EAC-PM Policy Note

What the Data Cannot Say: Small-Sample Inference and India's National Accounts

April 2026

Virat Singh & Aditya Sinha

Table of Contents

What the Data Cannot Say: Small-Sample Inference and India's National Accounts................. 3

Key Takeaways: 3

1. The GDP Misestimation Claim and Why It Matters. 4

2. How The Small Samples Produce Statistically Unreliable Estimates That Are Sensitive to Omitted Values, Noise & Sample Selection & Design Choices. 4

3. How Serial Correlation Reduce Effective Sample Sizes Further and Compounds the Problem. 7

4. The Data Do Not Identify FY2011-12 as a Break Point.. 7

5. Conclusion 8

References 9

Appendix 1...10

Appendix 2...16

Annex 16

What the Data Cannot Say: Small-Sample Inference and India's National Accounts

Why casual inference in underpowered studies mislead, and the policy costs of mistaking noise for signal - illustrated with an application to National Accounts Assessment.

Virat Singh and Aditya Sinha[1]

Key Takeaways:

- *The recent paper by Anand, Felman and Subramanian (2026) argues that India's January 2015 national accounts methodology revision caused GDP growth to be overstated by approximately 1.5-2 percentage points per year between FY2011-12 and FY2024-25. The paper presents evidence for this view using informal-sector survey data, cross-country regressions, and correlation analysis that is the focus of this note. It has gathered some attention in media, and in political discussions, recently.*

- *This note disputes the evidentiary weight placed by the authors on the correlation estimations. It shows that they cannot by themselves support strong quantitative claims such as how much GDP may have been misestimated. The sign and magnitude of correlation changes, with their small samples, are not even fair or robust qualitative indicators of direction, let alone actual interpretable quantitative estimates of magnitude.*

- *The findings of the paper are misleading because the small samples they use produce unstable and unreliable coefficients with large standard errors. The estimations cannot be used to make a quantitative assessment for a whole GDP methodology; the authors arguments are based on biased estimations that mislead in direction and size; where the coefficient can look negative even when the underlying relationship is not; and an observed change can look larger than it really is. The sample design and serial correlation statistically compound the problem. The correlation coefficients that AFS relies upon are indistinguishable from noise and should be treated as uncertain rather than as characterisations of the true relationship.*

---

[1] Virat Singh is Consultant, Economic Advisory Council to the Prime Minister and Aditya Sinha is former Officer on Special Duty (Policy & Research), Economic Advisory Council to the Prime Minister. We would like to place on record our sincere gratitude to Dr Shamika Ravi, Member, Economic Advisory Council to the Prime Minister and Dr Mudit Kapoor, Associate Professor, Economics & Planning Unit, Indian Statistical Institute, Delhi for their constant guidance, suggestions and comments for this Policy Note. However, the contents of the Note, including facts and opinions, are of the authors.

1. The GDP Misestimation Claim and Why It Matters.

The March 2026 Peterson Institute for International Economics (PIIE) working paper by Anand, Felman and Subramanian (2026) (hereafter AFS) argues that the January 2015 revision to India's national accounts methodology caused GDP growth to be overstated by approximately 1.5 to 2 percentage points per year over the FY2011-12 to FY2024-25 period. They claim that India's official GDP estimation is roughly 22 percentage points of economic output greater than the true observed GDP. This paper indicts not just a statistical methodology but a generation of policy decisions made on supposed false premises. Their claims have gained attention in media and political discussions, recently.

AFS is a good example of a well-meaning economic argument overlooking the statistical context and limitations of making inferences with small samples of data. The paper splits its sample at FY2011-12, yielding only 17 observations in Period 1 (spanning FY1995-96 to FY2011-12, under the old methodology of estimating GDP) and only 11 observations in Period 2 (FY2012-13 to FY2024-25, under the 2015 revised methodology). The authors exclude the pandemic years of FY2020-21 and FY2021-22, under the rationale of an exogenous COVID-induced economic shock and its mechanical effect on the data. For each of five macro indicators - real exports, the index of industrial production or IIP, real bank credit, real direct taxes, and electricity consumption - AFS reports the correlation coefficients with real Gross Value Added (GVA) growth separately for the two periods (figure 2 of the paper). They interpret the consistent (breakdown within the) directional pattern as evidence that the 2015 methodology weakened the relationship between measured GDP and observable economic activity.

The findings of the paper are misleading because the small samples in AFS produce unstable and unreliable correlation coefficients with large standard errors. Simply put, even large changes in the correlations are statistically indistinguishable from noise, given the large variance across a small dispersed sample of observations. The influence of any single data point on the estimated correlation is large by mathematical necessity. A true change of 0.15 in correlation units could easily be observed as 0.30 - 0.45; the sign and magnitude of correlation changes are unreliable even as qualitative indicators of direction, let alone quantitative estimates of magnitude.

2. How The Small Samples Produce Statistically Unreliable Estimates That Are Sensitive to Omitted Values, Noise & Sample Selection & Design Choices.

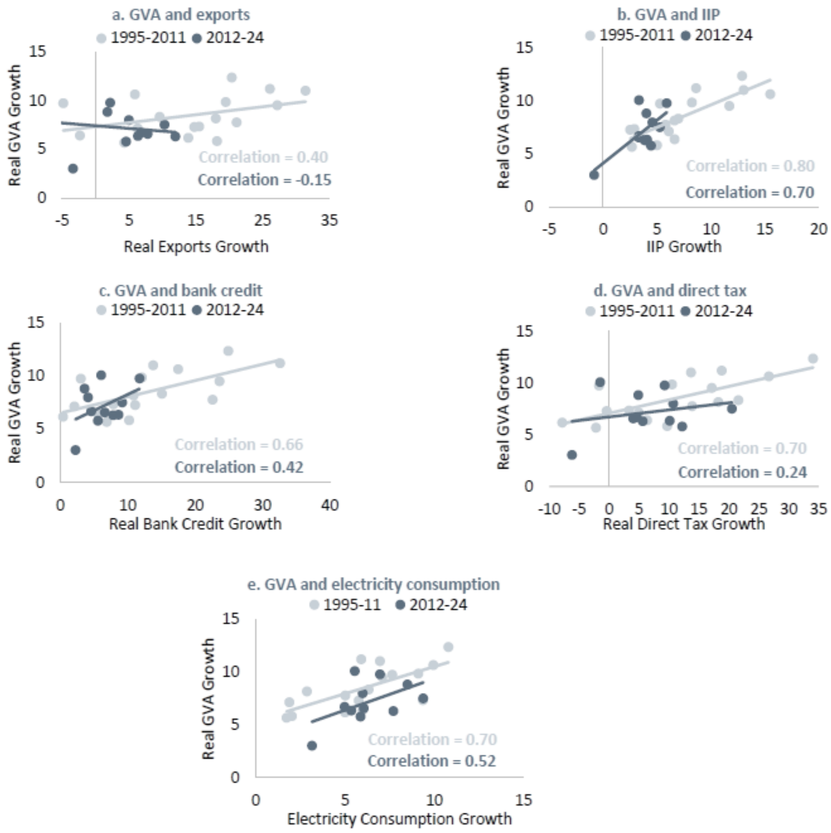
Panel (a) of Figure 2 of AFS (produced here as Figure 1) provides a near-textbook example of how a coefficient can look negative even when the underlying relationship is not. This is known as a Type S error. When the noise-to-signal ratio in the data is large, the sampling distribution of an estimation is very wide relative to the true effect. In small samples, pure noise can mechanically push the correlation estimate in the wrong direction. If the post 2012 period is treated in isolation, exports would appear to be a drag on growth - which is highly improbable - and that breakdown in the relationship is what AFS relies on to make their claims.

However, simply dropping a single value (FY2015-16) in Period 2 resolves the sign issue - the AFS correlation coefficient for GVA-Exports of (−0.15) changes to +0.22 (Table A.1.2). If the COVID year observations for FY2020-21 and FY2021-22, originally excluded by the authors, are included to the original sample (including FY2015-16), the AFS correlation coefficient resolves from (-0.15) to +0.49. Period 2 now shows a stronger correlation than the Period 1 (+0.40) which AFS argued was the true observed relationship between the indicators. In fact, Period 2 correlations across all indicators, not just exports, are all comparable or even stronger to Period 1 (Table A.1.2). This is also the case across almost all the other correlations in the paper (Annex 1).

Figure 1: Correlations Between Gross Value Added (GVA) and Core Macro Indicators (Figure 2 reproduced from AFS For Context). Data is available in [Appendix 1](#).

Figure 2

Correlation between Gross Value Added and core macro indicators



Note: Real bank credit and real direct taxes are obtained by deflating with CPI-IW core from 1995 to 2011 and with CPI core thereafter. For the period 1995–2011, real GVA and real exports are measured using the 2004 base year series; from 2011 onward, the 2011 base year series is used. 2020 and 2021 are excluded to avoid pandemic-related distortions arising from supply shocks, lockdowns, and policy interventions that obscure normal cyclical relationships. GVA excludes agriculture and public administration. Bank credit excludes food credit.

Sources: RBI, CEA, and CMIE.

Source: Anand, Felman and Subramanian (2026). The figure is reproduced here from the paper and is not the original work of the authors.

This is the nature of any computation using small samples. None of these coefficients quantitatively tell us anything substantive without examining how dispersed the data is, and if they are statistically significant to reliably make inferences from.

Standard errors show how reliable an estimate is as a characterisation of the true relationship across the data. A smaller standard error means a sample estimate is likely closer to the true value. A larger one means more uncertainty. With 17 observations in Period 1 and 11 in Period 2, the standard error of the difference in the correlations across the two periods is 0.44. This means that for a correlation change to be statistically significant at the conventionally accepted 5% level, using the Fisher r-to-z transformation (Appendix 2), the difference must exceed approximately 0.87 in correlation units. To put this concretely, the correlations between the two periods would need to change from +0.40 to roughly −0.50 before the test would consider it statistically distinguishable from zero, a null effect. None of the five annual correlation changes reported in AFS clear this threshold. None achieve statistical significance (Table A.1.4 and annex Table A.2.1a and Table A.2.1b).

More revealing are the 95% confidence intervals. Four of the five correlations span negative values. The GVA-Exports correlation has a confidence interval of [−0.69, +0.49]. The GVA-Bank Credit correlation of 0.42 has an interval of [−0.24, +0.81]. With intervals this wide, a negative correlation is statistically just as plausible as a strongly positive one (Figure A.1.1). The point estimates that AFS relies upon are indistinguishable from noise (Table A.1.4 and annex Table A.2.1a and Table A.2.1b). They should be treated as uncertain rather than as characterisations of the true relationship.

The negative baseline and strong swings in the coefficients are more than just a product of incorrect inference - they are also a product of two compounding design choices that has nothing to do with measurement quality or methodology. First is excluding the two years of strong positive co-movement. Second is retaining the years of sharpest mechanical divergence.

Statistical power - the probability that the test would detect a real change, if one existed - average roughly 15% across the five pairs of correlations. At that power level, the exaggeration ratio (what Gelman and Carlin, 2014, call the Type M error) is approximately 2 to 3[2]. This means that any observed change that happens to clear a significance threshold is likely two to three times larger than the true underlying effect. A real change of 0.15 correlation units could easily appear as 0.30 to 0.45 in these samples, and for some correlations even more. The estimated differences can't be treated as qualitative signals of direction, and definitely less so as quantitative measures of magnitude (Table A.1.4 and annex Table A.2.1a and Table A.2.1b).

So, which is the true correlation relationship between GVA and exports? -0.15? 0.22? 0.40 or 0.49? None of them. With such small samples, they are all, in essence, misleading and inconclusive to be basing an economic argument on.

AFS also presents cross-country regressions in which an India dummy coefficient of approximately 2.4 percentage points is interpreted as the magnitude of GDP misestimation. This coefficient falls by nearly one-quarter, to 1.8 percentage points, simply by adding the two excluded COVID observations back into the sample with no other change to the model (Table A.1.3). If two observations can reduce the paper's headline number by a quarter, the results are not robust. They are susceptible to the same small-sample fragility that undermines the correlation evidence.

The authors deduce the 1.5-2 pp overestimation through the entire period as the combination of two effects in Period 2. They claim that the revised methodology overestimates GDP from the period 2012-19 and 2022-25 and underestimates from 2020-22. Their argument is based on the revised GDP methodology using the Wholesale Price Index (WPI) to deflate prices. The WPI deflators are shown to understate production-price inflation when commodity prices fall. Each time the CPI outpaced the WPI, they show real GDP outpaced sales, and the official numbers overstated economic growth.[3]

It may just be that their arguments have merit.[4]. Ministry of Statistics and Program Implementation's (MoSPI) February 2026 revision addresses several of the specific qualitative issues that AFS identifies. However, the small sample problem persists throughout all the correlations (Annex tables). The casual interpretation of the coefficients quantitatively is misleading without examining the confidence intervals

[2] Type M (Magnitude) Error is a concept introduced by Gelman and Carlin (2014) to describe a systematic bias that arises in low-powered studies. When a test has low statistical power (meaning it is unlikely to detect a real effect even if one exists) the handful of results that do clear the significance threshold tend to be dramatically inflated relative to the true underlying effect. A ratio of 3 means that a statistically significant finding is, on average, thrice as large as the true effect it purports to measure.
[3] The correlations in the Annex show that the Sales-Macro indicators and GVA-Sales relationship, that AFS emphasises as relevant relationship to highlight over and under estimation of GDP, are also statistically not significant and susceptible to the same small sample problems highlighted here.
[4] AFS' strongest statistical result and the one most robust to the sensitivity tests are the correlations between changes in WPI, manufacturing deflators, and oil prices.

and standard errors of changes. If sample sizes of 17 and 11 do not provide stable estimates, how can one infer the underestimation for COVID years using 2 observations?

3. How Serial Correlation Reduce Effective Sample Sizes Further and Compound the Problem.

Serial correlation, also called autocorrelation, is when in time series data a point value at time t is affected by previous values at time t-1, t-2 etc. In macroeconomics, we see autocorrelation caused by 'momentum' in almost all variables. Serial correlation in macroeconomic time series compounds the small sample problem explained earlier**.** The Fisher test and statistical interpretation above assume independent observations. Macroeconomic growth rates are not independent. GVA, credit, and electricity growth are driven by multi-year business cycles, meaning successive observations carry similar information.

In Period 1, autocorrelation reduces the sample sizes for GVA from 17 to roughly 6; bank credit and electricity to 5; direct taxes to 9; and IIP to 12 (Table A.1.5). Period 1 captures a sustained boom, especially between 2002 to 2008, in which all indicators moved in sync. The co-movement inflates the nominal correlations, but the underlying information content is that of 5 to 9 independent data points, not 17. Under autocorrelation-adjusted effective sample sizes, the standard error of the Fisher test rises from 0.44 to between 0.75 and 0.90 (Table A.1.6). A correlation change would need to in fact exceed 1.5 to 1.8 correlation units, not just 0.87, as indicated earlier, to achieve significance. No observed change comes close.

AFS presents quarterly data as a robustness check, noting the larger nominal samples (28 observations for Period 1 and 45 observations for Period 2) as a confirmation of their results. The quarterly autocorrelation is at least as pronounced as the annual. Many of the variables indicate dependence across all four lags and an effective sample of as low as 2, as in the case of bank credit, for instance (Table A.1.7). The quarterly data, despite their larger nominal counts, do not materially improve the inferential position. The apparent precision of larger sample sizes is illusory when most of those observations are statistically redundant (for ease of viewing the comparison of nominal and effect sample sizes, see Figure A.1.3 and Figure A.1.4).

4. The Data Do Not Identify FY2011-12 as a Break Point

AFS assigns the break to FY2011-12, on the basis of the year on which the methodology was revised and rebased. A natural empirical question is whether the data themselves support this choice. Chow-type structural break tests across ten candidate break years, FY2007-08 to FY2016-17, find that FY2011-12 does not produce the strongest break statistic for the five indicator pairs (Table A.1.8). For many indicators, there is evidence of a break coinciding with the Global Financial Crisis, not the methodology revision. However, each of the indicators are being driven by an underlying structural break at different points across the time series of both periods. They are driven by macroeconomic factors and individual trends, as opposed to any methodology change at a pre-specified year.

An argument can be made that the breaks are not being identified because of the lack of measurement quality in the revised series - which affects the values for GVA. This can be corroborated in two ways. First, by comparing the correlations for the overlapping years where data are available for the revised and old GDP methodology. Second, through rolling correlations using only values of the revised series which provide a good idea of the underlying trend between the variables around the pre-specified break year.

Correlations computed over the overlap years, FY2004-11, where both the old and revised GVA series exist show materially similar values under both methodologies (Table A.1.9). The divergence that AFS attributes to the methodology revision is not present in the years where both series can be compared together. Moreover, rolling 8-year correlations between GVA and the macro indicators decline gradually from FY2013-14 onward, reach their lowest values in windows ending around FY2017-19, and then recover (Table A.1.10 and Figure A.1.5). This pattern - a gradual decline followed by recovery - maps onto the sequence of discrete real-economy shocks (the post-GFC slowdown, demonetisation, GST transition, the NBFC crisis), not a one-time measurement break. If the methodology revision were the cause, the correlations should deteriorate sharply at the break point and remain low. Instead, they drift downward over several years and then reverse - behaviour consistent with a business-cycle moderation and shocks, not a permanent measurement distortion.

The AFS computations observe a break in correlations around 2011-12 because they are synchronous with a business cycle and economic moderation, which begins from the Global Financial Crisis, and is experiencing other discrete economics shocks at the start of the sample that are causing the divergence between the macro indicators. However, the correlations normalise in the later period where the economic cycle and indicator movements replicate those Period 1. More importantly, by dropping the two observations with the greatest mechanical evidence of co-movement across the indicators, the authors retain a Period 2 sample that shows greater divergence from the start.

5. Conclusion

The AFS paper's claims of GDP misestimation and quantitative assertions are misleading and incorrectly inferred. The statistical case rests almost entirely on correlation comparisons across two pre-selected periods with small samples of data, each too small to support the conclusions drawn from them.

AFS estimations cannot be used to make a quantitative assessment for a whole GDP methodology. The casual findings of the paper are misleading because the small samples produce unstable and unreliable coefficients with large standard errors. None of the five annual correlation changes reported in Figure 2 of AFS, and almost none of the other correlations shown throughout the paper, are statistically significant. The confidence intervals span negative values, meaning a negative correlation coefficient is as statistically plausible as a positive one. The coefficients are susceptible to be mechanically affected by noise, rather than explaining the actual descriptive relationship observed across the series. The sample design and serial correlation statistically compound the problem.

References

Andrews D.W.K. (1993). "Tests for Parameter Instability and Structural Change with Unknown Change Point." Econometrica 61(4):821–856.

Anand A., Felman J. and Subramanian A. (2026). "India's 20 Years of GDP Misestimation: New Evidence." PIIE Working Paper 26-3.

Bai J. and Perron P. (1998). "Estimating and Testing Linear Models with Multiple Structural Changes." Econometrica 66(1):47–78.

Bartlett M.S. (1935). "Some aspects of the time-correlation problem in regard to tests of significance." J. Royal Statistical Society 98(3):536–543.

Bartlett M.S. (1946). "On the theoretical specification of sampling properties of autocorrelated time series." J. Royal Statistical Society B 8(1):27–41.

Fisher R.A. (1921). "On the probable error of a coefficient of correlation deduced from a small sample." Metron 1:3–32.

Gelman A. and Carlin J. (2014). "Beyond Power Calculations: Assessing Type S (Sign) and Type M (Magnitude) Errors." Perspectives on Psychological Science 9(6):641–651.

Goyal A. and Kumar A. (2020). "Indian Growth is Not Overestimated: Mr. Subramanian You Got it Wrong." Macroeconomics and Finance in Emerging Market Economies 13(1):29–52.

Hamilton J.D. (1994). Time Series Analysis. Princeton University Press.

Müller U.K. and Watson M.W. (2018). "Long-Run Covariability." Econometrica 86(3):775–804.

Nagaraj R., Sapre A. and Sengupta R. (2019). "Four Years After the Base-Year Revision." India Policy Forum (NCAER), Vol. 16:78–131.

Pyke D.A. and Thompson W.A. (1978). "Statistical analysis of survival and removal rate experiments." Ecology 59(3):492–496.

Stock J.H. and Watson M.W. (1996). "Evidence on Structural Instability in Macroeconomic Time Series Relations." Journal of Business & Economic Statistics 14(1):11–30.

Subramanian A. (2019). "India's GDP Mis-estimation: Likelihood, Magnitudes, Mechanisms, and Implications." CID Faculty Working Paper No. 354, Harvard University.

Methodological Note:

All computations use the official AFS replication data (Data1.xlsx, Table 2.xlsx), unmodified from the published replication package.

Appendix 1

Tables & Figures in The Note

Table A.1.1a: Period 1 annual data used in Figure 1 (Figure 2 of AFS) - year-on-year real growth rates (%)

Fiscal Year (FY)	GVA (old series %) pre-2015 methodology	Real Exports (%) YoY real growth	IIP (%) YoY growth	Bank Credit (%) YoY real growth	Direct Tax (%) YoY real growth	Electricity (%) YoY growth
1995-96	11.02	31.40	13.05	13.75	13.67	6.96
1996-97	7.14	6.29	6.09	2.07	4.94	1.88
1997-98	6.40	**−2.33**	6.62	5.61	6.33	5.99
1998-99	6.17	13.88	4.10	0.36	**−7.83**	5.00
1999-00	8.16	18.00	6.63	10.87	18.25	2.87
2000-01	5.83	18.15	5.01	10.20	9.71	2.03
2001-02	5.68	4.31	2.67	6.87	**−2.20**	1.73
2002-03	7.77	21.09	5.79	22.58	13.88	5.02
2003-04	8.34	9.58	6.98	15.08	21.60	6.32
2004-05	9.52	27.18	11.72	23.61	17.15	7.09
2005-06	11.20	26.09	8.62	32.62	18.80	5.91
2006-07	12.36	20.39	12.90	24.91	34.08	10.80
2007-08	10.64	5.87	15.52	17.47	26.69	9.95
2008-09	7.28	14.78	2.52	11.08	**−0.42**	5.77
2009-10	9.74	**−4.83**	5.28	3.01	**−1.70**	7.65
2010-11	9.85	19.48	8.23	12.08	10.50	9.10
2011-12	7.34	15.49	2.89	7.87	3.33	9.37

Source: AFS replication data, Data 1 excel file 'Figure 2' sheet. GVA uses column 10 (old pre-2015 series). Note: Indicators are CPI-deflated year-on-year real growth rates matching those plotted in Figure 2. Red = negative growth rates and breaks in co-movement across series.

Table A.1.1b. Period 2 annual data including COVID years (excluded by AFS in Figure 2 of the paper) - year-on-year real growth rates (%)

Fiscal Year (FY)	GVA (new series %) 2015 methodology	Real Exports (%) YoY real growth	IIP (%) YoY growth	Bank Credit (%) YoY real growth	Direct Tax (%) YoY real growth	Electricity (%) YoY growth
2012-13	6.69	6.81	3.33	4.57	4.59	4.98
2013-14	6.57	7.79	3.28	6.60	4.03	6.05
2014-15	8.82	1.78	4.02	3.58	4.86	8.50
2015-16	10.08	**−5.65**	3.33	6.06	**−1.52**	5.55
2016-17	7.99	4.98	4.58	4.09	10.69	5.99
2017-18	5.79	4.56	4.44	5.58	12.20	5.87
2018-19	6.31	11.93	3.84	7.82	5.60	7.70
2019-20	3.03	**−3.38**	**−0.84**	2.23	**−6.22**	3.15
2020-21	**−5.23**	**−6.96**	**−8.47**	**−0.01**	**−13.87**	**−1.43**
2021-22	10.80	29.60	11.40	2.60	31.19	7.04
2022-23	7.52	10.32	5.26	9.18	20.53	9.38
2023-24	9.78	2.20	5.91	11.73	9.28	6.96
2024-25	6.36	6.34	4.07	8.63	10.14	5.35

Source: AFS replication data, Data 1 excel file "Figure 2" sheet. GVA uses column 10 (old pre-2015 series). Note: 2020-21 and 2021-22 have been excluded by AFS from Period 2. The COVID years show the strongest directional co-movement of any observations. Red = negative growth rates and breaks in co-movement across series.

Table A.1.2: Correlations Sensitivity Experiments for Figure 1 (Figure 2 of AFS) GVA vs Macro Indicators

Indicator (GVA vs)	Period 1 Corr (old, n=17)	Period 2 Corr (AFS, n=11)	Period 2+COVID Corr (n=13)	COVID change	Sign restored?	Key observation	Period 2 but drop-observation (n=10)	Residual gap to P1
Real Exports	0.40	**−0.15**	0.49	+0.64	YES	FY2015-16	0.22	+0.18
IIP	0.80	0.70	0.92	+0.22	—	FY2017-18	0.76	+0.04
Bank Credit	0.66	0.42	0.52	+0.10	—	FY2014-15	0.56	+0.10
Direct Tax	0.70	0.24	0.69	+0.45	—	FY2015-16	0.53	+0.17
Electricity	0.70	0.52	0.85	+0.33	—	FY2015-16	0.68	+0.02

Note: Period 1 Correlations refers to the AFS correlation, old GVA series. Period 2 Correlations refers to the AFS baseline, new GVA series (COVID excluded); Period 2+COVID adds the excluded COVID years to the Period 2 correlations. Period 2 drop observation drops the single observations and runs the period 2 correlations.

Table A.1.3: AFS Regression For Period 2 With & Without COVID years.

Metric	Period 2 baseline (excl. 2020–22)	Period 2 + COVID (incl. 2020–212)	Change/ Comments
India dummy coefficient (γ)	2.364***	1.786***	−0.578 pp (−24%)
India dummy p-value	<0.001	<0.001	Both Significant
R^2 of regression	0.782	0.794	Negligible

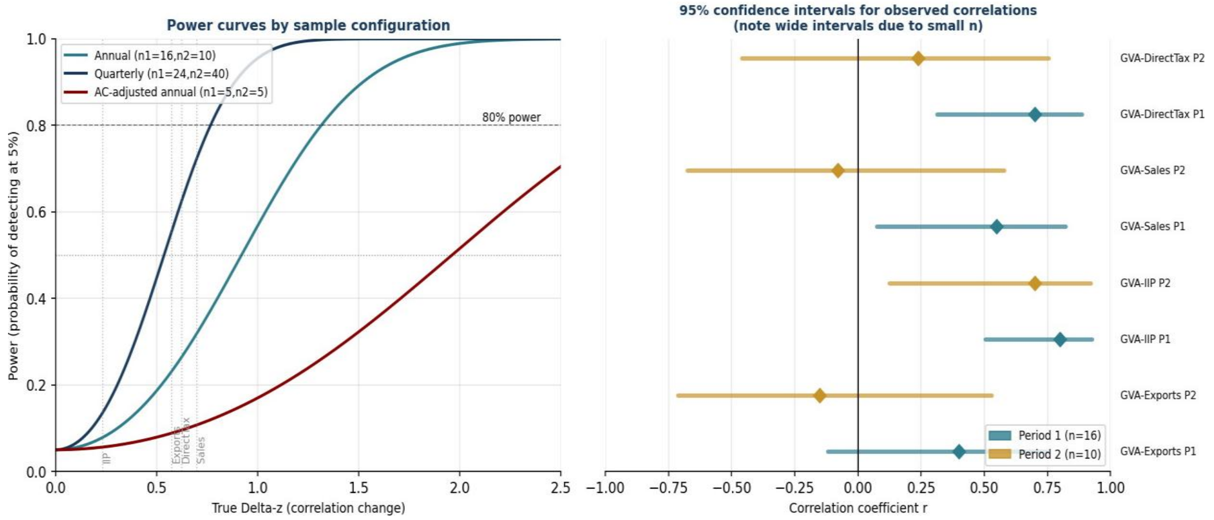
Note: The dummy falls by 0.58 pp when COVID years are included - to 1.79 pp.

Table A.1.4: Fisher Test For Period 1 & 2 of Figure 1 (Figure 2 of AFS).

Indicator	Period 1 Correlation (1995–2011-12)	n_1	Period 2 Correlation (2012–24 (excl COVID)	n_2	Change in z Stat Δz (z_1-z_2)	Standard Error (SE)	z-stat	p-value	Sig.	Power @5%	95% CI for Period 2 correlation	Type S risk
GVA & Exports	0.40	17	-0.15	11	0.575	0.443	1.297	0.195	n.s.	25.4%	[-0.69, 0.49]	Yes
GVA & IIP	0.80	17	0.70	11	0.231	0.443	0.522	0.602	n.s.	8.2%	[0.17, 0.92]	No
GVA & Bank Credit	0.66	17	0.42	11	0.345	0.443	0.779	0.436	n.s.	12.2%	[-0.24, 0.81]	Yes
GVA & Direct Tax	0.70	17	0.24	11	0.623	0.443	1.405	0.160	n.s.	29.0%	[-0.42, 0.73]	Yes
GVA & Electricity	0.70	17	0.52	11	0.291	0.443	0.656	0.512	n.s.	10.1%	[-0.12, 0.85]	Yes

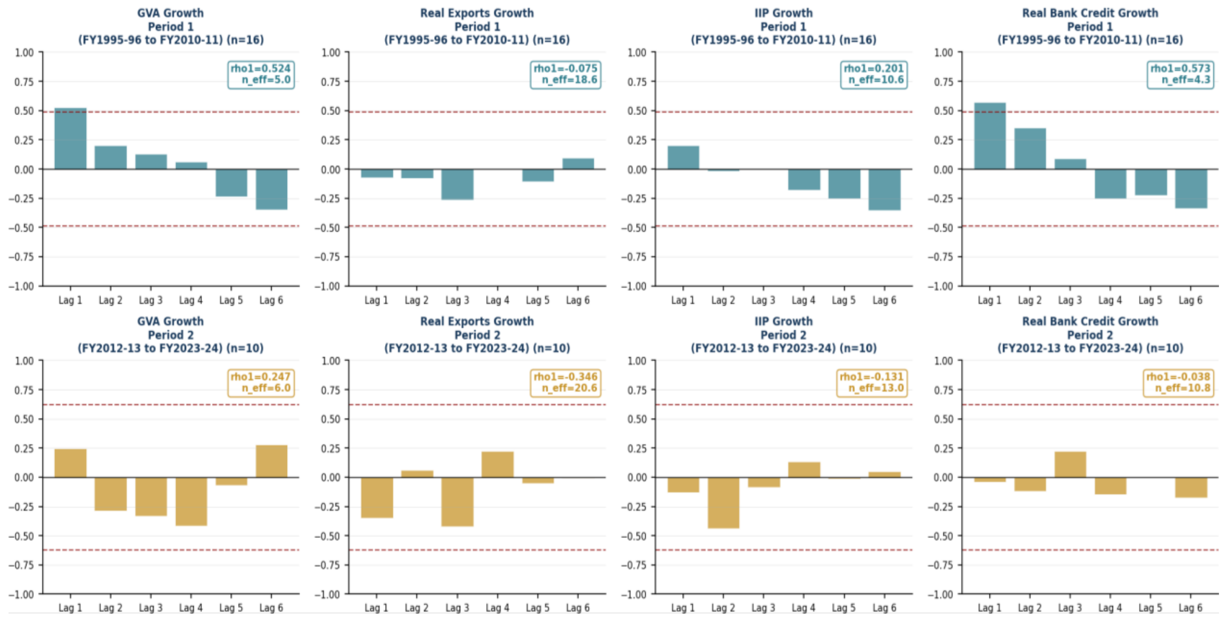
Note: n.s. = not significant. CI = confidence interval

Figure A.1.1: Statistical Power & Detection Intervals For Figure 2 AFS Correlations.



Note: Left: power curves at 5% significance for detecting correlation changes under three sample configurations. Vertical lines show observed Δz values. Right: 95% confidence intervals for individual period correlations (n1=17, n2=11).

Figure A.1.2: Autocorrelations Functions - Annual Growth Rate Series (95% confidence bands shown as dashed lines)



Note: Sample autocorrelation functions (lags 1-6) for annual year-on-year growth rates, Period 1 (blue) and Period 2 (gold). Dashed red lines show the 95% confidence band.

Table A.1.5. Sample autocorrelation (up to 4 lags) and effective sample sizes - annual year-on-year growth rate series.

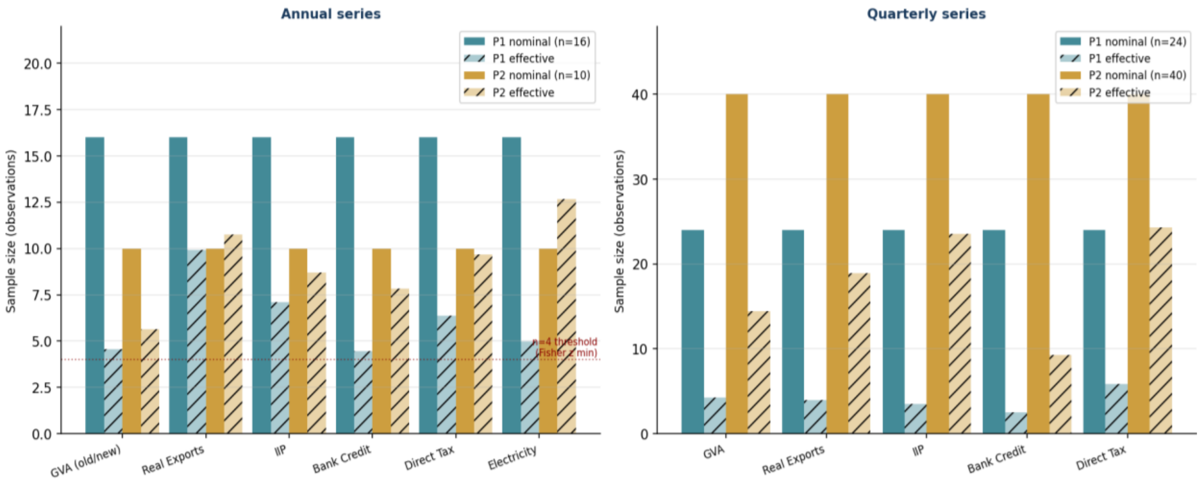
Series	P1 ρ_1	P1: ρ_2	P1: ρ_3	P1: ρ_4	P1: n nom	P1 n eff	P2: ρ_1	P2: ρ_2	P2: ρ_3	P2: ρ_4	P2 n nom	P2 n eff
GVA	0.487	0.171	0.147	0.026	17	5.9	0.187	-0.290	-0.236	-0.393	11	7.5
Real Exports	-0.070	-0.093	-0.264	-0.005	17	19.6	-0.353	0.113	-0.469	0.275	11	23.0
IIP	0.165	0.022	0.098	-0.311	17	12.2	-0.105	-0.413	-0.138	0.135	11	13.6
Bank Credit	0.566	0.394	0.104	-0.256	17	4.7	0.137	-0.003	0.095	-0.080	11	8.3
Direct Tax	0.298	-0.071	0.074	-0.034	17	9.2	-0.219	-0.204	-0.033	0.297	11	17.2
Electricity	0.517	0.300	0.307	0.267	17	5.4	-0.492	-0.095	0.088	0.111	11	32.3

All series are year-on-year real growth rates identical to those in Figure 2 of AFS. P1 = Period 1; P2 = Period 2; eff = effective; nom ; nominal.

Table A.1.6. Fisher z Tests Under Nominal and Autocorrelation-Adjusted Effective Sample Sizes (for Figure 2 of AFS)

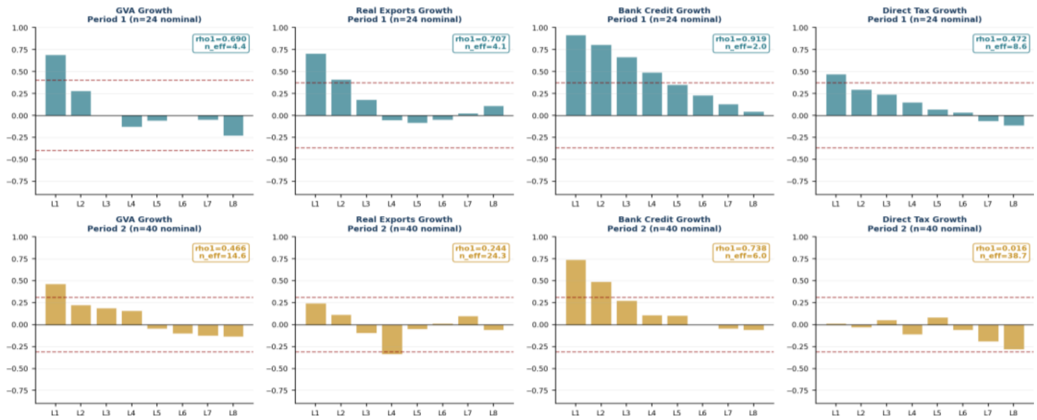
Pair	SE (nominal)	p- value (nominal)	n_1 Eff	n_2 Eff	SE (adjusted)	p (adjusted)	Power (adjusted)
GVA – Exports	0.443	0.195	5.9	7.5	0.755	0.446	11.9%
GVA – IIP	0.443	0.602	5.9	7.5	0.755	0.759	6.1%
GVA – Bank Credit	0.443	0.436	4.7	7.5	0.897	0.700	6.7%
GVA – Direct Tax	0.443	0.160	5.9	7.5	0.755	0.410	13.1%
GVA – Electricity	0.443	0.512	5.4	7.5	0.797	0.715	6.5%

Figure A.1.3: Nominal vs Autocorrelation-Adjusted Effective Sample Sizes (for data used in Figure 2 of AFS)



Note: Nominal (solid) vs autocorrelation-adjusted effective (hatched) sample sizes for annual (left) and quarterly (right) series. Period 1 in blue, Period 2 in gold. The gap between solid and hatched bars quantifies the informational loss from serial dependence. FY2024-25 is excluded in Period 2 and 2011-12 is excluded in Period 1.

Figure A.1.4: Autocorrelation Functions – Quarterly YoY Growth Rate Series (95% Confidence bands in dashed lines)



Note: Sample autocorrelation functions (lags 1-8) for quarterly year-on-year growth rates, Period 1 (blue) and Period 2 (gold). Note the high persistence in GVA and bank credit during Period 1.

Table A.1.7: Sample autocorrelation (1-4 lags) and effective sample sizes — quarterly year-on-year growth rate series.

Series	P1 ρ_1	P1: ρ_2	P1: ρ_3	P1: ρ_4	P1: N nom	P1 n eff	P2: ρ_1	P2: ρ_2	P2: ρ_3	P2: ρ_4	P2 N nom	P2 N eff
GVA	0.695	0.310	0.049	-0.090	28	5.0	0.466	0.212	0.155	0.119	45	16.4
Real Exports	0.688	0.342	0.045	-0.218	28	5.2	0.236	0.123	-0.085	-0.345	45	27.8
IIP	0.773	0.479	0.190	-0.087	28	3.6	0.047	0.036	-0.049	-0.099	45	41.0
Bank Credit	0.894	0.741	0.593	0.431	28	2.0	0.770	0.544	0.341	0.180	43	5.6
Direct Tax	0.442	0.378	0.295	0.150	28	10.8	-0.055	0.003	0.203	-0.232	45	50.2

Note: All series are year-on-year real growth rates identical to those in Figure 2 of AFS. P1 = Period 1; P2 = Period 2; eff = effective; nom ; nominal.

Table A.1.8: Chow-type structural break F-statistics across candidate break years.

Break Year	GVA-Exports F-stat / p-value		GVA-IIP F-stat/ p-value		GVA-Bank Credit F-stat/ p-value		GVA-Direct Tax F-stat / p-value		GVA-Electricity F-stat / p-value	
FY2007–08	4.520	0.022	1.399	0.266	0.431	0.655	1.065	0.360	4.597	0.020
FY2008–09	4.516	0.022	1.017	0.377	0.297	0.746	0.346	0.711	3.325	0.053
FY2009–10	2.368	0.115	0.866	0.433	0.057	0.944	0.284	0.755	1.513	0.240
FY2010–11	1.709	0.202	0.828	0.449	0.159	0.854	0.215	0.808	1.335	0.282
FY2011–12 (used by AFS)	0.894	0.422	2.380	0.114	0.409	0.669	0.252	0.779	0.827	0.449
FY2012–13	0.906	0.417	2.290	0.123	0.409	0.669	0.266	0.769	0.755	0.481
FY2013–14	0.916	0.414	2.256	0.127	0.468	0.632	0.295	0.747	0.628	0.542
FY2014–15	0.486	0.621	1.474	0.249	0.872	0.431	0.259	0.774	0.583	0.566
FY2015–16	0.397	0.677	1.607	0.221	1.551	0.233	1.180	0.324	1.569	0.229
FY2016–17	0.627	0.543	1.579	0.227	3.441	0.049	1.403	0.265	1.988	0.159
Potential Break	at FY2007–08 or 2008-09		FY2011–14 (none significant)		at FY2016–17		No evidence		at FY2007–08 or FY 2008-09	

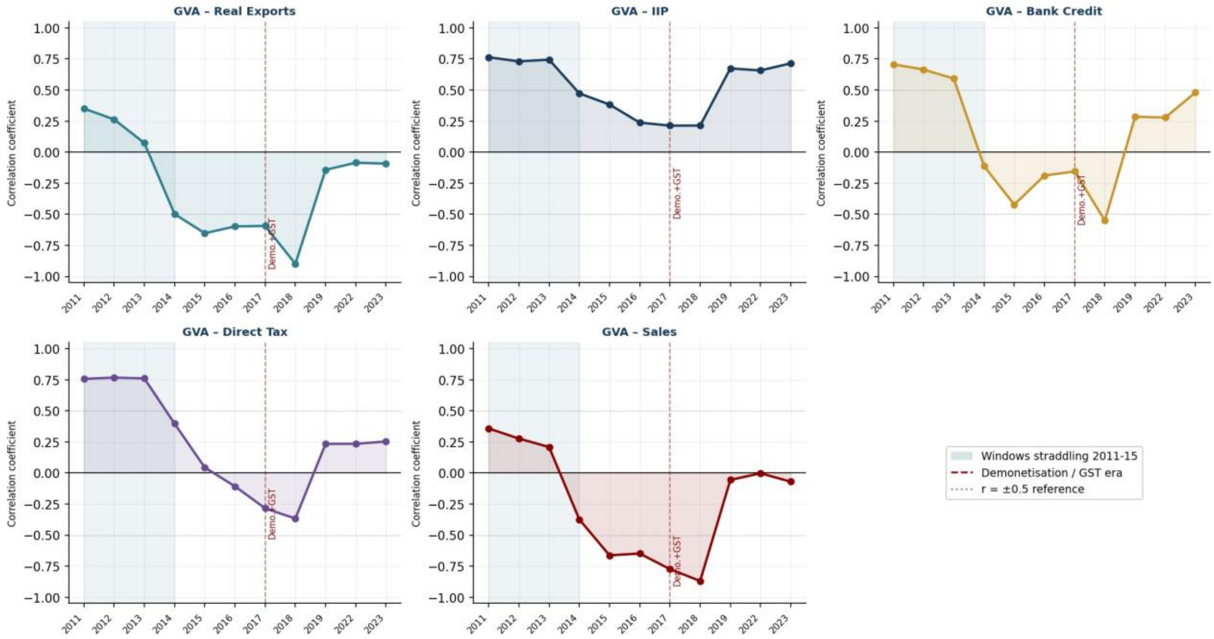
Note: No candidate break year clears the Andrews (1993) sup-F critical value at the 10% level for any pair.

Table A.1.9. Correlations Comparing the Old and Revised Overlapping GVA Series.

Indicator	Correlations (GVA old) 2004-11	Correlations (GVA new) 2004-11	Difference (New minus Old)
Real Exports	0.174	0.413	+0.239
IIP	0.705	0.686	-0.020
Bank Credit	0.573	0.654	+0.081
Direct Tax	0.816	0.732	-0.083
Electricity	0.628	0.438	-0.190

Note: Correlations for FY2004-05 to FY2010-11 using both GVA series over the same economic observations (n=7). Differences reflect methodology alone, not the economic period.

Figure A.1.5: Rolling 8-Year Correlations: GVA (Revised Series) against Macro-indicators



Note: Rolling 8-year correlations between GVA (new series) and macro indicators. Blue-shaded region marks windows straddling the 2011-15 methodology transition. The dashed vertical line marks the approximate timing of demonetisation and GST (2016-17). Correlations decline gradually from FY2013-14, reach their most negative values in windows ending around FY2017-19, then recover. COVID years are excluded to replicate AFS's sample. When included the recovery is unanimously stronger in the later years.

Table A.1.10. Rolling 8-year correlations: GVA (new series) vs macro indicators.

Window ends	GVA–Exports	GVA–IIP	GVA–Credit	GVA–Direct Tax	GVA–Sales
FY2011-12	0.351	0.764	0.707	0.757	0.358
FY2012-13	0.264	0.730	0.665	0.767	0.277
FY2013-14	0.073	0.744	0.592	0.761	0.208
FY2014-15	−0.498	0.473	−0.113	0.397	−0.374
FY2015-16	−0.652	0.383	−0.422	0.044	−0.663
FY2016-17	−0.598	0.238	−0.188	−0.108	−0.648
FY2017-18	−0.593	0.213	−0.155	−0.285	−0.774
FY2018-19	−0.899	0.214	−0.551	−0.367	−0.869
FY2019-20	−0.142	0.674	0.285	0.234	−0.056
FY2022-23	−0.085	0.658	0.279	0.234	−0.003
FY2023-24	−0.092	0.715	0.481	0.253	−0.070

Note: Revised GVA series throughout. Partial recovery visible in recent windows. COVID years are excluded to replicate AFS's sample. When included the recovery is unanimously stronger in the later years.

Appendix 2

Supplementary Equations for Computations Shown In the Note

The Fisher r-to-z test is the standard procedure in statistics for comparing two independent correlation coefficients to understand if they are statistically significant and their comparison is valid. The z statistic is calculated using the equation below:

$$z = \frac{1}{2}\ln\left(\frac{1+r}{1-r}\right) \tag{1}$$

The standard error for comparing two transformed correlations is:

$$SE = \sqrt{\frac{1}{n_1 - 3} + \frac{1}{n_2 - 3}} \tag{2}$$

With the number of observations n_1 = 17 (for Period 1) and n_2=11 (for Period 2), for two independent correlations to be statistically significant at the 5% level, the change in the z-statistic or $\Delta z = |z_1 - z_2|$ must exceed 0.87 correlation units (1.96 multiplied by the Standard Error).

For a series with first-order autocorrelation AR(1) denoted here as ρ_1, the standard correction to the effective sample size is:

$$n_{\mathrm{eff}} = n \times \frac{1-\rho_1}{1+\rho_1} \tag{3}$$

Here, n is the number of raw observations. When persistence is high, the effective sample size n_{eff} can drop sharply, so a period that appears to have 17 observations may statistically behave more like a sample of about 4 to 8 independent observations.

This adjustment applies to the AR(1) case and provides a lower bound on the effective sample: if significant autocorrelation extends beyond a single lag, as the autocorrelation function plots confirm it does in several series, the effective sample could be smaller still. Consequently, the adjusted standard error (1) would be larger.

Annex 1

Fisher r-to-z Tests and Sensitivities to Inclusion of COVID Year Observations

The annex tables show the results of the fisher tests for all the correlations computed in the AFS paper. All correlations are recomputed from the AFS replication data package including the file Data1.xlsx. Period 1 includes the years FY1995-96 to FY2011-12 (N=17, confirmed from CORREL formula rows 7–23). Period 2 includes the years FY2012-13 to FY2024-25 excluding COVID years. Including the COVID years increases the number of observations to 13.

Table A.2.1a: Figure 2 (AFS) Correlations GVA vs Macro Indicators (Annual) Excluding COVID years.

Indicator	Period 1 Correlation (1995–2011-12)	n_1	Period 2 Correlation (2012–24 (excl COVID)	n_2	Change in z Stat Δz (z_1-z_2)	Standard Error (SE)	z-stat	p-value	Sig.	Power @5%	95% CI for Period 2 correlation
GVA & Exports	0.40	17	-0.15	11	0.575	0.443	1.297	0.195	n.s.	25.4%	[-0.69, 0.49]
GVA & IIP	0.80	17	0.70	11	0.231	0.443	0.522	0.602	n.s.	8.2%	[0.17, 0.92]
GVA & Bank Credit	0.66	17	0.42	11	0.345	0.443	0.779	0.436	n.s.	12.2%	[-0.24, 0.81]
GVA & Direct Tax	0.70	17	0.24	11	0.623	0.443	1.405	0.160	n.s.	29.0%	[-0.42, 0.73]
GVA & Electricity	0.70	17	0.52	11	0.291	0.443	0.656	0.512	n.s.	10.1%	[-0.12, 0.85]

Note: All five pairs are not-significant in both specifications; four of the five pairs have negative confidence intervals (CI); n.s. = not significant

Table A.2.1b: Figure 2 (AFS) Correlations GVA vs Macro Indicators (Annual) Including COVID years.

Indicator	Period 1 Correlation (1995–2011-12)	n_1	Period 2 Correlation (2012–24 (incl COVID)	n_2	Change in z Stat Δz (z_1-z_2)	Standard Error (SE)	z-stat	p-value	Sig.	Power @5%	95% CI for Period 2 correlation
GVA & Exports	0.40	17	0.49	13	-0.112	0.414	-0.271	0.786	n.s.	5.8%	[-0.08, 0.82]
GVA & IIP	0.80	17	0.92	13	-0.490	0.414	-1.184	0.236	n.s.	22.0%	[0.75, 0.98]
GVA & Bank Credit	0.66	17	0.52	13	0.216	0.414	0.523	0.601	n.s.	8.2%	[-0.04, 0.83]
GVA & Direct Tax	0.70	17	0.69	13	0.019	0.414	0.047	0.963	n.s.	5.0%	[0.22, 0.90]
GVA & Electricity	0.70	17	0.85	13	-0.389	0.414	-0.939	0.348	n.s.	15.6%	[0.56, 0.95]

Note: All five pairs are not-significant in both specifications; 2 pairs have negative confidence intervals (CI); n.s. = not significant

Table A.2.2a for Figure 3 (AFS) - Sales vs Macro Indicators (Annual) Excluding COVID years.

Indicator	Period 1 Correlation (1995–2011-12)	n_1	Period 2 Correlation (2012–24 (excl COVID)	n_2	Change in z Stat Δz $(z_1 - z_2)$	Standard Error (SE)	z-stat	p-value	Sig.	Power @5%	95% CI for Period 2 correlation
Sales & Exports	0.70	17	0.88	11	-0.508	0.443	-1.147	0.251	n.s.	20.9%	[0.59, 0.97]
Sales & IIP	0.57	17	0.46	11	0.150	0.443	0.339	0.735	n.s.	6.3%	[-0.19, 0.83]
Sales & Bank Credit	0.73	17	0.39	11	0.517	0.443	1.166	0.243	n.s.	21.5%	[-0.27, 0.80]
Sales & Direct Tax	0.80	17	0.65	11	0.323	0.443	0.729	0.466	n.s.	11.3%	[0.08, 0.90]
Sales & Electricity	0.35	17	0.64	11	-0.393	0.443	-0.886	0.376	n.s.	14.4%	[0.07, 0.90]

Note: All five pairs are not-significant in both specifications; 2 pairs have negative confidence intervals (CI); n.s. = not significant

Table A.2.2b for Figure 3 — Sales vs Macro Indicators (Annual) including COVID years.

Indicator	Period 1 Correlation (1995–2011-12)	n_1	Period 2 Correlation (2012–24 (incl COVID)	n_2	Change in z Stat Δz $(z_1 - z_2)$	Standard Error (SE)	z-stat	p-value	Sig.	Power @5%	95% CI for Period 2 correlation
Sales & Exports	0.70	17	0.96	13	-1.079	0.414	-2.605	0.009	***	74.1%	[0.87, 0.99]
Sales & IIP	0.57	17	0.73	13	-0.281	0.414	-0.679	0.497	n.s.	10.4%	[0.30, 0.91]
Sales & Bank Credit	0.73	17	0.15	13	0.778	0.414	1.878	0.060	*	46.7%	[-0.44, 0.65]
Sales & Direct Tax	0.80	17	0.86	13	-0.195	0.414	-0.470	0.638	n.s.	7.6%	[0.59, 0.96]
Sales & Electricity	0.35	17	0.57	13	-0.282	0.414	-0.681	0.496	n.s.	10.5%	[0.03, 0.85]

Note: In the base case (excluding COVID), none of the five Sales–macro pairs reach statistical significance. With COVID included, sales-exports achieves strong significance (p=0.009) suggesting the sales–export relationship is at least as strong in the new-methodology period as in the old. Sales and bank credit achieve significance with a marginally weaker correlation coefficient but large standard errors (negative CI); n.s. = not significant.

Table A.2.3 for Figure 4 — GVA vs Sales (Annual).

Indicator	Period 1 Correlation (1995–2011-12)	n_1	Period 2 Correlation (2012–24)	n_2	Change in z Stat Δz $(z_1 - z_2)$	Standard Error (SE)	z-stat	p-value	Sig.	Power @5%	95% CI for Period 2 correlation
GVA & Sales (Period 2 excl. COVID)	0.51	17	-0.07	11	0.633	0.443	1.428	0.153	n.s.	29.8%	[-0.64, 0.55]
GVA & Sales (Period 2 incl. COVID)	0.51	17	0.47	13	0.053	0.414	0.127	0.899	n.s.	5.2%	[-0.11, 0.81]

Note: Neither specification of Figure 4 reaches statistical significance. Also note the negative confidence intervals (CI). n.s. = not significant

Table A.2.4 for Figure 10 — CPI–WPI Wedge Correlations.

Series	Correlations	Number of Observations (n)	t-stat	p-value	Sig.	95% CI for Correlations
Panel A: GVA-Sales Wedge vs CPI−WPI (excl. COVID)	0.87	11	5.294	0.0005	***	[0.56, 0.97]
Panel A: GVA-Sales Wedge vs CPI−WPI (incl. COVID)	0.90	13	6.848	0.0000	***	[0.69, 0.97]
Panel B: Discrepancy vs CPI−WPI (excl. COVID)	0.21	20	0.911	0.3740	n.s.	[-0.26, 0.60]
Panel B: Discrepancy vs CPI−WPI (incl. COVID)	0.41	22	2.010	0.0580	*	[-0.01, 0.71]

Note: Panel A is highly significant in both specifications (p<0.001, r = 0.87 and 0.90). This is the paper's strongest statistical result and the one most robust to the COVID exclusion decision. Panel B is not significant in the base case. The paper's reported r=0.31 may reflect a slightly different sample and differs from the replication here. With COVID included, Panel B rises to borderline significance. Note the negative confidence intervals (CI).

Table A.2.5 for Figure 13 - NSC vs Official GVA vs Sales.

Series	Correlations	Number of Observations (n)	t-stat	p-value	Sig.	95% CI for Correlations
Official Back-casted GVA vs Sales	0.45	17	1.952	0.0699	*	[-0.04, 0.77]
NSC (Mundle Committee) GVA vs Sales	0.53	17	2.421	0.0286	**	[0.07, 0.81]

Note: Both series results are statistically significant.