

Estimating the efficiency of GST: a counterfactual approach based on consumption expenditures with West Bengal as a case study

Dr. Sattwik Santra

Centre for Training and Research in Public Finance and Policy

(A Ministry of Finance, Government of India initiative)

Centre for Studies in Social Sciences, Calcutta



Contents

Preface	1
Introduction	2
Data and Organization	5
Methodology and Results	6
Remarks and conclusion	19
References	24
Appendix	25

Preface

This report provides an overview into a mechanism of estimating potential indirect taxes from consumption data. The objective of the report is to suggest various alternative models to quantify the efficiency of the Goods and Services Tax collected by the government with a case study of the state of West Bengal.

The outline of the report is as follows: the introduction highlights the importance of indirect taxes as a source of revenues for the government and elaborates on the aims and scope of this study. The sources and nature of data used in this study is discussed in the next section, along with the methods used to organize the data. The Methodology and Results section develops the underlying models used in the analysis and the results obtained therefrom. The concluding section presents the shortcomings of this study and offers some final remarks. Technical details and results from the empirical exercises are presented in the Appendix.

The report contains a number of figures, notes and tables. Notes are placed at relevant parts of the text and are boxed. Important tables and the main results of the article are boxed with thicker outlines. The cover art depicts the plots of the nation-wide counterfactual (in blue) and the official (in green) monthly GST figures.

The author would like to thank Prof. Jyotsna Jalan and Mr. Arjab Bose for their helpful comments and suggestions. Responsibility for all errors and omissions made lie solely with the author.

Sattwik Santra
April 26, 2023

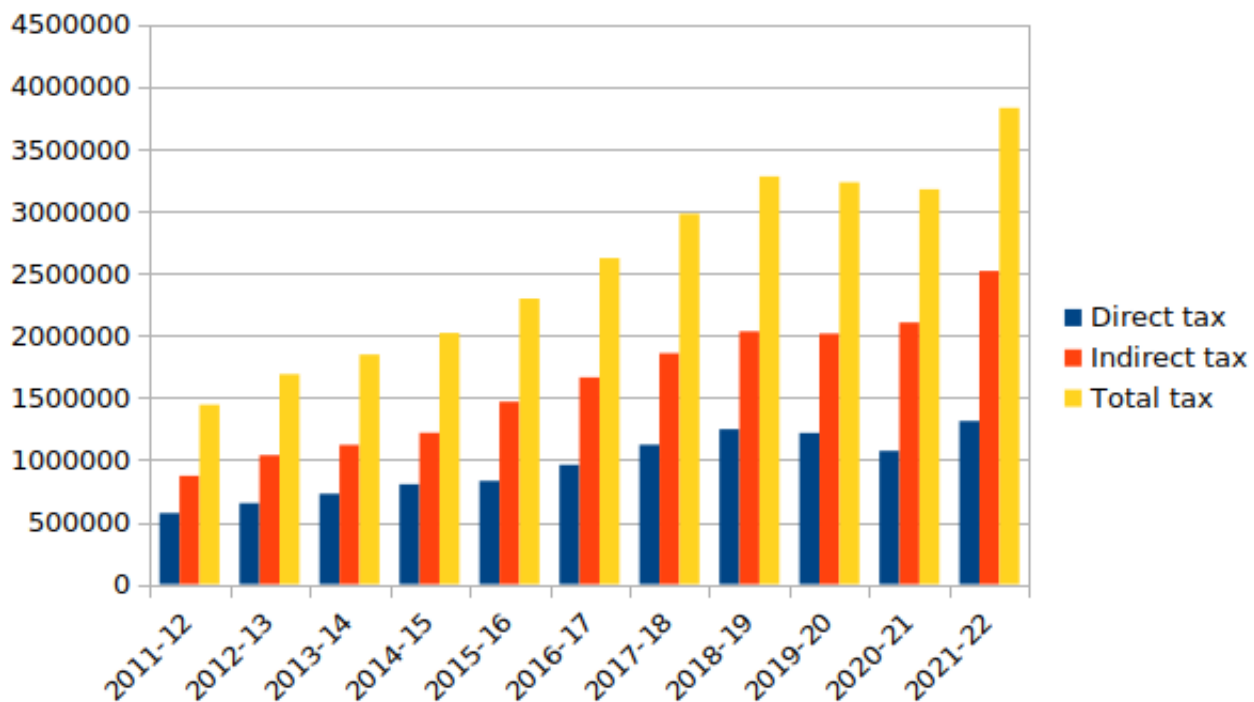
Introduction

Taxes are one of the most important sources of revenue for any government. In India, indirect tax (i.e., tax on goods and services) contributes on an average (from 2011 to 2021), 63% of the total taxes collected across the nation and is one of the primary sources of revenues for the state governments. The following table and graphs present some relevant nationwide macro-economic aggregates.

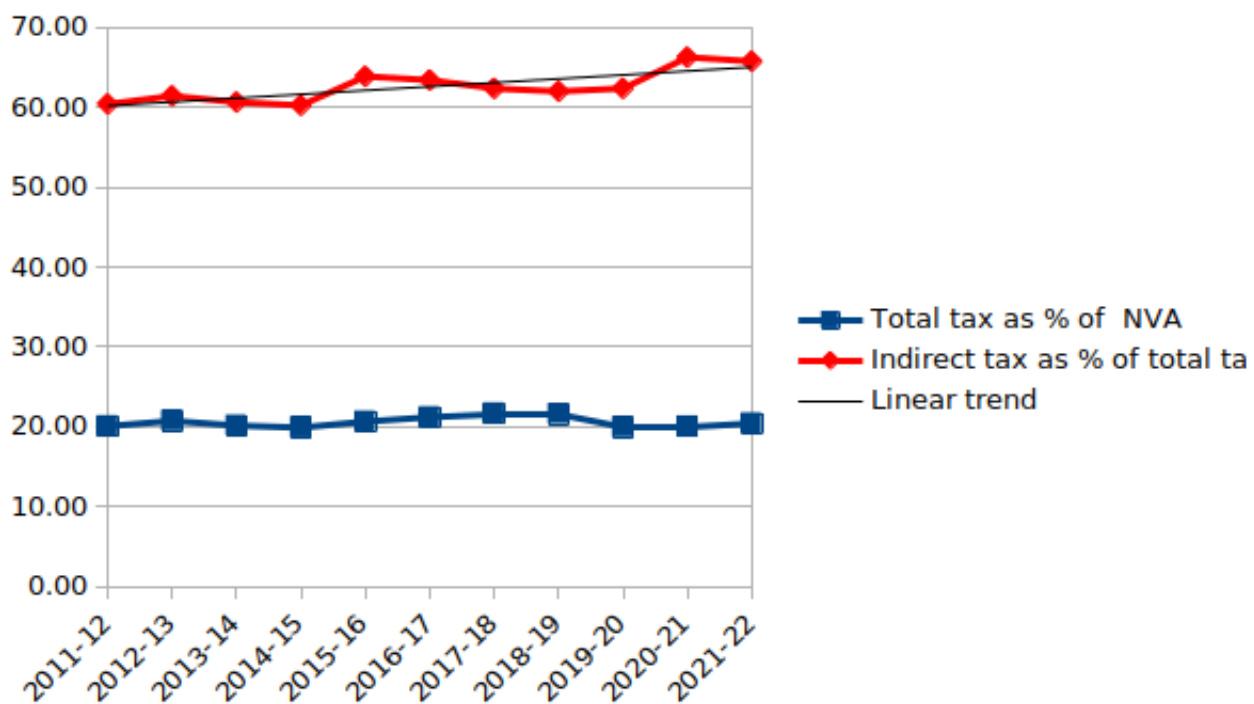
Year	Net Value Added (NVA)	Direct tax	Indirect tax	Total tax	Total tax as % of NVA	Indirect tax as % of total tax
2011-12	71.8977	5.7125	8.7151	14.4275	20.07	60.41
2012-13	81.4179	6.5123	10.3673	16.8796	20.73	61.42
2013-14	91.6718	7.2677	11.1977	18.4655	20.14	60.64
2014-15	101.6199	8.0344	12.1729	20.2073	19.89	60.24
2015-16	111.2480	8.3012	14.6698	22.9710	20.65	63.86
2016-17	123.7387	9.5963	16.6252	26.2215	21.19	63.40
2017-18	137.4085	11.2119	18.5695	29.7813	21.67	62.35
2018-19	151.9084	12.4608	20.3286	32.7895	21.59	62.00
2019-20	162.1751	12.1620	20.1538	32.3158	19.93	62.37
2020-21	158.4819	10.7155	21.0404	31.7559	20.04	66.26
2021-22	187.6443	13.1145	25.1844	38.2989	20.41	65.76

Table 1: Some relevant all-India macro-economic aggregates and some ratios

Notes: All figures in columns 2 to 5 are in rupees trillion. Data on Net Value Added retrieved from the Central Statistics Office (CSO) Government of India (<https://www.mospi.gov.in/data>). Data on Net Value Added for 2019-20, 2020-21 and 2021-22 are respectively the third, second and first revised estimates. Data on taxes are compiled from the Handbook of Statistics on the Indian Economy 2021-22, published by the Reserve Bank of India (<https://rbi.org.in>). Tax data for 2020-21 are revised estimates.



(a) Table 1 in graphs



(b) Table 1 in graphs

The above table and graphs clearly show that indirect tax is a major component of the total tax revenue. Moreover, its proportion in total tax collections shows an increasing linear trend (estimated coefficient: 0.4844, standard error: 0.13, p-value < 0.01).

At present, India has in place a system of Goods and Service Taxes (GST) that was put in place on the 1st of July, 2017 after over ten years of rigorous and widespread consultations and negotiations with the state governments. According to the Goods and Services Tax Council’s website, “...the Council has been witness to many improvements in the economy such as the increased formalization of the economy and widening of the tax base, as highlighted by the Economic Survey 2017-18... Adjusting the base for double and triple counting, the GST has increased the number of unique indirect taxpayers by more than 50 percent, a substantial 3.4 million...Monthly average of GST Collection over last” (2017-18) “financial year was Rs. 89,885 crores; This figure increased to an average of Rs. 97,000 crores in 2018-19 registering a growth rate of approximately 9% despite multiple rate cuts recommended by the Council”.

Against this backdrop, this report assesses the recent performances of GST by comparing the official GST collections against a methodologically formulated counterfactual with a particular emphasis on the state of West Bengal.

The counterfactual GST figures are computed from observed domestic consumption expenditures using officially reported GST rates. In the literature on tax capacity and tax efficiency, several authors (see for example Mukherjee (2020) and the citations therein or Huu (2019)) have used Gross (Net) Domestic Product/Value Added as a proxy for consumption base. However, the present study uses data on domestic consumption expenditures to estimate the difference between the potential and the realized tax revenues.

There are three advantages of using data on consumption expenditures over domestic products to explain tax collections. First, since GST is a consumption based tax, consumption expenditure is the better quantifier of the tax base. Secondly, since consumption data is available at a relatively greater level of commodity-wise disaggregation than data on domestic product/value added, it is possible to compute the prospective tax collections or estimate the tax responsiveness (the rate of change in the tax collected vis-a-vis a change in the tax base) with greater level of precision. Finally, data on domestic product/value added compiled by governmental agencies are collected from relatively formal sources that are easily scrutinized by such bodies. This suggests particularly in the context of developing economies, that the data excludes the relatively informal sectors. Any analysis based on such data would underestimate of the true tax base and the consequent estimates of potential taxes: which would not be the case for an analysis using data on consumption expenditures.

The literature on tax capacity and tax efficiency utilizes various empirical techniques; from simple linear models to more sophisticated models, to estimate the efficiency of the tax system (measured as the difference between the potential and the realized tax revenues). In keeping with this trend, the empirical analysis in this report adopts linear parametric and non-linear non-parametric methods to estimate the tax efficiency.

Data and Organization

The analysis uses data on consumption expenditures compiled by the Centre for Monitoring Indian Economy (CMIE) Pvt. Ltd.. This is a consumption expenditure survey that is continuously implemented on a panel (i.e. the same households are visited over time) of randomly selected households from twenty eight states and union territories (which shall all be referred to as ‘states’). The data from this survey is available from January 2014 at a monthly frequency where every sampled household is interviewed once every four months. The average monthly household sample size is 1,31,941. In the current exercise, data from April 2018 to August 2022 is considered. The data consists of household level observations on consumption expenditures of 136 commodities as well as various household characteristics like the location parameters of the households (state, district etc.), religion, social class etc.. This study uses the consumption data and the locational information of the households. The sample of observations from the states is adjusted by the respective states’ estimated population figures obtained from the Census of India 2011, populations projections data (Report of the Technical Group on Population Projections (2020)) to arrive at monthly state level commodity-wise consumption aggregates.

The States and Union Territories are:

Andhra Pradesh, Assam, Bihar, Chandigarh, Chhattisgarh, Delhi, Goa, Gujarat, Haryana, Himachal Pradesh, Jammu and Kashmir, Jharkhand, Karnataka, Kerala, Madhya Pradesh, Maharashtra, Odisha, Puducherry, Punjab, Rajasthan, Tamil Nadu, Telangana, Uttar Pradesh, Uttarakhand, West Bengal, Tripura, Meghalaya and Sikkim.

If x represents the value of total consumption of a commodity which includes the GST paid and t denotes the commodity’s GST rate, then $x \frac{t}{1+t}$ yields the amount of tax revenue generated by the consumption of that commodity.

The above formula is used together with the commodity-wise GST rates available at the Goods and Services Tax Council’s website and the consumption expenditure data from CMIE, to obtain the monthly commodity-wise and (thus) aggregate counterfactual GST estimates of the states. The list of relevant commodities and the respective GST rates considered, is summarized in Appendix table A1. To allow for inter-temporal comparisons of all the reported and estimated quantities (which are subject to price fluctuations), all the respective variables are adjusted by the monthly price indices obtained from the Consumer

Price Indices Warehouse of The Central Statistics Office (CSO) (Consumer Price Indices, 2022).

Methodology and Results

Figure 2 plots the 2018-2021 average of the annual consumption expenditures, Net State Domestic Products (NSDP) and officially reported GST collections of the respective states. The 2018-2019 and the 2018-2021 averages for the same set of variables across the states are reported in Table A2 in the Appendix. The yearly aggregates over the states are presented in the Table 2. The results reported in Figure 2 and Table 2 unequivocally

reveal an important aspect of the data: the consumption figures (and the potential tax computed therefrom) are almost always higher than the corresponding value added figures reported by the government agencies. This comes

Table 2: Yearly aggregates of consumption, counterfactual GST, NSDP and reported GST collections in rupees trillion

Year	Consumption	Counterfactual GST	NSDP	Officially reported GST collections
2018	217.7636	13.6590	166.7260	7.4222
2019	220.8774	13.8981	173.6271	7.9625
2020	169.5301	10.9693	164.9056	6.7923
2021	179.6902	11.5454	125.1059	8.1767

as no surprise given the presence of a significant informal sector. As already stated earlier, this sector remains largely uncaptured in official data but is adequately represented in a consumption survey thus leading to the discrepancy. Furthermore, as reported in Appendix table A3, the elasticity of officially reported GST collections is only statistically significantly

related to NSDP and not with consumption expenditures. This further indicates the disassociation between the GST and consumption expenditure figures and re-asserts the official sources' failure to account for the transactions undertaken in the informal economy. These observations highlight the importance of con-

Elasticity of a variable (y) with respect to another variable (x) is defined as the percentage change of the first variable (y) due to a unit percentage change in the second variable (x). In a linear regression with both dependent and independent variables considered in logarithms, respective coefficient estimates are synonymous to elasticity estimates.

sumption expenditure as a major macro-economic aggregate that should not only be used to quantify the tax base but should also be used to corroborate trends observed in other reported official macro-economic aggregates.

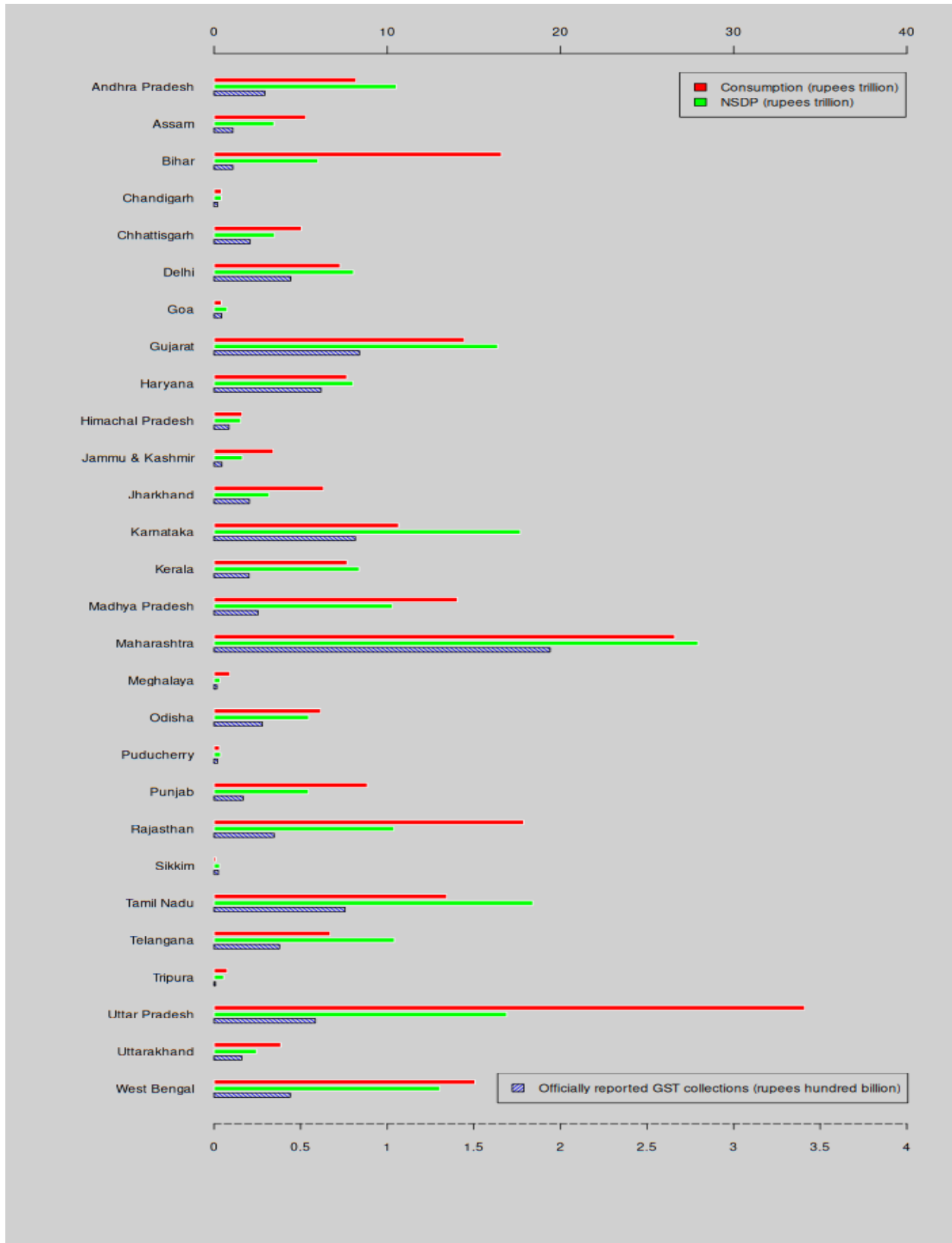
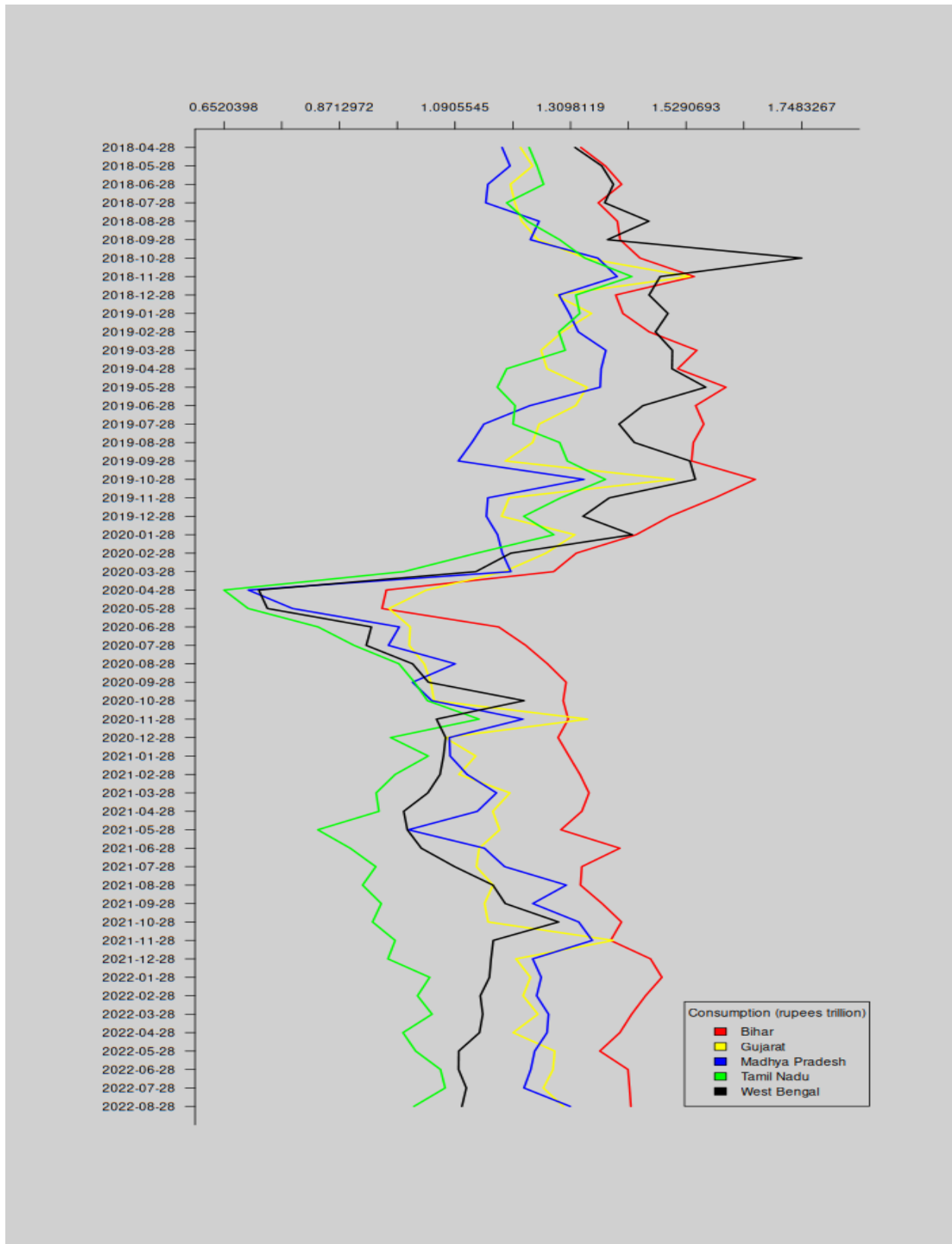


Figure 2: Consumption, NSDP and GST: 2018-2021 averages

Except for NSDP which is available only annually, the data allows one to explore further into the monthly trends of the other variables. Figures 3 shows the trends of the consumption expenditures and Figure 4 compares the counterfactuals and the officially reported GST figures of the selected states of Bihar, Gujarat, Madhya Pradesh, Tamil Nadu and West Bengal. These five states are selected owing to the similarity in terms of their 2018-2021 average consumption expenditures as indicated in Appendix table A2.

There are some salient features observed in these figures. There are similarities in the consumption patterns across the states. This is observed in the counterfactual GST as well. The official GST collections also manifest a similar inter-state temporal correlation but when compared to their respective counterfactual counterparts, exhibit in general, a lesser extent of inter-temporal variation. Another interesting observation is that all the reported quantities experience a ubiquitous drop in magnitude from around October 2019 that finally flattened at April 2020. The values recover to an extent from this all-time minimum but as of August 2022, pre-October 2019 levels are yet to be achieved. This drop in magnitude could be readily attributed to two factors. First (chronologically) is the economic slowdown that affected the economy (see for example, Gupta and Tyagi 2022) and the second is the emergence of COVID in the country and the resultant nation-wide lock-down that was enforced by the government to contain the spread of the infection.

The rest of the analysis focuses on a comparison between the counterfactual and the officially reported monthly GST figures for the state of West Bengal. In Figure 5, the consumption expenditures, the counterfactual GST figures and the officially reported GST collections for the state, are plotted. Given the stochastic nature of the time series, any statistic based on the month by month difference between the counterfactual and the officially reported monthly GST figures would not serve as a robust measure of tax efficiency. To arrive at a consistent measure of tax efficiency, the present study uses statistical techniques to smooth out the inter-temporal fluctuations. Appropriate statistics are then estimated from the smoothed series. This ensures two things: (i) it reduces the impact of outliers on the estimates and (ii), it enhances the correlation between the counterfactual and the officially reported monthly GST figures, if any. In what follows, two such smoothing methods are discussed along with the results obtained therefrom.



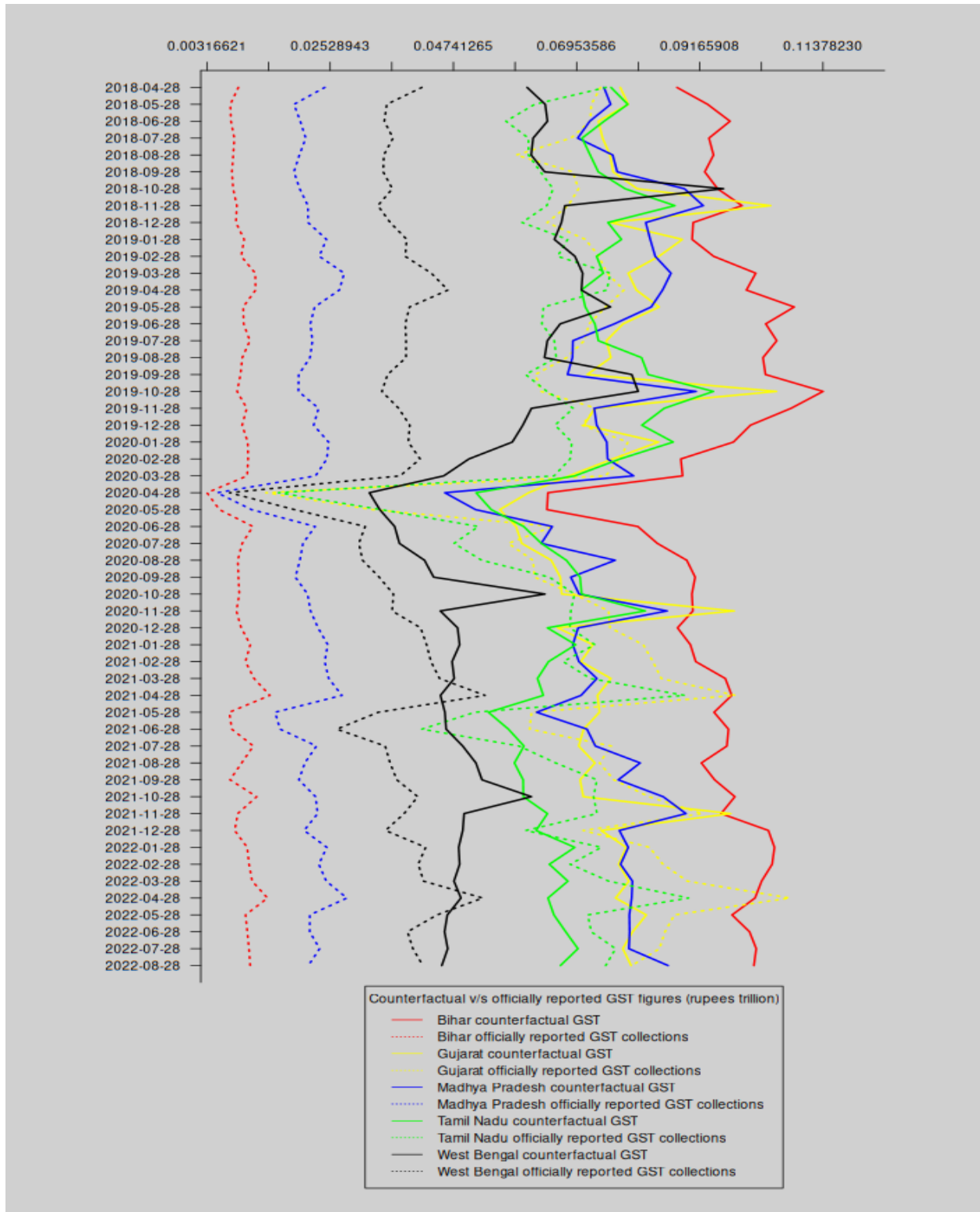


Figure 4: Counterfactual and officially reported monthly GST for selected states

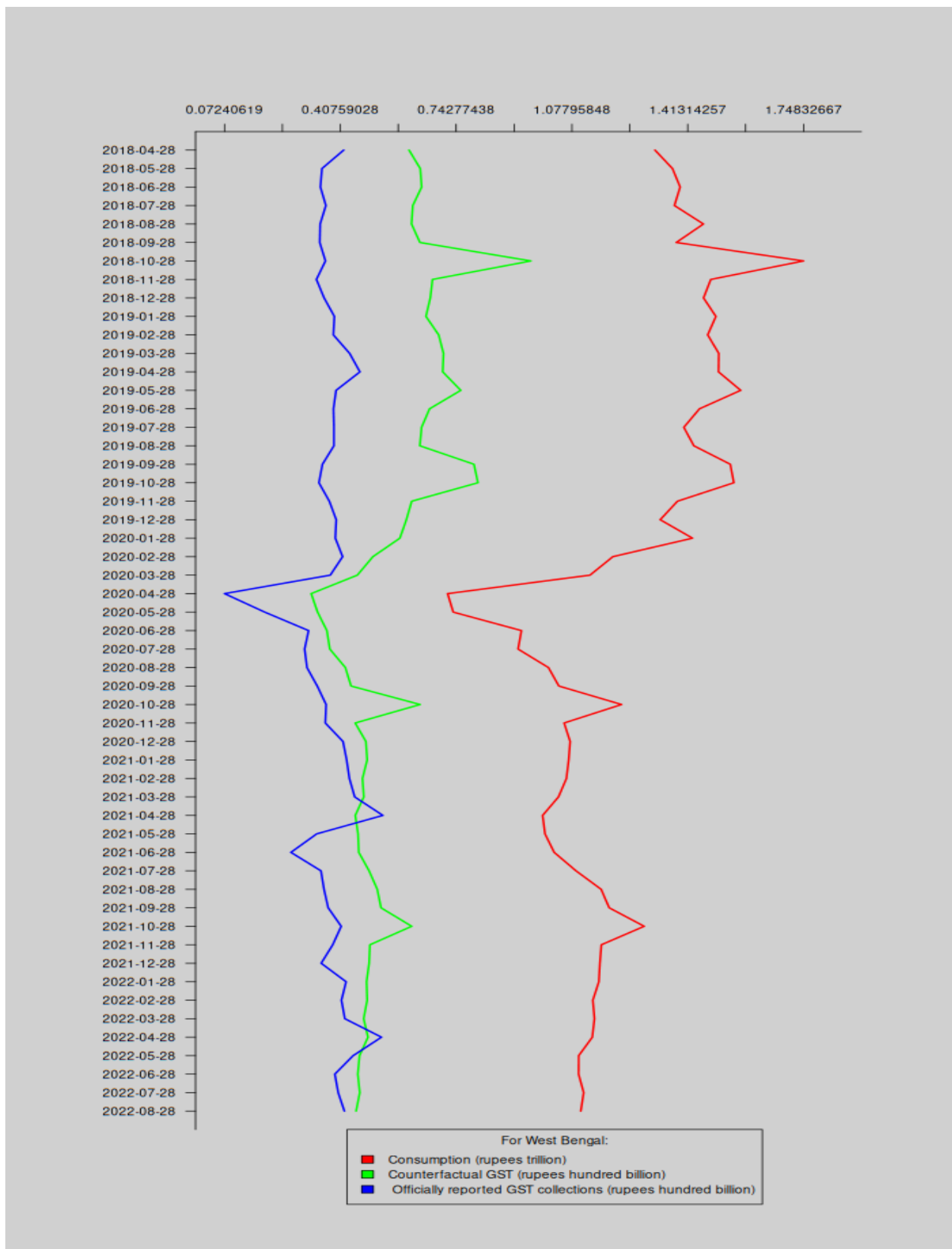


Figure 5: Counterfactual and officially reported monthly GST for West Bengal

First, a Locally Weighted Scatterplot Smoothing (LOWESS) (Cleveland, 1979; Cleveland and Devlin, 1988) is used to estimate the non-linear trends in the data.

LOWESS creates a smooth line through a sequence of data points. At each point ‘ x ’ in the range of the data set, a low-degree polynomial is fitted to a subset of the data that corresponds to a local neighborhood around x with the corresponding index serving as the explanatory variable. The fitted value corresponding to the index of x is considered as LOWESS smoothed value corresponding to x . The polynomial is fitted using weighted least squares, giving more weights to points near the point whose response is being estimated and less weights to points further away. The LOWESS fit is complete after the fitted values have been computed for each of the data points.

In the present context, a local linear fit is used. Two alternate formulations of ‘locality’ are considered here. The first formulation considers all data points that are within twelve months from the point whose LOWESS fit is estimated. The second formulation considers all data points that are within twenty four months. Also, a tri-cubic weighting scheme is used for the regression. The tri-cubic weighting function is given by: $(1 - |d|^3)^3$ where ‘ d ’ is the distance measured in units of

Although the fit is locally linear, globally, the estimates are non-linear and non-parametric.

month, scaled to lie in the range from 0 to 1. The LOWESS estimates of the trends of the counterfactual and the officially reported monthly GST figures are plotted alongside the actual variables in Figures 6 and 7. Table 3 reports the summary statistics of the fitted LOWESS trends among others.

Inspection of Figures 6 and 7 reveal the two features of LOWESS: its non-linearity and its ability to smooth out the effects of extreme observations. Thus in Figure 6 and especially in Figure 7 it is found that even with the quantitative drop in magnitude of the original variables concurrent with the economic slowdown of 2019 and the emergence of COVID in 2020, the smoothed trend values have remained relatively immune to such changes. Furthermore, the effects of local variations in the original time series on the estimated trends diminish as one moves further away in time. Given these properties of LOWESS, the tax efficiency is quantified as the difference between the smoothed counterfactual and the the smoothed officially reported GST figures. The summary statistics from this exercise are reported in Table 3. In this report, only the results for the twenty four months smoothed LOWESS results are discussed. However, the results for the twelve months smoothed LOWESS are also available in the table for the interested readers. Note that the descriptives for the twelve months smoothed LOWESS are not significantly different from the corresponding statistics for the twenty four months smoothed LOWESS.

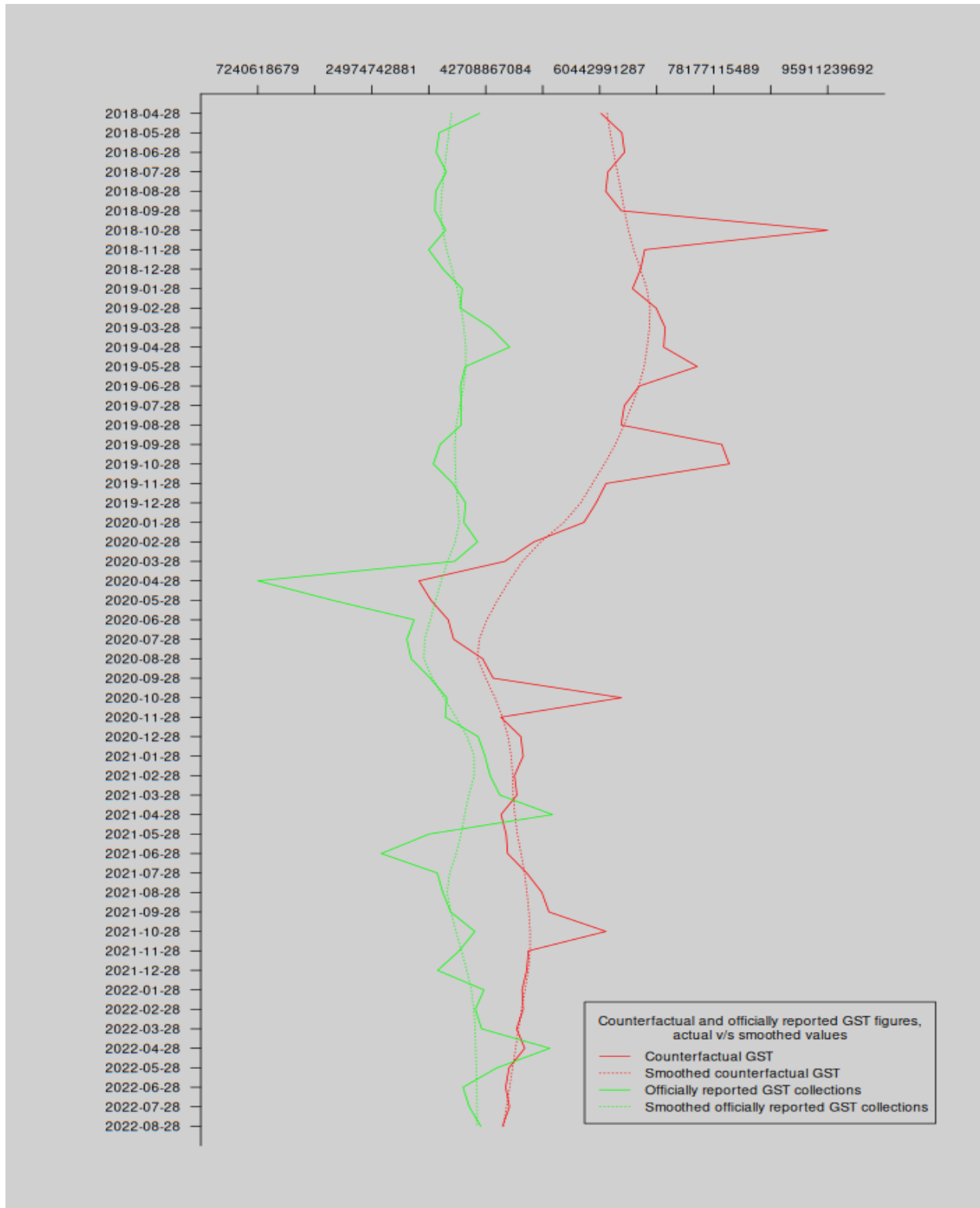


Figure 6: Counterfactual and reported GST for West Bengal with Locally Weighted Scatterplot Smoothing: twelve nearest neighbors

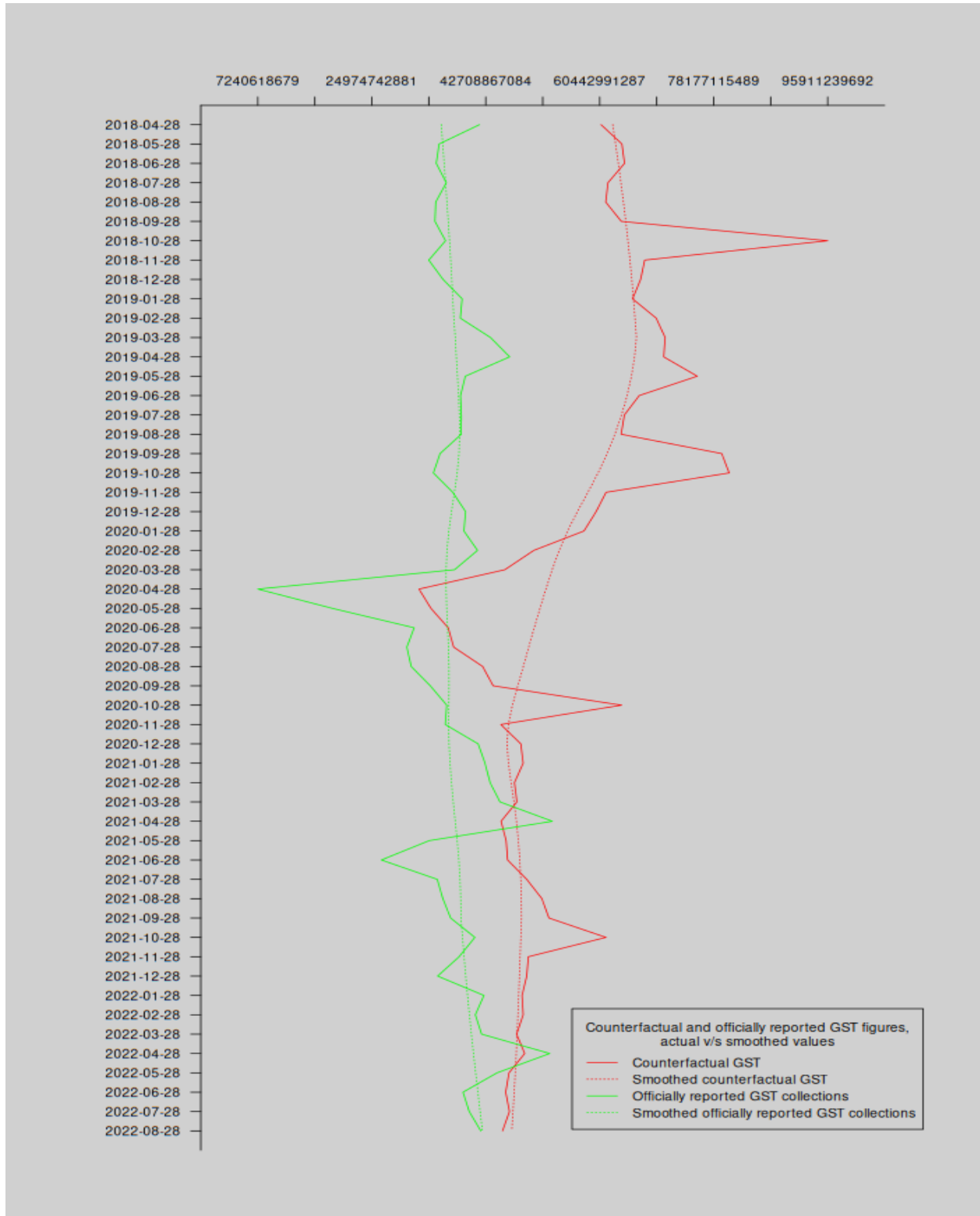


Figure 7: Counterfactual and officially reported monthly GST for West Bengal with Locally Weighted Scatterplot Smoothing: twenty four nearest neighbors

Table 3: Summary statistics for smoothed counterfactual and officially reported GST for West Bengal in rupees

	Mean	Median	Maximum	Minimum
<u>LOWESS estimates of</u> <u>(12 months)</u>				
Counterfactual monthly GST	5391,79,06,383	4933,28,65,498	6828,33,92,212	4142,27,56,291
Officially reported monthly GST	3812,51,21,211	3811,23,35,207	4135,81,00,760	3308,57,19,514
Difference between smoothed monthly counterfactual GST and smoothed officially reported monthly GST	1579,27,85,172	1162,82,76,382	2949,03,89,086	410,27,41,948
	Mean	Median	Maximum	Minimum
<u>LOWESS estimates of</u> <u>(24 months)</u>				
Counterfactual monthly GST	5444,93,02,442	5025,01,54,497	6615,53,49,330	4602,19,62,701
Officially reported monthly GST	3808,47,92,745	3775,06,89,151	4218,25,16,238	3579,32,80,598
Difference between smoothed monthly counterfactual GST and smoothed officially reported monthly GST	1636,45,09,697	1347,55,05,311	2819,30,90,965	457,34,38,144

From Table 3 it is apparent that for West Bengal, the average of the smoothed officially reported monthly GST collections stands at Rs. 3803 crores. The average of the smoothed counterfactual monthly GST figures stands at Rs. 5445 crores. The average of the differences between the smoothed counterfactual monthly GST figures and the smoothed officially reported monthly GST collections is Rs. 1636 crores which is 43% of the average smoothed officially reported monthly GST collection. A relatively conservative alternative (where average is replaced by the median which is more unresponsive to extreme observations) puts the estimate to Rs. 1348 crores or 36% of the corresponding smoothed median official monthly GST collection. The lower bound of the differences stands at 457 crores and the upper bound at Rs. 2819 crores.

The two features of LOWESS smoothing have some trade-offs. First, the non-linearity in the smoothed series does not lend itself to an easy interpretation of the trend in the dataset (as it differs continuously over time). It also further complicates the task of making out-of-sample forecasts. Second, LOWESS's ability to smooth out the time series makes it difficult to model structural breaks in the underlying data generating process. To alleviate these concerns, the current report adopts a segmented linear regression (SLR) (Muggeo, 2003) model to parameterize the underlying trend in the data.

Segmented linear regression is a regression technique that fits a piece-wise linear continuous trend over a multitude of break-points the locations of which are considered endogenous to the model. The number of breakpoints to be considered however, is exogenous and has to be pre-set for model estimation.

Given that the economic slowdown of 2019 and the emergence of COVID in 2020 have occurred in successive time periods with some degree of overlap, the number of breakpoints considered for the SLR model is set at two. This separates the entire time span into three regimes: a pre - economic slowdown / emergence of COVID (hereafter referred to as the 'dip') regime, a 'dip' regime and a post 'dip' regime. Positing that the 'dip' regime is characterized by instabilities (in demands and productions) wherein the standard economic relationships might fail, trends and estimates from this regime are not used for the analysis. As stated above, the exact temporal locations of the breakpoints are estimated from the model and are different for the two time series (of the counterfactual GST and the officially reported GST collections) implying that the 'dip' regions for the two time series are different. The differences in the estimated dates of the breakpoints raise a methodological problem since a comparison between the trends of the counterfactual GST and the officially reported GST collections requires these trends to pertain to the same (pre/post) regime with same temporal extents. To effect such a comparison, the period between the first and the ultimate break dates considering both the time series, is considered as the period for the 'dip'. Separate trends for the counterfactual GST and the officially reported GST collections are estimated for the pre-'dip' and post-'dip' regimes and the difference between linear trends for the respective regimes are computed to evaluate the efficiency of the tax system.

Figure 8 plots the the counterfactual GST and the officially reported GST collections, the estimated breakpoints for the two series, the respective linear trends for both the pre-'dip' and post-'dip' regimes. The summary statistics of the predicted linear trends for the two time series in the two regimes are delineated in Table 4.

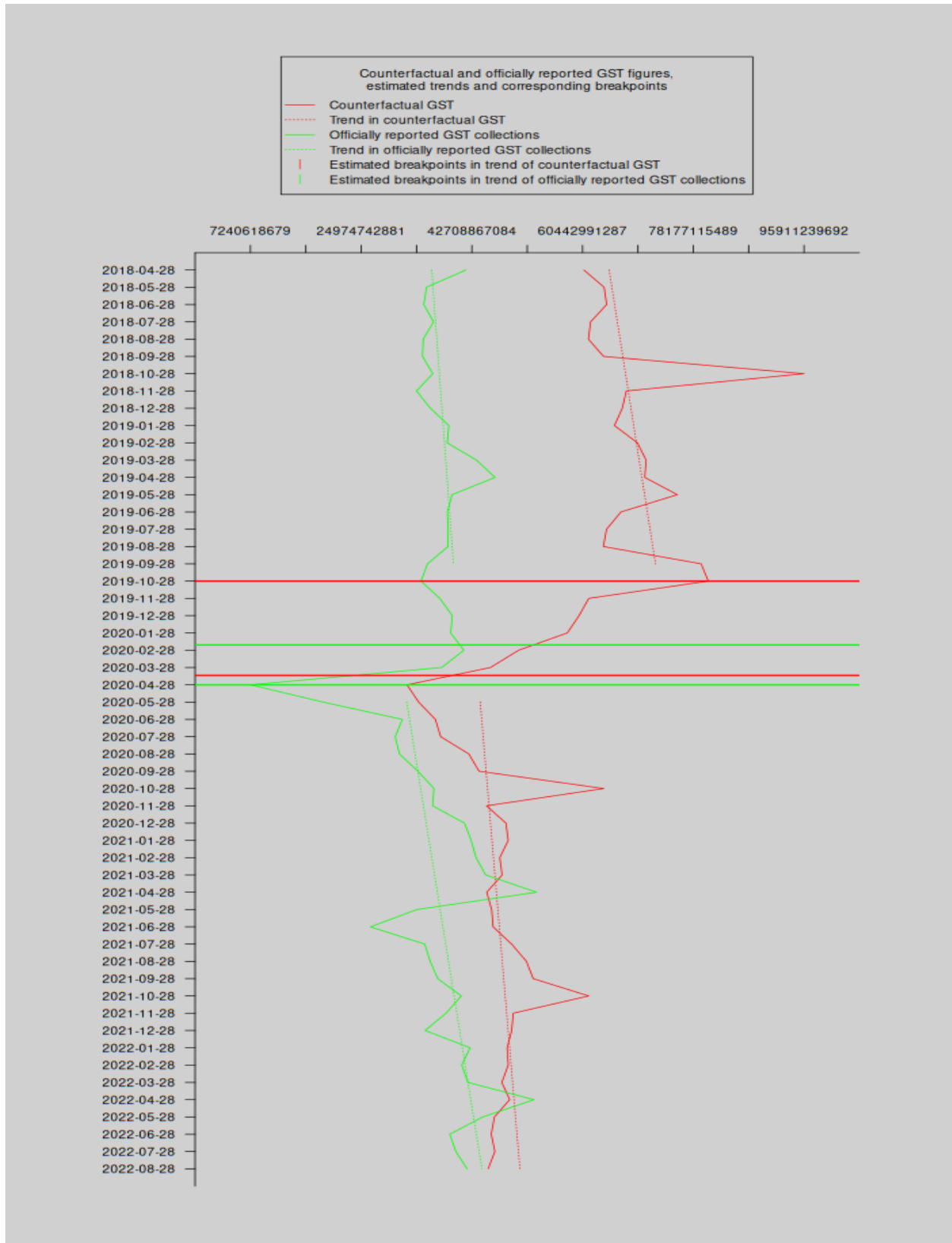


Figure 8: Counterfactual and reported GST for West Bengal with segmented trends

Table 4: Summary statistics for trends in counterfactual and officially reported monthly GST figures for West Bengal: trends estimated using breakpoints from segmented linear regression. Figures in rupees

	First break point	Second break point	
Counterfactual monthly GST	19.00 (October 2019)	24.45 (March 2020)	
Officially reported monthly GST	22.69 (February 2020)	25.00 (April 2020)	
<u>Before first break point</u> (month 19) linear estimates of	Mean	Maximum	Minimum
Counterfactual monthly GST	6841,12,26,288	7212,08,20,812	6470,16,31,764
Officially reported monthly GST	3801,22,17,595	3974,83,13,064	3627,61,22,126
Difference between linear estimates of counterfactual and officially reported monthly GST	3039,90,08,693	3237,25,07,748	2842,55,09,638
<u>After ultimate break point</u> (month 25) linear estimates of	Mean	Maximum	Minimum
Counterfactual monthly GST	4722,15,59,030	5040,05,05,507	4404,26,12,552
Officially reported monthly GST	3829,25,01,955	4431,34,86,789	3227,15,17,120
Difference between linear estimates of monthly counterfactual GST and monthly official GST	892,90,57,075	1177,10,95,432	608,70,18,718
	Before first break point (month 19)	After ultimate break point (month 25)	
Slope of linear estimates of counterfactual monthly GST	43,64,22,885	23,54,77,517	
Slope of linear estimates of officially reported monthly GST	20,42,46,526	44,59,98,877	

From Table 4 it is apparent that for West Bengal, before October 2019, the average of the linear estimates of officially reported monthly GST collections stands at Rs. 3801 crores. The average of the linear estimates of the counterfactual monthly GST figures stands at Rs. 6841 crores. The average of the differences between the linear estimates of the counterfactual and the officially reported monthly GST figures is Rs. 3040 crores which is 80% of the average of the linear estimates of officially reported monthly GST collections. The lower bound of the differences stands at Rs. 2843 crores and the upper bound at Rs. 3237 crores.

After April 2020, the average of the linear estimates of officially reported monthly GST collections stands at Rs. 3829 crores. The average of the linear estimates of the counterfactual monthly GST figures stands at Rs. 4722 crores. The average of the differences between the linear estimates of the counterfactual and the officially reported monthly GST figures is Rs. 892 crores which is 23% of the average of the linear estimates of officially reported monthly GST collections. The lower bound of the difference stands at 608 crores and the upper bound at Rs. 1177 crores.

Note that for a linear trend, the average coincide with the median and hence this statistic is not reported.

The linear trends also allow one to estimate the temporal rate of change of the predicted time series. For the pre-‘dip’ period, the linear estimates of the officially reported monthly GST collections increase by Rs. 20 crores per month, whereas the same for the linear estimates of the counterfactual monthly GST figures rise by Rs. 44 crores per month. This indicates a divergence (by Rs. 24 crores per month) between the two series given that the average difference between the linear estimates of the counterfactual and the officially reported monthly GST figures is positive.

For the post-‘dip’ period, the linear estimates of the officially reported monthly GST collections increase by Rs. 45 crores per month whereas the same for the linear estimates of the counterfactual monthly GST figures rise by Rs. 24 crores per month indicating a convergence (by Rs. 21 crores per month) between the two series. This follows given an average positive difference between the linear estimates of the counterfactual and the officially reported monthly GST figures.

Remarks and conclusion

This study presents different models to evaluate the tax efficiency using data on consumption expenditures. Using first, a LOWESS model to estimate the non-linear trends in the data and second, a segmented linear regression model to estimate the structural breaks of the

linear trends in the data, this study evaluates the tax efficiency of West Bengal during the period from April 2018 to August 2022. Comparing the results from these two models it is found that the average LOWESS smoothed values of both the counterfactual and the officially reported monthly GST figures are not significantly different from the post-‘dip’ average estimated trends for the respective series. From this observation it can be established with a certain degree of analytical certainty that at an aggregate level, a conservative estimate puts the rate of inefficiency in GST collections at 23%.

For the sake of completeness it should be stated that the present analysis has its own set of limitations. The reported estimates could be biased in either direction. The consumption survey on which the counterfactual GST estimates are based upon, excludes a number of items which are consumed by households. This may easily be seen by comparing the list of goods and services in the consumption survey and the list presented at the Goods and Services Tax Council’s website. Had these items been included in the survey, the counterfactual GST figures would be higher. Depending on the characteristics of the supply chain, the effective tax rates of certain commodities are subject to vary from the ‘book’ rate. For such commodities the estimates of the counterfactual GST figures could be biased in either direction. Similarly, the present analysis does not consider temporal variations in the GST rates of the commodities. Inclusion of this variation would lead to a bias which could go in either direction.

The analysis presented here only provides an estimate of the tax inefficiency but is unable to present the distribution of tax inefficiency along different dimensions. For example, given the unavailability of official data, this study is not able to discern the distribution of tax inefficiency across the commodities. Readers are however provided with an unabridged commodity-wise annual estimates of counterfactual GST for West Bengal in Table 5.

This study has obvious extensions and refinements. Rigorous time series modeling could not only provide a further means of estimating current trends but also could be used to predict future trends. Causal models could be used to access transmission of exogenous shocks from consumption expenditures to officially reported GST collections and could be used to estimate tax inefficiencies. Access to data at a further level of disaggregation would lead to more precise estimates of potential tax collections and inefficiencies and would further enable identification of the potential sources of such inefficiencies. Such disaggregation could be at the level of both consumption expenditures and official GST collections. Matching consumer side expenditures with producer side sales could provide further insights into this issue and could lead to significant policy conclusions.

Table 5: Annual counterfactual GST across commodities for West Bengal in rupees billion

Commodities	2018	2019	2020	2021
Airfare	0.1029	0.131	0.0112	0.0081
Artificial Jewellery	2.1495	4.443	1.6172	2.5044
Baby Food	1.0075	1.6266	1.0206	0.7841
Bags Wallets Watches Glasses	1.8127	2.3347	0.3443	0.0782
Bathing Soap	10.0277	9.5722	10.6546	12.468
Beverages	1.4087	3.4635	0.6505	0.2941
Biscuits	12.6819	13.4496	13.6157	13.5668
Bottled Water	0.5741	0.7873	0.6913	0.3118
Bus Train and Ferry	12.9025	17.5739	8.0122	7.722
Cable TV	23.1298	29.9965	32.2174	35.2647
Cell Phone	26.0283	28.7559	32.9369	42.1944
Chocolates and Cakes	4.9772	6.5276	4.2352	4.9308
Cigarettes	12.1487	15.1336	8.2694	7.6254
Clothing	38.971	46.4601	17.4456	21.3574
Clothing Accessories	3.8634	6.3725	1.5369	2.0469
Coffee	0.0855	0.1123	0.1025	0.079
Cooking Fuel	29.6222	28.8102	26.2833	28.3066
Cosmetic and Toiletries	112.0094	110.876	84.9527	98.6001
Cosmetics	50.4443	51.1189	31.8136	37.4682
Creams	8.1892	8.7222	3.6337	4.783
Deodorants and Perfumes	6.7569	6.9707	3.1249	3.5485
Detergent Bars	3.6953	3.8261	2.6559	3.3353
Detergent Liquids and Powder	18.7758	16.4576	14.6028	16.8152
Diesel	1.6958	1.5864	0.1971	0.4207
Dry Fruits and Saffron	0.2629	0.5969	0.3556	0.4563
Dry Spices	10.3788	11.0952	8.9338	10.1887
Electronic Storage Devices	0.0496	0.1314	0.0357	0.0087
Entertainment	1.3281	2.754	0.3561	0.0132
Face Wash	2.2991	2.5847	1.299	2.3035
Flakes Muesli and Oats	0.0239	0.0917	0.0194	0.0811

Table 5: Annual counterfactual GST across commodities for West Bengal in rupees billion ...continued

Commodities	2018	2019	2020	2021
Food In Restaurants	30.5746	36.504	14.0796	21.1637
Footwear	11.4415	14.9909	5.0212	5.016
Furniture and Furnishings	0.2624	0.0552	0.1332	0.3457
Gems and Jewellery	0.2937	0.4382	0.0433	0.0633
Ghee	2.2324	2.6002	1.6432	1.4723
Hair Oil	12.2901	11.1885	9.3823	10.2574
Health Supplements	1.2236	1.8344	1.9611	1.5821
Household Appliances	1.4269	0.9344	0.37	0.6274
Hygiene Products	2.248	5.1095	4.6869	5.5666
Ice Creams	2.329	3.466	0.5049	1.268
Internet	3.177	1.892	0.882	1.0823
Jam Ketchup Pickles	0.5687	0.8162	0.5328	0.6536
Kitchen Appliances	0.2923	0.7232	0.1134	0.0763
Landline Phone	0.9277	0.6654	0.2557	0.2356
Lighting	0.7566	1.1236	0.1869	0.1143
Liquor	1.8365	1.3099	0.4546	0.234
Liquor In Restaurants	0.3879	0.3333	0.0225	0.0368
Medical Tests	0.6885	0.7092	0.1251	0.1316
Medicines	19.3231	18.8697	12.925	15.2506
Milk and Milk Powder	9.5624	9.7785	8.8919	7.9595
Milk Products	1.6066	1.9897	1.8561	2.8658
Mobiles and Accessories	0.8834	1.6436	0.5635	0.1399
Noodles and Pasta	2.7018	3.6594	1.9604	1.3119
Other Beauty and Hygiene Products	5.2495	6.8323	3.0394	4.3447
Other Cosmetics	4.6258	5.8149	2.9561	4.2985
Other Hair Cosmetics	0.6237	1.0174	0.0833	0.0461
Other Tobacco Products	10.4283	15.4316	18.08	19.6448
Outstation Taxi Bus Train	0.0606	0.0503	0.0581	0.036
Painting and Renovation	7.0623	1.2375	2.0879	1.112
Petrol and CNG	10.2776	10.1965	6.8188	10.0522
Powder	4.7399	5.2658	3.176	3.7337
Processed Cereals and Pulses	11.6983	12.8432	10.8904	11.7585

Table 5: Annual counterfactual GST across commodities for West Bengal in rupees billion ...continued

Commodities	2018	2019	2020	2021
Ready to Eat Food	0.6537	1.8151	0.7924	1.4915
Recreation	2.7736	4.4124	0.4744	0.0738
Salty Snacks	4.7516	5.3656	4.6229	5.3667
Scourers	4.0598	4.5255	4.174	4.7813
Shampoo and Conditioner	7.3753	7.1193	6.7813	7.2908
Shaving Articles	3.5443	2.4355	1.3772	1.2067
Sugar and Other Sweeteners	6.1829	6.0394	5.3317	5.6361
Tea	2.8893	3.3583	3.5815	3.9415
Toothbrush	1.4208	2.3084	2.1488	3.6407
Toothpaste	11.7428	11.3106	10.7839	10.6515
Toothpowder	0.0729	0.0846	0.0534	0.0485
Toys	1.5373	2.0184	0.1628	0.0477
Utensils	0.7093	0.9363	0.1115	0.1312

References

- Cleveland, W. S. (1979). Robust locally weighted regression and smoothing scatterplots. *Journal of the American Statistical Association*, 74(368):829–836.
- Cleveland, W. S. and Devlin, S. J. (1988). Locally weighted regression: An approach to regression analysis by local fitting. *Journal of the American Statistical Association*, 83(403):596–610.
- Consumer Price Indices (2022). Consumer Price Indices Warehouse of The Central Statistics Office (CSO), Ministry of Statistics and Programme Implementation. <http://164.100.34.62:8080>. [Online; accessed 28-November-2022].
- Goods and Services Tax Council (2022). <https://gstcouncil.gov.in>. [Online; accessed 28-November-2022].
- Gupta, P. and Tyagi, A. (2022). Slowdown of the Indian Economy during 2019-20: An Enigma or an Anomaly. NCAER Working Papers 137, National Council of Applied Economic Research.
- Huu, N. (2019). Impact of Direct Tax and Indirect Tax on Economic Growth in Vietnam. *The Journal of Asian Finance, Economics and Business*, 6(4):129–137.
- Muggeo, V. M. R. (2003). Estimating regression models with unknown break-points. *Statistics in Medicine*, 22(19):3055–3071.
- Mukherjee, S. (2020). Goods and Services Tax efficiency across Indian States: panel stochastic frontier analysis. *Indian Economic Review*, 55(2):225–251.
- Report of the Technical Group on Population Projections (2020). Population Projections for India and States 2011 - 2036. Technical report, National Commission on Population, Ministry of Health and Family Welfare.

Appendix

Table A1: List of commodities in the CMIE consumption database and their associated GST rates

Name of Commodity	GST Rate	Name of Commodity	GST Rate	Name of Commodity	GST Rate
Airfare	5	Diesel	5	Liquor In Restaurants	28
Bus Train and Ferry	5	Dry Fruits and Saffron	5	Medical Tests	5
Outstation Taxi Bus Train	5	Dry Spices	5	Medicines	5
Artificial Jewellery	18	Edible Oils	5	Milk and Milk Powder	5
Baby Food	5	Electronic Storage Devices	18	Milk Products	5
Bags Wallets Watches Glasses	18	Entertainment	28	Mobiles and Accessories	12
Bathing Soap	18	Face Wash	28	Noodles and Pasta	12
Batteries	28	Flakes Muesli and Oats	5	Painting and Renovation	28
Beauty and Hygiene Products other	28	Food In Restaurants	18	Petrol and CNG	5
Beverages	28	Footwear	18	Powder	28
Biscuits	18	Furniture and Furnishings	28	Processed Cereals and Pulses	5
Bottled Water	12	Gems and Jewellery	3	Ready to Eat Food	18
Cable TV	18	Ghee	12	Recreation	18
Cell Phone	18	Hair Oil	28	Salty Snacks	12
Chocolates and Cakes	18	Hair Cosmetics (others)	28	Scourers	28
Cigarettes	28	Health Supplements	5	Shampoo and Conditioner	28
Clothing	12	Household Appliances	28	Shaving Articles	28
Clothing Accessories	12	Hygiene Products	28	Soft Drinks	28
Coffee	5	Ice Creams	18	Spectacles Contact Lenses and other Aids	12
Cooking Fuel	5	Internet	18	Sugar and Other Sweeteners	5
Cosmetic and Toiletries	28	Jam Ketchup Pickles	12	Tea	5
Cosmetics	28	Juices	12	Tobacco Products (others)	28
Cosmetics other	28	Juices Jams	12	Toothbrush	28
Creams	28	Kitchen Appliances	28	Toothpaste	28
Deodorants and Perfumes	28	Landline Phone	18	Toothpowder	28
Detergent Bars	28	Lighting	12	Toys	15
Detergent Liquids and Powder	28	Liquor	28	Utensils	12

Table A2: Annual average consumption, counterfactual GST, NSDP and officially reported GST in rupees trillion

State	2018-2019				2018-2021			
	Consumption	Counterfactual GST	NSDP	GST	Consumption	Counterfactual GST	NSDP	GST
Andhra Pradesh	7.0425	0.4349	8.1877	0.2427	6.5913	0.4167	8.4757	0.2364
Assam	4.7318	0.2696	2.8058	0.0835	4.178	0.2634	2.7357	0.0856
Bihar	14.8184	0.9863	4.9663	0.0889	13.9541	0.9483	5.0387	0.0915
Chandigarh	0.3872	0.0258	0.3651	0.0178	0.3512	0.0231	0.3502	0.0168
Chhattisgarh	4.465	0.2883	2.8797	0.1636	4.145	0.28	2.8641	0.1694
Delhi	6.614	0.4079	6.6803	0.3912	6.1458	0.3727	6.8012	0.3741
Goa	0.4714	0.0354	0.6524	0.0402	0.3559	0.0257	0.6378	0.0369
Gujarat	12.395	0.7795	13.3963	0.6817	11.7893	0.7403	13.3631	0.6847
Haryana	6.9929	0.4652	6.4324	0.51	6.2194	0.4168	6.5132	0.5009
Himachal Pradesh	1.7533	0.1125	1.3134	0.0768	1.3734	0.0914	1.3106	0.0725
Jammu & Kashmir	3.3802	0.2406	1.2965	0.0367	2.716	0.1879	1.2999	0.0359
Jharkhand	5.7112	0.4099	2.731	0.1753	5.2509	0.3723	2.6379	0.169
Karnataka	9.4731	0.6654	13.5074	0.6638	8.546	0.5898	14.1795	0.654
Kerala	7.4288	0.4526	6.9824	0.1636	6.2243	0.3699	6.7855	0.1632
Madhya Pradesh	11.4503	0.7212	7.742	0.2022	10.9205	0.6961	8.002	0.1981
Maharashtra	25.4637	1.6913	22.7587	1.5754	21.2964	1.4234	22.3709	1.5531
Meghalaya	0.77	0.0469	0.2946	0.0134	0.7628	0.0466	0.2896	0.0133
Odisha	5.4626	0.3621	4.4087	0.2095	5.0129	0.3306	4.4569	0.2272
Puducherry	0.2493	0.0145	0.3195	0.0179	0.245	0.0146	0.2986	0.0162
Punjab	8.2411	0.5569	4.5252	0.1369	7.1804	0.4748	4.4189	0.1359
Rajasthan	15.8449	0.9962	8.3058	0.2839	14.754	0.9448	8.5635	0.286
Sikkim	0.1027	0.0061	0.2585	0.017	0.0951	0.0052	0.2624	0.0195
Tamil Nadu	12.3287	0.7548	14.6654	0.6212	10.7562	0.6882	14.7399	0.6058
Telangana	5.4402	0.3221	8.0483	0.2972	5.2384	0.322	8.1349	0.2951
Tripura	0.7341	0.0332	0.4496	0.0055	0.614	0.0268	0.4486	0.0059
Uttar Pradesh	30.9154	1.8394	13.9618	0.4776	27.3214	1.6738	13.5403	0.4689
Uttarakhand	3.0034	0.2229	1.9993	0.1407	3.0148	0.23	1.9115	0.1248
West Bengal	13.6494	0.637	10.2424	0.3582	11.913	0.5427	10.3206	0.3481
All India	219.3206	13.7785	170.1765	7.6924	196.9655	12.5179	170.7517	7.5888

Table A3: Fixed effects regression of officially reported GST collections on consumption and NSDP

Officially reported GST collections (in logarithm)	All states	States in the bottom quan- tile consumption removed
	<u>Panel(a)</u>	
Consumption (in logarithm)	0.0544 (0.07)	-0.0810 (0.10)
NSDP (in logarithm)	1.0098*** (0.21)	0.9002*** (0.21)
Linear trend	0.0672*** (0.02)	0.0729*** (0.02)
Covid year	-0.1485*** (0.05)	-0.2035*** (0.05)
	<u>Panel(b)</u>	
Consumption (in logarithm)	0.0839 (0.16)	-0.0356 (0.21)
NSDP(in logarithm)	1.3336*** (0.43)	1.7231*** (0.38)
N	28	21
T	3-4	3-4

Notes: This table reports the results of regressions of officially reported GST collections considered in logarithms on the logarithms of consumption expenditures and NSDP. Also included in the regressions are linear trends and covid-year dummies. The observations for the variables are annual aggregates at the state level. Panel (a) assumes a common state-wide linear trend and a common state-wide shift in the intercept as indicated by the parameter of the covid-year dummy. Panel (b) relaxes the common state-wide parameters assumption and allows the linear trend and shift in the intercept (associated with covid year) to vary across states. The state specific trends and shifts in intercepts are not reported. Both these regressions assume state level fixed effects i.e. state-specific overall intercept terms. Levels of significance are indicated with asterisks where ‘***’: $p \leq .01$, ‘**’: $0.01 < p \leq .05$ and ‘*’: $0.5 < p \leq 0.1$.

