5	3.001	6	3.245	3	0.293	12	3.795	15	4.088
24	0	0.000	18	7.086	18	7.086	0	0.000	10	5.721	10	5.721	0	0.000	28	12.807	28	12.807
25	5	0.831	2	0.020	7	0.851	7	0.790	5	1.015	12	1.805	12	1.621	7	1.035	19	2.656
26	4	0.443	5	0.161	9	0.604	2	0.155	7	0.813	9	0.968	6	0.598	12	0.974	18	1.572
27	3	0.589	14	0.965	17	1.554	2	0.271	14	1.177	16	1.448	5	0.860	28	2.142	33	3.002
28	8	0.865	2	0.207	10	1.072	14	1.917	13	1.781	27	3.698	22	2.782	15	1.988	37	4.770
29	3	2.946	0	0.000	3	2.946	2	0.573	2	1.339	4	1.912	5	3.519	2	1.339	7	4.858
30	0	0.000	3	1.519	3	1.519	1	0.189	2	0.068	3	0.257	1	0.189	5	1.587	6	1.776
31	1	0.032	0	0.000	1	0.032	2	0.099	0	0.000	2	0.099	3	0.131	0	0.000	3	0.131
32	4	0.135	3	0.052	7	0.187	5	0.743	4	0.014	9	0.757	9	0.878	7	0.066	16	0.944
Sum	59	11.718	140	22.072	199	33.790	50	7.510	156	36.365	206	43.875	109	19.228	296	58.437	405	77.665

5.2.1 Stable, Moving and Residual seasonality

In general, identifying seasonality has several complications. The presence of 'identifiable seasonality' relies on the combining the F tests for stable and moving seasonality, along with a Kruskal-Wallis (KW) test for stable seasonality. Within the set of identifiable series, 42 series show presence of 'moving' seasonality. The number of seasonal items can also be arrived as $405 - 157 - 42 = 206$, where 157 is the count of the series that fail the key M7 quality check stat. (see Table 16). Of the remaining series, i.e. $199(405 - 206)$, 95 series returned the result 'identifiable seasonality *probably not present*' (See section 4.1). The presence of stable seasonality implies that seasonal factors for these series remain more-or-less constant thus making prediction and forecasting relatively easier. There is limited information loss between seasonal and non-seasonal months and the decomposition achieves almost all desirable results.

Moving seasonality on the other hand poses difficulties as the seasonal effect does not recur in a fixed pattern and thus distorts month-to-month effects. The presence of moving seasonality is detected using an F statistic that tests for the stability of the ratio of yearly to irregular component variation. Formally, this result is reported by the M7 stat in the X-13 ARIMA-SEATS (see Table 16) for details. The intuition behind moving seasonality is also relevant for understanding the nature of industrial production and data quality. Consider that the total seasonal and irregular component in an additive case is given by $(S+I)_{ij} = a_i + b_j + v_{ij}$ where a_i and b_j represent the contributions of due to the *effects of the ith year and jth month* and v_{ij} is a white noise irregular component. If the total variance is decomposed as $\sigma_y^2, \sigma_m^2, \sigma_r^2$, and the null hypothesis is stated as $a_1 = a_2 = \ldots a_n$, i.e. the year effect does not change relative to the irregular component, the $F_{df=11}$ is σ_y^2/σ_r^2.

Given this metric, it is meaningful to ask; does commodity production show moving seasonality? In the total of 405 commodities, 157 (38%) fail the M7 diagnostics implying that the irregular component variation is much larger than the yearly variation. Most of these commodities ($74/157 \approx 47\%$) are within the group of Chemicals, Pharmaceuticals, Rubber & Plastics and Basic Metals. The trends of production suggests that the result on moving seasonality remains inconclusive as it requires a much longer series to identify consistent patterns in data. While level shifts are detected, it remains inconclusive as to whether such shifts are data induced or are genuine changes in production patterns. The result requires a much detailed investigation with information on reporting factories and product schedules. In case of residual seasonality, the diagnostic stats show only 2 items to have residual seasonality at 5% level of significance. For the remaining 204 items, we find no residual seasonality thus conforming correct model specifications.

5.2.2 Household consumables vs. industrial products

We can separate items into two broad categories; household items and industrial products as the seasonal patterns are likely to be different. Household items include food, beverages, textiles, plastic items, electrical equipments, furniture, vehicles, among others as these items are final products. Industrial items typically serve as inputs but can also be final products, such as media/ transport equipments, machinery, etc. Within household items, food items (Sugar, oil seeds, tea, coffee, icecreams, ready to eat food, rice, aerated drinks, consumer durable such as TV, Washing machines, Air conditioners, etc.) show a marked seasonal pattern during post monsoon months with level dips during September - November, and a gradual rise thereafter. Similarly, industrial products such as; petrol products, Bitumen, rubber items, plastics products, etc. also show seasonal patterns. However, there are some peculiar aspects of this seasonality need to be taking into account. First, while production and sale data may show a lagged relation, the trend of production typically *precedes* the actual event such as festivals, arrival of seasons, etc. Second, the lagged adjustment varies as per the inventory cycle which typically varies between 15-45 days depending on the nature of commodities (See Pandey et. al. (2018)).

5.2.3 Festival effects

Combining the trends of production with festival effects provides a more useful way to understand month-on-month growth rates. Employing festival related level dummies do not provide a detailed insight if the seasonal adjustment is done at the aggregate index level.

Table 13: Count of items with significant festival effect grouped by NIC

	Significant Effect							Total			
NIC	N	Value	$+ve$	$-ve$	$\sum_p w_p$	$\sum_n w_n$	$\sum W$	Value	$+ve$	$-ve$	$\sum W$
10	16	0	0	16	-	1.838	1.838	0	11	42	5.304
11	2	0	0	2	-	0.160	0.160	0	2	5	1.036
12	-	-	-	-	-	-	-	1	2	1	0.799
13	8	0	0	8	-	2.118	2.118	3	4	14	3.292
14	1	1	1	0	0.038	-	0.038	4	2	2	1.323
15	2	1	0	2	-	0.436	0.436	2	0	3	0.502
16	4	1	0	4	-	0.175	0.175	2	1	5	0.193
17	2	1	0	2	-	0.484	0.484	1	0	3	0.872
18	2	1	0	2	-	0.397	0.397	4	1	5	0.680
19	3	0	2	1	2.811	0.174	2.985	0	7	3	11.775
20	18	0	1	17	0.046	3.251	3.297	2	21	44	7.878
21	3	3	0	3	-	0.222	0.222	20	3	17	4.982
22	16	2	1	15	0.019	1.176	1.195	4	3	22	2.425
23	6	2	0	6	-	2.487	2.487	3	6	9	4.088
24	5	0	1	4	1.349	0.911	2.260	0	9	19	12.807
25	8	5	0	8	-	1.225	1.225	12	3	16	2.656
26	7	1	0	7	-	0.903	0.903	6	3	15	1.572
27	10	1	0	10	-	1.159	1.159	5	6	27	3.002
28	9	3	0	9	-	1.726	1.726	22	7	30	4.770
29	5	3	0	5	-	4.504	4.504	5	0	7	4.858
30	1	0	0	1	-	1.363	1.363	1	1	5	1.776
31	1	1	0	1	-	0.051	0.051	3	1	2	0.131
32	4	2	0	4	-	0.327	0.327	9	4	12	0.944
Sum	133	28	6	127	4.263	25.087	29.350	109	97	308	77.665

For instance, in India Diwali is an important festival that usually falls within the months of October - November. At the item level, festival effects on production can be both positive and negative and given the simple regression technique, festival dummies have to do with previous months of production, rather than the month of the event. A similar exercise is done in Bhattacharya et.al (2016) based on the assumption that the level of activity changes on a date, before the event and remains at that level until the day before Diwali.

Using the same algorithm (See Sax and Eddelbuettel (2020)), Table 13 shows the festival impact of items grouped by major NIC. Statistically, 133/405 items show a significant impact with six items on the positive side and 127 on the negative side. In terms of weight, the negative impact is sizable as it adds up to nearly 32% (25/77) of the weight in the index. The pattern of the impact closely matches with the groups that show major outliers - level shifts and changes in pattern over the years. The festival effect thus can be understood as changes in activity in a previous month, leading up the event. The dip in output also partly on account of lesser number of man-days in the month of Diwali and the built up of inventory leading up the duration of the festival.

PART III

AFTER SEASONAL ADJUSTMENT

6 Building the seasonally adjusted IIP Index

Based on tests of identifiable seasonality, a total of 206 seasonal items were aggregated along with the remaining Non-SA series with their respective weights to form the SA-IIP manufacturing index. The index $A_t^I = \sum_{n=1}^{N} \omega_n a_{n,t}$ and ω_n is the fixed weight of the individual (a_n) items is reconstructed as; $A_t^I = \sum_i w_{si}.C_{ai} + \sum_j w_{nsj}.C_{nsj}$, where (s, ns) denote seasonal and non-seasonal items. For each series, post adjustment of outliers and time series effects, the seasonal procedure was achieved by a regression;

$$y_t = \mathbf{X}\beta + \underbrace{\textsc{arima}(0,1,1)(0,1,1)}_{z_t} \tag{3}$$

where $\mathbf{X}$ is a matrix of covariates that includes calendar and outlier effect regressors, and z the residual that is modeled using the ARIMA component in the X-13 ARIMA SEATS. Formally, the seasonal and non-seasonal (wherever applicable) can be described as;

$$\phi(B)\Phi(B^s)(1-B)^d(1-B^s)^D \underbrace{\left(y_t - \sum_i \beta_i x_{it}\right)}_{\approx z_t} = \theta(B)\Theta(B^s)z_t \tag{4}$$

where B is a back shift or lag operator $(B^p y_t = y_{t-p})$, $\phi(B)$ and $\Phi(B^s)$ are the non-seasonal and seasonal lag operators, $\theta(B)$ and $\Theta(B^s)$ are the non-seasonal and seasonal MA operators, and z_t is an *iid* white noise process with mean zero and variance as σ_z^2. The theoretical process is briefly described in Appendix A.

6.1 Comparison of IIP and Seasonally Adjusted (SA) IIP

The primary difference between IIP and SA-IIP is in the annual (Y-o-Y) growth rate figure. In case of IIP the annual growth rate is the difference between the average value of the index between period (t) and $(t-1)$, whereas post seasonal adjustment, we can compute the annual growth rate by summing the growth rate of 12 months. Figure 6 plots the actual and seasonally adjusted IIP levels to show the difference achieved by removing seasonality, which also gives us a timely (i.e. current year) measure of annual growth rate.

Table 14: Actual and Seasonally Adjusted
annual growth rate (April - March)

Year	2013-14	2014-15	2015-16	2016-17	2017-18	2018-19
IIP Actual	3.63	3.78	3.00	4.26	4.51	3.80
IIP SA	2.29	3.55	4.78	3.70	5.38	4.33
Diff (IIP –IIP-SA)	1.34	0.23	–1.78	0.56	–0.87	–0.53
Chg. (Δ) IIP Actual	–	0.15	–0.78	1.26	0.25	–0.71
Chg. (Δ) IIP SA	–	1.26	1.23	–1.08	1.68	–1.05

IIPSA is sum of 12 monthly growth rates

Figure 6: Actual and Seasonally Adjusted Manufacturing IIP

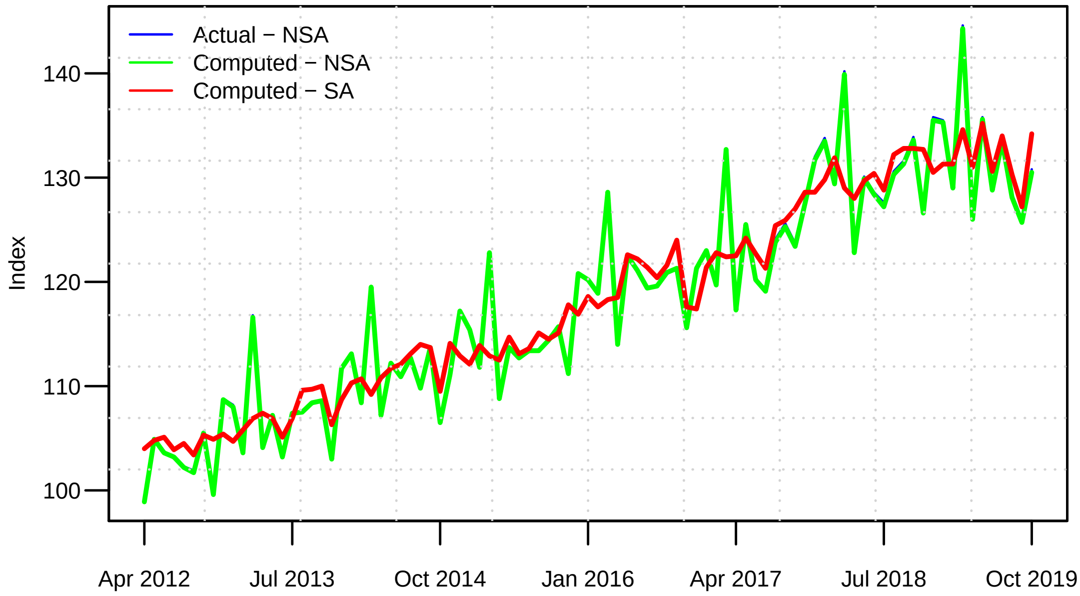
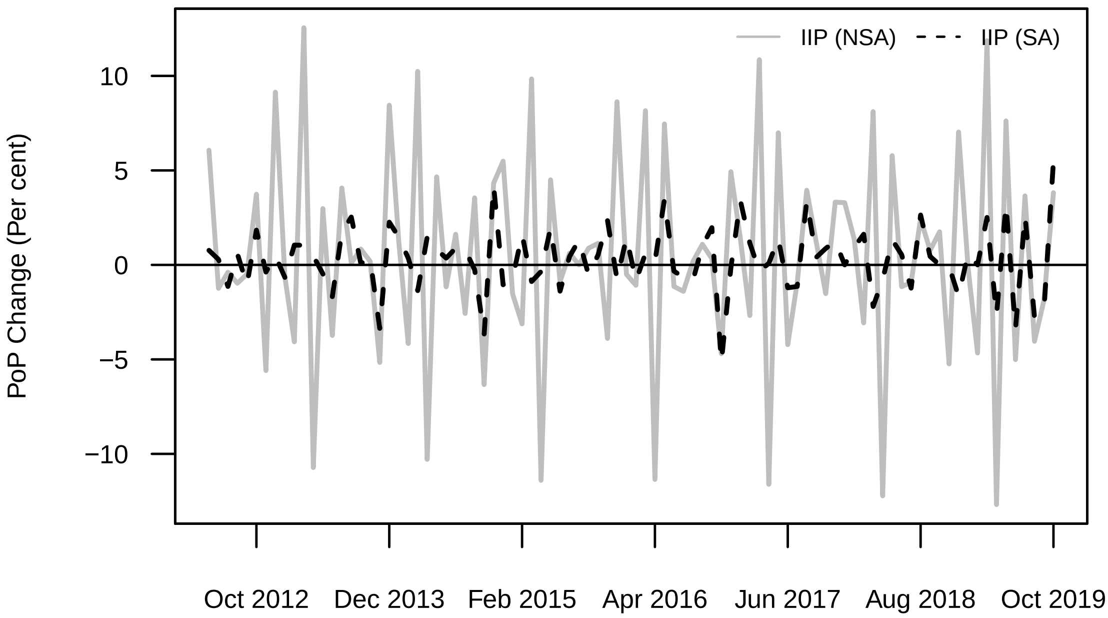


Figure 7: Trend of actual and seasonally adjusted M-o-M growth rate



The SA index shows minor fluctuations around an upward trend thus achieving a considerably smoothened series. To visualize the change in numbers, Table 14 describes the annual growth rates from both the indexes. For maintaining comparability, the P-o-P growth rates are summed up for the financial year as opposed to the calendar year. The Y-o-Y growth rate, on average, turns out to be lower than the P-o-P growth rate as monthly growth considers variation only within the year and is not affected by trends of the previous year.

The Y-o-Y growth is relative to the average position of the previous year as by construct, Y-o-Y growth 'averages' out seasonality. The other difference between the two indexes is on the direction of change. Monthly SA growth rates relate directly to current period output as compared to Y-o-Y growth rates. For 2015-16 and 2016-17, the IIP and SA-IIP show opposite growth trends (Table 14). Actual IIP shows an increase in growth rate from 3.00% to 4.26%, suggesting an improvement in growth performance, whereas SA shows a decline from 4.78% to 3.70%. The average has a bias towards the condition in the previous year and has a tendency to 'magnify' in case of low base in the previous year and 'subdued' in the reverse case. The SA growth rates thus have the distinct advantage that they are free of such base effects and capture current news by removing monthly fluctuations.

6.1.1 Lead indicator?

The practical issues of average vs. SA monthly growth are more relevant when IIP is used for extrapolation in the national accounts and is taken as a lead indicator of industrial activity. Can the SA-IIP serve as a better lead indicator? Since the purpose of the lead indicator is to provide an advance intimation of yearly production, it necessarily has to capture noise-free information of the current year. For example, value addition is estimated at the financial year-end which is an outcome of production 'during the year'. Thus, the sum of monthly growth rates is more appropriate to capture the year end position as opposed to using average growth between year (t) and $(t-1)$.

6.1.2 Revision cycle and linkages with NAS

The National Accounts has a long revision cycle for release of GDP aggregates. In its sequence of six estimates, the first three estimates, viz. 1st & 2nd Advance (AE) and Provisional Estimates (PE) are based on extrapolation of previous years' Provisional and 1st Revised Estimate. The process of extrapolation is based on representative high frequency indicators available for different length of time (See Sapre & Sengupta (2017) for details). In case of 1st and 2nd AE the data availability is of 7 and 8 months, whereas the PE has full 12 month data. To compile the AEs, first one has to forecast the high frequency indicator for the remaining 5 months in case of 1st AE and 4 months in case of 2nd AE.

Two aspects are worth noting; (i) data used in benchmark based extrapolation are not seasonally adjusted, and (ii) every year forecasting is for specific months of the year, i.e. November to March. In the first case non-seasonal data has noise and larger variation thus growth rates based on non-SA series will lead to higher (or lower) index levels depending on the volatility. Second, for extrapolation, only the aggregate index and forecasting does not take into account the seasonal variation either at the overall index or item level as it depends on past trends. Thus, in the initial stages of the revision cycle, the variation in extrapolated levels of sectoral GVA (where IIP is used) from 1st to 2nd AE and to PE is mainly on account of moving from forecast data to actual 12 month data. This variation can be contained to only convey 'news' if the seasonal fluctuation is removed with seasonal adjustment. In effect, using SA-IIP can contain revisions and the magnitude can be computed if the extrapolation techniques used by the statistical agency are well documented.

7 Making sense of seasonality

How do we make sense of the seasonally adjusted data? Table 15 tabulates the month-on-month actual and SA levels. The SA process achieves its first objective of lowering the standard deviation (SD) of the series by compressing peaks/falls. Second, comparing the levels of Nov. to March shows the advantage of the SA process. The SA series gives smoothed levels as compared to actual IIP levels and in the event of forecasting, the SA series will lead to growth rates that have reduced noise and are seasonal effects.

Table 15: Actual and Seasonally Adjusted monthly index values

	Actual IIP							SA IIP						
Mnt	2013	2014	2015	2016	2017	2018	2019	2013	2014	2015	2016	2017	2018	2019
Jan	108.20	113.20	115.40	120.10	123.10	133.80	135.50	104.70	110.30	112.10	118.60	121.40	129.80	131.30
Feb	103.70	108.50	111.90	118.90	119.70	129.70	129.30	105.80	110.70	113.90	117.60	122.80	131.90	131.30
Mar	116.80	119.60	122.90	128.50	132.70	140.20	144.60	106.90	109.20	112.90	118.30	122.40	129.00	134.60
Apr	104.20	107.20	108.70	114.00	117.30	123.10	126.20	107.40	110.80	112.50	118.50	122.50	128.00	131.00
May	107.30	112.20	113.60	122.40	125.60	130.10	135.80	106.90	111.70	114.70	122.60	124.20	129.70	135.20
Jun	103.40	111.00	112.70	121.10	120.30	128.60	129.00	105.10	112.10	113.10	122.20	122.70	130.40	130.60
Jul	107.50	112.80	113.40	119.40	119.30	127.60	133.70	106.90	113.10	113.60	121.40	121.30	128.80	134.00
Aug	107.50	109.90	113.40	119.60	124.10	130.60	128.40	109.60	114.00	115.10	120.40	125.40	132.20	130.30
Sep	108.50	113.70	114.40	121.00	125.60	131.60	126.00	109.70	113.70	114.50	121.60	125.90	132.80	127.20
Oct	108.60	106.60	115.70	121.30	123.70	133.90	130.80	110.00	109.50	115.10	124.00	127.00	132.80	134.20
Nov	103.10	111.20	111.20	115.70	127.70	126.80	-	106.30	114.10	117.80	117.60	128.60	132.70	-
Dec	111.80	117.30	120.70	121.40	132.00	135.80	-	108.70	112.90	116.90	117.40	128.60	130.50	-
μ	107.55	111.93	114.50	120.28	124.26	130.98	131.93	107.33	111.84	114.35	120.02	124.40	130.72	131.97
σ	3.92	3.83	3.92	3.56	4.84	4.52	5.66	1.80	1.74	1.72	2.29	2.65	1.71	2.49

7.1 When diagnostics fail

Seasonal algorithms can provide (i) adjusted data even when there is no seasonality present and (ii) adjusted data that may contain residual seasonality. Thus, validation of identifiable seasonality is a critical requirement in the entire process. Empirically, making sense of seasonality relies on diagnostics and quality checks. For example, in the indirect method, the requirements for a conducting a meaningful seasonal adjustment are much more elaborate. Apriori, it cannot be assumed that all or most series will pass diagnostic tests, nor each series is expected to pass each test. In most cases it is possible to identify the reason of test failure leading to two avenues; either the problem can be traced to data, or it remains beyond correction.

The X-13AS provides a list of final diagnostics that help us determine the performance and quality of seasonal adjustment done on any series and *what to infer* in case diagnostic tests fail. Diagnostic failure means test statistics are beyond the range of acceptance and hence the result is either inconclusive or the decision is left to the user's discretion. The checklist comprises of 11 test statistics, M1 to M11 each having a range of values between 0 and 3 and an acceptable region from 0 to 1. Diagnostics fail when a test value is over 1. The overall quality is given by a Q stat which is a weighted average of all M stats.

- **M1** tests for the relative contribution of the irregular component over a three month span. Changes in the test statistic are influenced by the length of the series. The test fails if the contribution is over 10%.

- **M2** is the relative contribution of the irregular component to the 'stationary portion of the variance. The test values can be misleading if the trend has several changes of direction or is not approximated by a straight line. The test fails if the contribution is over 10%.
- **M3** tests the month-to-month change in the irregular component as compared to the month-to-month change in the trend-cycle. If the monthly change is too high, it is considered that the variation is due to the irregular component and the test fails the value exceeds 3.
- **M4** tests for autocorrelation in the irregular component using a statistic based on the average duration of run. The test fails if the statistic value exceeds 1.
- **M5** stat is the number of months it takes the change in the trend-cycle to surpass the amount of change in the irregular component. A test value of 6 and above (or M5 >1) is considered unacceptably large. M3 and M5 diagnostics may fail if the series has a flat trend. The failure is because the change in the irregular is always more than the change in the trend. However, if the actual production were to be of that nature, then the procedure is not invalidated as no seasonal adjustment is required.
- **M6** tests for correct seasonal filter based on the moving seasonality ratio.
- **M7** is the key statistic that tests the amount of moving seasonality relative to the amount of stable seasonality. Moving seasonality causes problems in estimating different time series components. If the diagnostic fails, then there is no measure of identifiable seasonality in the series.
- **M8** tests for random movements or fluctuations in the seasonal component throughout the series. Larger components are a sign of problem in estimating the seasonal component.
- **M9** tests for linear movement in the seasonal component and checks if the seasonal component is changing rapidly. A high value of the test statistic indicates problems in estimating the seasonal factors.
- **M10 and M11** recompute the M8 and M9 only for a recent time period and not the whole series.

Table 16 shows the count of cases where the M stat failed for each individual series as it exceeded the threshold value. The table highlights a consistent pattern as observed from previous cases of outliers. Most series are in the groups of NIC 10, 13, 20-22, 24, 27-28 that also have the bulk of additive and level shift outliers. The highest count is for M1 that tests for the contribution of the irregular component. Most series have a sizable irregular component compared to the seasonal component which makes it difficult to separate the two components.

Further, if seasonality is *not present*, the year-on-year movements of a de-trended series are expected to be constant and centered within the time span. However, if seasonality is present and varying, there are two possible outcomes: one that changes in the same direction, or changes with random fluctuations. The first is measured with the M9 and M11 statistics, while the latter is tested with the M8 and M10 statistics. It is worth noting that nearly half of the series have

Table 16: Count of failure of test statistics (M stat > 1) grouped by NIC

NIC	M1	M2	M3	M4	M5	M6	M7	M8	M9	M10	M11	Sum	Val
10	31	16	18	10	19	8	14	21	9	25	24	195	0
11	3	2	2	2	4	1	0	2	2	2	2	22	0
12	3	3	2	2	3	0	1	3	2	3	3	25	1
13	15	10	4	2	4	2	8	9	7	11	10	82	3
14	3	1	1	0	1	1	2	1	1	2	1	14	4
15	1	1	1	0	0	1	1	1	1	1	1	9	2
16	5	2	1	0	1	1	3	4	2	5	5	29	2
17	2	0	0	1	0	0	2	2	1	3	3	14	1
18	5	2	2	0	3	1	2	2	1	2	2	22	4
19	5	6	2	0	2	1	1	2	2	3	3	27	0
20	56	48	24	8	30	8	30	32	19	36	34	325	2
21	19	13	8	1	9	2	16	12	7	15	14	116	20
22	20	10	3	1	6	2	7	11	3	14	13	90	4
23	12	5	2	0	2	1	6	6	3	7	6	50	3
24	26	18	11	2	11	5	12	17	8	17	17	144	0
25	14	7	6	1	3	5	6	5	2	8	7	64	12
26	13	6	4	1	5	0	9	9	6	10	10	73	6
27	25	17	10	1	11	3	16	16	11	21	20	151	5
28	27	20	18	5	17	5	9	14	4	20	17	156	22
29	4	2	0	1	0	0	3	3	3	4	4	24	5
30	4	2	2	1	3	1	3	3	3	3	3	28	1
31	3	1	0	0	0	0	1	3	0	3	3	14	3
32	16	9	6	2	6	2	5	5	2	11	9	73	9
Sum	312	201	127	41	140	50	157	183	99	226	211	–	109

failed diagnostics that detect seasonality in the recent past, as opposed to the entire series. This change is likely to be driven by (i) change in source data (factory) or (ii) interpolation of missing values. A systematic pattern in the M stats suggest underlying data problems as most of the irregular variation could be data induced on account of different type of outliers. Thus, except for series that are free of data problems, no meaningful conclusions can be drawn on item level seasonality.

The indirect seasonal adjustment process has revealed several issues with data that considerably limit the scope of quality seasonal adjustment and its inferences. One of the key quality check in data is monthly variation as discussed in section 3.2.1. In CSO (2014) this problem was described explicitly as follows;

> A number of shortcomings of the IIP data, which are essentially sporadic and reflect inconsistent variation in magnitude, have been noticed from time to time. First of all, a considerable number of commodities of the basket tend to show volatile behavior in production making the index behave erratically. Statisticians can't do much but to attribute such phenomena to sampling error. Secondly, most of the Govt. of India organizations, departments or ministries, which supply monthly IIP data suffer from problems of non-response in primary data collection, which may be due to lack of legislative or regulatory control over the production units. In such cases they are left with no option but to estimate the production of the non-responding units. (see CSO (2014))

7.1.1 Usefulness and implications for policy

To summarize, given data limitations, what is the usefulness of the SA series and what is the usefulness for policy?

- The SA process shows identifiable seasonality in 206/405 (50%) items and 63 within the 206 (30%) with stable seasonality. The highest number of seasonal items are within the broad NIC groups; 10 (Food), 20 (Chemicals), 22 (Rubber, Plastics), 24 (Basic metals), 25 (Metal products), 27 (Electrical items) and 28 (Machinery, equipments). Almost 50% of value based items show identifiable seasonality with a combined weight of 19% in the manufacturing index.
- The index has base effects owing to seasonality in months of March and April. Over the twelve month cycle, highest levels are observed during March and has a tendency to fall during summers and rise in winter months.
- Seasonal patterns are identifiable for few items; (level dip) for food during summer - monsoon months (May - August), (level rise) for beverages during summer months (March -May), (level rise) for textiles, leather, apparels during winters (Dec. - Feb), (level rise) for electronic items during autumn (Sept - October)
- Industrial production shows influence of festivals and the Diwali effect is seen for the months preceding the event. Production levels in Food, Chemicals, Plastics, Electrical equipments show a dip post the event-month.
- At the aggregate, seasonally adjusted levels provide a smooth and low fluctuation series that can be used for analyzing monthly variation and for extrapolation in the advance and provisional estimate stage.
- Individual series are useful insofar as they are free of data problems. Given the number of outliers and size of the irregular effect, detecting seasonality remains a challenge. Post seasonal adjustment, the series is useful only at the sub-index level.

8 Conclusion

In this paper we deal with seasonal adjustment (SA) of Indian Index of Industrial Production. We rebuild the SA-IIP using an indirect approach of first identifying seasonality at the commodity level and then aggregating seasonal and non-seasonal items to construct the index. Prior to doing seasonal adjustment, we deal with variation in individual series to understand the extent of level shifts, inexplicable outliers, changes in patterns and quality of data. We discuss pre-adjustments to account for outliers, calendar effects and deflators used to convert values based series to volume.

We use the X-13 ARIMA-SEATS algorithm to identify seasonal items in the entire list of 405 items in the manufacturing index. The SA process gives mixed results as the underlying data quality of several series leads to inconclusive results. While we find 50% of items (206/405) to have identifiable seasonality, the quality statistics reveal problems with source data. Majority of series (315/405) show a high presence of the irregular component and moving seasonality, thereby leading to failed diagnostics and inconclusive judgment on seasonality. Given the quality of seasonal adjustment, one can only highlight the fluctuations at the sub-group levels that show peaks and falls during summer and winter months. We also find that industrial production is influenced by festivals in the months prior to the event and production levels in food, chemicals, plastics, electrical equipments show a dip in the months after Diwali.

The pattern of high growth rates, level & additive outliers and diagnostics failure are concentrated in few groups; Food, Chemicals, Pharmaceuticals, Basic Metals, thus indicating data problems rather than genuine seasonal movements. The irregular movements in growth rates of several commodities despite having confidence bands on checking variation and performing seasonal adjustment show inexplicable trends in production. While most series show major level shifts and change in pattern post 2015, seasonal adjustment cannot address issues like change in source data, reporting factories or revision in past data. We find one limited advantage of seasonally adjusted index as it useful for extrapolation of GVA aggregates in some sectors in the national accounts. The SA levels provide current ‘news’ with reduced noise as they are based on aggregation of month-on-month growth rates as opposed to year-on-year changes.

In the past, several issues have already been raised over the reliability of the IIP index. The seasonal adjustment process equally reveals that while empirical methods can be applied to capture genuine fluctuations in production, they cannot overcome data problem. Lastly, although seasonal adjustment has clear advantages, it pre-supposes pristine data quality. Given the tends of production shown by several individual series, the IIP or its seasonally adjusted version is inadequate in providing any conclusive state of the growth performance of the manufacturing sector at the individual item level.

* * * * * * *

Appendix A Tables

Table 17: Maximum of Point-on-Point (PoP) and Year-on-Year (YoY) Growth rates found in various NIC groups (%), Apr. 2011 - Nov. 2019

		Non Seasonal Items		Seasonal Items	
NIC	Series in Groups	Max PoP	Max YoY	Max PoP	Max YoY
10	Food	346.73	15684.51	403.68	9351.43
11	Beverages	64.51	100.69	74.98	97.86
12	Tobacco	136.35	144.16		
13	Textiles	101.2	279.24	31.88	33.12
14	Apparels	95.6	384.89	46.77	60.5
15	Leather	44.11	41.42	29.54	51.02
16	Wood	141.08	355.9	61.33	184.24
17	Paper	75.88	107.11	22.15	35.93
18	Media	263.71	308.48	118.95	239.73
19	Petro	45.24	91.46	64.63	60.3
20	Chemicals	221.01	1166.67	186.27	573.53
21	Pharmaceuticals	261.74	708.8	151.06	1117.24
22	Rubber, Plastic	82.56	225.44	88.06	134
23	NM Minerals	73.91	158.19	44.52	64.85
24	Basic metals	195.77	518.18	112.03	126.67
25	Metal products	299.56	327.67	247.91	628.52
26	Computer electr.	627.77	2275.86	183.59	515.71
27	Electrical Equip.	529.57	80733.33	318.71	7975
28	Machinery Equip.	390.79	4219	231.94	436.64
29	Vehicles	80.51	193.74	59.32	94.44
30	Transport Equip	109.56	477.38	156.27	132.78
31	Furniture	64.16	186.44	63.14	119.06
32	Other Manuf.	329.14	778.95	159.95	550.5

Appendix B Theory and model setup

The theory of seasonal adjustment revolves around 'smoothing' the time series to extract a secular trend and capturing the irregular components around the trend. One of the simplest methods is to use month wise dummy variables to capture the effect of month-specific movements in a time series. Such a method involves specifying a regression equation of the type

$$x_t = \alpha_0 + \sum_i \beta_i m_t + z_t \tag{5}$$

where x_t is the time series on a monthly frequency and m_t denotes a set of 11 dummy variables for respective months [January, February, March, ... November], with December month being the base case, and z_t is an error term. The regression coefficients on the dummy variables capture the sign (direction) and the magnitude of 'shifts' in the level (captured by the intercept α_0) of the variable as compared to the base case month. The method, though simple, poses several challenges for analysis and inference. First, the method assumes a fixed pattern of seasonality i.e. month wise dummy variables do not distinguish the effect of calendar related events (seasons, festivals, holidays, etc.) that move between months. Second, the method also does not extract the information of cycles or irregular component in the time series. Extraction of such components requires time series filters or use of dynamic models. Third, the method does not take into account time series related problems such as serial correlation and non-stationarity. As a consequence, the method leaves out useful information in the error term z_t which if extracted can provide improved and robust results.

The next line of approach is to use ARIMA models based on the seminal works of Dagum (1975) and Box & Jenkins (1976). These models make use of a lag structure, differencing and moving averages of errors to model a time series. In principle, these three aspects account for serial correlation, non-stationarity and error structure of the time series. Typically, serial correlation (and Partial Auto Correlation (PAC)) forms a basis for determining the lag structure of the model. Since values across adjacent months may have high correlation, the lag structure aids in identifying past periods that may be useful in explaining variation in current time period. Differencing aims to bring stationarity in the time series, which is useful for robust estimates of parameters. The Moving Average (MA) component

The lag structure also plays a significant role in specifying an appropriate econometric model for estimation. Consider a monthly time series x_t, where $t = 0, 1, \ldots T$. Using logarithms, we can define $y_t = \ln(x_t)$ and specify (L) as a lag operator based on the frequency of the time series. In general, the lag operator for a monthly series can be decomposed as $(1-L)(1+L+L2+...+L11)$. Based on the lag operator we can obtain two kinds of growth rates, i.e. $(1 - L12)y_t$ yields the Year-on-Year (YoY) growth rate of y, whereas $(1 - L)y_t$ gives us the Point-on-Point (PoP) (month-on-month in case of a monthly series) growth rate. These two growth rates are the reference points of comparison after seasonal adjustment.

The estimation strategy utilizes the basic time series components such as; trend (T), cyclic (C), seasonal (S) and irregular (I). In several cases it is common to specify the three major components (C, S, I) using either an additive or multiplicative model. Specifically, if y denotes a time series, then we can write;

$$\text{Additive: } y = C + S + I$$
$$\text{Multiplicative: } y = C \times S \times I$$

If $\hat{s}$ is the estimate of the seasonal component, then the seasonally adjusted series is obtained as $y - \hat{s} = C + I$ in the additive case and $y/\hat{s} = C \times I$ in the multiplicative case. Statistically, seasonality in a multiplicative model moves in proportion with the trend. This implies that over time, if the trend component rises, the seasonal effects also rise, and vice-versa. However, in the additive case, the effect of seasonality is largely constant, irrespective of the movement of the trend component. These two aspect define the spectrum of seasonality in terms of its range, variation and stability.

B.1 X-13 ARIMA - SEATS

X-13 SIGNAL EXTRACTION IN ARIMA TIME SERIES (SEATS) (X-13 AS henceforth) is a seasonal adjustment method developed by the US Census Bureau. The X-13 version is an enhancement of the earlier X-11 and X-12 ARIMA methods originally developed by Shiskin, Young, & Musgrave (1967). X-13 AS has a similar construct to its previous version that uses regression models with errors that follow an ARIMA process. The method uses standard notation, i.e. $(p, d, q)(P, D, Q)s$ for ARIMA models, where (p, d, q) refers to the orders of the *non-seasonal* autoregressive (AR) component (p), number of differencing (d), and moving average (MA) operators (q) in the model. Similarly, $(PDQ)s$ refers to the orders of *seasonal* autoregressive, differencing, and moving average operators in the model, while (s) denotes the seasonal periods, e.g., $s = 12$ in case of monthly data.

A general multiplicative seasonal ARIMA model for a time series z_t can be specified as follows. Consider the expression

$$\theta(B)\Theta(B^s)a_t = \phi(B)\Phi(B^s)(1 - B)^d(1 - B^s)^D z_t \tag{6}$$

where

$\Phi(B^s) = (1 - \Phi 1B^s - \cdots - \Phi_P B^{Ps})$	is the *seasonal* AR operator, and
$\Theta(B^s) = (1 - \Theta_1 B^s - \cdots - \Theta_Q B^{Qs})$	is the *seasonal* MA operator
$\phi(B) = (1 - \phi 1B - \cdots - \phi_p B^p)$	is the *non-seasonal* autoregressive (AR) operator
$\theta(B) = (1 - \theta 1B - \cdots - \theta_q B^q)$	is the *non-seasonal* moving average (MA) operator
B	is a back shift operator, i.e. $(Bz_t = z_{t-1})$
s	is a seasonal period, $(s = 12)$ for a monthly series
a_t	are $i.i.d$s with zero mean and σ^2 as variance

In the expression the term $(1 - B)^d(1 - B^s)^D$ implies non-seasonal differencing of order d and seasonal differencing of order D. Further, as a simplification, if $d = D = 0$, i.e. there is no differencing on either non-seasonal or seasonal component it is possible to replace the error z_t by its deviations from its mean, that is, $z_t - \mu$ where $\mu = E[z_t]$. To make the model more useful and operational, a time varying mean function is modeled. Specifically, consider a simple linear regression equation that models a time series y_t as follows;

$$y_t = \sum_i \beta_i x_{it} + z_t \tag{7}$$

where y is a variable of interest, x is a set of observable covariates and βs are coefficients. As usual, the error z_t can be written as $z_t = y_t - \sum_i \beta_i x_{it}$, which in this particular case is assumed to follow an ARIMA process as specified in equation 6.

The basic difference in this approach as opposed to the conventional assumptions of the error term in a regression are that this model assumes that errors are autocorrelated over time. A standard regression assumes that errors are uncorrelated over time, which if applied in this case can lead to invalid results. The ARIMA process provides a remedy in this case by achieving multiple objectives; (i) making the variable stationary, (ii) extracting information by allowing autocorrelated errors in the model and (iii) producing empirically robust results. The procedure is achieved by substituting equation 7 in equation 6 and combining to obtain

$$\phi(B)\Phi(B^s)(1-B)d(1-B^s)^D \underbrace{\left(y_t - \sum_i \beta_i x_{it}\right)}_{\approx z_t} = \theta(B)\Theta(B^s)a_t \tag{8}$$

where the terms are defined as earlier and z_t now corresponds to the term in parenthesis. Using this expression, the estimation process can be understood as follows. First, the model allows for a mean function $\sum_i \beta_i x_{it}$ based on the observed x_ts. Second, errors (i.e. z_t) now follow an ARIMA process depending on lag and differencing specifications of the model. Third, the effects of coefficients obtained from the regression are subtracted from the original series y_t to generate a zero mean series of z_t. Fourth, based on the differencing specifications, the z_t series is differenced to obtain a stationary series w_t. The new series w_t is assumed to follow an ARMA process (since the series has already been differenced) which is captured by the terms $\phi(B)\Phi(B^s)w_t = \theta(B)\Theta(B^s)a_t$. Rewriting the equation in its reduced form, we have;

$$(1-B)^d(1-B^s)^D y_t = \sum_i \beta_i (1-B)^d(1-B^s)^D x_{it} + w_t \tag{9}$$

where w_t follows the procedure outlined above. The final outcome is achieved by differencing both y_t and x_t by the operator $(1-B)^d(1-B^s)^D y_t$. The model can be extended to include lagged terms of x_t, i.e. x_{t-1}, x_{t-2}, etc. depending on the purpose at hand (See US Bureau of the Census (2013)).

* * * * * * *

References

Astolfi R., Ladiray D., & Mazzi G. L (2001) Seasonal adjustment of European aggregates: Direct versus indirect approach, Working Document 14/2001, Theme 1, *General Statistics*, EuroStat

Bhattacharya R., Pandey R., Patnaik I., Shah A. (2016) Seasonal adjustment of Indian macroeconomic time-series, *NIPFP Working Paper*, No. 160, Jan, 2016

CSO (Undated) *Study on internal consistency of All India Index of Industrial Production data*, Economic Statistics Division, Central Statistics Office, Ministry of Statistics and Programme Implementation, New Delhi

CSO (Working Group (WB)) (2014) *Report of the working group for the development of methodology for compilation of the all India Index of Industrial Production with base year 2009-10/ 2011-12*, Central Statistics Office, Ministry of Statistics and Programme Implementation, New Delhi

Dagum, E. B. (1988) *The X-11-ARIMA/88 Seasonal Adjustment Method: Foundations and User's Manual*, Ottawa: Statistics Canada

ESS (2015) *ESS guidelines on seasonal adjustment, EuroStat Manuals and Guidelines*, 2015 Version, European Union: Luxembourg

Findley, D. F., Monsell, B. C., Bell, W. R., Otto, M. C., and Chen, B. C. (1998) New Capabilities and Methods of the X-12-ARIMA Seasonal Adjustment Program, *Journal of Business and Economic Statistics*, Vol. 16, pp. 127176

Gomez, V., & Maravall, A. (1997a) Programs TRAMO and SEATS, instructions for user (Beta Version: September 1996), *Working Paper 9628, Bank of Spain*

Gomez, V., and Maravall, A. (1997b) Guide for Using the Programs TRAMO and SEATS, Beta Version. Madrid: Banco de Espaa

Granger, C. W. J. (1979) Seasonality: Causation, interpretation, and implications. *In A. Zellner (Ed.), Seasonal analysis of economic time series.* New York: National Bureau of Economic Research.

Ladiray, D., and Quenneville, B. (2001) *Seasonal Adjustment with the X-11 Method*, New York: Springer-Verlag

Medel, Carlos A. (2018) A comparison between Direct and Indirect Seasonal Adjustment of the Chilean GDP 19862009 with X-12-ARIMA, *Journal of Business Cycle Research*, Vol. 14, pp. 47-87

Mazumder, Sandip and Soumya Chakraborty (2013) Study on seasonal adjustment of IIP using x12-arima, *The Journal of Industrial Statistics*, Vol. 2(1), pp. 3767

Manna, G.C (2010) Current status of industrial statistics in India: Strengths and weaknesses, *Economic & Political Weekly*, Vol. XLV, No. 46, pp. 67-76

Manna, G.C (2015) Exploring an alternative Index of Industrial Production, *Journal of Industrial Statistics*, Vol. 4, No. 2, pp. 169-177

Nagaraj, R (2002) How to improve India's industrial statistics, *Economic and Political Weekly*, Vol. 37, No. 10 (Mar. 9-15, 2002), pp. 966-970

Nagaraj, R (1999) How Good Are India's Industrial Statistics? An Exploratory Note, *Economic and Political Weekly* Vol. 34, No 6, pp. 6-12

Pandey, Radhika, Amey Sapre & Pramod Sinha (2018) What does the new 2011-12 IIP series tell about the Indian manufacturing sector? *India Growth and Development Review*, Vol. 11, No. 2

Sax, Christopher and Dirk Eddelbuettel (2020) Seasonal: An R package for X-13 ARIMA-SEATS Seasonal Adjustment, available in Sax C, Eddelbuettel D (2018) Seasonal Adjustment by X-13 ARIMA-SEATS in R., *Journal of Statistical Software*, 87(11), 1-17

Scheiblecker, Marcus (2014) Direct versus Indirect approach in Seasonal Adjustment, *WIFO Working Papers*, No. 460, WIFO

Singhi, M.C (2000) *Index of Industrial Production and Annual Survey of Industries*, Working Paper, Unpublished

Shastri, NS (2011) *Report on Statistical Audit of Index of Industrial Production and framework for future statistical audits*, Central Statistics Office, Ministry of Statistics and Programme Implementation, New Delhi

Shiskin, J. A, Young, & J. Musgrave (1967) The X-11 variant of the census method II seasonal adjustment program, Bureau of the Census, Technical Paper No. 15, US Dept. of Commerce, 1967

UNSD (2010) International Recommendations for the Index of Industrial Production 2010, Series F - No. 7, Department of Economic and Social Affairs, United Nations Statistical Division: 2010

US Bureau of the Census (2013) *X-13 ARIMA-SEATS Reference Manual*, Version 1.1. US Bureau of the Census, Washington, DC

MORE IN THE SERIES

- Bhattacharya, R., Sen Gupta, A., and Sikdar, S., (2020). Building Infrastructure to Promote Inclusive Growth, W.P. No. 321 (September).
- Bose, S., Bera M., and Ghosh, P., (2020). Centre-State Spending on Elementary Education: Is it Complementary or Substitutionary? W.P. No. 320 (September).
- Patnaik, I., and Sengupta, R., (2020). Impact of Covid-19 on the Indian Economy: An Analysis of Fiscal Scenarios, W.P. No. 319 (September).

Radhika Pandey is Fellow, NIPFP
Email: radhika.pandey@nipfp.org.in

Amey Sapre is Assistant Professor, NIPFP
Email: amey.sapre@nipfp.org.in

Pramod Sinha is Fellow, NIPFP
Email: Pramod.sinha@nipfp.org.in

National Institute of Public Finance and Policy,
18/2, Satsang Vihar Marg, Special Institutional Area (Near JNU),
New Delhi 110067
Tel. No. 26569303, 26569780, 26569784
Fax: 91-11-26852548
www.nipfp.org.in