Representing the Unreported
Experiences in compilation of accounts for Non-Financial Private Corporate (NFPC) Sector*

Brijendra Singh, Meera A. P., Saumya Mishra

From receipt of only Global PUC figures to receiving entire Frame of Active companies, MGT-7 data, CIN change history etc. , the increasing collaboration with Ministry of Corporate Affairs has aided not only in improving the estimates of NFPC sector through appropriate classification, duplicate detection etc. but has also enabled examination of the present methodology as NAD now has access to unit level data for both the reporting(detailed financial information) as well as non reporting (only basic details) companies. Instead of relying on a single Global Paid Up Capital (PUC) figure, made available in initial days, receipt of disaggregated data has provided an opportunity to factor in many dimensions in estimating the contribution of non reporting companies. This paper is an attempt to investigate some of these along with their likely impact on the estimates.

**NFPC in the present context excludes Quasi Corporations and only includes entities registered under Companies Act*

1. Background

Use of MCA 21 data pertaining to companies was one of the major changes brought about in the new series (2011-12) of the national accounts of India as data for lakhs of companies became available instead of estimates compiled on the basis of financial results of about 2500 companies, made available by RBI, in the earlier series (2004-05). Improved coverage of non financial private corporate sector resulted in an increase of about 10% in the GVA estimates for 2011-12[1] (Annexure – I). As per the guidelines of the Sub Committee on Private Corporate Sector[a] (report available in public domain at https://mospi.gov.in/sites/default/files/publication_reports/final_Report_Goldar_subcommittee2mar15.pdf) it was decided that Paid Up Capital of all active companies may be used in scaling up the estimates of reporting companies, to account for the contribution of active companies that had not filed their returns even after about a year and half of the end of the financial year. Since then, following formula has been used for scaling up the estimates (for second revised estimates):

$Aggregate_{i\_pop}$ (Population estimate) = $Aggregate_{i\_Rep}$ (Reporting Companies)*Global PUC/PUC of Reporting Companies

[$Aggregate_1$= Output, Aggregate $_2$= Intermediate Consumption ………$Aggregate_n$= Savings]

In absence of availability of any information on aggregate related to the population of active companies, the estimates were compiled on the basis of the global paid up capital made available by the Ministry of Corporate Affairs. Even though, the global PUC figures were made available by MCA for a few broad categories of industry (Industry disaggregation used in NAS is much more detailed), only overall PUC, for public and private limited companies, was used as a scaling up factor. This was necessitated as the unit level data of industry classification for the population wasn't made available and much deviation was found (initially through manual profiling and later through use of other databases like ASI, MGT-7, MGT-9 etc.) in the industry of operation of companies vis a vis the industry indicated at the time a company was incorporated. The latter resulted in a possibility of the same company classified differently in global PUC than the industry classification of reporting company which was scrutinised, at least in case of bigger companies. Hence, while overall Global PUC figures were thought to be robust, those at industry level weren't used. As at the time of second revised estimates about 85-90% of companies (in terms of size as indicated by PUC) had filed their Annual Reports, it was felt that the results would not be unduly impacted by the methodology used.

Table 1: Scaling Up factor[2]

	2012-13			2013-14			2014-15			2015-16		
	PLC	*PTC*	*Total*	*PLC*	*PTC*	*Total*	*PLC*	*PTC*	*Total*	*PLC*	*PTC*	*Total*
Scaling Up factor	**1.13**	**1.17**	**1.15**	**1.15**	**1.13**	**1.14**	**1.15**	**1.18**	**1.17**	**1.15**	**1.11**	**1.13**

	2016-17			2017-18			2018-19			2019-20		
	PLC	*PTC*	*Total*	*PLC*	*PTC*	*Total*	*PLC*	*PTC*	*Total*	*PLC*	*PTC*	*Total*
Scaling Up factor	**1.20**	**1.14**	**1.17**	**1.12**	**1.13**	**1.12**	**1.12**	**1.14**	**1.13**	**1.09**	**1.13**	**1.11**

Note: PLC: Public Limited Company, PTC: Private Limited Company

The results obtained as per above methodology were corroborated with growth observed in top companies comprising sizeable chunk in each Industries. Hence even though some like Nagaraj & Srinivasan[b] have argued that the number of companies for which accounts are available varies from year to year, leading to changes in "blowing-up" or "scaling-up" factor and correspondingly, thus greatly affecting the final estimates, we find that the seemingly huge variation in the count of companies filing the return cited by them pertains to very small ones who do not have much contribution in overall GVA, thus the estimates are not significantly impacted and the scaling up factor has not changed significantly compared to changes in the coverage as reflected by the count of companies.

The discussion around estimates of Private corporate sector being affected by inclusion of shell companies[$] seems to be more an issue of data validation rather than about estimation methodology. It is also felt that in view of recent initiatives being undertaken by Ministry of Corporate Affairs, the issue of shell companies will be significantly reduced. Till 2020-21 about 3.8 lakh shell companies[3] have been struck off u/s 248 (1) of the Companies Act, under the special drives undertaken by Registrar of Companies. In case shell companies majorly featured in non reporting ones, it may be argued that scaling up method would unduly scale up the estimates to account for them. However, as shell companies do not have significant physical presence etc. their contribution in the PUC of non reporting companies might be insignificant. The same is corroborated by the fact that 2017-18 and 2018-19 the years that saw unusually high number of striking off of companies u/s 248, the paid up capital of such companies was only about 13.5 Thousand Crore on the average[4], accounting for about 0.5% of the global paid up capital of Rs over 25 lakh Crore. These figures of striking off would also include companies other than those identified as shell companies struck off for other reasons besides the special drive of MCA.

$ Organization for Economic Cooperation and Development (OECD) defines a Shell Company as a company which is formally registered or otherwise legally organized in an economy but which does not conduct any operation in that economy other than in a pass through capacity.

As to the issue of self selection pointed out by Nagaraj & Srinivasan, use of scaling up factor at a more disaggregated level as indicated in the paper would largely address the issue. In any case the companies that did not file for any given FY can at the most be best represented by similar company in terms of industry and size. However, from analysis of late filing of companies from FRE to SRE and SRE to TRE(latter analysed only for one year in view of hardly any increase in coverage) it appears that there is no significant trend positive or negative in the performance of such companies , which may lead to a bias.

Regarding the blow up factor in case of 85% coverage in terms of PUC being 1.15, the same is 1.176 (100/85) as suggested by Sapre and Sinha[c] in the working paper series of NIPFP and not 1.15. Also, use of Industry wise growth rate as suggested by Sapre and Sinha is largely used in case of FRE. However, instead of applying the same at unit level as suggested in the paper, the method could only be applied to aggregates at industry level for the want of disaggregated data of non-reporting companies. Also, when we tried populating the frame of active companies for the recent year (i.e 2020-21) by the filings available at least once in last three years (i.e. 2017-18 to 2020-21) (as suggested in the paper in accordance with definition of active companies), it was found that more than 4 lakh companies could still not be populated. It might happen as a company could turn from dormant or inactive status to active, might be newly registered in last couple of years etc. without filing any annual return subsequently. Change in CIN could also contribute to the same , as CIN is normally used as unique identifier for all such operations. Hence, unit level imputation might be not be feasible for large number of companies. Compared to the same the industry wise growth of common companies applied to the aggregates estimated during SRE, the method followed in FRE, is simpler, but is not very sensitive to the company demographics like births and deaths etc. observed during the FY . Though the size of companies that cease to exist may not be that large, the additions to the active frame each year is not that small as indicated by the tables below. New companies accounted for 0.6 % and 1.3% of GVA and PUC of the reporting companies during 2019-20.

Table 2: Number of Active Non Government Companies[3]

Years	2016-17	2017-18	2018-19	2019-20	2020-21
Count	958549	1024664	1092460	1177200	1235405
Difference		66115	67796	84740	58205

The data in respect of Non-Government Companies Registered and Struck off u/s 248 during last 3 Financial Years is presented in Table 3.

Table 3[2&4]

Financial Year	Share (%) of Authorized Capital of companies registered during the year in the Aggregate PUC of Active Companies	Share (%) of PUC of companies struck off u/s 248 during the year in the Aggregate PUC of Active Companies
2019-20	7.66	0.21
2020-21	2.56	0.04
2021-22	5.99	0.20

Data Source: Annual Reports of MCA and Master Frame of Active Companies

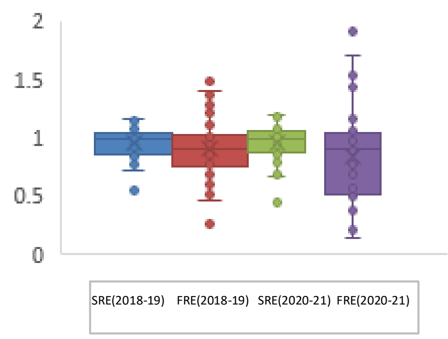
As indicated by the tables above, the issue of increase in case of active companies is significant though the same is addressed by including companies with zero values in previous years for the common companies set. Moreover, the data received during FRE is much smaller compared to SRE, leaving aside scope of future revision.

In the 2011-12 series, different methodology was used for estimating contribution of NFPC sector during FRE and SRE as the coverage across industries was much more divergent compared to SRE, making application of a uniform scaling up factor across all industries less accurate in industry level estimation.

Table 4: Summary of Industry Wise Coverage in Terms of PUC (Major industries)[2]

	Second Revised Estimates			First Revised Estimates		
	2018-19	2019-20	2020-21	2018-19	2019-20	2020-21
Coverage in terms of PUC(normalized)						
Min	0.8	0.9	0.8	0.5	0.3	0.4
Max	1.1	1.1	1.2	1.4	2.5	1.9
Range	0.3	0.3	0.4	0.9	2.1	1.5
Coverage in terms of PUC						
Min	69.7	72.4	64.3	29.6	11.9	17.7
Max	97.1	95.0	97.5	88.1	86.0	82.7
Overall	84.5	83.6	81.3	62.8	34.9	43.4
Range	27.4	22.	33.2	58.6	74.1	64.9
Count(lakh)	6.7	7.	7.2	5.0	0.8	1.4

Graph 1: Industry Wise Representation in FRE and SRE[2]



Also, Provisional Paid Up Capital Figures (Global) provided at the time of FRE were sometimes significantly revised at the stage of SRE, reducing the sanctity of application of global PUC figures as a scaling up factor at the stage of FRE. During the last three years for which both set of figures are available, i.e. 2018-19 - 2020-21 the difference between the Provisional and Final Global PUC Figures, on the average was around Rs 1.2 lakh Cr or about 4 % vis a vis average PUC figures of Rs 28.5 lakh Cr. However, industry wise growth served as a pointer subsequently at the stage of SRE wherein only overall shortfall was uniformly accommodated across all industries. While the methodology of FRE, being based on growth of common companies did not adequately address the issue of company demographics, the same for the last FY was already accommodated through SRE.

In case common panel of companies is considered across years, as suggested by some to overcome the problem of self selection, the same would result in considerable loss of data .

Table 5[2]

Filings and number of companies common over the years (Excluding PSUs)					
	2011-12	2012-13	2013-14	2014-15	2015-16
Count of companies	557559	557559	605213	596102	628329
Count of pairwise common companies			486631	522428	524214
	2016-17	2017-18	2018-19	2019-20	2020-21
Count of companies	704564	623691	721658	767198	768073
Count of pairwise common companies	557376	521194	521194	605477	684967
Companies Common across all years	235196				

2. Testing the assumptions in present methodology :

2.1 Impact of Industry -wise Coverage

Even if the application of an overall scaling up factor to account for the non reporting companies worked reasonably well for the overall economy (keeping in view the small size being estimated), it would not provide good estimates at industry level unless nearly equal fraction of companies (equal to overall shortfall) did not report across all industries. In case the reporting fraction for any given industry was more than the overall reporting fraction , application of a uniform overall scaling up factor would lead to an over estimate for the industry and vice versa. In case the trend was sustained across all years, there would be a likelihood of an upward or downward bias for the given industry in general. However, if the fraction oscillated sometime above the overall fraction, sometimes below it, the same would distort the industry level growth rate, though it might not lead to a biased estimate over the years. Observations of the industry (selected) wise reporting fraction over the years vis a vis the overall coverage of reporting companies (in terms of PUC) is given in the Table 6[2] below :

Table 6[2]

Selected Industries	Codes	Ratio of Industry-wise coverage of reporting companies to aggregate coverage of reporting companies (in terms of PUC)				Share (%) of Industry in 2019-20 overall NFPC (excluding Quasi corporations) GVA Estimates	Impact/ Behaviour
		2017-18	2018-19	2019-20	2020-21		
Agriculture, forestry & fishing	**A1, A2, A3, B1**	0.80	0.91	0.91	0.79	0.44	**Under Represented**
Mining	**C1, C2, C3**	0.82	0.86	0.93	0.95	1.28	
Other Business Activities	**K5**	0.86	0.87	0.92	0.92	8.89	
Manufacturing	**D1**	0.98	1.01	1.03	1.06	42.86	**Nearly accurate**
Retail Trade	**G2**	0.98	1.04	1.00	1.01	2.60	**Nearly accurate**
Construction	**F1**	1.06	0.99	0.96	0.96	5.65	**Nearly accurate**
Telecommunication	**I5**	1.09	1.07	1.09	1.05	3.22	**Over Represented**
Post and Courier	**IP**	1.25	1.16	1.17	1.06	0.16	
Computer and Related Activities	**K3**	1.12	1.05	1.06	1.12	18.37	
Land Transport	**I1**	0.82	0.84	1.13	0.98	0.89	**Fluctuating**
Wholesale Trade	**G1**	1.03	1.01	1.01	0.84	3.32	**Fluctuating**

It is evident from the data given at Table 6 that in case of Industries like Agriculture, forestry & fishing, Mining and Other Business Activities, the industry -wise coverage is less than aggregate coverage and hence it is likely that scaling up might not be adequately accounting for the non reporting companies. At the same time in case of Telecommunication, Post and Courier and Computer and related activities there are chances that the shortfall is more than compensated for. Some industries like Land Transport, Wholesale Trade indicated fluctuating trend, indicating that year to year growth rates for the industry might be more prominently impacted. However, Manufacturing Industry with largest share in the

NFPC sector in terms of Gross Value Added (GVA) had coverage close to the overall coverage increasing the likelihood of nearly accurate estimates for this sector. Same is the case with industries Retail Trade and Construction.

Apart from the likely impact of the current methodology at the industry level, it might be worthwhile to examine the overall impact in the economy as well. Even in case the representation isn't uniform across industries, present methodology might work well in case GVA to PUC ratio is nearly equal across all industries as shown below (an assumption that is less likely to hold given the fact that capital intensity(Capital required per unit of GVA) is expected to vary across industries).

Considering the case of common blow up,

$$GVA_{pop.} = \frac{\sum_i \mathbf{GVA_{i\_rep.}}}{\sum_i \mathbf{PUC}_{i\_rep.}} \text{ x } PUC_{pop.} \quad ; \text{ i= industries}$$

If $\frac{\mathbf{GVA_{i\_rep.}}}{\mathbf{PUC_{i\_rep.}}}$ = k (constant) $\Longrightarrow$ $GVA_{i\_rep.} = \mathbf{k \text{ x } PUC_{i\_rep.}}$

$$\Longrightarrow \quad \sum_i GVA_{i\_rep.} = k\ X \sum_i PUC_{i\_rep.}$$

$$\Longrightarrow \quad GVA_{pop.} = \frac{k\ X \sum_i PUC_{i\_rep.}}{\sum_i PUC_{i\_rep.}} \text{ X } PUC_{pop.} = k \text{ X } PUC_{pop} \longrightarrow (1)$$

Considering the case of industry-wise blow up,

$$GVA_{pop.} = \sum_i \frac{\mathbf{GVA_{i\_rep.}}}{PUC_{i\_rep}} \text{ x } PUC_{i\_pop.} \quad ; \text{ i= industries}$$

$$\frac{\mathbf{GVA_{i\_rep.}}}{\mathbf{PUC_{i\_rep.}}} = k \text{ (constant)} \Longrightarrow GVA_{pop.} = k \text{ x } \sum_i \mathbf{PUC_{i\_pop.}} = k \text{ X } PUC_{pop} \longrightarrow (2)$$

Above formula indicates that in case capital intensity is uniform across industries, estimates obtained from a single scaling up factor would provide reasonably good estimate at the overall level, irrespective of the fact whether representation across industries was nearly uniform of not.

Table 7: Illustrative Example- uniform GVA/PUC Ratio (1.7) and unequal coverage

Population			Reporting			Est Using Overall Blow Up	Est Using Ind Wise Blow Up
Industry	**GVA (1)**	**PUC (2)**	**GVA (3)**	**PUC (4)**	**Coverage (5)**	**GVA (6)=(3)*1.9448**	**GVA (7) =(3)*(2)/(4)**
A	125	75	100	60	0.8	194.48	125
B	50	30	25	15	0.5	48.62	50
C	750	450	525	315	0.7	1021.02	750
D	555	333	111	66.6	0.2	215.87	555
Total	1480	888	761	456.6	0.5	1480	1480
Overall Scaling Up factor(888/456.6)=1.9448							

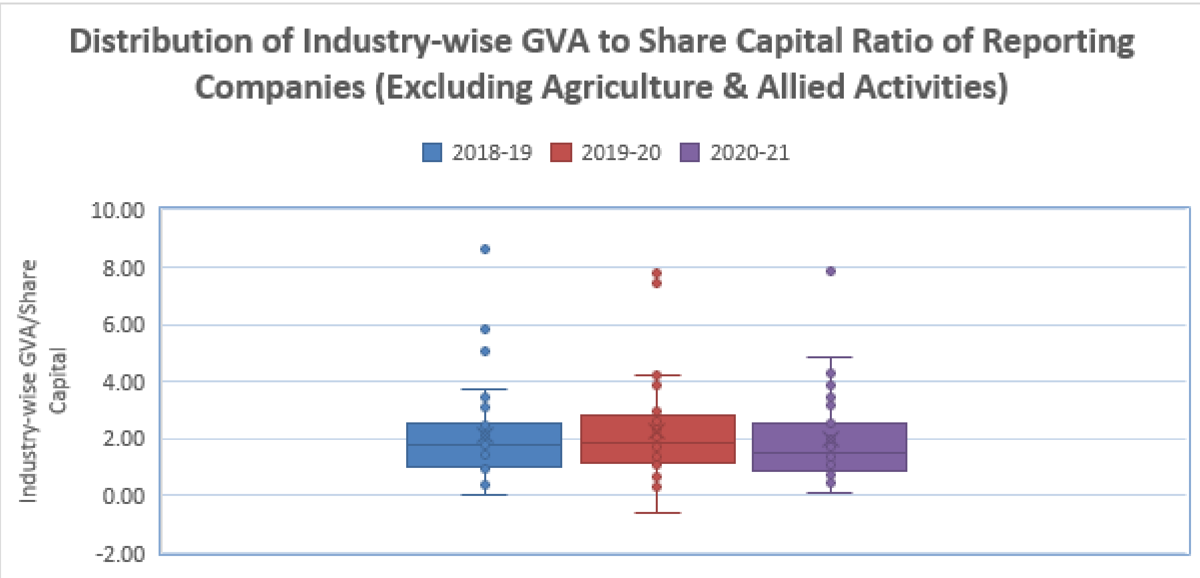
However, the assumption of uniform capital intensity is less likely to hold compared to the assumption of nearly uniform representation across industries discussed earlier in the context of application of uniform scaling up factor. Accordingly, the extent of departure was examined in the data and the same is

indicated in the Table 8[2] below.

Table 8: The time series of Industry-wise Ratio of GVA to Share Capital of reporting NFPCs belonging to selected Industries[2]

Industry code description	Industry Code	2018-19	2019-20	2020-21
Electricity	E1	0.46	0.55	0.45
Real estate	K1	0.34	0.91	0.75
Wholesale trade	G1	1.01	1.13	1.51
Construction	F1	1.18	1.19	1.06
Telecommunication	I5	1.09	1.30	1.35
Retail trade	G2	1.49	1.61	1.39
Manufacturing	D1	2.68	2.45	2.21
Other business activities	K5	2.46	3.25	2.52
Research and development	K4	3.42	3.89	4.31
Activities of membership etc.	O2	5.85	7.82	4.87
Computer and related activities	K3	8.62	7.45	7.82
Total		**1.87**	**2.15**	**2.00**

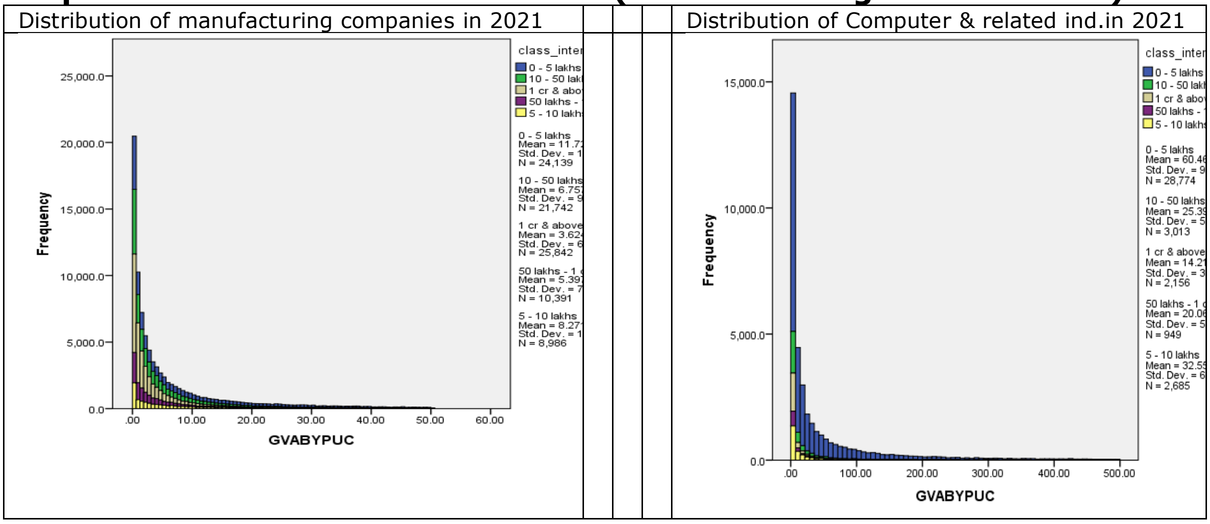
Graph 2[2].



As expected, above table and graph indicate that the second assumption is much weaker as the capital intensity (GVA to PUC ratio) across industries varies much more than the coverage in terms of PUC. A likely repercussion of the same would be that over representation of high GVA to PUC ratio industry companies would have a tendency to increase the estimates whereas over representation of low GVA to PUC ratio industry companies in the reporting set would have a tendency to pull it down.

Also GVA to PUC ratio for any given industry varies significantly across different size class within the same industry as indicated in the graph for a couple of industries below. It is found that the ratio is higher in small size class compared to the Bigger size classes.

Graph 3: Size class wise distribution (after removing a few outliers)[2]



3. Comparison of different methods

Different methods were compared across years to study the impact on estimation. The comparison was possible only after 2017-18 as the frame was made available for the first time during the year. Strictly speaking only methods 2 to Method 6 were comparable as the figures of Global PUC made available by MCA (used in method 1) were generally smaller compared to those arrived at using the frame resulting in smaller scaling up factor in the first method. Amongst other reasons, frame being supplied at a later date contained information as on date of extraction in absence of any time stamp. For method 6, to make the methods comparable, levels arrived at by using method 3 in 2017-18 were moved forward.

Table 9: Differences in estimates (Rs Lakh Cr) of different method from the average (Average- Mi)[2]

Year	M1	M2	M3	M4	M5	M6
2018-19	2.7	0.5	1.4	-1.4	-0.3	-0.3
2019-20	4.5	-1.1	2.6	-2.6	-0.1	1.1
2020-21		-0.9	3.8	-3.3	-0.4	0.8

M1: Global PUC based overall scaling up factor

M2: Frame PUC Based overall Scaling up Factor

M3: Frame Based Industry Wise Scaling Up factor

M4 :Frame Based Size Class Wise Scaling Up Factor

M5: Frame Based Industry X Size Class (Annexure II)Wise Scaling Up Factor

As M1 is based on a different data set (Global PUC) estimates may not be strictly comparable and have been excluded in calculating average	**M6:** Growth of Common Companies

Average = Average (M2:M6)

Above table reveals that compared to method 2, method 3 usually results in smaller estimates. The same is explained largely on account of differences in manufacturing and computer & related industries which together accounted for about 90 percent of the variation. As both these industries were over represented in the sample, scaling up factor decreased when industry wise scaling up method was used instead of overall scaling up method. Further, as manufacturing industry had a bigger size with bigger GVA to PUC ratio compared to the overall average, and Computer Industry though not as big as manufacturing, had significantly higher GVA to PUC ratio compared to the average, both resulted in significant decrease when smaller scaling up factor was applied due to industry level method. However, the comparison is slightly different during 2017-18 as manufacturing industry is slightly underrepresented (manufacturing coverage/overall coverage is 0.98). Industry level variation is accounted for in method 3, 5 & 6.

Table 10[2]

	Scaling Up		**GVA-PUC Ratio**	
Industry	2019-20	2020-21	2019-20	2020-21
Manufacturing	1.16	1.16	2.45	2.21
Computer & Related Ind.	1.13	1.10	7.45	7.82
Overall	1.20	1.23	2.15	2

The estimates, however, increased on applying the scaling up factor at a more disaggregated level i.e. industry x size class in method 5, compared to application of industry level scaling up factor in method 3.

This was largely on account of divergence in estimates of manufacturing (D1) to a lesser extent and Computer and related Industries (K3) & Other Business services(K5) to a larger extent. It was observed that generally GVA to PUC ratio was bigger in smaller size classes compared to the bigger ones and the level of reporting in the overall sample within the industry was lesser in the smaller size classes. The comparatively higher share and the significant divergence led to D1, K3 and K5 accounting for most of the divergence between estimates derived from method 3 and method 5. Table below indicates the effect for three years 2018-19, 2019-20 and 2020-21. Further, as indicated in the Graph before and in the table below, the divergence across size class is much more in K3 and K5, leading to more contribution in the overall divergence between method 3 & method 5 even though their share is less than D1.

Table 11: Size Class wise Scaling Up factor and GVA to PUC ratio for D1, K3 & K5 as a ratio of industry level scaling up and GVA to PUC ratio for the same[2]**.**

Normalized Scaling up factor	**2018-19**	**0 - 5 lakhs**	**5 - 10 lakhs**	**10 - 50 lakhs**	**50 lakhs - 1 cr**	**1 cr & above**
	D1	1.5	1.3	1.2	1.1	1.0
	K3	1.6	1.4	1.2	1.2	1.0
	K5	1.4	1.3	1.0	1.0	1.0
Normalized GVA to PUC ratio	D1	16.5	6.7	3.7	2.8	0.9
	K3	21.6	6.8	4.5	3.8	0.8
	K5	20.6	6.0	4.8	2.9	0.5
Normalized Scaling up factor	**2019-20**	**0 - 5 lakhs**	**5 - 10 lakhs**	**10 - 50 lakhs**	**50 lakhs - 1 cr**	**1 cr & above**
	D1	1.5	1.4	1.2	1.2	1.0
	K3	1.5	1.4	1.2	1.2	1.0
	K5	1.4	1.3	1.0	1.0	1.0
Normalized GVA to PUC ratio	D1	16.3	7.1	4.3	3.0	0.9
	K3	21.6	6.4	4.9	4.0	0.8
	K5	30.6	8.9	6.5	3.9	0.5
Normalized Scaling up factor	**2020-2021**	**0 - 5 lakhs**	**5 - 10 lakhs**	**10 - 50 lakhs**	**50 lakhs - 1 cr**	**1 cr & above**
	D1	1.9	1.6	1.2	1.1	1.0
	K3	2.0	1.7	1.2	1.2	1.0
	K5	1.6	1.5	1.0	1.0	1.0
Normalized GVA to PUC ratio	D1	18.7	8.4	4.6	3.2	0.9
	K3	26.9	6.8	5.0	3.9	0.8
	K5	34.9	10.8	6.3	4.6	0.5

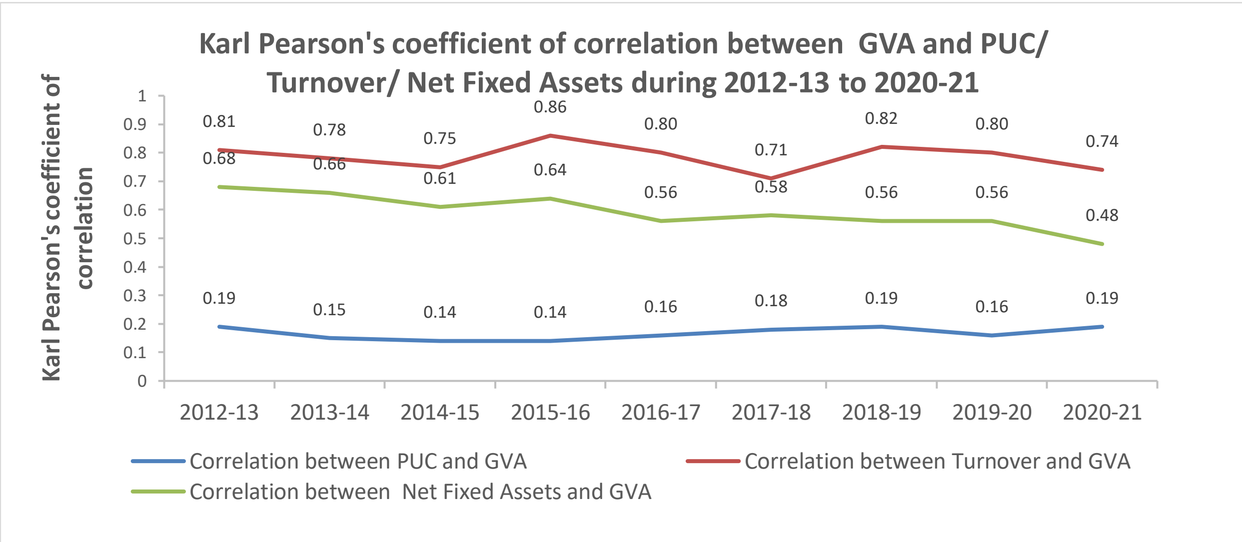
On the average, during last three years (i.e. 2018-19 to 2020-21) in case, the scaling up is only carried out for **size classes** ignoring industry wise differentials, the estimates are scaled up more (with the estimates at Rs. 45, 51 & 53 lakh Cr respectively) than industry X size class wise estimates (with the estimates at Rs. 44, 48 & 50 lakh Cr respectively). Thus considering only one dimension i.e. industry wise scaling up (with estimates at Rs 43, 45 & 46 lakh Cr) or size class wise scaling up, estimates vary much more, with former providing lower estimates on the average and latter providing higher estimates) compared to the estimates scaled at overall (Rs. 44, 49 & 51 lakh Cr) and Industry X size class levels lying in between. Estimates scaled up overall are more closely aligned with most disaggregated scaled up estimates i.e. Industry X Size class, with differences being positive and negative in different years whereas compared to only one dimension scaling up (i.e. industry or size class wise) the difference (with Industry X Size class scaling) is bigger in magnitude and is unidirectional i.e. either positive or negative.

Method of common companies growth (method 6) moved more in tandem with method 5. The differences could be on account of truncated size (common companies) and the company demography i.e. births and deaths of companies, which were missed in common companies growth. As the impact of births outweighed that of deaths, exclusion of the these is expected to result in slightly lower estimates in method 6 with cumulative effect.

4. Alternatives to overall PUC based scaling-up Factor

The present PUC based blow-up method relies on the assumption that PUC and GVA have a linear relation. Hence, in case of incomplete data, PUC of the available companies can be used to infer the value addition of the un-reporting active companies. In order to check the validity of this basic assumption a time series analysis was done using the MCA reporting data and the results are presented in Graph below. The exercise was also carried out for other alternative indicators like fixed asset and turnover (which has a more direct relationship with GVA) to explore if they could be better substitutes for PUC. Initially PUC was used as it was the only variable for which population figures were available. Also, since it was being published by MCA for some years, it was felt that the data quality would be acceptable.

Graph 4[2].



Above chart reveals, as expected, that correlation with GVA improves significantly if turnover (vis a vis PUC) is considered (as turnover has direct relationship with GVA), with correlation between asset and GVA lying in between. Considering other variables related to capital, it is found that during 2020-21 correlation improves marginally (from 0.19 to .35) on considering long term borrowing as well (besides equity) and increases significantly (from 0.19 to .55). in case of capital employed (Total assets – current liabilities = Equity + Noncurrent liabilities) which is more comprehensive measure of size of capital. However, the choice of variables would also depend on data quality & ease of interpretation. For example, turnover as a scaling up variable is more volatile and might require adjustments during years of upturn or downturn. Further, as the share of services sector is increasing, and the GVA of services sector (computer & related industries K3, other business services K5) isn't that closely linked to fixed assets as is the case with manufacturing(D1), the correlation coefficient between asset and GVA reflects a declining trend in the graph above. (Application of Spearmans rank correlation, to discount for the effect of extreme values also resulted in similar result Year 2020-21 Corr with PUC :0.26, turnover :0.86, NFA: 0.54, Capital employed : 0.48).

The strength of correlation in respect of seven major Industries (viz. Manufacturing, Construction, Wholesale trade, Retail trade, Telecommunication, Computer and related activities and Other business activities, comprising around 85% GVA for 2019-20 excluding Quasi Corp.) during 2012-13 to 2020-21 is given in Table. In alignment with the correlation at overall level, more industries have strong correlation with GVA in case of turnover compared to net fixed assets with PUC exhibiting moderate to weak correlation for most of the industries. The industry wise correlation between capital employed & GVA is more aligned than that of fixed assets.

Table 12: Correlation between PUC/ Turnover/ Net Fixed Assets and GVA of selected Industries during 2012-13 to 2020-21[2]

<table>
<tr><th>Strength of Correlation</th><th colspan="2">Between PUC and GVA</th><th colspan="3">Between Turnover and GVA</th><th colspan="3">Between Net Fixed Assets and GVA</th></tr>
<tr><td>strong
(> 0.75)</td><td rowspan="2"></td><td></td><td>D1, F1, I5, K3</td><td></td><td rowspan="2">G2, K5</td><td rowspan="2">D1, G2, I5</td><td></td><td rowspan="2"></td></tr>
<tr><td>Moderate
(> 0.5 and <= 0.75)</td><td rowspan="2">G2, I5</td><td rowspan="2"></td><td rowspan="2">G1</td><td rowspan="2">F1, K3, K5</td></tr>
<tr><td>weak
(< =0.5)</td><td>D1, F1, G1, K3, K5</td><td></td><td></td><td>G1</td></tr>
</table>

Note:

Industry codes	*Description*	*Industry codes*	*Description*
D1	*Manufacturing*	*I5*	*Telecommunication*
F1	*Construction*	*K3*	*Computer and related activities*
G1	*Wholesale trade*	*K5*	*Other business activities*
G2	*Retail trade*		

In order to ascertain the impact of PUC as scaling up factor, companies contributing to 10% PUC coverage were randomly missed in 9 (PUC up to 10 Cr) out of the 14 size classes (Annexure III) in line with the observation that non reporting is more in smaller size classes. The scaling up factor was applied at different disaggregation (i.e. overall, size class wise, Industry Wise, Industry x Size Class wise) to account for variability in two dimensions i.e. across industries (37) & size class (14). It was found that use of size class, resulted in significant improvement in overall accuracy even if industry wise distribution wasn't considered.

Table 13: Results of 10 iterations of scaling up factor applied to 2020-21 data[2]

	Overall (Public/Private)	Industry Wise	Size Class Wise	Industry x Size Class Wise
Average Accuracy	73.7	75	93.2	90.4
Min	72.9	74.2	90.5	86.8

Max	74.4	75.7	96.0	99.0
Std Deviation	0.5	0.5	1.8	4.3

The pattern of variability in the actual data however, indicates that considering both industry & size class matters with impact of latter being more.

The overall scaling up resulted in under estimation, overestimation or random fluctuations in case the companies were randomly missed only from lower size classes(1-9) ,Upper size classes (13 & above) or randomly across all size classes respectively. Missing from medium size class only also generally resulted in underestimation.

Comparison of overall scaling up with different choice of indicators indicates that indicators like Capital employed which is more comprehensive measure of size compared to PUC & fixed assets provide more accurate estimates even at overall level along with more direct indicators like turnover. However, the latter is more volatile to economic conditions (necessitating adjustments) whereas capital employed is stable.

Table 14: Results of 10 iterations of scaling up factor (overall) applied to 2020-21 data[2]

	PUC Based	Fixed Assets Based	Capital Employed Based	Turnover Based
Average Accuracy	73.7	75.4	92.6	89.2
Min	72.9	73.1	88.2	85.4
Max	74.4	78.9	99.7	91.9
Std Deviation	0.5	2.0	4.4	2.2

5. Conclusion

Scaling up factor method, though preferable, as it provides independent estimates for each year without any loss of data and taking cognizance of company demography, is influenced by disaggregation at which it is applied (overall, industry and size class) and the choice of scaling up variable per se. Common companies growth method skirts some of the issues like data quality of an additional variable (which is otherwise required for the population in case of scaling up), its relationship with target variable etc., but requires adjustments for changes in companies demography, which might otherwise get compounded over the years, in cases where such impact is significant. In view of the violation of assumptions of uniform non reporting across industries and size class and differences in capital intensity across industries and size class, it might be more robust to apply paid up capital, as a scaling up factor, at a more disaggregated level (i.e. Industry x Size Class), in case the same variable is continued with.

Contrary to the initial perception, it was generally found that considering only one dimension (industry or size class, with former providing lower estimates on the average and latter providing higher estimates) estimates diverged much more, with estimates scaled at overall and Industry X size class levels lying in between.

Also, use of other variables like Capital Employed, Turnover etc. provided much improved estimates even at overall level of scaling up. However, data quality of the variable, its stability over the years (turnover being more volatile might require adjustments during years of upturn and downturn) and its availability for populating the frame would also affect the choice of variable.

6. Disclaimer

Though the authors are working in the National Accounts Division, National Statistical Office, MoSPI, Government of India, the views expressed are personal and do not necessarily reflect the position of Government of India.

Data Sources

1. Changes in methodology and data sources in the new series of National Accounts Base Year 2011-12, National Accounts Division, Ministry of Statistics and Programme Implementation, Government of India, New Delhi
2. MCA 21 data
3. Ministry of Corporate Affairs
4. Annual Report of Ministry of Corporate Affairs

References

a. Final Report of the Sub-Committee on Private Corporate Sector including PPPs, National Accounts Division, Ministry of Statistics and Programme Implementation, Government of India, New Delhi

b. Measuring India's GDP Growth: Unpacking the Analytics and Data Issues behind a Controversy That Has Refused to Go Away, R. Nagaraj, Indira Gandhi Institute of Development Research T. N. Srinivasan, Yale University.

c. Some areas of concern about Indian Manufacturing Sector GDP estimation, NIPFP Working Paper Series, No. 172, 22 Aug16 Amey Sapre and Pramod Sinha.

Annexure I

GVA for non-financial private corporate sector excluding quasi - corporate sector in 2011-12[1]

(Rs. crore)

Sl. No.	Industry	2004-05 Series	2011-12 Series	% Difference
1.	Agriculture, forestry & fishing	35591	8878	-75.1
2.	Mining & quarrying	23001	39159	70.2
3.	Manufacturing	761593	980452	28.7
4.	Electricity, gas, water supply and other utility services	19658	52252	165.8
5.	Construction	101355	138242	36.4
6.	Trade, repair, hotels & restaurants	274582	100578	-63.4
7.	Transport, storage, communication & services related to broadcasting	91705	155495	69.6
8.	Real estate, ownership of dwellings and professional services	321750	397932	23.7
9.	Other Services	143796	74001	-48.5
Total Non-financial Corporations		**1773031**	**1946989**	**9.8**

Annexure II

Classes	I	II	III	IV	V	VI
PUC Value (Rs.)	0 - 5 lakhs	5 - 10 lakhs	10 - 50 lakhs	50 lakhs - 1 Crore	1 Crore – 5 Crore	Above 5 Crore

Annexure III

PUC value (Rs.)	**Classes**
Less than 1 Lakh	1
1 Lakh - 5 Lakhs	2
5 Lakhs - 10 Lakhs	3
10 Lakhs - 25 Lakhs	4
25 Lakhs - 50 Lakhs	5
50 Lakhs - 1 Crore	6
1 Crore - 2 Crores	7
2 Crores - 5 Crores	8
5 Crores - 10 Crores	9
10 Crores - 25 Crores	10
25 Crores - 100 Crores	11
100 Crores - 500 Crores	12
500 Crores - 1000 Crores	13
Above 1000 Crores	14