STAFF STUDIES

CENTRAL BANK OF SRI LANKA

ISSN 1391 – 3743 **Vol. 51 - 2021**

STAFF STUDIES

CENTRAL BANK OF SRI LANKA

Volume 51 – 2021

The views presented in the papers are those of the authors and do not necessarily indicate the views of the Central Bank of Sri Lanka.

For all correspondence about the Staff Studies Journal, please contact:

Director of Economic Research
Central Bank of Sri Lanka
30, Janadhipathi Mawatha
Colombo 01, Sri Lanka
Email: directorerd@cbsl.lk

ISSN 1391 – 3743

Published by the Central Bank of Sri Lanka, Colombo 01, Sri Lanka.

Director of Economic Research, Central Bank of Sri Lanka

Dr. S Jegajeevan

Research Advisory Panel of the Economic Research Department

Dr. P K G Harischandra (Co-Chair)
Dr. S Jegajeevan (Co-Chair)
Dr. L R C Pathberiya
Dr. V D Wickramarachchi
Dr. M M J D Maheepala
Dr. S P Y W Senadheera
Dr. P A H E Paranavithana
Mr. E G J K Edirisinghe
Dr. K D Pathirage
Dr. V S Jayawickrema
Mr. W G P R Wathugala
Mr. W S N Perera
Mrs. M M L K Wijerathna
Mrs. P Ratnavadivel
Mr. A J Wijayawardhana
Mr. C H Kulawardena

List of External Reviewers

Prof. K Amirthalingam (University of Colombo)
Dr. Rahul Anand (International Monetary Fund)
Prof. Prema-chandra Athukorale (Australian National University)
Prof. Sunil Chandrasiri (University of Colombo)
Prof. Sirimewan Colombage (Open University of Sri Lanka)
Prof. Priyanga Dunusinghe (University of Colombo)
Dr. Manuk Ghazanchyan (International Monetary Fund)
Dr. Ananda Jayawickrama (University of Peradeniya)
Dr. Rajith Lakshman (Institute of Development Studies – UK)
Dr. Koshy Mathai (International Monetary Fund)
Prof. Nelson Perera (University of Wollongong)
Dr. Victor Pontines (The SEACEN Centre)
Mr. K D Ranasinghe (Central Bank of Sri Lanka, retired)
Mr. Adam Remo (International Monetary Fund)
Dr. Ole Rummel (The SEACEN Centre)
Mr. Ananda Silva (Central Bank of Sri Lanka, retired)
Dr. T Vinayagathasan (University of Peradeniya)
Dr. Missaka Warusawitharana (Federal Reserve System –USA)
Dr. Janaka Fernando (University of Sri Jayewardenepura)

Editorial Committee

Dr. S Jegajeevan (Editor)
Mr. S M Rajakaruna (Assistant Editor)
Mrs. I E M Gunawardhana (Assistant Editor)

Cross–Temporal Coherent Forecasts for Gross Domestic Product

G I Rathanayke [1]

Abstract

Timely and accurate forecasts aligning different views of economic agents are of utmost importance in macroeconomic forecasting to facilitate effective policy decisions. Thus, this study investigates the ability of a reconciliation approach to align different viewpoints regarding forecasts and thereby increasing the forecast performance specifically related to GDP forecasting. The proposed methodology is based on forecasting hierarchical time series which is a collection of time series that follow an inherent aggregation structure. The aggregation constraints can be cross-sectional or temporal dimension. Thus, this method attempts to reconcile forecasts so that they follow aggregation constraints in both dimensions. This property is referred to as cross-temporal coherency. As the initial step forecasts are obtained for each of the series in the cross-temporal hierarchy. These are referred to as base forecasts and are often incoherent. These forecasts are then revised so that they become cross-temporally coherent. This is referred to as cross-temporal forecast reconciliation. Empirical applications based on disaggregated economic activities of the production approach for Sri Lankan GDP reveal that this approach brings improvement in forecast accuracy by blending different viewpoints in a data driven way. These cross-temporal coherent forecasts align decisions within an organisation transparently towards one number by aligning short term forecasts with long term forecasts and aligning views at different levels within the GDP hierarchy. As the proposed method is independent of forecasting models different short term forecasting models and long term forecasting models can be used to reflect different viewpoints.

***Key Words**: Cross-sectional aggregation, Temporal aggregation, Forecast combinations, Hierarchical time series, Forecast reconciliation*

***JEL Classification:** C53; N15; F43*

1 The author is currently serving as a Senior Assistant Director of the Macroprudential Surveillance Department. Corresponding email: gayani_ishara@cbsl.lk. The views presented in this paper are those of the author and do not necessarily indicate the views of the Central Bank of Sri Lanka

1. Introduction

1.1. Background

Forecasting macroeconomic variables (especially Gross Domestic Product (GDP) and inflation) is a leading research topic in the current macroeconometric literature as the challenges faced evolve over time. Macroeconomic forecasts are of utmost importance for policy makers to make informed decisions. Particularly, to take proactive decisions rather than reactive decisions. For instance, an early forecast of a recession would assist the government to move towards an expansionary fiscal policy to mitigate the impact of a severe economic downturn. Moreover, a forecast of inflation dropping under the target level of a central bank would give them an early indication to go for easing of monetary policy to stimulate the economy to bring the inflation rate back to the target at the right time. The timing of policy decisions is crucial as it is well known and universally accepted that the impact of monetary policy and fiscal policy decisions are transmitted with a lag. This highlights the importance of accurate forecasts as policy decisions must be timed in such a way that their impact is transmitted to the economy when it is required in order to obtain the intended results. In other words, policies are implemented today for forecasted future economic situations. Thus, improving the reliability and accuracy of macroeconomic forecasts is vital at this stage. Ample sophisticated forecasting models have been developed over time in macroeconomic forecasting literature, both in univariate and multivariate settings. The most prominent models include Dynamic Stochastic General Equilibrium models (DSGE), Dynamic factor models, VAR, and Bayesian VAR. In these models, GDP is commonly modeled in aggregate form. Focus on modeling and forecasting disaggregated subcomponents of GDP either based on the demand side or the production side is very limited. However, this area has recieved growing attention in recent years with studies such as Hahn and Skudelny (2008); Barhoumi et al. (2012); Esteves (2013); Higgins (2014); and Heinisch and Scheufele (2018) which mainly focus on exploring and comparing the accuracy gain of direct GDP forecasting and disaggregated GDP forecasting using a bottom-up approach. This approach involves forecasting the most disaggregated series and simply adding them to form forecasts of the aggregated series. The bottom-up approach has the strength in a way that it does not lose information due to aggregation. However, it only uses information from a single level of aggregation and ignores any correlations between the components and aggregates. In addition, this will perform poorly if the disaggregated series have low signal to noise ratio.

1.2. Hierarchical time series

A collection of time series that follows an aggregation constraint is referred to as a hierarchical time series (Hyndman and Athanasopoulos (2018)). For example, contemporaneous aggregation of GDP components in the production front which is a supply oriented decomposition of the value added by economic activities based on the national accounting methodology (European Commission, International Monetary Fund, Organisation for Economic Co-operation and Development, United Nations, and World Bank (2009)) is a cross-sectional hierarchy with aggregation constraints imposed via national accounting identities. There is growing literature which focuses on forecasting such a collection of hierarchical time series with the aim of ensuring that forecasts adhered to the aggregation constraints across the hierarchy. That is, the sum of the disaggregates should be equal to the corresponding aggregates. If we consider cross-sectional dimensions in the context of the GDP hierarchy, the sum of the forecasts of the economic activities should add up to the forecast of GDP. This property is referred to as coherency (Kourentzes and Athanasopoulos (2019)). Forecasts that are generated separately for each series in the hierarchy are base forecasts. These forecasts may not follow the aggregation constraints of the hierarchy except in the case where forecasts are generated by a simple naïve method. The process that adjusts these incoherent base forecasts to be in line with the aggregation constraints in the hierarchy is known as forecast reconciliation. Forecast reconciliation with cross-sectional hierarchies will align lower-level operational forecast with strategic forecast at higher levels.

1.3. Temporal hierarchies

As explained by Athanasopoulos et al. (2017) temporal hierarchy can be computed for any time series by using non overlapping temporal aggregations. For example, if GDP series is observed in quarterly frequency, we can compute semi-annual and annual levels to form the temporal hierarchy. Forecast reconciliation with temporal hierarchies will align short term forecasts with long term forecasts.

These forecast reconciliation methods have been proven to produce coherent forecasts that adhere to the aggregation structure and improve forecasting accuracy (Hyndman et al. (2011); Athanasopoulos, Ahmed, and Hyndman (2009); Hyndman, Lee, and Wang (2016); Wickramasuriya, Athanasopoulos, and Hyndman (2019); Athanasopoulos et al. (2017); Kourentzes and Athanasopoulos (2019); Athanasopoulos et al. (2020)). However, most of these studies focus on cross-sectional reconciliation or temporal reconciliation separately. To the best of my knowledge, the only studies that consider both these dimensions of reconciliation are those of Kourentzes and Athanasopoulos (2019) which introduce a two step method to generate cross-temporal coherent forecasts for Australian tourism, and Spiliotis et al. (2020) which attempts to sequentially combine multiple temporal aggregation with cross-sectional hierarchies related to electricity consumption.

1.4. Forecasting cross-temporal hierarchical time series

Forecasting cross-temporal hierarchical time series is challenging as forecasts need to adhere to both cross-sectional and temporal aggregation constraints. This is referred to as cross-temporal coherency. This property is important as it enables the forecasts to reflect real features of data. Further, coherent forecasts will enable aligned policy direction with one unique view.

In the context of GDP forecasting, it is vital to have forecasts that adhere to both cross sectional and temporal aggregation constraints for aligned decision making with one unique view on the future economic path. A recent study by Athanasopoulos et al. (2020) focuses on the application of cross-sectional forecast reconciliation using income and expenditure approach national accounting identities. However, to the best of my knowledge, no study has explored the accuracy gains and aligned decision making that would result in using cross-temporal reconciliation in the context of GDP forecasting. This research attempts to address this gap by proposing an alternative direct approach to the two step cross-temporal reconciliation approach introduced by Kourentzes and Athanasopoulos (2019).

1.5. Objectives

The main objective of this research is to explore the application of the cross-temporal forecast reconciliation methodology in the context of GDP forecasting. In this regard, we consider an empirical application which focusses on Sri Lankan production approach real GDP to obtain coherent forecasts while improving forecast accuracy. The motivation of this application is to explore the ability of this method to produce coherent forecasts which improve forecast accuracy compared to traditional bottom-up and direct approaches.

The contribution of this study to existing literature is significant in several aspects. First, it extends the cross-temporal forecast reconciliation methodology to macroeconomic forecasting. Further, it will strengthen current forecast models with the addition of this novel approach to GDP forecasting. Moreover, it will produce GDP forecasts which are coherent across all the sub activities as well as across time. This will align the short term quarterly projections with long term annual forecasts and facilitate the exploration of detailed sub activities which are drivers behind the forecasted GDP growth. It provides a better understanding of the current situation. This will facilitate policymakers to identify economic activities which have significant impact and focus on specialised policies to address specific economic activities under consideration. Methodologically, the exploration of the alternative direct approach to the two step cross-temporal reconciliation approach introduced by Kourentzes and Athanasopoulos (2019) will extend the current literature in this area.

1.6. Outline

Section 2 provides a detailed review of the literature on cross-sectional and temporal hierarchical forecasting approaches developed over time. Section 3 elaborates on the current methodology of forecast reconciliation and introduces the direct cross-temporal forecast reconciliation approach developed in this research study. Section 4 focusses on the empirical application of cross-temporal hierarchical forecasting for Sri Lankan GDP. Finally, section 6 summarises the conclusions of this study.

2. Literature review

2.1. Approaches in forecasting hierarchical time series

Earlier approaches in forecasting hierarchical time series mainly focused on selecting a single level of aggregation and then these were combined in a linear manner to generate coherent forecasts for the hierarchical structure. Top-down and bottom-up are two approaches prominent in literature (Syntetos et al. (2016)). The bottom-up approach involves forecasting the most disaggregated bottom-level series at the lowest level in the hierarchy and using simple aggregation to obtain forecasts at higher levels of the hierarchy (Hyndman and Athanasopoulos (2018)). The top-down approach starts with the forecast for the most aggregated top-level and disaggregates the forecast for the lower levels in the hierarchy as needed. The disaggregation can be based on weights derived from historical data as suggested by Gross and Sohl (1990). However, historical proportions do not reflect the dynamic changes in proportions over time. Athanasopoulos, Ahmed, and Hyndman (2009) propose using proportions based on forecasts to overcome this issue. Another less prominent approach uses a combination of bottom-up and top-down approaches. This is referred to as the middle out approach as it chooses an intermediate middle level to forecast and then aggregating bottom-up, as well as disaggregating top-down (Syntetos et al. (2016)).

Relative comparison of top-down and bottom-up methods in different fields in literature is rather inconclusive on the superiority of any method as conclusions depend on the characteristics of the empirical problem considered. Research that favours top-down approaches argue that disaggregate data are error prone and would produce imprecise forecasts due to high volatility and noise and hence top-down will result in better performance as it focuses on forecasting a smooth aggregated series which can reduce specification error (Grunfeld and Griliches (1960)). Research that favours bottom-up argues that information loss is substantial when aggregating series in a top-down approach (Dunn, Williams, and DeChaine (1976)); Weatherford, Kimes, and Scott (2001). Another set of researchers argues that the best approach depends on the correlation among the time series (Fliedner (1999)) or the underlying data generation process (Zotteri, Kalchschmidt, and Caniato (2005); Zotteri and Kalchschmidt (2007)).

The methods discussed so far all have a common limitation. They only consider one aggregation level and do not incorporate information from the entire hierarchical structure. Furthermore, as highlighted by Kourentzes, Barrow, and Petropoulos (2019) overreliance on a single model for all forecasts may increase model selection risk. On the other hand, if forecasts are generated independently for each level in the hierarchy as a simple method to use information from all levels, they may not be coherent and would fail to account for inherent correlation structure.

2.2. Forecast reconciliation methods

To overcome these limitations in traditional methods in forecasting hierarchical time series Hyndman et al. (2011) introduced a forecast reconciliation method. As explained above, if we forecast each of the time series in the hierarchical structure independently, it will not guarantee that the forecast generated will be coherent. In this context, forecast reconciliation can be considered as a process of adjusting forecasts to make them coherent. The basic idea of the methodology introduced by Hyndman et al. (2011) is to first forecast each time series in the hierarchical structure independently, which they term as "base forecasts". Then, to use a regression model to optimally combine and reconcile these forecasts to produce coherent forecasts. The Ordinary Least Squares (OLS) approach introduced in this paper computes reconciliation weights that only depend on the hierarchical structure and they are completely independent of the data. Hyndman et al. (2011) and Athanasopoulos, Ahmed, and Hyndman (2009) show that this method outperforms the commonly used top-down and bottom-up approaches. Extending this concept, Wickramasuriya, Athanasopoulos, and Hyndman (2019) show that reconciled forecast may be improved by using the information on the variance covariance matrix of the reconciled forecast errors. They further strengthen this approach by providing theoretical justification and introduce a new forecast reconciliation method which they refer to as minimum trace (MinT) reconciliation. In this method they produce an optimal forecast reconciliation approach by minimising the mean squared error of the coherent forecasts across the entire collection of time series which are given by the trace of the variance covariance matrix of the reconciled forecast errors under the assumption of unbiasedness.

The focus of all the above methods was limited to a cross-sectional forecast reconciliation setting. Athanasopoulos et al. (2017) extends this reconciliation approach in the time dimension with the introduction of the Temporal Hierarchical Forecasting (THieF) approach. Temporal aggregations can be constructed for any time series by computing non-overlapping temporal aggregates. In this reconciliation approach, the forecasts produced at all aggregation levels are combined to produce temporally reconciled, accurate and robust forecasts. The strength of this concept is based on combining information and borrowing strength from various levels of temporal aggregation of a time series, to generate forecasts. Apart from enabling aligned decision making in different planning horizons Athanasopoulos et al. (2017) show both in simulations and multiple empirical settings that the THieF approach results in improved forecast accuracy in all forecast horizons.

In literature there are only a limited number of attempts which focus on combining temporal and cross-sectional reconciliation. Kourentzes and Athanasopoulos (2019) combine these two concepts, namely the temporal hierarchical forecasting which align different planning horizons and cross-sectional hierarchical forecasting which align the forecast across the cross-sectional structure to produce forecast which are reconciled in both dimensions. This provides greater transparency as forecasts will align in one direction when different viewpoints within the organisation are considered. Apart from this transparency in decision making, Kourentzes and Athanasopoulos (2019) show that this method improves accuracy when forecasting Australian tourism demand. Highlighting the challenge of dimensionality that would result if the cross-temporal reconciliation is performed in one step, they propose an alternative two step procedure.

Another approach to produce cross-temporally reconciled forecast is presented in the work by Spiliotis et al. (2020) where they attempt to apply cross-sectional and temporal hierarchical forecast reconciliation sequentially. Further, they emphasise that multiple temporal aggregation enables to reduce model uncertainty and combining this with cross-sectional hierarchies result in substantial gains in forecast accuracy. However, the sequential nature of this approach does not guarantee coherent forecast across all dimensions.

2.4. Hierarchical forecasting methods for GDP forecasting

National accounting methodologies present three main disaggregation approaches in computing headline GDP. These are namely, expenditure, production, and income approaches. The expenditure approach is a demand side view which uses the national accounting identity that production equals domestic expenditure made on final goods and services. The production approach is a supply oriented decomposition of value added by economic activities. The income approach measures GDP as the sum of factor income flows (European Commission, International Monetary Fund, Organisation for Economic Co-operation and Development, the United Nations, and World Bank (2009)).

In the context of GDP forecasting, the direct approach is dominant in empirical literature. Modelling and forecasting disaggregated subcomponents of GDP based on either the demand side or production side is limited in recent literature. The focus has been on the debate on whether direct GDP forecasting or bottom-up GDP forecasting produce better results. An early contribution in this topic is the study by Fair and Shiller (1990) which compares direct and bottom-up GDP forecasting for the USA. They use a VAR model to forecast aggregated Gross National Product (GNP) directly. Then they use Autoregressive Component (AC) models separately to forecast each of the disaggregated component of GNP and sum up the forecasts based on the GNP identity to produce the final GNP forecast. They find that the disaggregated AC model improves forecasting accuracy compared to the direct approach. Hahn and Skudelny (2008) extend the bottom-up approach to the production side to derive forecasts for Euro area real GDP growth but do not provide a comparison with the direct approach. Barhoumi et al. (2012) produce forecasts for GDP growth in France by aggregating component forecasts from both the supply and demand sides using bridge equation models. They emphasise that disaggregated forecasting produces more background information to build up the story around the forecasts. Moreover, GDP growth seems to be more precisely forecasted using the supply side approach. Heinisch and Scheufele (2018) compare bottom-up and direct GDP forecasting for Germany using an indicator based approach and conclude that the direct approach outperforms the bottom-up approach. Furthermore, the comparison of the performance of the production side disaggregated forecasting to the demand side revealed that using the production approach generates more accurate forecasts. Esteves (2013) studies the question of direct or bottom-up approaches for GDP forecasting using a different perspective. He emphasises that the choice of the approach is not dependent on the forecast performance but the level of analysis that forecasters wish to perform and on their expertise. In particular, the institution that focuses on short term forecasts will opt for a bottom-up approach as they must be able to explain the reasons behind the forecasts and identify current developments to help build the medium term forecasts.

In forecast reconciliation literature, only research that attempts to employ reconciliation methods in the context of GDP forecasting is that of Athanasopoulos et al. (2020). They focus on the application of cross-sectional forecast reconciliation using both income and the expenditure approach national accounting identities for Australian GDP. The study concludes that forecast reconciliation produces coherent forecasts and improves the overall forecast accuracy compared to a bottom-up approach when simple ARIMA models are used to derive the base forecasts.

This review of existing literature in this area indicates that to the best of my knowledge that no study has explored the application of cross-temporal reconciliation in the context of GDP forecasting. When it comes to GDP forecasting, coherent forecasts are of utmost importance to align policy direction. To achieve this objective, coherence in temporal dimension as well as cross-sectional dimension is important. Temporal coherence will ensure that short term policy direction is aligned with long term policy direction. Cross-sectional coherence will enable to identify economic activities which contributed to the forecasts. Therefore, it is valuable to investigate the application of cross-temporal forecast reconciliation for GDP forecasting, and this research aims to address this gap in literature.

3. Methodology

3.1. Hierarchical time series

Following a notation similar to Kourentzes and Athanasopoulos (2019), let $\mathbf{y}$ be an n- dimensional vector containing observations of the complete hierarchical structure and $\mathbf{b}$ be an m-dimensional vector of the most disaggregated times series which is often referred to as the bottom-level time series. We can write the aggregation constraints in any hierarchy as,

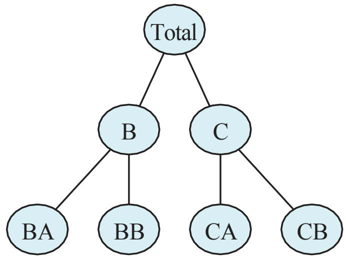
$$\mathbf{y} = \mathbf{S}\mathbf{b} \qquad (1)$$

where $\mathbf{S}$ is the summing matrix of order $n \times m$ which contains the linear aggregation constraints in the hierarchical structure in terms of bottom-level series.

For example, consider a simple two-level hierarchical time series either in the cross-sectional or temporal dimension which is represented in Figure 1. Level 0 is the most aggregated level, level 1 is the first level of disaggregation, and level 2 is the most disaggregated time series.

Let $\boldsymbol{y_{Total}}$ denote the observation of the most aggregated level 0 and $\boldsymbol{y_i}$ the observation corresponding to the node $\mathbf{i}$ of the levels below the top-level. The aggregation constraints

Figure 1: Two-level hierarchical structure



for this hierarchy in terms of the most disaggregated bottom-level time series can be represented by,

$$\begin{aligned} y_{Total} &= y_B + y_C \\ &= y_{BA} + y_{BB} + y_{CA} + y_{CB} \\ y_B &= y_{BA} + y_{BB} \\ y_C &= y_{CA} + y_{CB} \end{aligned} \qquad (2)$$

For this example, n which is the total number of nodes in the hierarchy is 7 and m which is the number of bottom-level series is 4. $\boldsymbol{y} = [y_{Total}, y_B, y_C, y_{BA}, y_{BB}, y_{CA}, y_{CB}]'$ and $\boldsymbol{b} = [\, y_{BA}, y_{BB}, y_{CA}, y_{CB}]'$ and the summing matrix is given by,

$$\boldsymbol{S} = \begin{bmatrix} \mathbf{1} & \mathbf{1} & \mathbf{1} & \mathbf{1} \\ \mathbf{1} & \mathbf{1} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{1} & \mathbf{0} \\ \square & \boldsymbol{I_4} & \square & \square \end{bmatrix}$$

where $\mathbf{I}_4$ is 4 x 4 identity matrix. Each aggregation constraint is represented by a row in the summation Matrix **S**. Thus, the same notation can be applied to represent any complex hierarchical structure.

3.2. Forecast reconciliation

The first step in forecast reconciliation is to generate h-steps ahead base forecasts for the complete hierarchy. Any forecasting method can be used to produce these forecasts, even multivariate models. However, these forecasts almost certainly will not be coherent. In other words, these will not follow hierarchical aggregation constraints other than in the case where a simple model such as naïve is used to generate base forecasts.

Let $\hat{\boldsymbol{y}}_{\boldsymbol{h}}$ be the h-step ahead base forecasts stacked in the same order as data **y**. Then, linear reconciliation methods can be written as,

$$\hat{y}_h = SG\hat{y}_h \qquad (3)$$

An appropriately selected matrix **G** of order m × n linearly maps base forecasts $\hat{\boldsymbol{y}}_{\boldsymbol{h}}$ to bottom-level forecasts. Then **S** sums these up to a set of reconciled forecasts $\hat{\boldsymbol{y}}_{\boldsymbol{h}}$ which are coherent. Thus, **SG** is referred to as the reconciliation matrix.

In traditional methods **G** only uses information from a single level from base forecasts which is a major drawback as highlighted earlier. For example, in the bottom-up approach $\boldsymbol{G} = [\boldsymbol{0}_{\boldsymbol{m}\times(\boldsymbol{n}-\boldsymbol{m})}|\boldsymbol{I}_{\boldsymbol{m}}]$ where $\boldsymbol{0}_{\boldsymbol{m}\times(\boldsymbol{n}-\boldsymbol{m})}$ is a null matrix of order m × (n − m) and $\boldsymbol{I}_{\boldsymbol{m}}$ is an identity matrix of order m × m. Thus, **G** only extracts bottom-level base forecasts from $\hat{\boldsymbol{y}}_{\boldsymbol{h}}$ and then these are summed by **S** to return the bottom-up coherent forecasts for the entire hierarchy.

Hyndman et al. (2011) show that if the base forecasts are unbiased the reconciled forecasts will preserve that unbiasedness if **SGS=S**. This holds for the bottom-up but not for top-down approaches. Therefore, this study will only focus on the bottom-up method for comparison. The identification of appropriate **G** which uses information from all levels within the hierarchy and which is also unbiased is important for the better performance of the forecast reconciliation method.

3.3. Optimal MinT reconciliation

Wickramasuriya, Athanasopoulos, and Hyndman (2019) frame the problem of finding appropriate **G** as an optimisation problem. They show that the variance covariance matrix of the h-step-ahead coherent forecast errors is given by,

$$V_h = var[y - \tilde{y}_h] = SGW_hG'S' \qquad (4)$$

Where $\boldsymbol{W}_{\boldsymbol{h}} = E[\hat{\boldsymbol{e}}_{\boldsymbol{h}}\hat{\boldsymbol{e}}_{\boldsymbol{h}}{}']$ is a positive definite covariance matrix of the base forecast's errors $\hat{\boldsymbol{e}}_{\boldsymbol{h}} = \boldsymbol{y} - \hat{\boldsymbol{y}}_{\boldsymbol{h}}$. Then the error variances of the coherent forecast are on the diagonal of the matrix $\mathbf{V}_\mathrm{h}$. Hence, the sum of all the error variances is given by the trace of this matrix. Wickramasuriya, Athanasopoulos, and Hyndman (2019) shows that the form of the matrix **G** that minimises the trace of $\mathbf{V}_\mathrm{h}$ subject to **SGS=S** is given by,

$$G = (S'W_hS)^{-1}S'^{W_h^{-1}} \qquad (5)$$

This would give the best (minimum variance) linear unbiased reconciled forecasts and is referred to as MinT (minimum trace) reconciliation. Substituting **G** into Equation 3, reconciled forecasts from the MinT approach are given by,

$$\tilde{y}_h = S(S'W_hS)^{-1}S'^{W_h^{-1}}\,\hat{y}_h \qquad (6)$$

The MinT approach has the ability of incorporating the full correlation structure of the hierarchy through $\boldsymbol{W}_{\boldsymbol{h}}$. However, the challenge in this approach is to estimate $\boldsymbol{W}_{\boldsymbol{h}}$ which is the variance covariance matrix of the base forecast which is of the dimension n × n. Thus, several alternative estimators for $\boldsymbol{W}_{\boldsymbol{h}}$ are used in literature.

3.4. OLS reconciliation

Set $\boldsymbol{W_h} = \boldsymbol{k_h I_n}$ where $\boldsymbol{k_h} > 0$ is a proportionality constant and $\boldsymbol{I_n}$ is n × n identity matrix. This will reduce the form of the MinT estimator to the OLS estimator proposed by Hyndman et al. (2011). This simplified assumption has performed well in practice (Hyndman et al., 2011; Athanasopoulos, Ahmed, and Hyndman, 2009). In this approach **G** only depends on **S**. Thus, this method can be used with forecasts generated from any forecasting method, such as judgmental forecasting. However, even though this is easy to apply, it ignores the correlations across series and the scale differences between the levels of the hierarchy due to aggregation.

3.5. Variance scaling

Set $\boldsymbol{W_h} = \boldsymbol{k_h diag(\widehat{W}_1)}$ for all h where $\boldsymbol{k_h} > 0$ and $\boldsymbol{\widehat{W}_1} = \boldsymbol{\frac{1}{T}\sum_{t=1}^{T} \hat{e}_t \hat{e}_t'}$ where $\boldsymbol{\hat{e}_t}$ is the in-sample one-step ahead forecast errors of the base forecasts. This is referred to as a weighted least squares (WLS) estimator as it scales the base forecasts using the variance of in-sample residuals. This will account for heterogeneity within aggregation levels as well as across aggregation levels.

3.6. Structural scaling

Athanasopoulos et al. (2017) proposed to set $\boldsymbol{W_h} = \boldsymbol{k_h \Lambda}$ for all h where $k_h > 0$, $\boldsymbol{\Lambda} = \mathbf{diag(S1)}$ where **1** is a unit vector of dimension n. This is specifically applicable in the context of temporal hierarchies as it assumes that each of the bottom-level base forecasts has equal error variance $\boldsymbol{k_h}$ and are uncorrelated. In this approach error variances of the higher levels are taken as the sum of error variances that contributing to that aggregation level. As the weight scheme only depends on the aggregation structure, this is referred to as structural scaling. In contrast to the OLS approach this only assumes equal forecast error variances at the bottom level of the structure and not across all levels. Furthermore, as this does not require an estimate of variances of forecast errors it can be used with forecasts generated from any forecasting method, such as judgmental forecasting where sample residuals may not be available.

3.7. Sample covariance estimate for MinT

Set $\boldsymbol{W_h} = \boldsymbol{k_h \widehat{W}_1}$ for all h where $\boldsymbol{k_h} > 0$. This assumes $\boldsymbol{W_h}$ to be proportional to unrestricted sample covariance estimator for h=1. This is relatively simple to obtain and provides a good estimate for small hierarchies. However, when the number of bottom-level series (m) is larger compared to the length of the series T, this will not provide reliable results (Wickramasuriya, Athanasopoulos, and Hyndman (2019); Athanasopoulos et al. (2020)).

3.8. Shrinkage covariance estimator for MinT

Set $\boldsymbol{W_h} = k_h \boldsymbol{\widehat{W}^*_{1D}}$ for all h where $\boldsymbol{k_h} > 0$ and $\boldsymbol{\widehat{W}^*_{1D}} = \boldsymbol{\lambda \widehat{W}_{1D} + (1 - \lambda)\widehat{W}_1}$. This estimator shrinks the sample covariances to the diagonal target matrix $\boldsymbol{\widehat{W}^*_{1D}}$ which comprises of the diagonal elements of $\boldsymbol{\widehat{W}_1}$. Thus, off diagonal elements of $\boldsymbol{\widehat{W}_1}$ are shrunk towards zero. As proposed by Schäfer and Strimmer (2005) the shrinkage intensity parameter **λ** is set to,

$$\hat{\lambda} = \frac{\sum_{i \neq j} Var(\hat{r}_{ij})}{\sum_{i \neq j} \hat{r}_{ij}} \qquad (7)$$

where $\boldsymbol{\hat{r}_{ij}}$ is the **ij**th element of $\boldsymbol{\widehat{R}_1}$, one-step ahead sample correlation matrix.

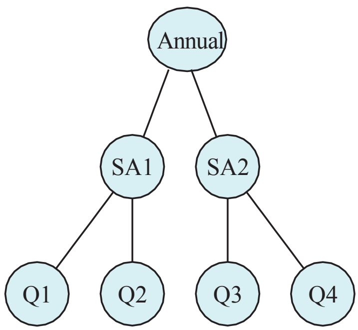
3.9. Cross-sectional forecast reconciliation

A cross-sectional hierarchy can be defined as a collection of time series that follows an aggregation constraint as shown in Figure 1. For example, consider a case where several geographical regions add up to give the total number for the whole country. In this setting the time series within each level and across each level represent different entities. Thus, we must account for heterogeneity within the levels and across the levels. Therefore, when estimating $\boldsymbol{W_h}$ more suitable estimators would be Variance scaling and Shrinkage MinT.

3.10. Temporal forecast reconciliation

The concept of temporal forecast reconciliation was introduced by Athanasopoulos et al. (2017). A temporal hierarchy can be developed for any time series by creating non overlapping temporal aggregates which do not introduce non-integer seasonality. If m is the highest frequency observed per year of a series, then each of the temporal aggregates created should be a factor of m. For example, if a series is observed in quarterly frequency then a temporal hierarchy can be constructed as shown in Figure 2, where the bottom level comprises of four quarterly observations (Q1, Q2, Q3, Q4) which adds up to the two semi-annual series (SA1, SA2) in the intermediate level which adds up to the total annual at the top level.

Figure 2: Temporal hierarchy for quarterly data



In contrast to cross-sectional forecast reconciliation the forecast horizon at each aggregation level will differ and it will depend on the specific aggregation level. For example, if we consider 4 quarters ahead forecasts, then the forecast horizon will be 4 when we consider the quarterly series, while it will be 1 and 2 for annual and semi-annual frequencies, respectively. In general, if h^* is the maximum required forecast horizon at the most disaggregated level and m is the highest frequency observed per year, then we would require $h = \lceil h^*/m \rceil$ forecasts at the most aggregated level. Then for each aggregation level k, we must generate $M_k h$ step ahead forecasts conditional on $\lfloor T/k \rfloor$ observations, where M_k is the number of observations per year for the k^{th} aggregation level and T is the length of the time series based on the highest frequency.

In this setting as forecasts for each level are created by one series, it is reasonable to assume homogeneous forecast errors within each level. Therefore, when estimating $\boldsymbol{W_h}$ assumptions behind the structural scaling estimator are justifiable in this situation.

3.11. Cross-temporal forecast reconciliation

In order to construct a cross-temporal hierarchy, a cross-sectional hierarchy needs to be combined with a respective temporal hierarchy. To illustrate this let us consider the simple cross-sectional hierarchy with one levels shown in Figure 3, where the two series B and C add up to the total and the temporal hierarchy for quarterly data shown in Figure 2.

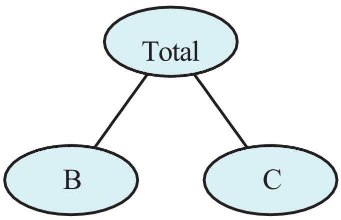
To develop a cross-temporal hierarchy we must consider the temporal aggregation at each of the cross-sectional nodes as shown in Figure 4. In this cross-temporal hierarchy, there are m = 8 bottom-level series, which comprise of four quarterly series at each of the two cross-sectional nodes. Further, with seven temporal aggregates at each cross-sectional node and with three cross-sectional nodes, there are $n = 7 \times 3 = 21$ nodes in the total cross-temporal hierarchy.

To create the cross-temporal Summation matrix (**S**) we need to combine the cross-sectional summation matrix ($\boldsymbol{S_C}$) and the temporal summation matrix ($\boldsymbol{S_T}$). In this regard, each of the elements in cross-sectional $\boldsymbol{S_C}$ need to be replaced with temporal $\boldsymbol{S_T}$.

Mathematically this is given by the Kronecker product of $\boldsymbol{S_C}$ with $\boldsymbol{S_T}$;

$$S = S_C \otimes S_T \qquad (8)$$

Figure 3: Simple cross-sectional hierarchy



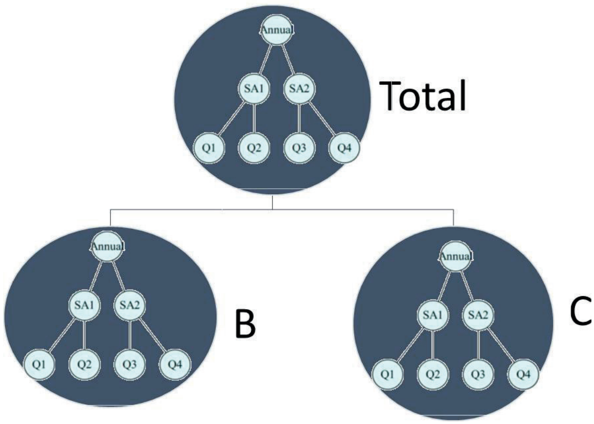
For example, the cross-sectional summation matrix ($\boldsymbol{S_C}$) corresponding to the hierarchy in Figure 3 in terms of the two bottom-level series B and C is,

$$S_C = \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}_{3\times 2}$$

The temporal summation matrix for the temporal hierarchy in Figure 2 in terms of the four quarterly observations in the bottom-level is given by,

$$S_T = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ \square & I_4 & \square & \square \end{bmatrix}_{7\times 4}$$

Figure 4: A cross-temporal hierarchy with quarterly data



Thus, the corresponding cross-temporal summation matrix for the cross-temporal hierarchy in Figure 4 would be,

$$\mathbf{S} = \mathbf{S_C} \otimes \mathbf{S_T}$$

$$= \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ \multicolumn{4}{c}{I_4} & \multicolumn{4}{c}{I_4} \\ 1 & 1 & 1 & 1 & & & & \\ 1 & 1 & 0 & 0 & \multicolumn{4}{c}{0_{7\times4}} \\ 0 & 0 & 1 & 1 & & & & \\ \multicolumn{4}{c}{I_4} & & & & \\ & & & & 1 & 1 & 1 & 1 \\ \multicolumn{4}{c}{0_{7\times4}} & 1 & 1 & 0 & 0 \\ & & & & 0 & 0 & 1 & 1 \\ & & & & \multicolumn{4}{c}{I_4} \end{bmatrix}_{21\times8}$$

If we stack all the series in the cross-temporal hierarchy in vector **y** and all the bottom-level series in vector **b** then aggregation constraints in any cross-temporal hierarchy can be also represented by Equation 1. In the case of the above example. We let,

$$\boldsymbol{y} = [y_{Total_A}, y_{Total_{SA1}}, y_{Total_{SA2}}, y_{Total_{Q1}}, y_{Total_{Q2}}, y_{Total_{Q3}}, y_{Total_{Q4}}, y_{B_A}, y_{B_{SA1}}, y_{B_{SA2}},$$
$$y_{B_{Q1}}, y_{B_{Q2}}, y_{B_{Q3}}, y_{B_{Q4}}, y_{C_A}, y_{C_{SA1}}, y_{C_{SA2}}, y_{C_{Q1}}, y_{C_{Q2}}, y_{C_{Q3}}, y_{C_{Q4}}]'_{21\times1}$$
$$\boldsymbol{b} = [y_{B_{Q1}}, y_{B_{Q2}}, y_{B_{Q3}}, y_{B_{Q4}}, y_{C_{Q1}}, y_{C_{Q2}}, y_{C_{Q3}}, y_{C_{Q4}}]'_{8\times1}$$

Using the reconciliation matrix **SG** with **G** specified in Equation 5, optimal MinT reconciliation for the cross-temporal hierarchies can be computed directly using the same procedure as explained in Section 3.3. However, estimating $\boldsymbol{W_h}$ will be more challenging compared to considering cross-sectional and temporal dimensions separately as its dimension will become very large very quickly.

OLS reconciliation and structural scaling estimates of $\boldsymbol{W_h}$ can be directly applied to cross- temporal hierarchy with the developed $\mathbf{S}$ matrix as it does not require an estimate of forecast error variance. If we consider the cross-temporal hierarchy given in Figure 4 as it has 21 nodes, OLS reconciliation estimator of $\boldsymbol{W_h = k_h I_{21}}$, where I_{21} is 21 × 21 identity matrix. This is referred to as OLS in the results that follow. The structural scaling estimator for $\boldsymbol{W_h}$ with the assumption that equal forecast error variance at the bottom-level series is given by,

$$W_h = k_h diag(8, 4, 4, 2, 2, 2, 2, 4, 2, 2, 1, 1, 1, 1, 4, 2, 2, 1, 1, 1, 1)$$

The 8 at the top of the diagonal matrix represents that 8 bottom-level series are used to construct the top-level annual series. This is referred to as Struc in results to follow. Even though the assumptions behind these estimators are highly restrictive they are the only estimators that are applicable when in-sample forecast error variances are not available (e.g. with judgmental forecasts).

The variance scaling estimator of $\boldsymbol{W_h}$ for the cross-temporal hierarchies can be computed in a similar way as explained in Section 3.5 with in-sample residuals of the base forecasts stacked in the same way as the data. For example, the resulting estimator for the cross-temporal hierarchy in Figure 4 is given by,

$$\boldsymbol{W_h} = k_h diag(\sigma^2_{Total_A}, \sigma^2_{Total_{SA1}}, \sigma^2_{Total_{SA2}}, \sigma^2_{Total_{Q1}}, \sigma^2_{Total_{Q2}}, \sigma^2_{Total_{Q3}}, \sigma^2_{Total_{Q4}}, \sigma^2_{B_A}, \sigma^2_{B_{SA1}}, \sigma^2_{B_{SA2}}, \sigma^2_{B_{Q1}}, \sigma^2_{B_{Q2}}, \sigma^2_{B_{Q3}}, \sigma^2_{B_{Q4}}, \sigma^2_{C_A}, \sigma^2_{C_{SA1}}, \sigma^2_{C_{SA2}}, \sigma^2_{C_{Q1}}, \sigma^2_{C_{Q2}}, \sigma^2_{C_{Q3}}, \sigma^2_{C_{Q4}})$$

where σ_i^2 is the estimated variance of the in-sample residuals corresponding to each time series. This is referred to as VAR in the results that follow. The variance scaling estimator using the diagonal of the sample covariance matrix requires fewer error variances to be estimated as compared to sample covariance estimate for MinT. However, the sample available to estimate each variance is limited to $\lfloor T/m \rfloor$. This will create stability problems with time series with limited history. Therefore, an alternative variance scaling estimator was also considered similar to the series variance scaling estimator introduced by Athanasopoulos et al. (2017). This assumes a common variance within the same temporal aggregation level in each of the cross-sectional nodes. This assumption is not unreasonable as the base forecast errors within the same aggregation level are for the same series in that particular frequency (i.e., semi-annual or quarterly). For example, the resulting estimator $\boldsymbol{W_h}$ for the cross-temporal hierarchy of the Figure 4 is given by,

$$\boldsymbol{W_h} = k_h diag(\sigma^2_{Total_A}, \sigma^2_{Total_{SA}}, \sigma^2_{Total_{SA}}, \sigma^2_{Total_Q}, \sigma^2_{Total_Q}, \sigma^2_{Total_Q}, \sigma^2_{Total_Q}, \sigma^2_{B_A}, \sigma^2_{B_{SA}}, \sigma^2_{B_{SA}}, \sigma^2_{B_Q}, \sigma^2_{B_Q}, \sigma^2_{B_Q}, \sigma^2_{B_Q}, \sigma^2_{C_A}, \sigma^2_{C_{SA}}, \sigma^2_{C_{SA}}, \sigma^2_{C_Q}, \sigma^2_{C_Q}, \sigma^2_{C_Q}, \sigma^2_{C_Q})$$

i.e., four quarterly forecast error variances for each year for each series will be replaced by one common quarterly forecast error variance, and two semi-annual forecast error variances for each year for each series will be replaced by one common semi-annual forecast error variance. This is referred to as SVAR in the results that follow. The shrinkage MinT estimator for the cross-temporal hierarchy can be computed as explained in Section 3.8. For example, the diagonal target matrix $\boldsymbol{\widehat{W}_{1D}}$ which comprises of diagonal elements of in-sample one-step ahead forecasts residual matrix $\boldsymbol{\widehat{W}_1}$ for the cross-temporal hierarchy of Figure 4 is given by,

$$\widehat{\boldsymbol{W}}_{\boldsymbol{1D}} = diag(\sigma^2_{Total_A}, \sigma^2_{Total_{SA1}}, \sigma^2_{Total_{SA2}}, \sigma^2_{Total_{Q1}}, \sigma^2_{Total_{Q2}}, \sigma^2_{Total_{Q3}}, \sigma^2_{Total_{Q4}}, \sigma^2_{B_A}, \sigma^2_{B_{SA1}},$$
$$\sigma^2_{B_{SA2}}, \sigma^2_{B_{Q1}}, \sigma^2_{B_{Q2}}, \sigma^2_{B_{Q3}}, \sigma^2_{B_{Q4}}, \sigma^2_{C_A}, \sigma^2_{C_{SA1}}, \sigma^2_{C_{SA2}}, \sigma^2_{C_{Q1}}, \sigma^2_{C_{Q2}}, \sigma^2_{C_{Q3}}, \sigma^2_{C_{Q4}})$$

The shrinkage intensity parameter λ was estimated using the method proposed by Schäfer and Strimmer (2005) which is implemented in the SHIP package (Jelizarow and Guillemot (2015)) for R (R Core Team (2020)). This is referred to as Shrk in the results that follow.

The sample covariance estimator for MinT was not considered for cross-temporal hierarchy. Even though it is straight forward to apply, estimates are highly unstable with the increasing dimensionality.

4. Results and discussion

GDP is the total value of goods and services produced within the boundaries of a country in a particular period. The System of National Accounts (SNA) (European Commission, International Monetary Fund, Organisation for Economic Co-operation and Development, United Nations, and World Bank (2009)) presents an internationally agreed standard set of recommendations on how to compile measures of economic activity including GDP. As defined in this framework, "GDP is derived from the concept of value added. Gross value added (GVA) is the difference between output and intermediate consumption. GDP is the sum of gross value added of all resident producer units plus that part (possibly the total) of taxes on products, less subsidies on products, that is not included in the valuation of output." Based on this methodology there are three approaches of computing GDP, which are: the Production approach, Income approach, and Expenditure approach. These approaches compute GDP as an aggregate of various economic variables. This forms a natural cross-sectional hierarchy. Thus, using a hierarchical approach to forecasting will enable us to improve forecasting accuracy, preserve coherency of the forecasts as well as provide aligned information on the contributors of the forecasts generated.

4.1. Sri Lankan GDP

Sri Lankan National Accounts are currently compiled by the Department of Census and Statistics (DCS) in compliance with guidelines given in SNA 2008 using 2010 as the base year. This case study uses the production approach of GDP by economic activity at constant prices from 2010-Q1 to 2019-Q4. I restrict my attention to the Production approach which is also known as the Output approach as it presents the supply-side decomposition of value added by economic activities. It allows the tracking of the overall performance of the whole economy. This approach provides data for the analysis of the productivity of each economic activity and changes in the structure of the economy. Furthermore, it allows policy makers to analyse the performance of specific economic activities against the industry averages (Viet (2009))

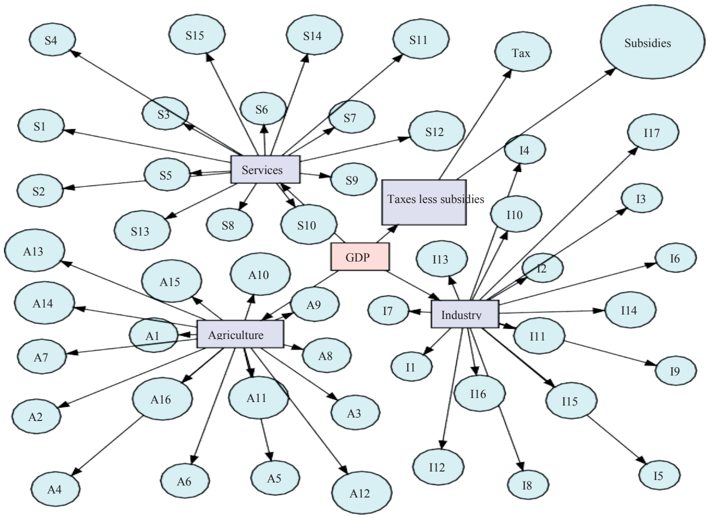
GDP is defined by the production approach as the sum of the GVA at basic prices of all resident producers plus taxes on products payable less subsidies on products receivable (European Commission, International Monetary Fund, Organisation for Economic Co-operation and Development, United Nations, and World Bank (2009)).

$$\textbf{GDP} = \textbf{GVA at basic prices} + \textbf{all taxes on products} - \textbf{all subsidies on products} \qquad (9)$$

The GVA is an aggregated value added based on value added generated by economic activities which are classified according to Sri Lanka Standard Industry Classification based on International Standard Industry classification - Rev.4.

The most detailed dissemination table provides 48 economic activities which are categorised into 3 main streams: 16 activities related to Agriculture, forestry and fishing activities, 17 activities related to Industry activities and 15 activities relating to Services activities.

Figure 5: Hierarchical structure of the income approach for GDP



Note: The Pink cell contains GDP which the most aggregated series purple cells contain intermediate-level series and blue cells contain bottom-level series.

Figure 5 shows the full hierarchical structure capturing all components aggregated to form GDP using the production approach. This hierarchy has three levels. The most aggregated top-level of the hierarchy, which is level 0, comprises of the GDP. Level 1 comprises of GVA generated by three main activities and the component tax less subsidies. The bottom level has 50 series. Thus, in total this hierarchy has n = 55 series. These are summarised in Table 1.

Figure 6 shows some of the time series in the production approach. The top panel shows the most aggregated time series which is the total GDP as well as, Level 1 series namely: Agriculture, Industry and Services activities along with the component of taxes less subsidies on products (TaxLessSubsi). The bottom panel shows some selected series in the most disaggregated bottom level. Each series shows diverse dynamics with some series showing prominent seasonality while others simply showing a trend. This highlights the need to account for the different dynamics observed to produce a better model for forecasting each series.

Table 1: Number of time series per level of hierarchy

Hierarchy	Number of series
Level 0 (top-level)	1
Level 1	4
Level 4(bottom-level)	50
Total	55

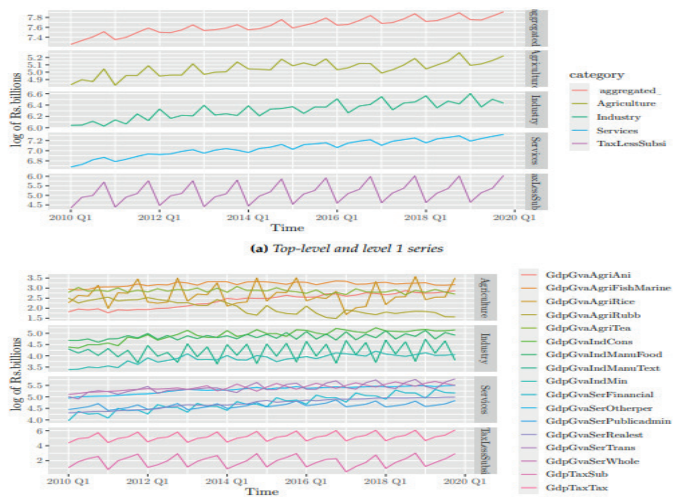
Figure 7 plots the same hierarchy as in Figure 6 but now in the annual frequency. As expected, series are now much smoother with a prominent trend, as seasonality is filtered out. Therefore, different temporal aggregation levels capture different features of the times series. Thus, these features could be extracted to improve forecast accuracy with temporal reconciliation. The cross-sectional reconciliation will enable to extract diverse dynamics of each of the series within the hierarchy. Moreover, using cross-temporal reconciliation will enable to extract these diverse signals from both cross-sectional and temporal dimensions to improve overall forecast accuracy.

4.2. Empirical application methodology

The data are quarterly from 2010-Q1 to 2019-Q4. As an only limited history is available, the last 8 quarters (2 years) will be considered as the test set to evaluate the forecast accuracy of competing approaches and to identify the potential of cross-temporal reconciliation to improve forecast accuracy.

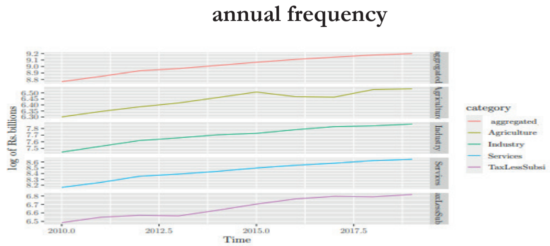
The cross-temporal structure is not currently supported in an R package. Thus, I expand on the base implementations of cross-sectional hierarchical structure facilitated in the fpp3 package (Hyndman, Athanasopoulos, and O'Hara-Wild (2020)) for R (R Core Team (2020)). The code developed for this can be shared if requested.

Figure 6: Time plots for series from different levels of production approach hierarchy in quarterly frequency

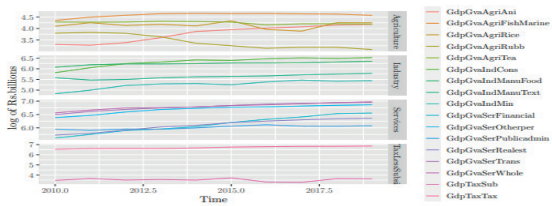


(a) *Top-level and level 1 series*

(b) *Selected bottom-level series in each main activity*



(a) *Top-level and level 1 series*



(b) *Selected bottom-level series in each main activity*

4.3. Forecasting models

The first step in forecast reconciliation is to obtain base forecasts for all the series in the hierarchy. The cross-sectional aggregation structure comprises 55 series and with 3 temporal aggregation levels. Thus, the cross-temporal hierarchy has $55 \times 3 = 165$ different series. To develop base forecasts for each of these I consider, two classes of forecasting models namely ExponenTial Smoothing (ETS) and AutoRegressive Integrated Moving Average (ARIMA) models as implemented in the ARIMA and ETS functions in the fable package (Hyndman, Athanasopoulos, and O'Hara-Wild (2020)) for R (R Core Team (2020)). The appropriate ETS and ARIMA models are chosen by minimising the Akaike Information Criterion (AIC) corrected for small sample sizes.

ETS models are commonly used in empirical research as they perform well with limited data and are relatively simple to build (Kourentzes and Athanasopoulos (2019)). ETS captures time series as the total of four fundamental components of a time series which are level, trend, seasonality, and the error process, where these components are combined additively or multiplicatively. Forecasts produced using exponential smoothing methods are weighted averages of past observations, with the weights decaying exponentially as the observations get older. ARIMA models aim to describe the autocorrelations in the data as opposed to ETS models which are based on a description of the trend and seasonality in the data. The Autoregressive component of the ARIMA model captures the habitual elements in the time series by regressing the variable of interest using a linear combination of past values of the variable after the series is difference as required to make it stationary. The moving average component regresses the variable of interest using a linear combination of past forecast errors of the stationary time series to smooth out the inherent noise in the data (Hyndman and Athanasopoulos (2018)).

For this application ETS forecasts were on average more accurate than the ARIMA forecasts (Refer AppendixA.2 TableA.4 to TableA.6) and using ARIMA models had minimal impact on conclusion of this study. Thus, I will only present the results obtained using ETS models. The results obtained for ARIMA models are given in AppendixA.3 TableA.7 to TableA.9.

Apart from these univariate models, other sophisticated multivariate models such as VAR models or indicator based regression type models can also be used for specific series to generate these base forecasts as its completely flexible and independent of the reconciliation methodology which is an advantage of forecast reconciliation.

The base forecasts do not adhere to the aggregation constraints in the cross-temporal hierarchy, and they also do not consider information available in other temporal or cross- sectional aggregation levels. Hence, cross-temporal coherent forecasts are generated reconciling the base forecasts as per the reconciliation Equation 3. The cross-temporal summation matrix was compiled according to the process explained in Section 3.11. The cross-sectional GDP hierarchy as summarised in Table1 has n = 55 series in total with m = 50 bottom-level series. Thus, the cross-sectional summation matrix is of order 55 × 50. As the series are observed in quarterly frequency the corresponding temporal summation matrix will be a matrix of order 7 × 4. Therefore, the corresponding cross- temporal summation matrix complied by taking the kronecker product of cross-sectional and temporal summation matrices will be a large matrix of order 385 × 200 with m = 200 bottom-level series.

The first set of cross-temporally coherent forecasts were generated using the bottom-up method which only use the information from the bottom-level of the hierarchy. This is referred to as BU in the results to follow and provides the natural benchmark to assess the benefit of generating forecasts at all aggregation levels (Athanasopoulos et al. (2017)). Three sets of alternative reconciled forecasts were also generated using OLS reconciliation (OLS), Structural scaling (Struc) and the Series Variance scaling (SVAR). The Variance scaling estimator and Shrinkage covariance estimator were also used but due to limited length of the series, forecasts error variances estimated for certain series were close to zero and it created problems in using these approaches. Reconciled forecasts were also computed using only cross-sectional reconciliation to compare the accuracy gain of using cross-temporal reconciliation.

4.4. Forecast accuracy evaluation

The forecast accuracy was evaluated using Mean Absolute Percentage Error (MAPE), Root Mean Square Error (RMSE) and Mean Absolute Error (MAE). Several accuracy measures are considered in this empirical application to calibrate the results and to evaluate whether forecast performance depends on the accuracy measure used. For a particular series $\mathbf{j}$ in a particular aggregation level, for h-steps ahead forecast:

$$RMSE_j = \sqrt{\frac{1}{h}\sum_{i=1}^{h}(y_{ij} - \hat{y}_{ij})^2} \qquad (9)$$

$$MAE_j = \frac{1}{h}\sum_{i=1}^{h}\left|y_{ij} - \hat{y}_{ij}\right| \qquad (10)$$

$$MAPE_j = \frac{1}{h}\sum_{i=1}^{h}\left|\frac{100(y_{ij} - \hat{y}_{ij})}{y_{ij}}\right| \qquad (11)$$

where $\boldsymbol{y_{ij}}$ and $\widehat{\boldsymbol{y}}_{\boldsymbol{ij}}$ are actual and forecast values for the series $\mathbf{j}$ in the period $\mathbf{i}$. RMSE and MAE are the most commonly used accuracy measures, but they have the disadvantage of being scale dependent. However, they are useful in evaluating different methods applied to the same data set. MAPE has an advantage of being independent of scale and frequently used to compare the forecast accuracy of different data sets (Hyndman and Athanasopoulos (2018)). There are also certain issues in MAPE such as being unidentified or infinite if $\boldsymbol{y_{ij}}$ is zero or close to zero, assuming a meaningful zero and imposing a heavier penalty on positive errors than on negative errors (Hyndman and Koehler (2006)). However, in this application $\boldsymbol{y_{ij}}$ has a meaningful zero and it is not close to zero. Further, over estimation of growth may be more harmful than under estimation so imposing heavier penalty on positive errors can be justifiable. The summary accuracy measures in the tables that follow are the arithmetic mean of these accuracy measures calculated for each of the time series in the dimension considered.

It is common in the forecasting literature to express the accuracy measures in terms of a skill score (Wheatcroft (2019)), which is defined as,

$$skill\ score = \frac{A_f - A_r}{A_p - A_r} \qquad (12)$$

where A_p is the value of the accuracy measure if the outcome is known perfectly and A_f and A_r are the values of the accuracy measure using the method of interest and reference method, respectively. A_p is zero for the forecast accuracy measures considered in this application and incoherent base forecasts are taken as the reference forecasting method. Skill score can be interpreted as the proportional increase in accuracy of the forecasting method of interest compared to base forecasts. Thus, if the skill score is positive, it represents an improvement in forecasting accuracy over the base forecasts while negative values represent a deterioration. The summary measures in the tables that follow are the skill scores calculated based on arithmetic mean of the accuracy measures in the dimensions considered.

4.5. Results

Table 2 summarises the skill scores calculated based on average MAPE of the all cross- sectional series in the temporal dimension considered, where MAPE was computed based on forecasts up to and including the forecast horizon h. The results are presented for the complete hierarchy, bottom-level series, and top-level series (i.e., GDP) separately. Furthermore, results are presented for each temporal aggregation level (i.e., annual, semi-annual, and quarterly) separately together with an average measure across all temporal aggregation levels. The incoherent base forecasts were taken as the reference method. The Table 2 summaries the resulting skill scores of coherent forecasts obtained from the classical method bottom-up and the reconciliation methods. It should be noted that for cross- sectional reconciliation VAR referred to Variance scaling (Refer Section 3.5) and in cross- temporal reconciliation SVAR refers to Series variance scaling (Refer Section 3.11). The measures are summarised for cross-sectional and cross-temporal reconciliation separately to evaluate the accuracy gains of using cross-temporal reconciliation. The colored cells show the best performing method in each row (i.e., the temporal aggregation level). The darker the shade, the higher the improvement across the temporal aggregation levels. Skill scores calculated based on MAE and RMSE are given in Appendix A.1 Table A.2 and Table A.3. The conclusion based on these measures was also similar to that of MAPE.

The first panel refers to results summarised over all series, the second panel refers to top-level GDP series, and the last panel refers to the bottom-level. Reported figures are skill scores computed based on average MAPE over the entire test set of h=1 to 8. A positive (negative) entry shows a relative improvement (loss) over the base incoherent forecasts. Skill scores are summarized for cross-sectional and cross-temporal reconciliation separately to evaluate the accuracy gains of using cross-temporal reconciliation. The coloured cells show the best performing method in each row (i.e. the temporal aggregation level). The darker the shade, the higher the improvement across the temporal aggregation levels. Skill scores calculated based on MAE and RMSE are given in Appendix A.1 TableA.2 and Table A.3. The conclusion based on these measures was also similar to that of MAPE.

Table 2: Skill scores for point forecasts from alternative methods (with reference to incoherent base forecasts) using MAPE for Sri Lankan production approach

All-levels									
		Cross-Sectional			Cross-Temporal				
Temporal level	h	BU	OLS	Struc	VAR	BU	OLS	Struc	SVAR
Annual	2	0.00	-0.42	-0.11	-0.01	0.13	-0.24	0.08	**0.17**
Semi-annual	4	0.02	-1.86	-1.24	0.02	0.11	-0.44	-0.13	**0.16**
Quarterly	8	0.01	-2.00	-0.99	0.00	0.01	-0.72	-0.37	**0.04**
Average		0.01	-1.47	-0.80	0.00	0.08	-0.48	-0.15	**0.12**
Top-level									
		Cross-Sectional			Cross-Temporal				
Temporal level	h	BU	OLS	Struc	VAR	BU	OLS	Struc	SVAR
Annual	2	-0.05	0.00	-0.02	-0.04	0.88	0.57	0.76	**0.94**
Semi-annual	4	0.72	-0.02	0.37	0.52	0.75	0.18	0.54	**0.89**
Quarterly	8	-0.40	0.01	0.16	0.17	-0.40	-1.55	-0.54	**0.28**
Average		0.16	-0.01	0.12	0.15	0.71	0.23	0.56	**0.86**
Bottom-level									
		Cross-Sectional			Cross-Temporal				
Temporal level	h	BU	OLS	Struc	VAR	BU	OLS	Struc	SVAR
Annual	2	0.00	-0.44	-0.12	-0.01	0.10	-0.27	0.05	**0.15**
Semi-annual	4	0.00	-1.95	-1.30	0.01	0.09	-0.47	-0.15	**0.14**
Quarterly	8	0.00	-2.06	-1.03	-0.01	0.00	-0.74	-0.39	**0.03**
Average		0.00	-1.53	-0.85	0.00	0.06	-0.51	-0.18	**0.11**

First, we compare cross-sectional reconciliation with cross-temporal reconciliation. It can be clearly seen from the Table 2, that in general using cross-temporal reconciliation has improved the forecast accuracy in all the cross- sectional and temporal levels considered irrespective of the reconciliation method. It shows that extracting and sharing information from different temporal aggregation levels to supplement the signals extracted from the cross-sectional hierarchy improves the forecast accuracy of all the reconciliation methods considered. As shown in Figure 5 and Figure 6 the seasonal component of the series dominates at quarterly frequency, possibly concealing the trend when it comes to model selection and estimation. At the annual level trends become dominant but estimation efficiency will be low due to limited sample size. Therefore, using temporal aggregation with cross-sectional aggregation will extract seasonal information and estimation efficiency to annual level and extract the trend information from annual level to the quarterly level. Thus, cross-temporal reconciliation gives a better view of the data in different angles which allows to bring in more information and improve the overall forecast accuracy.

The strength of cross-temporal reconciliation is not limited to accuracy gains. Another gain is the cross-temporally coherent forecasts which align the decision making and provide transparency within the organization. The short term view will align with the long term view while the disaggregated activity level forecasts will align with the country level GDP forecasts. This will facilitate consistent, transparent and align policy implementation.

It is interesting to note that, although all cross-temporal reconciliation alternatives perform better than cross-sectional reconciliation, cross-temporal SVAR reconciliation forecasts are consistently the most accurate in every scenario considered. Further, all the skill scores of SVAR forecasts are positive indicating that these are more accurate compared to the incoherent base forecasts, which is taken as the reference method. In general, OLS and Struc have failed to perform better than the conventional bottom-up method and even worse than the incoherent base forecasts for this application. However, it is noteworthy to highlight that accuracy gains are positive for OLS and Struc if performance is evaluated based on MAE and RMSE (refer AppendixA.1 TableA.2 and TableA.3). According to Table 2 the accuracy of OLS and Struc based on MAPE are worse at the bottom-level series, this was not evident in the skill scores based on RMSE and MAE. Therefore, there is some indication that OLS and Struc are performing relatively poorly at some low base series in the bottom level which result in overall loss in accuracy when evaluated based on MAPE.

The results are also disaggregated to top-level and bottom-level of the hierarchy for further investigation. These results are presented in the 2nd and the 3rd panels of Table 2. The accuracy gains of the top-level are higher than the bottom level. This indicates that additional information received at the top-level from the bottom level is arguably more influential than the additional information received at the bottom level.

5. Conclusion

This study investigates a direct cross-temporal hierarchical forecasting approach specifically in a macroeconomic setting relating to the forecasting of GDP. The main aim was to produce a set of forecasts which are cross-temporally coherent so that it will facilitate aligned policy decisions directed towards a one number forecast. This study proposes a direct approach in combining cross-sectional reconciliation and temporal reconciliation to get the maximum information available in the hierarchical structure as an alternative method to the two-step approach introduced by Kourentzes and Athanasopoulos (2019).

The results of the empirical applications revealed that cross-temporal reconciliation can further improve the forecasting accuracy obtained through cross-sectional reconciliation. This can be attributable to the valuable information provided by temporal hierarchies within the cross-sectional structure. As highlighted by Athanasopoulos et al. (2017) the source of forecast improvement in using temporal hierarchies is that it can strengthen the signal to noise ratio and reduced outlier effect at the aggregated lower frequencies of the time series, while mitigating loss of information and estimation efficiency at higher frequencies. Accuracy gains are greater for the top-level single series GDP compared to the bottom-level series. In addition, gains received at the lower frequencies are higher than the gains received at the higher frequencies.

Evaluation of alternative reconciliation methods revealed that cross-temporal SVAR, which is the series variance scaling reconciliation method yields the highest improvement in forecast accuracy in forecasting the Sri Lankan GDP.

Cross-temporal reconciliation aligns decisions within an organization towards one number. Temporal reconciliation aligns short term forecasts with more strategic long term forecasts. Cross-sectional reconciliation will align the view of the decision makers at different levels within the hierarchy. This is possible as reconciliation methods are model free, so judgmental forecasts produced at strategic levels can also be combined with data driven forecasts at more operational bottom-level in a transparent data driven method. It should be highlighted that this feature is not available with the classical bottom-up method. Furthermore, this will facilitate the alignment of the overall policy direction. This is very important specifically in GDP forecasting as policy decisions should be taken to direct the country towards one direction. To achieve this objective, short-term forecasts should align with long term forecasts. In addition, forecasts of the disaggregated economic activities should also align with the overall GDP forecasts.

In developing cross-sectional forecasts within the GDP hierarchy, reconciliation methods provide the benefit of using different models for different scenarios as the concept is independent of models used. This gives the opportunity to combine different specialised or in other words judgmental forecasts for certain economic activities with data driven sophisticated forecasting models. This is an advantage as for some disaggregated economic activities and at lower frequencies, availability of data or indicator variables will be limited to develop multivariate models. This ability to reconcile different views in a transparent method to enhance efficiency in managerial decision making is the main outcome of this cross-temporal reconciliation approach. In addition, the concept of forecast reconciliation involves forecasting GDP through disaggregated economic activities. This has an additional benefit over direct GDP forecasting which is commonly used in GDP forecasting literature as it has the ability of identifying economic activities which contributed to the overall projected GDP growth. Thus, policymakers can identify any issues at the bottom-levels and design specialised policies to address them.

References

Athanasopoulos, G, RA Ahmed, and RJ Hyndman (2009). "Hierarchical forecasts for Australian domestic tourism." *International Journal of Forecasting,* vol. **25** no.1, pp.146–166.

Athanasopoulos, G, P Gamakumara, A Panagiotelis, RJ Hyndman, and M Affan (2020). "Hierarchical forecasting". *Macroeconomic Forecasting in the Era of Big Data*, edited by Peter Fuleky. Springer, pp.689–719.

Athanasopoulos, G, RJ Hyndman, N Kourentzes, and F Petropoulos (2017). "Forecasting with temporal hierarchies." *European Journal of Operational Research.* Vol. **262**, no.1, pp. 60–74.

Australian Bureau of Statistics (2012). *Australian System of National Accounts: Concepts, Sources and Methods*. Vol. Cat 5216.0.

Barhoumi, K, O Darné, L Ferrara, and B Pluyaud (2012). "Monthly GDP forecasting using bridge models: Application for the French economy." *Bulletin of Economic Research,* vol. **64**, s53–s70.

Dunn, D, W Williams, and T DeChaine (1976). "Aggregate versus subaggregate models in local area forecasting." *Journal of the American Statistical Association,* vol. **71** no.353, pp. 68–71.

Esteves, PS (2013). "Direct vs bottom–up approach when forecasting GDP: Reconciling literature results with institutional practice." *Economic Modelling,* vol. **33**, pp. 416–420.

European Commission, International Monetary Fund, Organisation for Economic Co- operation and Development, United Nations, and World Bank (2009). *System of national accounts 2008*.

Fair, RC and RJ Shiller (1990). "Comparing information in forecasts from econometric models." *The American Economic Review*, pp. 375–389.

Fliedner, G (1999). "An investigation of aggregate variable time series forecast strategies with specific sub aggregate time series statistical correlation." *Computers & operations research,* vol. **26,** no.10-11, pp. 1133–1149.

Gross, CW and JE Sohl (1990). "Disaggregation methods to expedite product line forecasting." *Journal of forecasting,* vol. **9,** no.3, pp. 233–254.

Grunfeld, Y and Z Griliches (1960). "Is aggregation necessarily bad? ", *The Review of Economics and Statistics*, pp. 1–13.

Hahn, E and F Skudelny (2008). "Early estimates of euro area real GDP growth: a bottom up approach from the production side." European Central Bank.

Heinisch, K and R Scheufele (2018). "Bottom-up or direct? Forecasting German GDP in a data-rich environment." *Empirical Economics,* Vol. **54**, no.2, pp.705–745.

Higgins, PC (2014). "GDPNow: A Model for GDP'Nowcasting." Federal Reserve Bank of Atlanta.

Hyndman, R, G Athanasopoulos, and M O'Hara-Wild (2020). *fpp3:Forecasting: Principles and Practice(3rd Edition).* R Package version 0.3.

Hyndman, RJ, RA Ahmed, G Athanasopoulos, and HL Shang (2011). "Optimal combination forecasts for hierarchical time series. "*Computational statistics & data analysis,* vol. **55**, no.9, pp. 2579– 2589.

Hyndman, RJ and G Athanasopoulos (2018). *Forecasting: principles and practice*. OTexts.

Hyndman, RJ and AB Koehler (2006). "Another look at measures of forecast accuracy." *International journal of forecasting,* vol. **22**, no.4, pp. 679–688.

Hyndman, RJ, AJ Lee, and E Wang (2016). "Fast computation of reconciled forecasts for hierarchical and grouped time series." *Computational statistics & data analysis,* vol. **97**, pp. 16–32.

Jelizarow, M and V Guillemot (2015). *SHIP:SHrinkage covariance Incorporating Prior knowledge*. R Package version 1.0.2.

Kourentzes, N and G Athanasopoulos (2019). :"Cross-temporal coherent forecasts for Australian tourism." *Annals of Tourism Research,* vol.**75**, pp.393–409.

Kourentzes, N, D Barrow, and F Petropoulos (2019). "Another look at forecast selection and combination: Evidence from forecast pooling." *International Journal of Production Economics,* vol. **209**, pp.226–235.

Park, M and M Nassar (2014). "Variational Bayesian inference for forecasting hierarchical time series." In: *ICML Workshop*. Citeseer.

Pennings, CL and J van Dalen (2017). "Integrated hierarchical forecasting." *European Journal of Operational Research,* vol. **263**, no.2, pp. 412–418.

R Core Team (2020). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria.

Schäfer, J and K Strimmer (2005). "A shrinkage approach to large-scale covariance matrix estimation and implications for functional genomics." *Statistical applications in genetics and molecular biology, vol.* **4**, no.1.

Spiliotis, E, F Petropoulos, N Kourentzes, and V Assimakopoulos (2020). "Cross-temporal aggregation: Improving the forecast accuracy of hierarchical electricity consumption." *Applied Energy*, vol. **261**, no. 114339.

Syntetos, AA, Z Babai, JE Boylan, S Kolassa, and K Nikolopoulos (2016). "Supply chain forecasting: Theory, practice, their gap and the future." *European Journal of Operational Research,* vol. **252**, no.1, pp. 1–26.

Viet, VQ (2009). "GDP by production approach: A general introduction with emphasis on an integrated economic data collection framework." *Published as United Nations' training material for Statistical Capacity Development in China and Other Developing Countries in Asia*, pp. 1–137.

Weatherford, LR, SE Kimes, and DA Scott (2001). "Forecasting for hotel revenue management: Testing aggregation against disaggregation." *Cornell hotel and restaurant administration quarterly,* vol.**42**, no.4, pp. 53–64.

Wheatcroft, E (2019). Interpreting the skill score form of forecast performance metrics. *International Journal of Forecasting,* vol. **35**, no.2, pp.573–579.

Wickramasuriya, SL, G Athanasopoulos, and RJ Hyndman (2019). Optimal forecast recon- ciliation for hierarchical and grouped time series through trace minimization. *Journal of the American Statistical Association,* vol. **114**, no.526, pp. 804–819.

Zotteri, G and M Kalchschmidt (2007). "A model for selecting the appropriate level of aggregation in forecasting processes." *International Journal of Production Economics,* vol. 108, no. 1- 2, pp.74–83.

Zotteri, G, M Kalchschmidt, and F Caniato (2005). The impact of aggregation level on forecasting performance. *International Journal of Production Economics,* vol. **93**, pp.479–491.

Appendices

Table A.1: Detailed Economic activities in Production Approach

Variables	Detailed economic activities	Main Activity
GdpGvaAgriCereal	Growing of Cereals (except rice)	Agriculture, Forestry and Fishing
GdpGvaAgriRice	Growing of Rice	Agriculture, Forestry and Fishing
GdpGvaAgriVege	Growing of Vegetables	Agriculture, Forestry and Fishing
GdpGvaAgriSugar	Growing of Sugar cane, tobacco and other non-perennial crops	Agriculture, Forestry and Fishing
GdpGvaAgrFruits	Growing of fruits	Agriculture, Forestry and Fishing
GdpGvaAgriOle	Growing of Oleaginous Fruits (Coconut, king coconut, Oil palm)	Agriculture, Forestry and Fishing
GdpGvaAgriTea	Growing of Tea (Green leaves)	Agriculture, Forestry and Fishing
GdpGvaAgriBeve	Growing of other beverage crops (Coffee, Cocoa	Agriculture, Forestry and Fishing
GdpGvaAgriSpice	Growing of spices, aromatic, drug and pharmaceutical crops	Agriculture, Forestry and Fishing
GdpGvaAgriRubb	Growing of rubber	Agriculture, Forestry and Fishing
GdpGvaAgriPere	Growing of other perennial crops	Agriculture, Forestry and Fishing
GdpGvaAgriAni	Animal Production	Agriculture, Forestry and Fishing
GdpGvaAgriPlant	Plant propagation and agricultural supporting activities	Agriculture, Forestry and Fishing
GdpGvaAgriForest	Forestry and Logging	Agriculture, Forestry and Fishing
GdpGvaAgriFishMarine	Marine fishing and Marine Aquaculture	Agriculture, Forestry and Fishing
GdpGvaAgriFishInland	Fresh water fishing and Fresh water Aquaculture	Agriculture, Forestry and Fishing
GdpGvaIndMin	Mining and quarrying	Industry
GdpGvaIndManuFood	Manufacture of food, beverages and Tobacco products	Industry
GdpGvaIndManuText	Manufacture of textiles, wearing apparel and leather related products	Industry
GdpGvaIndManuWood	Manufacture of wood and of products of wood and cork, except furniture	Industry
GdpGvaIndManuPaper	Manufacture of paper products, printing and reproduction of media products	Industry
GdpGvaIndManuCoke	Manufacture of coke and refined petroleum products	Industry
GdpGvaIndManuChemi	Manufacture of chemical products and basic pharmaceutical products	Industry
GdpGvaIndManuRubb	Manufacture of rubber and plastic products	Industry
GdpGvaIndManuNonmet	Manufacture of other non- metallic mineral products	Industry
GdpGvaIndManuMetal	Manufacture of basic metals and fabricated metal products	Industry
GdpGvaIndManuMachin	Manufacture of machinery and equipment	Industry
GdpGvaIndManuFurni	Manufacture of furniture	Industry
GdpGvaIndManuOther	Other manufacturing, and Repair and installation of machinery and equipment	Industry
GdpGvaIndElectri	Electricity, gas, steam and air conditioning supply	Industry
GdpGvaIndWater	Water collection, treatment and supply	Industry
GdpGvaIndSewerage	Sewerage, Waste, treatment and disposal activities	Industry
GdpGvaIndCons	Construction	Industry
GdpGvaSerWhole	Wholesale and retail trade	Services
GdpGvaSerTrans	Transportation of goods and passengers including Warehousing	Services
GdpGvaSerPostal	Postal and courier activities	Services
GdpGvaSerAccom	Accommodation, Food and beverage service activities	Services
GdpGvaSerProgram	Programming and broadcasting activities and audio video productions	Services
GdpGvaSerTele	Telecommunication	Services
GdpGvaSerIT	IT programming consultancy and related activities	Services
GdpGvaSerFinancial	Financial Service activities and auxiliary financial services	Services
GdpGvaSerInsurance	Insurance, reinsurance and pension funding	Services
GdpGvaSerRealest	Real estate activities, Including Ownership of dwelling	Services
GdpGvaSerProfess	Professional services	Services
GdpGvaSerPublicadmin	Public administration and defense; compulsory social security	Services
GdpGvaSerEdu	Education	Services
GdpGvaSerHealth	Human health activities, Residential care and social work activities	Services
GdpGvaSerOtherper	Other personal service activities	Services

Table A.2: ***Skill scores for point forecasts from alternative methods (with reference to incoherent base forecasts) using RMSE for Sri Lankan production approach GDP***

All-levels

		Cross-Sectional				Cross-Temporal			
Temporal level	h	BU	OLS	Struc	VAR	BU	OLS	Struc	SVAR
Annual	2	-0.03	0.01	0.00	-0.02	0.53	0.46	0.57	**0.61**
Semi-annual	4	0.34	-0.08	0.16	0.19	0.36	0.21	0.38	**0.45**
Quarterly	8	-0.02	-0.08	0.04	0.13	-0.02	-0.15	0.07	**0.15**
Average		0.07	-0.02	0.05	0.05	0.43	0.34	0.47	**0.52**

Top-level

		Cross-Sectional				Cross-Temporal			
Temporal level	h	BU	OLS	Struc	VAR	BU	OLS	Struc	SVAR
Annual	2	-0.05	0.00	-0.02	-0.04	0.86	0.57	0.77	**0.94**
Semi-annual	4	0.74	0.00	0.38	0.52	0.73	0.25	0.59	**0.88**
Quarterly	8	-0.55	0.01	0.08	0.08	-0.55	-1.33	-0.46	**0.11**
Average		0.10	0.00	0.07	0.09	0.78	0.44	0.68	**0.89**

Bottom-level

		Cross-Sectional				Cross-Temporal			
Temporal level	h	BU	OLS	Struc	VAR	BU	OLS	Struc	SVAR
Annual	2	0.00	0.02	0.02	0.00	0.12	**0.31**	0.30	0.20
Semi-annual	4	0.00	-0.22	-0.09	-0.10	0.02	0.11	**0.14**	0.09
Quarterly	8	0.00	-0.16	-0.08	0.06	0.00	0.04	**0.09**	**0.09**
Average		0.00	-0.07	-0.03	-0.02	0.07	0.21	**0.22**	0.15

The first panel refers to results summarised over all series, the second panel refers to top-level GDP series, and the last panel refers to the bottom level. Reported figures are skill scores computed based on average MAPE over the entire test set of h=1 to 8. A positive (negative) entry shows a relative improvement (loss) over the base incoherent forecasts.

Table A.3: *Skill scores for point forecasts from alternative methods (with reference to incoherent base forecasts) using MAE for Sri Lankan production approach GDP*

All-levels									
		Cross-Sectional				Cross-Temporal			
Temporal level	h	BU	OLS	Struc	VAR	BU	OLS	Struc	SVAR
Annual	2	-0.03	0.01	0.00	-0.02	0.55	0.46	0.58	**0.62**
Semi-annual	4	0.34	-0.07	0.21	0.18	0.37	0.20	0.38	**0.45**
Quarterly	8	0.03	-0.10	0.13	0.04	0.03	-0.16	0.09	**0.17**
Average		0.07	-0.02	0.06	0.04	0.46	0.34	0.48	**0.54**
Top-level									
		Cross-Sectional				Cross-Temporal			
Temporal level	h	BU	OLS	Struc	VAR	BU	OLS	Struc	SVAR
Annual	2	-0.05	0.00	-0.02	-0.04	0.88	0.57	0.76	**0.94**
Semi-annual	4	0.73	-0.01	0.37	0.52	0.76	0.18	0.55	**0.89**
Quarterly	8	-0.38	0.01	0.17	0.18	-0.38	-1.56	-0.55	**0.29**
Average		-0.62	-0.80	-0.67	-0.65	0.72	0.24	0.56	**0.86**
Bottom-level									
		Cross-Sectional				Cross-Temporal			
Temporal level	h	BU	OLS	Struc	VAR	BU	OLS	Struc	SVAR
Annual	2	0.00	0.03	0.02	0.01	0.13	**0.32**	0.31	0.21
Semi-annual	4	0.00	-0.21	-0.08	-0.07	0.01	**0.16**	**0.16**	0.10
Quarterly	8	0.00	-0.19	-0.10	0.04	0.00	0.06	**0.12**	0.09
Average		0.00	-0.07	-0.02	-0.01	0.08	0.24	**0.24**	0.16

The first panel refers to results summarised over all series, the second panel refers to top-level GDP series, and the last panel refers to the bottom level. Reported figures are skill scores computed based on average MAPE over the entire test set of h=1 to 8. A positive (negative) entry shows a relative improvement (loss) over the base incoherent forecasts.

Table A.4: *Average MAPE for Sri Lankan production approach GDP hierarchy*

					All-levels					
ETS										
			Cross-Sectional				Cross-Temporal			
Temporal level	h	Base	BU	OLS	Struc	VAR	BU	OLS	Struc	SVAR
Annual	2	6.89	6.86	9.78	7.63	6.95	6.00	8.52	6.33	5.69
Semi-annual	4	7.64	7.51	21.87	17.08	7.50	6.82	10.97	8.61	6.42
Quarterly	8	7.83	7.77	23.52	15.61	7.83	7.77	13.50	10.73	7.51
Average		7.45	7.38	18.39	13.44	7.42	6.86	11.00	8.56	**6.54**
ARIMA										
			Cross-Sectional				Cross-Temporal			
Temporal level	h	Base	BU	OLS	Struc	VAR	BU	OLS	Struc	SVAR
Annual	2	9.95	10.14	20.49	17.64	9.54	6.25	15.18	11.08	6.20
Semi-annual	4	7.94	7.80	13.57	10.47	8.07	6.85	16.78	12.13	6.91
Quarterly	8	7.89	7.72	18.85	14.94	7.93	7.72	17.65	13.06	7.82
Average		8.59	8.56	17.64	14.35	8.51	**6.94**	16.54	12.09	6.98
					Top-level					
ETS										
			Cross-Sectional				Cross-Temporal			
Temporal level	h	Base	BU	OLS	Struc	VAR	BU	OLS	Struc	SVAR
Annual	2	4.28	4.49	4.29	4.35	4.46	0.51	1.85	1.03	0.25
Semi-annual	4	2.25	0.62	2.29	1.42	1.09	0.57	1.86	1.03	0.25
Quarterly	8	0.73	1.02	0.73	0.62	0.61	1.02	1.87	1.12	0.53
Average		2.42	2.05	2.44	2.13	2.05	0.70	1.86	1.06	**0.34**
ARIMA										
			Cross-Sectional				Cross-Temporal			
Temporal level	h	Base	BU	OLS	Struc	VAR	BU	OLS	Struc	SVAR
Annual	2	1.88	5.90	1.95	3.24	3.19	0.61	0.15	0.24	0.71
Semi-annual	4	2.14	1.32	2.09	1.80	1.67	0.61	0.31	0.35	0.70
Quarterly	8	2.62	0.80	2.55	1.91	1.75	0.80	0.71	0.67	0.82
Average		2.21	2.67	2.20	2.32	2.20	0.67	**0.39**	0.42	0.75
					Bottom-levels					
ETS										
			Cross-Sectional				Cross-Temporal			
Temporal level	h	Base	BU	OLS	Struc	VAR	BU	OLS	Struc	SVAR
Annual	2	7.19	7.19	10.36	8.02	7.27	6.48	9.15	6.82	6.13
Semi-annual	4	8.06	8.06	23.75	18.58	8.02	7.36	11.82	9.31	6.91
Quarterly	8	8.37	8.37	25.64	16.99	8.43	8.37	14.58	11.62	8.08
Average		7.87	7.87	19.92	14.53	7.90	7.40	11.85	9.25	**7.04**
ARIMA										
			Cross-Sectional				Cross-Temporal			
Temporal level	h	Base	BU	OLS	Struc	VAR	BU	OLS	Struc	SVAR
Annual	2	10.57	10.57	22.19	19.10	10.13	6.75	16.48	12.05	6.64
Semi-annual	4	8.36	8.36	14.56	11.23	8.59	7.41	18.22	13.20	7.41
Quarterly	8	8.34	8.34	20.41	16.18	8.48	8.34	19.16	14.20	8.39
Average		9.09	9.09	19.05	15.50	9.07	7.50	17.95	13.15	**7.48**

The 1st and 2nd panels refers to results summarised over all series for ETS and ARIMA models respectively, the 3rd and 4th panel refers to results summarised over the top-level GDP series for ETS and ARIMA models respectively, and the 5th and 6th panels refers to the bottom level results summary for ETS and ARIMA models respectively.Reported figures are average MAPE of the all cross-sectional series in the level considered, where MAPE was computed based on forecasts up to and including the forecast horizon h.Bold figures are the lowest error in the panel.

Table A.5: *Average RMSE for Sri Lankan production approach GDP hierarchy*

Temporal level	h	Base	Cross-Sectional				Cross-Temporal			
			BU	OLS	Struc	VAR	BU	OLS	Struc	SVAR
All-levels										
ETS										
Annual	2	25.62	26.29	25.42	25.66	26.15	12.02	13.73	10.90	9.87
Semi-annual	4	10.37	6.85	11.15	8.67	8.35	6.68	8.17	6.43	5.75
Quarterly	8	4.17	4.25	4.48	4.01	3.64	4.25	4.78	3.88	3.55
Average		13.39	12.46	13.69	12.78	12.71	7.65	8.89	7.07	**6.39**
ARIMA										
Annual	2	22.62	33.42	22.33	22.74	24.34	10.17	13.15	10.23	11.54
Semi-annual	4	10.44	8.60	10.63	9.70	9.64	5.76	7.29	5.83	6.43
Quarterly	8	5.21	3.65	5.45	4.67	4.74	3.65	4.35	3.65	3.91
Average		12.76	15.23	12.80	12.37	12.91	**6.53**	8.26	6.57	7.29
Top-level										
ETS										
Annual	2	434.14	454.23	434.81	440.83	450.63	59.39	184.76	101.34	27.63
Semi-annual	4	122.82	31.81	122.31	75.99	58.35	33.00	92.38	50.89	14.59
Quarterly	8	21.03	32.50	20.76	19.35	19.45	32.50	48.93	30.71	18.64
Average		192.66	172.85	192.62	178.72	176.14	41.63	108.69	60.98	**20.29**
ARIMA										
Annual	2	198.75	591.06	204.50	329.39	322.34	63.43	14.39	28.67	78.80
Semi-annual	4	118.06	72.14	115.69	99.93	92.43	31.94	16.47	19.17	41.14
Quarterly	8	71.52	25.07	70.02	53.62	49.53	25.07	19.59	19.63	26.88
Average		129.45	229.42	130.07	160.98	154.77	40.15	**16.82**	22.49	48.94
Bottom-levels										
ETS										
Annual	2	10.67	10.67	10.44	10.48	10.64	9.38	7.40	7.48	8.51
Semi-annual	4	5.27	5.27	6.41	5.76	5.79	5.16	4.67	4.52	4.78
Quarterly	8	2.95	2.95	3.44	3.19	2.78	2.95	2.84	2.70	2.69
Average		6.30	6.30	6.76	6.48	6.40	5.83	4.97	**4.90**	5.33
ARIMA										
Annual	2	13.08	13.08	13.01	11.78	13.12	7.78	9.77	8.43	8.36
Semi-annual	4	6.10	6.10	6.36	6.18	6.35	4.45	5.37	4.73	4.73
Quarterly	8	2.65	2.65	2.97	2.80	2.98	2.65	3.04	2.74	2.79
Average		7.28	7.28	7.45	6.92	7.48	**4.96**	6.06	5.30	5.29

The 1st and 2nd panels refers to results summarised over all series for ETS and ARIMA models respectively, the 3rd and 4th panel refers to results summarised over the top-level GDP series for ETS and ARIMA models respectively, and the 5th and 6th panels refers to the bottom level results summary for ETS and ARIMA models respectively.Reported figures are average RMSE of the all cross-sectional series in the level considered, where RMSE was computed based on forecasts up to and including the forecast horizon h.Bold figures are the lowest error in the panel.

Table A.6: *Average MAE for Sri Lankan production approach GDP hierarchy*

All-levels										
ETS										
			Cross-Sectional				Cross-Temporal			
Temporal level	h	Base	BU	OLS	Struc	VAR	BU	OLS	Struc	SVAR
Annual	2	24.68	25.34	24.43	24.68	25.17	11.04	13.25	10.42	9.39
Semi-annual	4	9.31	6.12	9.98	7.39	7.64	5.87	7.41	5.82	5.08
Quarterly	8	3.56	3.45	3.90	3.09	3.42	3.45	4.13	3.25	2.95
Average		12.52	11.64	12.77	11.72	12.07	6.78	8.26	6.50	**5.81**
ARIMA										
			Cross-Sectional				Cross-Temporal			
Temporal level	h	Base	BU	OLS	Struc	VAR	BU	OLS	Struc	SVAR
Annual	2	21.63	32.48	21.31	21.83	23.57	9.47	12.75	9.65	10.90
Semi-annual	4	9.41	7.65	9.62	8.72	8.66	4.99	6.70	5.18	5.75
Quarterly	8	4.36	2.94	4.65	3.90	3.92	2.94	3.68	3.03	3.23
Average		11.80	14.36	11.86	11.48	12.05	**5.80**	7.71	5.95	6.63
Top-level										
ETS										
			Cross-Sectional				Cross-Temporal			
Temporal level	h	Base	BU	OLS	Struc	VAR	BU	OLS	Struc	SVAR
Annual	2	419.87	440.61	420.56	426.79	436.90	50.31	181.44	100.79	24.67
Semi-annual	4	111.23	29.59	112.71	69.86	53.49	26.95	90.72	50.39	12.33
Quarterly	8	17.69	24.34	17.60	14.75	14.55	24.34	45.36	27.36	12.59
Average		101.87	164.85	183.62	170.46	168.31	28.79	77.76	44.43	**14.24**
ARIMA										
			Cross-Sectional				Cross-Temporal			
Temporal level	h	Base	BU	OLS	Struc	VAR	BU	OLS	Struc	SVAR
Annual	2	184.68	578.53	191.31	317.83	313.19	59.93	14.36	23.49	70.22
Semi-annual	4	105.67	65.16	103.24	89.20	82.66	29.97	14.80	17.23	35.11
Quarterly	8	64.52	19.36	62.93	47.18	43.11	19.36	17.32	16.38	20.51
Average		118.29	221.01	119.16	151.40	146.32	36.42	**15.50**	19.03	41.95
Bottom-levels										
ETS										
			Cross-Sectional				Cross-Temporal			
Temporal level	h	Base	BU	OLS	Struc	VAR	BU	OLS	Struc	SVAR
Annual	2	10.25	10.25	9.99	10.04	10.17	8.90	7.02	7.12	8.10
Semi-annual	4	4.74	4.74	5.74	5.14	5.09	4.69	3.99	3.99	4.27
Quarterly	8	2.52	2.52	3.00	2.78	2.41	2.52	2.37	2.23	2.29
Average		5.84	5.84	6.24	5.98	5.89	5.37	4.46	**4.45**	4.89
ARIMA										
			Cross-Sectional				Cross-Temporal			
Temporal level	h	Base	BU	OLS	Struc	VAR	BU	OLS	Struc	SVAR
Annual	2	12.59	12.59	12.49	11.30	12.69	7.20	9.40	8.03	7.86
Semi-annual	4	5.59	5.59	5.87	5.69	5.84	3.86	4.88	4.29	4.26
Quarterly	8	2.19	2.19	2.53	2.36	2.49	2.19	2.58	2.30	2.34
Average		6.79	6.79	6.96	6.45	7.01	**4.42**	5.62	4.87	4.82

The 1st and 2nd panels refers to results summarised over all series for ETS and ARIMA models respectively, the 3rd and 4th panel refers to results summarised over the top-level GDP series for ETS and ARIMA models respectively, and the 5th and 6th panels refers to the bottom level results summary for ETS and ARIMA models respectively.Reported figures are average MAE of the all cross-sectional series in the level considered, where MAE was computed based on forecasts up to and including the forecast horizon h. Bold figures are the lowest error in the panel.

Table A.7: Skill scores for point forecasts from alternative methods (with reference to incoherent base forecasts) using MAPE for Sri Lankan production approach GDP with ARIMA models

		All-levels							
		Cross-Sectional				Cross-Temporal			
Temporal level	h	BU	OLS	Struc	VAR	BU	OLS	Struc	SVAR
Annual	2	-0.02	-1.06	-0.77	0.04	0.37	-0.53	-0.11	**0.38**
Semi-annual	4	0.02	-0.71	-0.32	-0.02	**0.14**	-1.11	-0.53	0.13
Quarterly	8	0.02	-1.39	-0.89	-0.01	**0.02**	-1.24	-0.66	0.01
Average		0.00	-1.05	-0.67	0.01	**0.19**	-0.92	-0.41	**0.19**
		Top-level							
		Cross-Sectional				Cross-Temporal			
Temporal level	h	BU	OLS	Struc	VAR	BU	OLS	Struc	SVAR
Annual	2	-2.14	-0.04	-0.72	-0.70	0.68	**0.92**	0.87	0.62
Semi-annual	4	0.38	0.02	0.16	0.22	0.71	**0.85**	0.84	0.67
Quarterly	8	0.69	0.03	0.27	0.33	0.69	**0.73**	**0.74**	0.69
Average		-0.21	0.01	-0.05	0.00	0.70	**0.82**	0.81	0.66
		Bottom-level							
		Cross-Sectional				Cross-Temporal			
Temporal level	h	BU	OLS	Struc	VAR	BU	OLS	Struc	SVAR
Annual	2	0.00	-1.10	-0.81	0.04	0.36	-0.56	-0.14	**0.37**
Semi-annual	4	0.00	-0.74	-0.34	-0.03	**0.11**	-1.18	-0.58	**0.11**
Quarterly	8	0.00	-1.45	-0.94	-0.02	**0.00**	-1.30	-0.70	-0.01
Average		0.00	-1.10	-0.71	0.00	0.17	-0.98	-0.45	**0.18**

The first panel refers to results summarised over all series, the second panel refers to top-level GDP series, and the last panel refers to the bottom level. Reported figures are skill scores computed based on average MAPE over the entire test set of h=1 to 8. A positive (negative) entry shows a relative improvement (loss) over the base incoherent forecasts.

Table A.8: Skill scores for point forecasts from alternative methods (with reference to incoherent base forecasts) using RMSE for Sri Lankan production approach GDP with ARIMA models

All-levels

		Cross-Sectional				Cross-Temporal			
Temporal level	h	BU	OLS	Struc	VAR	BU	OLS	Struc	SVAR
Annual	2	-0.48	0.01	-0.01	-0.08	**0.55**	0.42	**0.55**	0.49
Semi-annual	4	0.18	-0.02	0.07	0.08	**0.45**	0.30	**0.44**	0.38
Quarterly	8	0.30	-0.05	0.10	0.09	**0.30**	0.17	**0.30**	0.25
Average		-0.19	0.00	0.03	-0.01	**0.49**	0.35	**0.49**	0.43

Top-level

		Cross-Sectional				Cross-Temporal			
Temporal level	h	BU	OLS	Struc	VAR	BU	OLS	Struc	SVAR
Annual	2	-1.97	-0.03	-0.66	-0.62	0.68	**0.93**	0.86	0.60
Semi-annual	4	0.39	0.02	0.15	0.22	0.73	**0.86**	0.84	0.65
Quarterly	8	0.65	0.02	0.25	0.31	0.65	**0.73**	0.73	0.62
Average		-0.77	0.00	-0.24	-0.20	0.69	**0.87**	0.83	0.62

Bottom-level

		Cross-Sectional				Cross-Temporal			
Temporal level	h	BU	OLS	Struc	VAR	BU	OLS	Struc	SVAR
Annual	2	0.00	0.01	0.10	0.00	**0.41**	0.25	0.36	0.36
Semi-annual	4	0.00	-0.04	-0.01	-0.04	**0.27**	0.12	0.22	0.22
Quarterly	8	0.00	-0.12	-0.06	-0.12	**0.00**	-0.15	-0.03	-0.05
Average		0.00	-0.02	0.05	-0.03	**0.32**	0.17	0.27	0.27

The first panel refers to results summarised over all series, the second panel refers to top-level GDP series, and the last panel refers to the bottom level. Reported figures are skill scores computed based on average MAPE over the entire test set of h=1 to 8. A positive (negative) entry shows a relative improvement (loss) over the base incoherent forecasts.

Table A.9: Skill scores for point forecasts from alternative methods (with reference to incoherent base forecasts) using MAE for Sri Lankan production approach GDP with ARIMA models

		All-levels							
		Cross-Sectional				Cross-Temporal			
Temporal level	h	BU	OLS	Struc	VAR	BU	OLS	Struc	SVAR
Annual	2	-0.50	0.01	-0.01	-0.09	**0.56**	0.41	0.55	0.50
Semi-annual	4	0.19	-0.02	0.07	0.08	**0.47**	0.29	0.45	0.39
Quarterly	8	0.33	-0.07	0.11	0.10	**0.33**	0.16	0.31	0.26
Average		-0.22	-0.01	0.03	-0.02	**0.51**	0.35	0.50	0.44
		Top-level							
		Cross-Sectional				Cross-Temporal			
Temporal level	h	BU	OLS	Struc	VAR	BU	OLS	Struc	SVAR
Annual	2	-2.13	-0.04	-0.72	-0.70	0.68	**0.92**	0.87	0.62
Semi-annual	4	0.38	0.02	0.16	0.22	0.72	**0.86**	0.84	0.67
Quarterly	8	0.70	0.02	0.27	0.33	0.70	**0.73**	0.75	0.68
Average		-0.87	-0.01	-0.28	-0.24	0.69	**0.87**	0.84	0.65
		Bottom-level							
		Cross-Sectional				Cross-Temporal			
Temporal level	h	BU	OLS	Struc	VAR	BU	OLS	Struc	SVAR
Annual	2	0.00	0.01	0.10	-0.01	**0.43**	0.25	0.36	0.38
Semi-annual	4	0.00	-0.05	-0.02	-0.04	**0.31**	0.13	0.23	0.24
Quarterly	8	0.00	-0.16	-0.08	-0.14	**0.00**	-0.18	-0.05	-0.07
Average		0.00	-0.03	0.05	-0.03	**0.35**	0.17	0.28	0.29

The first panel refers to results summarised over all series, the second panel refers to top-level GDP series, and the last panel refers to the bottom level. Reported figures are skill scores computed based on average MAPE over the entire test set of h=1 to 8. A positive (negative) entry shows a relative improvement (loss) over the base incoherent forecasts.

Total Factor Productivity Losses Resulting from Capital and Labour Misallocation in Sri Lanka

Ranpati Dewage Thilini Sumudu Kumari[1]

Abstract

Aggregate productivity can largely be determined by how production factors are allocated across heterogeneous firms. Although, the existing literature documents the contribution of misallocation in capital and labour to the Aggregate Total factor Productivity (TFP), studies on the relative roles of labour and capital misallocation affecting productivity in a single economy are limited. Using annual firm-level survey data and a static model, this paper contributes to the literature by estimating the cross-firm misallocation of labour relative to capital and their impact on aggregate productivity loss for Sri Lanka during 1994-2017. The results suggest that relative to the counterfactual efficient allocation of capital and labour, misallocation of both capital and labour has been rising and entails sizable negative impacts on TFP. The misallocation of labour has steeply been rising and 2.4-fold of that for capital misallocation. The results further suggest that firms can hardly grow bigger due to firm size-dependent capital and labour misallocation. The results specifically suggest that labour misallocation can be a binding constraint for business expansion preventing Sri Lanka from moving to a more productive economy.

Key Words: Capital misallocation, Firm-level distortion, Labour misallocation, Sri Lanka, TFP

JEL Classification: D24, E24, O47, O53

[1] The author is currently serving as a Deputy Director of the Central Bank of Sri Lanka. Corresponding emails: thilini@cbsl.lk; thilini.ranpatidewage@research.uwa.edu.au; thilinisk@gmail.com. Author acknowledges the support from the Department of Census and Statistics of Sri Lanka. She appreciates the feedback from the participants at the 4th CBSL-ADBI-APAEA Macroeconomics Conference held in the Central Bank of Sri Lanka, Colombo in October 2023. The author also appreciates the anonymous referees for their important comments on the earlier version of this article. The views presented in the paper are those of the author and do not necessarily indicate the views of the Central Bank of Sri Lanka. The author bears the responsibility for any remaining errors in the paper.

1. Introduction

Significant differences in per capita income among countries are mainly attributed to the differences in Total Factor Productivity (TFP) across such countries (Restuccia and Rogerson, 2017). Low technology diffusion and misallocation of resources at most disaggregated levels lower the aggregate productivity (Hsieh and Klenow, 2009). Misallocation of resources is a situation where countries or production units are less efficient in allocating the factors of production to their best use consequent on the impairment of the allocation function due to market frictions, specifically instigated by policy distortions, inefficient institutions, and information asymmetry. Misallocation prevents economies from achieving the optimum level of output. Reallocating resources from less to more productive units increases the aggregate output.

In a pioneering study, Hsieh and Klenow (2009) examined the resource misallocation effect on aggregate TFP using manufacturing sector firm-level data in China, India and the US by adopting a monopolistic competition model with heterogeneous firms. Subsequently, many studies have followed the methodology adopted by them to measure the degree of resource misallocation in different countries and come up with varying magnitudes and different underlying sources of misallocation. Misallocation may originate from firm-specific distortions, either in output or in production factors (Restuccia and Rogerson, 2017). The studies on factor distortions mainly focus on misallocation of capital[2] and labour.[3] The literature suggests that the policies and institutions generating misallocation are rampant in many developing countries (Restuccia, 2019).

Against this backdrop, Sri Lanka's macroeconomic performance, policies, and institutions present an interesting case study of misallocation. Since opening up its economy, Sri Lanka has undergone a structural transformation, evolving from an agriculture based economy to a manufacturing and services based economy (Athukorala, 2017). However, despite the transformation in the economy, aggregate TFP growth as a share of GDP growth has steadily declined over the past four decades (Kumari and Tang, 2024). Also, the manufacturing sector's contribution to the GDP continues to stagnate less than 20% of the country's GDP for the past six decades (World Bank, 2022). Although, following the liberalisation, Sri Lanka's manufacturing sector experienced a high rate of TFP growth, since the 1990s, manufacturing TFP growth has continued to decline, in spite of the capital accumulation made into the sector (IMF, 2018). One reason for the significantly low and declining manufacturing productivity in Sri Lanka is the misallocation of resources at most disaggregated level, i.e., at firm-level (Kumari et al., 2021). This poses a serious question as to which factor of production, i.e., capital or labour has been more misallocated and severely impacted on Sri Lanka's productivity loss.

Sri Lanka's policies and institutions on capital and labour present an interesting circumstance relating to regulation and misallocation. Particularly, the prevalence of distortions in both capital and labour markets such as strict regulations applicable for financial institutions, favourable treatments made by the state owned financial institutions (Thilakaweera, 2016), multiple and disconnected labour regulations

[2] For examples of capital misallocation, see: Caballero et al. (2008); De Mel et al. (2008); Song et al. (2011); Midrigan and Xu (2014); Gopinath et al. (2017); and Ranasinghe and Restuccia (2018).

[3] For examples of labour misallocation, see: Bloom et al. (2012); Caselli and Gennaioli (2013); Bai and Cheng (2016); Hsieh et al. (2019); López and Torres (2020); and Ranasinghe (2020).

and institutions, minimum wages imposed in certain industries, stringent regulations with respect to termination of employment that restrict the labour mobility across sectors and firms (Center for International Development, 2018), and other financial and labour market frictions, prevents the optimal factor allocation in the economy.

In this setting, this paper broadly examines the negative impact of cross-firm misallocation of capital and labour on aggregate productivity in Sri Lanka. To the best of this researcher's knowledge, the magnitudes of labour and capital misallocation and the consequential productivity losses have not been computed and compared in relation to Sri Lanka. Additionally, it investigates the relation between observable firm characteristics and factor misallocation.

This study covers the 24-year period from 1994 to 2017 during which, Sri Lanka had a liberalised-economy-regime in place, underwent some structural reforms, was caught up in a three-decade-long war, and faced with several external/internal natural/manmade shocks. For the empirical analysis, this study employs the monopolistic competition model with heterogeneous firms adopted by Hsieh and Klenow (2009) with necessary modifications.[4] To calibrate the model, a manufacturing firm-level dataset at four-digit International Standard Industrial Classification is sourced from the Sri Lanka's Annual Survey of Industries for 1995-2018 (ASI 1995-2018) conducted by the Department of Census and Statistics of Sri Lanka (DCS). Annual survey data are available for around 120 industries comprising about 14,000 manufacturing firms per year.

The results show that labour misallocation in Sri Lanka is more severe than that of capital misallocation. When capital and labour are hypothetically allocated to equalise Total Factor Productivity Revenue (TFPR) across firms in a given industry or when resource misallocation is eliminated, the average aggregate TFP gains from removing labour misallocation is 57.6% relative to that from removing capital misallocation of 24% suggesting that labour misallocation in Sri Lanka's manufacturing sector is severe and around 2.4 times that of capital misallocation. The results further show that allocative efficiency of both labour and capital in manufacturing has been deteriorating over time, though labour misallocation has steeply been rising.

The results further show that both capital and labour are more misallocated in firms that are located outside the Western province, non-textile oriented and unincorporated firms, relative to their counterfactual groups. However, in each category labour is more misallocated relative to the capital. Evidence further suggests that both labour and capital distortions are firm size-dependent. However, labour distortion shows a stronger relation with firm size in terms of value added having a covariance coefficient of 1.7% and a correlation coefficient of 0.7% relative to the covariance and correlation coefficients of 0.9% and 0.3%, respectively between capital distortions and firm size, showing that firms face more labour constraints when growing bigger.

The results also show that more productive firms face higher distortions both in capital and labour, but a stronger positive relationship between those firms and labour distortions relative to capital distortions. These results show that labour is more misallocated relative to capital.

[4] This paper normalises the model by using output distortion whereas Hsieh and Klenow (2009) normalise the model by using labour distortions.

The robustness checks confirm that the baseline results are strong enough to make the conclusion that labour is more misallocated relative to capital in Sri Lanka. Overall, the results suggest that labour misallocation exerts a larger negative impact on aggregate TFP and economic growth despite labour's relative abundance to capital in Sri Lanka.

These findings thus facilitate policymakers to formulate appropriate resource reallocation policies and shed light on the link between resource misallocation and low productivity in other economies from Sri Lanka's perspective. Further, findings would pose the main economic challenge that countries may have to address when embarking on liberalisation reforms, given the limited factor endowments within their specific economies. However, further research is needed to make a concrete inference on underlying sources of labour and capital misallocation in Sri Lanka.

The rest of the paper proceeds as follows. Section 2 discusses the relevant previous studies and Section 3 outlines the model for measuring distortion and misallocation. Section 4 presents the dataset and the data cleaning process. Empirical results are discussed in Section 5 along with the robustness results. Section 6 conducts further analyses. Finally, Section 7 concludes.

2. Literature Review

This study is related to the literature on misallocation of resources and resulted aggregate productivity losses. In 2008, Restuccia and Rogerson illustrated how policy distortions generate resource misallocation and how misallocation leads to sizable losses in productivity and output. In a seminal work, Hsieh and Klenow (2009) constructed a framework to quantify the effect of resource misallocation on aggregate productivity. By using manufacturing sector firm level data for the US, China, and India, they found sizable productivity gains for these counties when the resources are optimally allocated across firms. Since the work by Hsieh and Klenow (2009), growing literature has been attempted to measure the aggregate productivity losses due to misallocation of resources for different countries (Brandit et al., 2013; Calligaris, 2015; Chen, 2017; Kumari et al., 2021)

Literature has also examined the underlying sources of misallocation. Restuccia and Rogerson (2017) suggest that resource misallocation originates from firm specific distortions in output or in production factors. The studies on misallocation of production factor mostly focus on the capital and labour inputs. Some emphasise credit market imperfections and capital misallocation. Specifically, the literature on financial frictions highlights wedges between the marginal product of capital and borrowing rates (De Mel et al., 2008), 'zombie lending' practices (Caballero et al., 2008), the impact of capital on entry and technology adoption (Midrigan and Xu, 2014), Euro adoption and decrease in interest rates (Gopinath et al., 2017), policy distortions (Brandt et al., 2013), credit subsidy policies (Jo and Senga, 2019), and finance and productivity growth (Levine and Warusawitharana, 2021). The capital distortion can be persistent (Banerjee and Moll, 2010) and a key barrier to development (Banerjee and Duflo, 2005; Ranasinghe and Restuccia, 2018).

Other studies emphasise the misallocation of labour, including differences in the quality of managerial practices (Bloom et al., 2012; Caselli and Gennaioli, 2013), effects of race and gender on talent misallocation (Hsieh et al., 2019; Ranasinghe, 2020), firm characteristics and allocation of talent (López and Torres, 2020), and labour misallocation over time (Bai and Cheng, 2016), among many others.

However, studies that calculate and compare the misallocation of labour and capital in a single analysis are limited. JG Bun and de Winter (2022) using firm-level data for 2001-2017 found a combination of steeply rising capital misallocation and relatively stable labour misallocation in the Netherlands. They used the dispersion in marginal products of capital and labour to measure the extent of misallocation. This paper aims to fill the research gap by investigating the misallocation of resources by calculating and comparing misallocation of capital and labour from a perspective of a developing country, by using Sri Lanka's manufacturing sector firm-level data for 1994-2017.

3. Methodology and Calibrations

This section discusses the model derivation, variable measurements, parameter calibrations, measurement of factor distortions and the procedure for calculating the TFP gains.

3.1 Model

Largely based on Hsieh and Klenow (2009), the model quantifies the aggregate manufacturing TFP gains when misallocation in capital and labour are eliminated.[5] The manufacturing sector is comprised of S industries, indexed by subscript $s = 1, \ldots, S$. In ASI from 1995 to 2018, s refers to the four-digit industry level. A single final good Y is produced using a Cobb-Douglas technology:

$$Y = \prod_{s=1}^{S} Y_s^{\theta_s}, where \ \sum_{s=1}^{S} \theta_s = 1, \qquad (1)$$

where θ_s is the value-added share of the sector $s = 1, \ldots, S$. The correspondent cost minimization is:

$$P_s Y_s = \theta_s P Y, \qquad (2)$$

where, P_s and P are the price of the industry output (Y_s) and the price of the final manufacturing output (Y), respectively. At industry s level, the output Y_s is a CES aggregate of M_s differentiated products.

$$Y_s = \left(\sum_{i=1}^{M_s} Y_{si}^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}, \qquad (3)$$

where σ is the elasticity of substitution across different inputs. Finally, the output of firm i in industry s is produced according to a two-factor Cobb-Douglas technology:

$$Y_{si} = A_{si} K_{si}^{\alpha_s} L_{si}^{1-\alpha_s}, \qquad (4)$$

where A_{si}, K_{si}, and L_{si} are firm-level TFP, capital stock and, labour stock respectively, whilst α_s, and $(1-\alpha_s)$ are industry-specific shares of capital and labour, respectively. Accordingly, in this setting, $\alpha_s + (1-\alpha_s) = 1$. Each firm faces two types of firm-specific distortions, i.e., in capital $\left(\tilde{\tau}_{K_{si}}\right)$, and labour $\left(\tilde{\tau}_{L_{si}}\right)$. The objective of firm si is profit maximisations by choosing P_{si}, Y_{si} and taking factor prices, distortions and the output demand curve as given.

$$\pi_{si} = P_{si} Y_{si} - \left(1 + \tilde{\tau}_{K_{si}}\right) R K_{si} - (1 + \tilde{\tau}_{L_{si}}) w L_{si}, \qquad (5)$$

[5] This paper normalises the model by using output distortion whereas Hsieh and Klenow (2009) normalise the model by using labour distortions.

where π_{si} is profit, $P_{si}Y_{si}$ is value-added, R is the rental price of capital and w is effective wage. Following Hsieh and Klenow (2009), firm-specific revenue productivity $(TFPR_{si})$, which is proportional to a geometric average of the firm's revenue product of capital and labour can be derived:

$$TFPR_{si} = P_{si}A_{si} = \frac{P_{si}Y_{si}}{K_{si}^{\alpha_s}L_{si}^{1-\alpha_s}} = \frac{\sigma}{\sigma-1}\left(\frac{R\left(1+\tau_{K_{si}}\right)}{\alpha_s}\right)^{\alpha_s}\left(\frac{w\left(1+\tau_{L_{si}}\right)}{1-\alpha_s}\right)^{1-\alpha_s},$$

$$TFPR_{si} \propto (MRPK_{si})^{\alpha_s}(MRPL_{si})^{1-\alpha_s}\,, \tag{6}$$

If there are no distortions (i.e., when $\tilde{\tau}_{K_{si}} = \tilde{\tau}_{L_{si}} = 0$), there would be no variation in $MRPK_{si}$, $MRPL_{si}$ and $TFPR_{si}$ within each industry. Following (6), revenue productivity for industry s is:

$$TFPR_s = \frac{P_sY_s}{K_s^{\alpha_s}L_s^{1-\alpha_s}}, \tag{7}$$

where $P_sY_s = \sum_i P_{si}Y_{si}$.

$TFPR_s$can also be written as:

$$TFPR_s = \frac{\sigma}{\sigma-1}\left[\frac{R}{\alpha_s\sum_{i=1}^{M_s}\frac{P_{si}Y_{si}}{\left(1+\tau_{K_{si}}\right)P_sY_s}}\right]^{\alpha_s}\left[\frac{w}{(1-\alpha_s)\sum_{i=1}^{M_s}\frac{P_{si}Y_{si}}{\left(1+\tau_{L_{si}}\right)P_sY_s}}\right]^{1-\alpha_s}, \tag{8}$$

which depends on the weighted average of the firm specific marginal product of capital and labour within each industry (by using firm's value-added share as its weight) within each industry.

Industry s physical productivity is defined as:

$$TFP_s = \frac{Y_s}{K_s^{\alpha_s}L_s^{1-\alpha_s}} = \left[\sum_{i=1}^{M_s}\left(A_{si}\frac{TFPR_s}{TFPR_{si}}\right)^{\sigma-1}\right]^{\frac{1}{\sigma-1}}. \tag{9}$$

In the absence of distortions, efficient TFP in industry s is:

$$TFP_s^{efficient} = \overline{A_s} = \left(\sum_{i=1}^{M_s}A_{si}^{\sigma-1}\right)^{\frac{1}{\sigma-1}}. \tag{10}$$

The physical productivity for the entire manufacturing sector is aggregated as:

$$TFP = \prod_{s=1}^{S}TFP_s^{\theta_s}. \tag{11}$$

The Cobb-Douglas aggregator gives the ratio between actual (Y^{actual}) and efficient output $(Y^{efficient})$ in the aggregate manufacturing sector:

$$\frac{Y^{actual}}{Y^{efficient}} = \prod_{s=1}^{S}\left(\frac{TFP_s}{\overline{A_s}}\right)^{\theta_s} = \prod_{s=1}^{S}\left[\sum_{i=1}^{M_s}\left(\frac{A_{si}}{\overline{A_s}}\frac{TFPR_s}{TFPR_{si}}\right)^{\sigma-1}\right]^{\frac{\theta_s}{\sigma-1}}. \tag{12}$$

Then, the potential reallocation gains are calculated by equalising total factor productivity revenue across firms in a given industry, as in Hsieh and Klenow (2009):

$$TFP\ Gain(removing\ all\ distortions) = \left(\frac{Y^{efficient}}{Y^{actual}} - 1\right) \times 100\%. \quad (13)$$

Lastly, the main measure of misallocation, i.e., the TFP gains are calculated by conducting counterfactual exercises in the following two cases: (1) from the actual TFP with all distortions to the TFP without capital distortion and (2) from the actual TFP with all distortions to the TFP without labour distortion, as in equations (14) and (15), respectively:

$$TFP\ Gain\ (removing\ \tau_{K_{si}}) = \left(\frac{TFP\left(\tau_{K_{si}}=0, \tau_{L_{si}}=actual\right)}{TFP^{actual}\left(\tau_{K_{si}}, \tau_{L_{si}}\right)} - 1\right) \times 100\%, \quad (14)$$

$$TFP\ Gain\ (removing\ \tau_{L_{si}}) = \left(\frac{TFP\left(\tau_{K_{si}}=actual, \tau_{L_{si}}=0\right)}{TFP^{actual}\left(\tau_{K_{si}}, \tau_{L_{si}}\right)} - 1\right) \times 100\%. \quad (15)$$

3.2 Measurement of Variables

Variables at the firm-level are obtained from Sri Lanka's ASI from 1995 to 2018. Fixed capital stock in the dataset is comprised of land, buildings and other constructions, plant and machinery, transport equipment, computers and accessories, furniture and other office equipment, intangible fixed assets.[6] *Capital* is the year-end book value of fixed capital stock. Annual year-end capital stock is arrived at by adding the gross additions[7] during the year and deducting the deprecations during the year to/from the opening stock of the year by using ASI data. Number of paid persons engaged (employees) and value of salaries and wages paid to such employees extracted from the dataset are used as *labour* stock and *wage bill*. Corresponding industry aggregates are calculated by summing over firms in each four-digit industry.[8]

3.3 Calibrations

Following the related literature, including Hsieh and Klenow (2009) and Calligaris (2015), I set $\sigma = 3$ and $R = 10\%$, where the real interest rate is 5%, and the depreciation rate is 5%. Adopting Kumari et al. (2021), the firm-level wage bill is adjusted in accordance with the macro-level labour share, which is 0.3 times the output, as in Penn World Tables 9.1 (Lederman et al., 2017). Then, the annual, average, effective wage in aggregate manufacturing sector w is calculated as:

$$w = \frac{\sum_{s=1}^{S}\sum_{i=1}^{M_S} WageBill_{si}}{\sum_{s=1}^{S}\sum_{i=1}^{M_S} L_{si}}, \quad (16)$$

[6] Leased or rented assets have not been included from the physical capital stock by DCS.

[7] This is defined as the total of the costs of new and second hand fixed assets acquired during the year and alterations, renovations and improvements purchased, less the value of sales of used fixed assets.

[8] Key variable definitions and data sources are listed in Table A.1 in Appendix A.

where $WageBill_{si}$ and L_{si} are the firm-level total wage bill and number of employees, respectively. The four-digit industry capital share (α_s) is calculated by deducting the labour share in the corresponding industry from unity. The four-digit industry labour share $(1-\alpha_s)$ is calculated by dividing the four-digit industry wage bill by the corresponding industry value-added, as below:

$$(1-\alpha_s) = \frac{\sum_{i=1}^{M_S} WageBill_{si}}{\sum_{i=1}^{M_S} P_{si}Y_{si}}. \tag{17}$$

Finally, each four-digit industry value-added share (θ_s) is calculated as below:

$$\theta_s = \frac{\sum_{i=1}^{M_S} P_{si}Y_{si}}{\sum_{s=1}^{S}\sum_{i=1}^{M_S} P_{si}Y_{si}}. \tag{18}$$

3.4 Measuring Distortions

Mostly following Hsieh and Klenow (2009), firm-level normalised distortions in capital, and labour are calculated as below:

$$1+\tau_{K_{si}} = \frac{\sigma-1}{\sigma}\frac{\alpha_S P_{si}Y_{si}}{RK_{si}}, \tag{19}$$

$$1+\tau_{L_{si}} = \frac{\sigma-1}{\sigma}\frac{(1-\alpha_S)P_{si}Y_{si}}{wL_{si}}, \tag{20}$$

3.5 Procedures for Calculating TFP Gains

Largely based on the procedure in Kumari et al. (2021), I calculate annual counterfactual TFP gains in the absence of distortions following the steps below.

1. Set $\sigma = 3$ and R = 10%.
2. Calculate w, α_s, and θ_s using (16), (17) and (18), respectively.
3. Calculate distortions in capital and labour using (19) and (20), respectively.
4. Following Hsieh and Klenow (2009), calculate firm-level physical productivity by using:

$$TFPQ_{si} = A_{si} = \frac{(P_{si}Y_{si})^{\frac{\sigma}{\sigma-1}}}{K_{si}^{\alpha_S}L_{si}^{1-\alpha_S}}. \tag{21}$$

5. Calculate firm-level revenue productivity $TFPR_{si}$ and the four-digit industry level revenue productivity $TFPR_s$ by using (6) and (8), respectively.
6. Calculate the four-digit industry distorted physical productivity TFP_s and efficient physical productivity $TFP_s^{efficient}$ by using (9) and (10), respectively.
7. Calculate the manufacturing sector efficiency level from (12).
8. Deduct the manufacturing sector efficiency level from unity to obtain the aggregate misallocation losses.
9. Conduct the counter-factual exercises to gauge TFP gains by removing all distortions, capital distortion and labour distortion (one at a time) by using equations (13), (14) and (15), respectively.

4. Data

This analysis covers a 24-year period, from 1994 to 2017, during which Sri Lanka had a liberalised-economy-regime in place, implemented reforms in factor markets, engaged in a three-decade-long destructive war that led to the displacement of people and firms and faced with internal and external shocks.

This study uses manufacturing firm-level data sourced from Sri Lanka's ASI from 1995 to 2018.[9] DCS, which is the statistical office under the Ministry of Finance, conducts the ASI annually. The ASI covers all firms with five or more persons engaged.[10] All firms with 100 or more persons are fully enumerated, whilst firms with 5-99 persons are covered by a sample. Usually, the ASI extends to the entire country, and the DCS' census is used to design its sample frame.[11]

The ASI is the main source of industrial statistics for the formal manufacturing sector in Sri Lanka. It contains information under four main industry divisions: mining and quarrying; manufacturing; electricity, gas, steam, and air conditioning supply; and water supply, sewerage, waste management and remediation activities. This study considers the manufacturing subcategory only, which has around 120 industries at the four-digit United Nations' International Industrial Classification level, comprising around 14,000 firms per year.[12] The ASI provides firm-level statistical information related to value of production, value of intermediate inputs, value-added, salaries and wages, employment stock, value of capital stock and some other firm characteristics, such as entry year, location and the ownership structure.[13]

The cross-sectional dataset for the sample period from 1994 to 2017 initially had around 258,308 manufacturing firms, but after the data were cleaned in line with Kumari et al. (2021), that number reduced to 113,827 firms.[14] All estimations in the study are based on the cleaned-sample dataset. Table 1 reports summary statistics for those variables included in the analysis.

---

[9] The ASI for any given year contains data and information for the previous year.

[10] Number of persons engaged consists of both paid workers (employees) and unpaid workers, mostly the family members.

[11] The industrial sample frame generated from the Census of Industry 1983 was used to conduct the ASI from 1984 to 2003. The frame, prepared by using 2003 Census of Industries, was used for the ASI from 2004 to 2012. The Economic Census conducted in 2013/14 was the base for the new frame of industries, which was used for the ASI from 2015 to 2018.

[12] During the sample period, there were three revisions of industrial classifications for industry allocation. Hence, to enable the comparability across years, data compiled before 2015 under United Nations' International Industrial Classification Revision-2 (1995-2002) and Revision-3.1 (2003-2013) are reclassified with Revision-4, the latest classification.

[13] Entry year and ownership information are available only from 2006 survey, as the relevant questions in the survey were included from only that year onward.

[14] First, observations with empty cells and observations with non-positive values for capital, labour, output and value-added or wage bill are dropped. Second, 1% of all variables, including firm's labour share, from both tails are dropped. Third, industries that had one or more labour share(s) are dropped. Fourth, 1% of tails from both tails of $\log(A_{si}/\overline{A_s})$ and $\log(TFPR_{si}/\overline{TFPR_s})$ are dropped, following Hsieh and Klenow (2009). Fifth, the industries that are left with a single firm are dropped to ensure the possibility of resource reallocation among firms within an industry. Surveys were not conducted in 2004 or 2014, as those were census years. 2005 and 2006 are dropped due to lower coverage.

Table 1: Summary statistics (1994-2017)[1].

Variable (unit)	Mean	Std. Dev.	Min.	P1[2]	P50[2]	P99[2]	Max.
Labour (Number of employees)	136	226	5	1,615	5	49	1,162
Capital (LKR thousands)	47,800	102,000	32	851,000	66	9,220	549,000
Value-added (LKR thousands)	52,800	116,000	42	1,190,000	100	10,000	599,000
Wage bill (LKR thousands)	40,000	94,200	18	1,620,000	54	7,220	458,000

Notes.

1. Summary statistics are for the cleaned dataset. The sample size is 25,537, representing 113,827 firms.
2. P1, P50 and P99 are the 1st, 50th (median) and 99th percentiles, respectively.

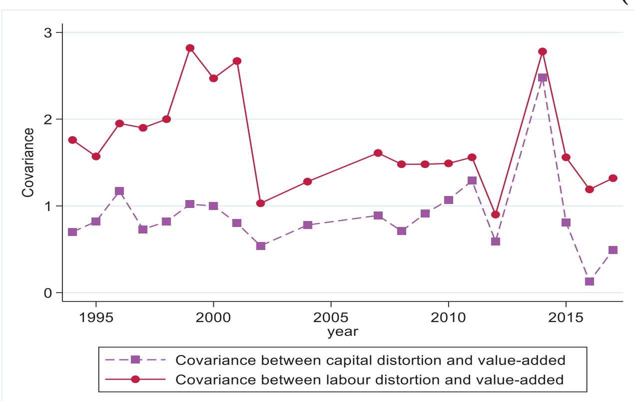
5. Empirical Analysis

This section discusses the main findings of the study: covariance and correlation between factor distortions and firm size and the TFP gains by removing capital and labour distortions relative to TFP gains with all distortions, along with the necessary robustness tests.

5.1 Factor Distortions and Firm Size

One measure of misallocation is the covariance coefficient between factor distortions and firm size. This coefficient indicates whether distortions and firm size move in the same direction or in the opposite direction. As seen in Figure 1, the covariance coefficient is positive for both capital and labour distortions, indicating that when firms become larger, they face more factor distortions. However, the covariance coefficient for labour distortion is higher than that for capital indicating a stronger relationship for labour distortion and firm size showing a higher labour constant for bigger firms. The average covariance coefficient between capital distortion and firm size and labour distortion and firm size during 1994-2017 are 0.89 and 1.74, respectively.[15]

Figure 1: Covariance coefficient between factor distortions and firm size (value-added)

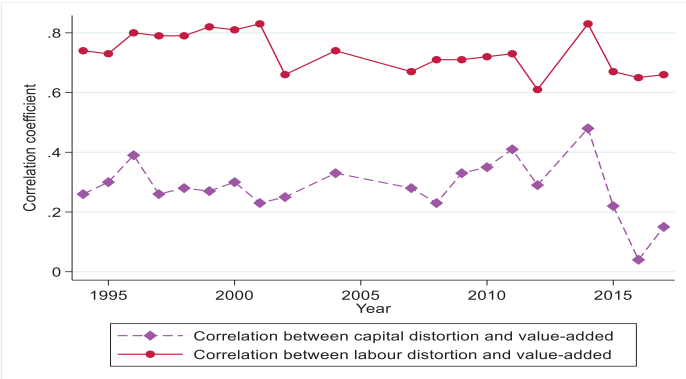


Notes. Entries are the covariance coefficient between capital/labour distortion and firm size in terms of value-added. Annual values are calculated by weighted averaging industry-level covariance, using the industry value-added share in each year. Industry-level covariance is calculated using the firm-level covariance.

[15] Detailed results are given in Table B.1 in Appendix B.1.

Figure 2 reinforces the covariance results and shows the results of the correlation coefficient that measures the correlation between factor distortions and firm size. The coefficients are positive for both distortions showing positive correlations between firm size and distortions. The average correlation for the whole sample period between capital distortion and firm size and labour distortion and firm size are 0.29 and 0.73, respectively. However, when firms become larger, they face more labour distortions relative to capital distortions.[16]

Figure 2: Correlation coefficient between factor distortions and firm size (value-added)



Notes. Entries are the correlation coefficient between capital/labour distortion and firm size in terms of value-added. Annual values are calculated by weighted averaging industry-level correlation, using the industry value-added share in each year. Industry-level correlation is calculated using the firm-level correlation.

5.2 TFP Gains by Removing Distortions

Table 2 presents the main measure of misallocation, which is the potential TFP gains by removing one factor distortion at a time relative to having all distortions. Column (1) reports the efficient TFP with no distortions, while column (2) reports the actual TFP with all firm-level distortions (i.e., capital and labour). Columns (3) and (4) show the TFP results after removing capital distortion and labour distortion, respectively.[17] The results show that, on average, log actual TFP after removing labour distortion is higher (at 20.9) relative to log actual TFP after removing capital distortion (at 20.7). Hence, in all years, after removing labour distortion, potential TFP gains relative to the baseline are higher than the gains from removing capital distortion, as shown in columns (6) and (5), respectively. Figure 3 also depicts clear evidence of greater labour misallocation. The period average TFP gain by removing labour distortion is much higher at 57% than that for capital at 24%.

[16] The detailed results are in Table B.2 in Appendix B.2.

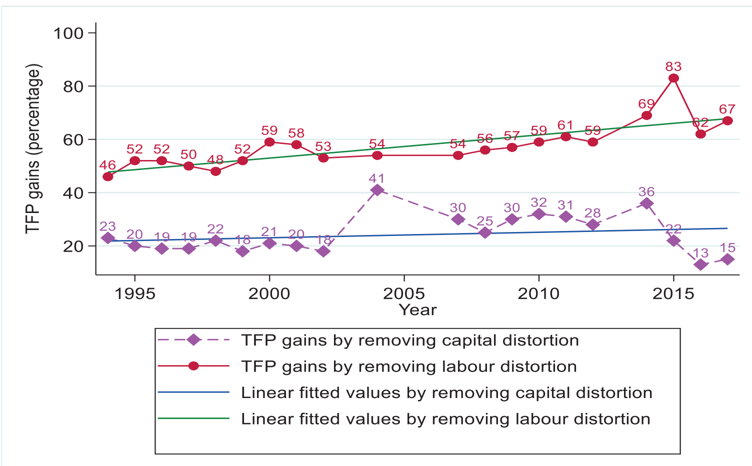
[17] Figure B.1 in Appendix B.3 depicts movements of log efficient and log actual TFP over time.

Table 2: TFP levels and TFP gains (%) by removing factor distortions

Year	Log efficient TFP	Log actual TFP			TFP gains (%)	
	All 3 scenarios	Baseline	$\tau_K = 0$	$\tau_L = 0$	$\tau_K = 0$	$\tau_L = 0$
	(1)	(2)	(3)	(4)	(5)=[exp(3-2)-1]*100	(6)=[exp(4-2)-1]*100
1994	19.6	19.0	19.2	19.3	22.7	46.0
1995	20.1	19.4	19.6	19.8	19.8	52.0
1996	20.2	19.6	19.8	20.0	19.2	52.0
1997	20.2	19.6	19.7	20.0	18.9	49.9
1998	19.8	19.2	19.4	19.6	22.1	48.4
1999	20.2	19.6	19.8	20.0	17.5	52.4
2000	20.4	19.7	19.9	20.2	20.5	58.8
2001	20.6	19.9	20.1	20.4	19.9	58.4
2002	20.9	20.3	20.5	20.7	17.8	52.8
2004	20.6	19.9	20.2	20.3	40.8	54.2
2007	21.3	20.5	20.8	21.0	30.1	53.9
2008	21.8	21.1	21.3	21.6	24.8	55.8
2009	21.4	20.7	20.9	21.1	29.6	57.4
2010	21.9	21.1	21.4	21.6	31.9	59.0
2011	21.7	20.9	21.2	21.4	31.0	61.3
2012	22.1	21.4	21.6	21.8	27.9	58.6
2014	21.8	21.0	21.3	21.5	36.0	68.8
2015	22.5	21.7	21.9	22.3	21.5	83.1
2016	22.9	22.3	22.4	22.8	12.8	61.6
2017	22.7	22.1	22.2	22.6	14.7	66.5
Average	21.1	20.4	20.7	20.9	24.0	57.6

Notes. The entries in column (1) are the log of annualised efficient TFP, obtained from equation (10), while columns (2), (3) and (4) are the log of annualised actual TFP, obtained from equation (9). In columns (5) and (6), TFP gains (%) = 100× [exp (Log TFP removing one distortion - Log actual TFP with all distortions) - 1].

Figure 3: TFP gains (%) by removing factor distortions



Notes. Figure 3 depicts the results in columns (5) and (6) of Table 2. TFP gains (%) = 100× [exp (Log TFP removing one distortion - Log actual TFP with all distortions) - 1].

Overall, the results suggest that the labour of Sri Lanka's manufacturing is more misallocated than its capital. Hence, the contribution of labour misallocation to aggregate TFP is larger than that of capital. More precisely, the contribution of labour misallocation to TFP is almost 2.4-fold that of capital, with 57.6% TFP gains by eliminating labour misallocation relative to 24% TFP gains by eliminating capital misallocation. The results also suggest rising factor misallocation with a steeper rise for labour misallocation. For instance, the potential TFP gains by removing labour distortions increased to 83.1% in 2015, relative to 46% in 1994, whereas the TFP gains by removing capital distortions marginally increased to 36% in 2014 compared to 23% that was in 1994.

5.3 Robustness Checks

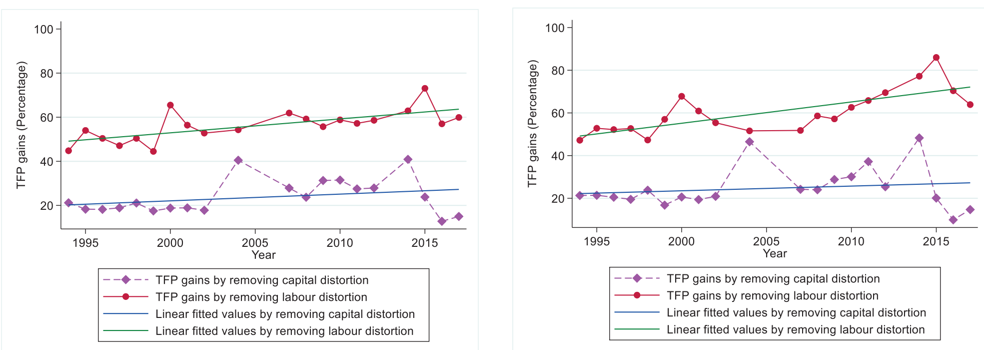
This section includes the sensitivity analyses I conducted to recalculate TFP gains by changing variables and model parameter values.

The TFP gain results presented in Table 2 could be sensitive to data quality, calibrations, and parameter values. Hence, four robustness tests are conducted by dropping very small firms, dropping very large firms, increasing the elasticity of substitution, and inflating the wage bill.[18]

The sensitivity results show that TFP gains after removing labour distortion range between 56.2% and 71.5%, whereas TFP gains after removing capital distortion range only between 16.7% and 30.4%. Figure 4 also reinforces the clear evidence of greater and sharply rising labour misallocation in all four scenarios; indeed, the line representing TFP gains from removing labour distortion is above that of capital; and the line representing TFP gains from removing labour distortion having a steeper slop. The robustness results endorse the baseline findings, suggesting that the contribution of labour misallocation to TFP is around 2.4 times that of capital and labour misallocation had been rising at a higher rate relative to that of capital. Thus, there is more and steeply rising industrial labour misallocation in Sri Lanka than capital misallocation.

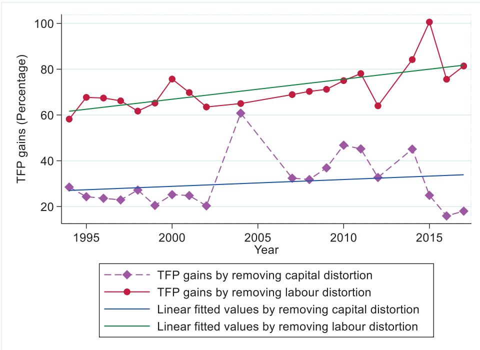
Figure 4: TFP gains (%) by removing factor distortions

(A) Employees ≥ 10 | (B) Employees ≤ 1,500

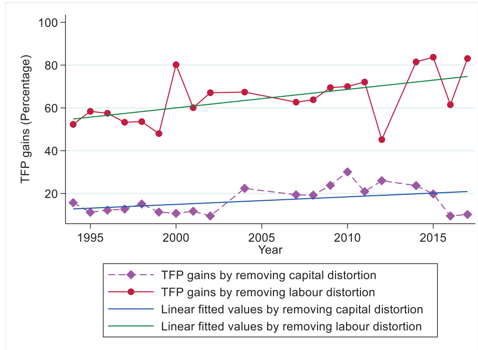


[18] Detailed results are in Table B.3 in Appendix B.4.

(C) δ = 4



(D) Wage/output = 0.35



Notes. Panels (A), (B), (C) and (D) depict the results in columns (1), (2), (3) and (4) of Table B.3 in Appendix B.4, respectively. TFP gains (%) = 100× [exp (Log TFP removing one distortion - Log actual TFP with all distortions) - 1].

6. Further empirical analysis

This section (1) conducts counterfactual analyses for different sub-groups of firms and (2) examines the relation between factor distortions and productivity.

6.1 Firm Characteristics and Misallocation

The counterfactual analyses are performed to quantify the TFP gains for different types of firms, by dividing the firms into subgroups according to some observable characteristics to see which factor in which types of firms is highly misallocated. Accordingly, TFP gains are quantified for different subsets of firms based on their geographical location, export/production orientation, incorporation status and age by removing one factor distortion, i.e., capital or labour in each subset at a time. Table 3 presents the summary results for each sub-group of firms.[19] Column (1) reports the magnitude of misallocation by only removing capital distortion with the period average TFP gain over the whole sample period, column (2) reports the same after removing the labour distortion only for each type of firm, whereas column (3) reporting the results by dividing the TFP gain results in column (2) by the TFP gains results in column (1).

The period average results in column (1) show that the magnitude of capital misallocation is higher for firms that are located outside Western Province, export oriented, non-textile, unincorporated and old firms relative to their counterparts. Column (2) shows that the period average labour misallocation is high for all sub-categories relative to the misallocation of capital as in column (1). Column (3) shows that, although in most of the sub categories, labour misallocation is more than two times of the capital misallocation, relative to capital, it is largely misallocated in non-export oriented firms (3.4 times of capital), whereas in un-incorporated firms, labour misallocation is 1.5 times of capital misallocation.

[19] Detailed results are given in Table C.1 in Appendix C.

Table 3: Firm Characteristics and TFP gains (%) by removing one factor distortion at a time

Firm Characteristics	$\tau_K = 0$ (1)	$\tau_L = 0$ (2)	τ_L/τ_K (3)=(2)/(1)
Location (1994-2015)			
Western	20.3	45.3	2.2
Outside Western	35.0	62.4	1.8
Production Orientation (1994-2017)[1]			
Export Oriented	25.3	59.3	2.3
Non-export Oriented	20.8	70.5	3.4
Product Category (1994-2017) [1]			
Textile	21.8	58.7	2.7
Non-Textile	23.6	69.3	2.9
Incorporation Status (2007-2017) [2]			
Incorporated	21.4	58.0	2.7
Un-incorporated	42.6	65.4	1.5
Age (2007-2017) [2]			
Young	27.3	66.2	2.4
Old	29.2	50.6	1.7

Notes. TFP gains (%) = 100× [exp (Log TFP removing one distortion - Log actual TFP with all distortions) - 1].

1. Classifications are made by using four-digit industry codes; hence those are approximations.
2. Ownership and entry year information are available only from 2007.

6.2 Factor Distortions and Productivity

To ascertain how factor distortions are associated with productivity, firm-level capital and labour distortions are separately regressed on firm-level productivity by using pooled data from 1994 to 2017 and equation (22).

$$Log(1+\tau_{X_{ist}}) = \beta_X Log(TFPQ_{ist}) + \lambda_s + \eta_t + \varepsilon_{\tau X_{ist}}, \qquad (22)$$

where, $X = K, L$. $Log(1+\tau_{X_{ist}})$ and $Log(TFPQ_{ist})$ are log of firm-level distortions for input i.e., capital or labour and log of firm-level productivity, respectively. Subscripts *i, s* and *t* denote firm, industry and year, respectively. β_X is the key coefficient of interest, and it measures how firm-level distortions in capital and labour respond to firm-level productivity changes. λ_s captures the industry fixed effects, while η_t captures the year fixed effects.

Results are shown in Table 4. Both β_K for capital distortion β_L for labour distortion are positive and both the coefficients are statistically significant at the 1% level. When firm-level productivity increases by 1%, capital distortion increases by 0.435%, whereas labour distortion increases by a larger percentage at 0.544%. The results show that highly productive firms face higher distortions both in capital and labour, but a stronger positive relationship is present between those firms and labour distortions relative to capital distortions. These results show that labour is more misallocated in highly productive firms relative to capital.[20]

[20] The same exercise is conducted using annual data, but without year fixed effects. The details are given in Appendix D, and the results are in Table D.1. Those results also confirm the findings in Table 4.

Table 4: TFPQ and distortions: Pooled data

Period	Capital Distortion β_K	Labour Distortion β_L
1994-2017	0.435***	0.544***
	(0.004)	(0.002)

Note: β_K and $\beta_L$are regression coefficients of log TFPQ with industry and year fixed effects. Standard errors clustered at the firm-level are displayed in parentheses. *** 1% significance.

7. Concluding Remarks

This paper quantifies the misallocation of capital and labour among firms and their relative contribution to aggregate TFP by using a firm-level dataset for Sri Lanka's manufacturing sector. Adopting a static production model used by (Hsieh and Klenow, 2009), the study finds that labour distortion, relative to that of capital, exhibits both positive and greater covariance and correlation with firm size suggesting that capital and labour misallocations are firm-size dependent. The results show that the aggregate TFP gains from removing labour distortion is 57.6% and it is 2.4 times of the TFP gains from removing capital distortion, which is only 24%. The results also show that both capital and labour misallocation were rising over time, though labour misallocation risen at a higher rate. The results further show that both capital and labour are more misallocated in firms that are located out-side Western province, non-textile oriented and unincorporated, relative to their counterfactual groups. However, in each category, labour is more misallocated relative to the capital. The findings also suggest that more productive firms face more factor misallocation and the relationship between productivity and misallocation is stronger for labour. The robustness results also confirm high and steeply rising labour misallocation relative to that for capital.

The findings of this study are in line with previous studies, which show that accessing labour is a major constraint to growth in Sri Lanka. Indeed, rigid labour regulations and disconnected labour institutions can be crucial barriers for business expansion and economic development relative to capital in developing countries. Some potential policy suggestions to improve the efficiency in labour reallocation would be, amending the legislations to reduce labour market protection, increasing labour mobility though relaxed hiring and firing rules, removal of minimum wage standers and the wage boards, liberalising the regulations relating the employing female workers, improving labour market governance, and enhancing labour market information for making informed employment choices.

This study contributes to the literature along three dimensions. First, this study is the first that examines the relative importance of labour and capital misallocations among firms and how these misallocations affect the country's aggregate manufacturing TFP in a developing country perspective. Second, it modified the model of Hsieh and Klenow (2009) by normalizing it using the output distortion. Third, this study is conducted using an annual survey dataset that spans over two decades, from 1994 to 2017.

There are some caveats to this paper. The ASI dataset does not include informal or very small manufacturing firms with less than five persons. Also, there is a possibility of miscalculation of TFP gains due to model specifications (Gong and Hu, 2016). Similarly, this paper does not focus on the impact of specific policy distortions or incentive problems, which are possibly relevant. Hence, further research is needed to understand specific policies and institutions underlying the factor misallocation in developing countries that lowers the aggregate productivity. These issues are left for future research on Sri Lanka.

References

Athukorala, Prema-chandra (2017). "Manufacturing exports from Sri Lanka: opportunities, achievements and policy options."

Bai, Peiwen and Wenli Cheng (2016). "Labour misallocation in China: 1980–2010". *Applied economics,* vol. 48, pp. 2321-2332.

Banerjee, Abhijit V. and Esther Duflo (2005). "Growth theory through the lens of development economics." *Handbook of economic growth,* vol. 1, pp. 473-552.

Banerjee, Abhijit V. and Benjamin Moll (2010). "Why does misallocation persist?" *American Economic Journal: Macroeconomics,* vol. 2, pp. 189-206.

Bloom, Nicholas, Christos Genakos, Raffaella Sadun and John Van Reenen (2012). "Management practices across firms and countries." *Academy of management perspectives,* vol. 26, pp. 12-33.

Brandt, Loren, Trevor Tombe and Xiaodong Zhu (2013). "Factor market distortions across time, space and sectors in China." *Review of Economic Dynamics,* vol. 16, pp. 39-58.

Caballero, Ricardo J., Takeo Hoshi and Anil K. Kashyap (2008). "Zombie lending and depressed restructuring in Japan." *American Economic Review,* vol. 98, pp. 1943-1977.

Calligaris, Sara (2015). "Misallocation and Total Factor Productivity in Italy: Evidence from Firm-Level Data." *Labour,* vol. 29, pp. 367-393.

Caselli, Francesco and Nicola Gennaioli (2013). "Dynastic management." *Economic Inquiry,* vol. 51, pp. 971-996.

Center for International Development (2018). Sri Lanka Growth Diagnostic, Harvard University.

Chen, Shawn Xiaoguang (2017). "VAT rate dispersion and TFP loss in China's manufacturing sector." *Economics Letters,* vol. 155, pp. 49-54.

De Mel, Suresh, David McKenzie and Christopher Woodruff (2008). "Returns to capital in microenterprises: evidence from a field experiment." *The Quarterly Journal of Economics,* vol. 123, pp. 1329-1372.

Gong, Guan and Guanliang Hu (2016). "The role of returns to scale in measuring frictions in resource allocation: Revisiting misallocation and manufacturing TFP in China." *Economics Letters,* vol. 138, pp. 26-29.

Gopinath, Gita, Şebnem Kalemli-Özcan, Loukas Karabarbounis and Carolina Villegas-Sanchez (2017). "Capital Allocation and Productivity in South Europe." *The Quarterly Journal of Economics,* vol. 132, pp. 1915-1967.

Hsieh, Chang-Tai, Erik Hurst, Charles I. Jones and Peter J. Klenow (2019). "The allocation of talent and US economic growth." *Econometrics,* vol. 87, pp. 1439-1474.

Hsieh, Chang Tai and Peter J. Klenow (2009). "Misallocation and Manufacturing TFP in China and India." *The Quarterly Journal of Economics,* vol. 124, pp. 1403-1448.

IMF (2018) "Sri Lanka : Selected Issues." In International Monetary Fund. Asia and Pacific Dept (ed.), *Country Report No. 18/176*.

JG Bun, Maurice and Jasper de Winter (2022). "Capital and labor misallocation in the Netherlands." *Journal of Productivity Analysis,* vol. 57, pp. 93-113.

Jo, In Hwan and Tatsuro Senga (2019). "Aggregate consequences of credit subsidy policies: Firm dynamics and misallocation." *Review of Economic Dynamics,* vol. 32, pp. 68-93.

Kumari, Ranpati Dewage Thilini Sumudu and Sam Hak Kan Tang (2024). "Identifying the Sources of Economic Growth in Sri Lanka under Trade Liberalization: The Primal and Dual Total Factor Productivity Growth." *Journal of the Asia Pacific Economy*.

Kumari, Ranpati Dewage Thilini Sumudu, Sam Hak Kan Tang, Shawn Xiaoguang Chen and Bei Li (2021). "Misallocation and productivity slowdown over two decades: evidence from Sri Lanka." *Applied economics,* vol. 53, pp. 4417-4435.

Lederman, Daniel, Justin T Lesniak, Robert C Feenstra, Robert Inklaar and Marcel P Timmer (2017) "The Next Generation of the Penn World Table", in Open and Nimble: Finding Stable Growth in Small Economies, The World Bank Group.

Levine, Oliver and Missaka Warusawitharana (2021). "Finance and productivity growth: Firm-level evidence." *Journal of Monetary Economics,* vol. 117, pp. 91-107.

López, José Joaquín and Jesica Torres (2020). "Size-dependent policies, talent misallocation, and the return to skill." *Review of Economic Dynamics*, vol. 38, pp. 59–83.

Midrigan, Virgiliu and Daniel Yi Xu (2014). "Finance and misallocation: Evidence from plant-level data." *American Economic Review,* vol. 104, pp. 422-458.

Ranasinghe, Ashantha (2020). "Misallocation across Establishment Gender." *University of Alberta Department of Economics Working Paper*.

Ranasinghe, Ashantha and Diego Restuccia (2018). "Financial frictions and the rule of law." *Journal of development economics,* vol. 134, pp. 248-271.

Restuccia, Diego (2019). "Misallocation and aggregate productivity across time and space." *Canadian Journal of Economics/Revue canadienne d'économique,* vol. 52, pp. 5-32.

Restuccia, Diego and Richard Rogerson (2017). "The causes and costs of misallocation." *Journal of Economic Perspectives,* vol. 31, pp. 151-174.

Song, Zheng, Kjetil Storesletten and Fabrizio Zilibotti (2011). "Growing like china." *American Economic Review,* vol. 101, pp. 196-233.

Thilakaweera, Bolanda Hewa (2016). "Efficiency and productivity in Sri Lanka's banking sector: Evidence from the post-conflict era." Doctor of Philosophy thesis, School of Accounting, Economics and Finance, University of Wollongong.

World Bank (2022). *World Development Indicators*. World Bank.

Appendices

Appendix A. Measurement of Variables

Table A.1. Definitions of Key Variables and Data Sources

Variable	Definition	Source
Employees/Labour	Number of paid persons engaged as at end year	Firm level variables from *Annual Survey of Industries* 1995-2018
Capital	Year-end book value of capital stock	
Wage Bill	Value of salaries and wages paid to employees during the year	
Value-added	Value of the output minus intermediate consumptions during the year	
Sigma (σ)	Elasticity of substitution within each four-digit industry ($\sigma = 3$)	Industry level variables from Hsieh and Klenow (2009)
Rental Rate of Capital (R_K)	10%	

Notes. Corresponding industry aggregates are calculated by summing over firms in each four-digit industry.

Appendix B. Results: Distortions and Productivity Gains

B.1 Covariance between Factor Distortions and Firm Size

Table B.1. Covariance between factor distortions and firm size

Year	Capital distortion	Labour distortion
1994	0.70	1.76
1995	0.82	1.57
1996	1.17	1.95
1997	0.73	1.90
1998	0.82	2.00
1999	1.02	2.82
2000	1.00	2.47
2001	0.80	2.67
2002	0.54	1.03
2004	0.78	1.28
2007	0.89	1.61
2008	0.71	1.48
2009	0.91	1.48
2010	1.07	1.49
2011	1.29	1.56
2012	0.59	0.90
2014	2.48	2.78
2015	0.81	1.56
2016	0.13	1.19
2017	0.49	1.32
Average	0.89	1.74

Notes. Annual values are calculated by weighted averaging industry-level covariance, using the industry value-added share in each year. Industry-level covariance is calculated using the firm-level covariance.

B.2 Correlation Coefficients between Factor Distortions and Firm Size

Table B.2. Correlation coefficients between factor distortions and firm size

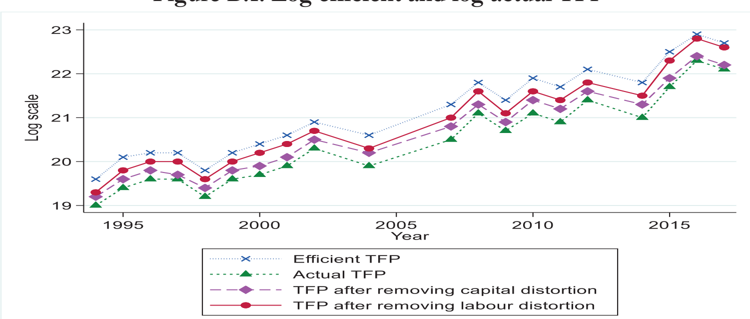
Year	Capital distortion	Labour distortion
1994	0.26	0.74
1995	0.30	0.73
1996	0.39	0.80
1997	0.26	0.79
1998	0.28	0.79
1999	0.27	0.82
2000	0.30	0.81
2001	0.23	0.83
2002	0.25	0.66
2004	0.33	0.74
2007	0.28	0.67
2008	0.23	0.71
2009	0.33	0.71
2010	0.35	0.72
2011	0.41	0.73
2012	0.29	0.61
2014	0.48	0.83
2015	0.22	0.67
2016	0.04	0.65
2017	0.15	0.66
Average	0.28	0.73

Notes. Annual values are calculated by weighted averaging industry-level correlation coefficients, using the industry value-added share in each year. Industry-level correlation coefficients are calculated using the firm-level correlation coefficients.

B.3 Log Efficient and Log Actual TFP

Figure B.1 depicts movements of log efficient and log actual TFP over time. All four series show increasing trends, along with the expansion of the economy over the period. Theoretically, log efficient TFP should place above the actual TFP, as it is the best possible efficiency, with zero distortions. The line representing log actual TFP after removing labour distortion is above the line that represents log actual TFP after removing capital distortion. The actual TFP line illustrates that the manufacturing sector is least productive when all distortions are considered.

Figure B.1. Log efficient and log actual TFP



Note. The entries are columns (1), (2), (3) and (4) of Table 2.

B.4 Robustness Tests: TFP Gains by Removing Distortions

Table B.3 presents the results of robustness tests for the calculations of TFP gains. Column (1) reports the reduced sample results by removing small firms that had less than 10 employees. In contrast, column (2) presents results after dropping very large firms that had more than 1,500 employees. In column (3), the elasticity of substitution increased from 3 to 4.[21] The results after inflating the wage bill to be equal to 0.35 times the output are in column (4).

Table B.3. Robustness results: TFP gains (%) by removing distortions

Year	Employees≥10		Employees≤1500		Sigma=4		Wage/Output=0.35	
	$\tau_K = 0$	$\tau_L = 0$	$\tau_K = 0$	$\tau_L = 0$	$\tau_K = 0$	$\tau_L = 0$	$\tau_K = 0$	$\tau_L = 0$
	(1)		(2)		(3)		(4)	
1994	21.2	44.8	21.3	47.2	28.5	58.2	15.7	52.3
1995	18.3	54.0	21.4	52.8	24.3	67.7	11.2	58.4
1996	18.2	50.4	20.5	52.2	23.6	67.4	12.2	57.5
1997	18.9	47.1	19.5	52.7	22.9	66.2	12.7	53.3
1998	21.1	50.4	23.8	47.3	27.2	61.7	15.1	53.6
1999	17.5	44.5	16.8	57.0	20.5	65.2	11.3	48.0
2000	18.8	65.5	20.6	67.8	25.2	75.7	10.7	80.2
2001	18.9	56.4	19.4	60.9	24.8	69.8	11.7	60.1
2002	17.8	52.8	20.9	55.4	20.3	63.5	9.5	67.1
2004	40.5	54.3	46.5	51.6	60.8	65.0	22.4	67.4
2007	27.9	61.9	24.2	51.8	32.4	68.9	19.4	62.7
2008	23.7	59.2	24.0	58.6	31.8	70.3	19.2	63.8
2009	31.3	55.7	28.8	57.2	36.9	71.2	23.8	69.5
2010	31.5	58.8	30.1	62.6	46.8	75.0	30.1	70.0
2011	27.5	57.2	37.2	65.8	45.2	78.1	20.9	72.1
2012	27.9	58.6	25.4	69.5	32.7	64.0	26.0	45.2
2014	40.9	62.9	48.3	77.2	45.1	84.2	23.7	81.5
2015	23.8	73.1	20.1	86.0	24.9	100.6	19.8	83.7
2016	12.8	57.0	9.9	70.4	15.9	75.6	9.5	61.5
2017	15.0	59.9	14.7	63.9	18.0	81.4	10.2	83.1
Average	23.7	56.2	24.7	60.4	30.4	71.5	16.7	64.5

Notes. To obtain the results, the same methodology used to obtain baseline results in columns (5) and (6) in Table 2 is followed. However, Table B.3 summarises only TFP gain results. Accordingly, TFP gains (%) = 100× [exp (Log TFP removing one distortion - Log actual TFP with all distortions) - 1].

Appendix C. Firm Characteristics and TFP Gains

Table C.1. Firm Characteristics and TFP gains (%) by removing distortions

Year	Western		Outside Western		Export Oriented		Non-export Oriented		Textile		Non-Textile		Incorporated		Un-incorporated		Young		Old	
	$\tau_K = 0$	$\tau_L = 0$	$\tau_K = 0$	$\tau_L = 0$	$\tau_K = 0$	$\tau_L = 0$	$\tau_K = 0$	$\tau_L = 0$	$\tau_K = 0$	$\tau_L = 0$	$\tau_K = 0$	$\tau_L = 0$	$\tau_K = 0$	$\tau_L = 0$	$\tau_K = 0$	$\tau_L = 0$	$\tau_K = 0$	$\tau_L = 0$	$\tau_K = 0$	$\tau_L = 0$
	(1)		(2)		(3)		(4)		(5)		(6)		(7)		(8)		(9)		(10)	
1994	20.1	39.9	21.9	46.7	48.6	80.7	18.0	55.2	22.7	46.0	19.2	57.0								
1995	18.6	48.6	21.1	68.0	19.8	52.0	18.1	55.8	19.8	52.0	17.9	63.2								
1996	19.2	40.7	21.8	57.8	19.2	52.0	15.1	59.6	19.2	52.0	18.7	59.4								
1997	17.2	40.3	20.4	54.9	18.9	49.9	12.4	62.5	18.9	49.9	20.2	66.0								
1998	17.9	47.8	24.7	69.6	22.1	48.4	15.0	53.7	22.1	48.4	22.5	60.4								
1999	18.2	35.1	35.3	56.8	17.5	52.4	14.5	76.5	17.5	52.4	19.2	65.8								
2000	13.3	47.5	39.4	89.5	20.5	58.8	16.8	70.7	20.5	58.8	19.9	76.2								
2001	17.9	40.1	35.0	51.1	19.9	58.4	12.3	89.5	19.9	58.4	17.8	73.8								
2002	15.4	43.8	38.9	44.2	17.8	52.8	9.9	83.8	17.8	52.8	14.0	72.5								
2004	41.8	53.2	38.5	54.2	40.8	54.2	42.2	61.2	40.8	54.2	43.7	56.9								
2007	29.7	44.0	34.4	48.1	30.1	53.9	22.7	78.0	29.6	57.4	23.1	72.8	30.8	52.8	32.4	41.1	27.0	64.6	33.5	36.4
2008	17.7	63.1	30.7	36.2	24.8	55.8	21.3	68.7	31.9	59.0	25.9	65.7	20.7	64.8	30.6	36.0	25.7	54.9	21.7	52.0
2009	21.9	56.9	39.2	46.7	29.6	57.4	25.6	62.8	31.0	61.3	31.1	67.1	28.5	50.4	51.2	55.6	32.1	51.0	27.9	57.1
2010	17.4	58.7	39.6	51.5	31.9	59.0	26.2	67.2	27.9	58.6	27.0	76.3	32.3	57.5	35.9	66.3	35.6	66.5	27.4	54.0
2011	26.0	43.8	46.3	67.5	31.0	61.3	41.4	80.2	36.0	68.8	32.2	80.1	38.2	45.9	33.1	84.2	25.1	54.5	38.1	66.4
2012	20.8	23.9	17.4	96.7	27.9	58.6	36.9	88.1	21.5	83.1	28.0	82.9	14.2	54.8	6.7	100.8	24.5	74.2	19.3	43.6
2014	16.0	49.5	97.6	77.4	36.0	68.8	40.0	48.5	12.8	61.6	50.5	59.1	13.7	68.3	127.1	45.6	54.0	92.4	70.9	38.2
2015	16.5	39.2	27.7	107.0	21.5	83.1	14.4	115.6	14.7	66.5	18.5	99.7	14.1	60.7	56.0	76.2	19.0	59.5	27.2	56.5
2016					12.8	61.6	4.2	77.3	4.2	77.3	10.5	69.3	12.0	56.8	17.0	62.2	13.8	80.1	12.4	44.7
2017					14.7	66.5	8.2	54.5	8.2	54.5	12.2	60.7	9.8	68.5	35.7	86.1	16.3	64.7	14.0	57.0
Average	20.3	45.3	35.0	62.4	25.3	59.3	20.8	70.5	21.8	58.7	23.6	69.3	21.4	58.0	42.6	65.4	27.3	66.2	29.2	50.6

Note. TFP gains (%) = 100× [exp (Log TFP removing one distortion - Log actual TFP with all distortions) - 1].

Appendix D. Productivity and Factor Distortions

As done in Section 5, the same regression analyses are performed but by using annual data with no year fixed effects and equations (D.1).

$$log(1+\tau_{ist}^{X}) = \beta_X log(TFPQ_{ist}) + \lambda_s + \varepsilon_{\tau_{ist}^{X}}. \quad \text{(D.1)}$$

Results are shown in Table D.1. In all years except for 2004, β_L for labour distortion is higher than β_K for capital. For all years β_L and β_K are positive. The results show that labour is more misallocated relative to capital.

Table D.1. Annual TFPQ and factor distortions

Year	Capital distortion β_K	Land distortion β_L
1994	0.452***	0.546***
	(0.016)	(0.006)
1995	0.478***	0.520***
	(0.167)	(0.006)
1996	0.454***	0.538***
	(0.186)	(0.006)
1997	0.448***	0.534***
	(0.018)	(0.007)
1998	0.461***	0.542***
	(0.168)	(0.007)
1999	0.428***	0.554***
	(0.193)	(0.008)
2000	0.436***	0.567***
	(0.020)	(0.008)
2001	0.407***	0.567***
	(0.019)	(0.007)
2002	0.478***	0.578***
	(0.031)	(0.011)
2004	0.560***	0.550***
	(0.026)	(0.008)
2007	0.422***	0.503***
	(0.022)	(0.009)
2008	0.448***	0.520***
	(0.170)	(0.000)
2009	0.475***	0.530***
	(0.187)	(0.008)
2010	0.466***	0.544***
	(0.020)	(0.008)
2011	0.485***	0.532***
	(0.209)	(0.009)
2012	0.401***	0.578***
	(0.057)	(0.021)
2014	0.502***	0.569***
	(0.023)	(0.009)
2015	0.443***	0.555***
	(0.258)	(0.102)
2016	0.284**	0.513**
	(0.029)	(0.010)
2017	0.293***	0.503***
	(0.023)	(0.009)

Notes. β_K and β_L are regression coefficients of log TFPQ with industry fixed effects, using equation (D.1). Standard errors clustered at the firm-level are displayed in parentheses. *** and ** 1% and 5%, significance, respectively.

Construction of Residential Property Price Indices Using the Hedonic Approach: An Application to the Residential Real Estate Market in Sri Lanka[1]

S R C L Gunawardhana[2]

Abstract

Monitoring real estate property prices is crucial for policy makers all over the world, particularly for central banks due to their interconnections with the monetary and financial system stability of an economy. As a result, compiling property price indices has increasingly gained attention from policymakers. However, compiling property price indices is believed to be difficult due to the highly heterogeneous nature of properties, requiring reliable data sources and a methodological approach that is different to those used in compiling other price indices. Against this backdrop, this paper attempts to compile price indices for residential properties in Sri Lanka with a view to supporting policymakers to monitor the price movements in the real estate sector. Considering the relatively less heterogenous nature of condominiums and stages of the buying and selling process, price indices are first developed for advertised condominiums and new condominiums using the Hedonic Regression based Rolling Window Time Dummy method, which is identified as the most suited property price index compilation method in the Sri Lankan context. Further, internationally accepted model specification improvement techniques and index smoothing techniques are also used in the study. Upon successful compilation of price indices for condominiums, the compilation process has been extended to cover the house and land markets of Sri Lanka.

Key Words: *Residential Property Price Index, Hedonic Regression, Asking Prices, Constant Quality Indices*

JEL Classification: *C2, C10, C43, C81, C82, C83, D12, E31, R21, R30, R31*

---

[1] The author is thankful to Mr. N O'Hanlon, Statistics Department of the International Monetary Fund (IMF) for providing technical assistance for this index compilation process and the Data for Decisions (D4D) fund of IMF for funding the technical assistance missions. The author is also thankful to Mr. A R K Wijesekara, Former Assistant Governor, Central Bank of Sri Lanka, Mr. G D P D Jayathilake, Former Director, Statistics Department of the Central Bank of Sri Lanka, Dr. (Mrs.) H K J Ekanayake, Director, Statistics Department of the Central Bank of Sri Lanka, Mrs. D G D I Ekanayake, Additional Director, Statistics Department of the Central Bank of Sri Lanka, Ms. R M M K Rathnayake, Deputy Director, Statistics Department of the Central Bank of Sri Lanka, Ms. P I Rupasinghe, Head of Economic Indicators and Statistical Investigations Division, Statistics Department of the Central Bank of Sri Lanka for their guidance and support extended throughout this study. In addition, the author would like to express her sincere thanks to all the staff members of the Economic Indicators and Statistical Investigations Division of Statistics Department of the Central Bank of Sri Lanka for contributing to the index compilation process in various stages. The views expressed in this paper are the author's own and do not necessarily reflect those of the Central Bank of Sri Lanka.

[2] Senior Assistant Director, Statistics Department, Central Bank of Sri Lanka (chamili@cbsl.lk)

1. Introduction

Developments in the real estate sector have an impact on the price and financial system stability due to its highly capital intensive nature. Thus, key interest rate changes can have an intermediate effect on the economy via real estate prices and might play a decisive role in the monetary policy transmission. In addition, affordability and price changes of properties have a direct impact on households, in terms of spending and indebtedness. Against this backdrop, real estate property price indices have gained importance, particularly for central banks, in making policy decisions. This is apparent in the case of Sri Lanka with the recent upswing in vertical developments in urban areas, especially condominiums. The unavailability of reliable indicators of real estate price movements in Sri Lanka was identified as a constraint to monitor developments in the real estate sector. To fill this gap, a condominium price index was initiated on experimental basis in 2016, and over time the methodology was improved while extending the compilation of price indices to house and land markets.

In some countries, administrative data collected by government agencies on property transactions provide suitable data sources, including information on individual property characteristics and prices, to underpin compilation of a price index. In others, however, the biggest challenge in compiling property price indices is the unavailability of required data. This is the case in Sri Lanka, where it is not possible to obtain timely granular data due to the absence of automated IT systems. Thus, surveys with property developers and real estate property advertisements published in property websites were identified as suitable alternative data sources to obtain sales information of properties. Accordingly, to extract advertised (asking) prices and other property characteristics published in property advertisements, web scraping techniques were implemented from January 2019 onwards on a monthly basis.

The expansion of property price indices compilation over time with various index compilation methods and different data sources are detailed in this paper. The data processing techniques and real estate property price index compilation methodologies applied in developing the indices were carried out based on internationally accepted best practices[3]. Accordingly, four residential property price index compilation methodologies, namely, the Hedonic Characteristic Laspeyres Price Index method, Hedonic Double Imputation Fisher Price Index method, Hedonic Imputation Fisher Price Index method, and the Hedonic Regression based Rolling Window Time Dummy method were used to experiment in developing the indices. It was observed that the Hedonic Regression based Rolling Window Time Dummy method is most suited for the thin real estate market in Sri Lanka. Further, it was identified that property price indices can be developed using prices available at different stages of its buying and selling process. Even though there are variations in price levels, the long term price trends are aligned to similar trends. In terms of price indices compiled for condominiums in Sri Lanka, the asking price index compiled using the advertised prices and the price index for new condominiums compiled using the prices at the point of the sale follows the same trend over time indicating a highly positive correlation. Thus, the outputs of this study fulfil the need of property price indices for Sri

[3] Indices were developed using the learnings gathered from three training sessions on developing Residential Property Price Indices (RPPIs) held at the Singapore Training Institute (STI) of the International Monetary Fund (IMF) and two technical assistance missions held at the Statistics Department of the Central Bank of Sri Lanka by an IMF expert. In addition, the methodologies introduced in the Handbook on RPPIs published by Eurostat of the European Commission and the relevant research papers published internationally were also referred.

Lanka to monitor the real estate sector prices and developments.

The rest of the paper is structured as follows: Section 2 contains the literature review on related past studies, Section 3 outlines the data collection, index compilation methodologies and the process used in the study, Section 4 consists of results and discussion while Section 5 concludes the study.

2. Literature Review

Securing access to a reliable data source and selecting a suitable index compilation methodology are crucial in compiling real estate property price indices. It is generally agreed that accurately reported transaction prices, collected at the completion of the transaction process, including granular information on individual property characteristics provide the most appropriate observations. However, property price information can be captured at different stages of the property buying and selling process, such as, at the point of advertising for sale, at the point of buyer-seller agreeing for a price, at the time of deed transformation and at the time of obtaining a loan if the buyer obtain a mortgage. In selecting a data source to compile the index, the availability of price determining characteristics information and the possibility of timely data collection are the most important factors to consider. In countries where official information systems are developed, administrative data sources can be used to obtain data for index compilation. But when such systems are underdeveloped, the compilers have to consider alternative data sources such as property agents and advertisements. In previous studies, it has been observed that the researchers have used different price types from different sources in compiling property price indices.

Shimizu *et al.* (2011) consider four types of property prices at different stages of its buying/selling process from different data sets. Accordingly, they use asking prices at which properties are initially listed for sale, prices when an offer is eventually made, contract prices reported by realtors, and finally the registry prices in the Greater Tokyo Area. Their study suggests that the prices collected at different stages of the buying/selling process exhibit substantial differences between price distributions but are still comparable since they follow similar trends. Therefore, different price types can be used in constructing house price indices, as long as they are quality adjusted in an appropriate manner. Further, the property characteristics that are listed in advertisements enable compilers to collect all required granular level data timely when advertisement information is used. Shimizu *et al.* (2010) use individual listings in a widely circulated real estate advertisement magazine in Tokyo, and Lyons (2019) examines the relationship between listed and transaction prices during Ireland's recent turbulent housing market cycle. The Irish study found that online listings represent a rich potential data source across economies and listed prices are an adequate substitute for transaction prices, even in relatively unstable market conditions.

In terms of index compilation methodologies, there are four standard methodologies that are discussed in the Handbook on Residential Property Price Indices (RPPIs) for real estate price index compilation. They are, (i) Stratification or Mixed Adjustment method, (ii) Hedonic Regression method, (iii) Repeat Sales method, and (iv) Appraisal-Based method. Among these, hedonic regressions are considered the most promising approach to control for changes in the quality of characteristics of properties transacted from period to period. Song and Wilhelsson (2010) also emphasise the necessity of using quality adjusted techniques in constructing house price indices due to their heterogeneous nature in conducting their study on compiling a price index for condominiums in Sweden.

Diewert *et al.* (2007) conducted a comparison between two main and quite distinct approaches to the measurement of hedonic price indices: time dummy hedonic indices and hedonic imputation indices. The study discusses that there may be reasons to prefer one against the other and accordingly, the hedonic imputation method is inherently more flexible since it constrains the parameters on the characteristic variables to be the same over the two periods under consideration. At the same time, it is likely to have less confidence in an index based on constraining the coefficients to be the same. In this sense the concern over the use of time dummy hedonic indices is warranted. Meanwhile, Hill *et al.* (2017) construct weekly hedonic house price indices using Rolling Time Dummy (RTD) method for Sydney and Tokyo. They note that RTD method tends to perform well with smaller data sets and computing high frequency indicators. Patrick (2017) also uses the rolling time dummy hedonic method for compilation of the official index for the republic of Ireland. De Haan (2004) discusses how the time dummy method fits into the matched model methodology in compiling these kinds of price indices. In addition, Diewert (2011) constructs house price indices for a small Dutch town following four alternative methods, namely, stratification method, time dummy hedonic regression method, hedonic regression imputation method, and additive hedonic regression method. He notes that the problem of change in historical results when new data become available, which is identified with many hedonic regression models, is addressed by the rolling window hedonic regression methodology. Shimizu *et al.* (2010) also discuss alternative hedonic housing price index compilation approach for condominiums in Tokyo Metropolitan Area and under its structurally restricted approach, they use the hedonic rolling window time dummy method for index compilation.

Thus, as per previous literature, the hedonic regression based rolling window time dummy method is recognised as an acceptable method in compiling residential property price indices. It not only corrects price changes for changes in the quality of items, but also allows the indices to incorporate matched and unmatched models. The hedonic regression based rolling window time dummy method has become popular with national statistical institutes in European countries such as Ireland, Luxembourg, Cyprus, and Malta in constructing the official House Price Index. Interestingly, each of these countries would be considered to have relatively thin housing markets.

Shimizu *et al.* (2010) discuss the different behaviours of house price indices depending on the estimation method. In their comparison between Hedonic and Repeat Sales Measures, it is found that there exists a substantial discrepancy in terms of turning points between hedonic and repeat sales indices, even though the hedonic index is adjusted for structural changes and the repeat sales index is also adjusted appropriately. The repeat sales measure is found to signal turning points later than the hedonic measure and this discrepancy cannot be fully removed even if the repeat sales index is adjusted for depreciation.

It is also important to look at the price determining characteristics considered in previous studies, in order to select suitable variables in compiling the property price indices. The explanatory variables used in the regression equations include price determining characteristics of properties in previous studies depending on the data availability. Accordingly, Diewert (2011) considers the age of the house, floor space area and land area in developing house price indices for a small Dutch town. In addition to these variables, Diewert and Shimizu (2014) consider the location of the property and number of bedrooms when compiling property price indices for Tokyo. Further, Shimizu *et al.* (2011) use the distance to the nearest station and travel time to the terminal station when observing prices of condominiums traded in the Greater Tokyo Area. When compiling property price indices for Sydney, Hill *et al.* (2017) consider explanatory variables such as the property type (i.e., detached or semi), number of bedrooms, number of bathrooms, land area, postcode, and exact address (longitude and latitude). To simplify the computations, they merge the number of bathrooms and number of bedrooms to broader groups (one, two, and three or more bathrooms; one or two, three, four, five or more bedrooms). When compiling a price index for condominiums in Sweden, Song and Wilhelsson (2010) consider factors such as size, number of rooms, location, floor level, whether the property has a balcony or not and whether the property have an elevator or not. Further, Patrick (2017) uses the floor area, number of bedrooms, and dwelling type to control for constant quality when compiling the property price index for Ireland. Accordingly, in most cases property characteristics and location attributes are considered as price determining factors of properties.

3. Methodology

The price index compilation process was initiated for condominiums in the Colombo district in 2016 using manually collected advertisement data from property websites and newspapers. The index was compiled according to the Hedonic Characteristic Laspeyres Price Index method from August 2016 to January 2017. Thereafter, a Hedonic Double Imputation Fisher Price Index methodology was applied in 2017. Meanwhile, a survey was introduced with condominium property developers in September 2017 on a quarterly basis mainly with the objective of collecting actual transaction information for new condominiums. Using this information, a separate index was also compiled for new condominiums following the Hedonic Double Imputation Fisher Price Index method until 2019. However, both indices compiled were highly volatile and seemed not to reflect market sentiments.

Against this background, experimental work was carried out following three alternative index compilation methods[4], and the Rolling Window Time Dummy (RWTD) method was selected for compiling indices for condominiums since this method is more suitable for markets with less data points as per the global practices (Hill *et al.* (2017)). Meanwhile, in January 2019, collection of property advertisements through web scraping, instead of the manual data collection, was initiated. Since 2019, an asking price index for existing condominiums in the Colombo district has been compiled using web scraped data, rather than manually collected data. In 2020, data preparation and index compilation processes using the RWTD method for both new and advertised condominiums were improved continuously. Subsequently, following the RWTD methodology, property price index compilation was expanded to cover land and housing markets of the Colombo district using advertisement information collected through web scraping.

3.1 The different residential property price index compilation methodologies used

In constructing a suitable residential property price index for Sri Lanka, four index compilation methodologies, namely, Hedonic Characteristic Laspeyres Price Index, Hedonic Double Imputation Fisher Price Index, Hedonic Imputation Fisher Price Index and Hedonic Regression based Rolling Window Time Dummy methods were taken into consideration and these methodologies are detailed in this section.

[4] i.e., Hedonic Characteristic Laspeyres Price Index method, Hedonic Double Imputation Fisher Price Index method, Hedonic Imputation Fisher Price Index method and Hedonic Regression based Rolling Window Time Dummy method.

3.1.1 The Hedonic Characteristic Laspeyres Price Index

A regression equation is fitted for each t^{th} period considering price *(P)* as the dependent variable and *K* number of selected price determining factors as explanatory variables *(X),* where *β* represents their coefficients.

$$\boldsymbol{P^t = \sum_{K=0}^{K} \beta_K^t X_K^t + \varepsilon^t} \tag{1}$$

Then using the forward selection method, the regression equation with k (k<K) number of variables is identified as the best regression considering the statistical significance of variables. Coefficients derived for the explanatory variables and the characteristics values related to a selected average property are applied to the best fitted model to impute a price for each period.

$$P^t = \sum_{k=0}^{k} \beta_k^t x_k^t \tag{2}$$

Based on these imputed prices for both the based period and the t^{th} period, Laspeyres Price Index P^{0t} is calculated for period "t", compared with the base period "0".

$$P^{0t} = \frac{\sum_{k=0}^{k_2} \beta_k^t \, x_k}{\sum_{k=0}^{k_1} \beta_k^0 \, x_k} \tag{3}$$

3.1.2 The Hedonic Double Imputation Fisher Price Index

The Hedonic Double Imputation Fisher Price Index is computed using the Hedonic Double Imputation Laspeyres Price Index and Hedonic Double Imputation Paasche Price Index. A regression equation is fitted for each t^{th} period considering log value of price *(P)* as the dependent variable and k number of selected price determining factors as explanatory variables *(X),* where *β* represents their coefficients.

$$ln(P^t) = \sum_{k=0}^{k} \beta_k^t X_k^t + \varepsilon^t \tag{4}$$

Based on the coefficients, a price is imputed for each condominium unit due to the unavailability of matched prices to be considered in index compilation. For example, a property sold in period 0 does not have a matched sale price in period t. Therefore, a matched price for the period t is imputed using period 0 characteristics and the period t regression coefficients. Accordingly, a constant quality index can be compiled.

The equation used to impute log price for n^{th} unit in t^{th} period:

$$ln(P^t) = \sum_{k=0}^{k} \beta_k^t x_{nk}^t \tag{5}$$

To impute price for a property, β values and x values need to be assigned appropriately and the exponential value of the right side of the equation 5 needs to be calculated. Those price estimates are used for index compilation as follows.

The Hedonic Double Imputation Laspeyres (HDIL) Price Index

Under this method, prices are imputed for base period properties for both period t and the base period. HDIL price index indicates the effect on prices of base period properties at period t, compared to their base period prices.

Hedonic Double Imputation Laspeyres Price Index

$$(P_{HDIL}^{0t}) = \frac{Exp((\sum_{n=1}^{N_0} \sum_{k=0}^{k} \beta_k^t x_{nk}^0)/N_0)}{Exp((\sum_{n=1}^{N_0} \sum_{k=0}^{k} \beta_k^0 x_{nk}^0)/N_0)} \times 100 \tag{6}$$

The Hedonic Double Imputation Paasche (HDIP) Price Index

Under this method, prices are imputed for properties available/sold during period t for both period t and the base period. HDIP price index indicates the effect on prices of properties at period t, compared to their base period prices.

Hedonic Double Imputation Paasche Price Index

$$(P_{HDIP}^{0t}) = \frac{Exp((\sum_{n=1}^{N_t} \sum_{k=0}^{k} \beta_k^t x_{nk}^t)/N_t)}{Exp((\sum_{n=1}^{N_t} \sum_{k=0}^{k} \beta_k^0 x_{nk}^t)/N_t)} \times 100 \tag{7}$$

The Hedonic Double Imputation Fisher (HDIF) Price Index

HDIF price index is calculated by taking the geometric mean of HDIL price index and HDIF price index as follows.

$$(P_{HDIF}^{0t}) = [P_{HDIL}^{0t} \cdot P_{HDIP}^{0t}]^{1/2} \tag{8}$$

3.1.3 The Hedonic Imputation Fisher Price Index

The Hedonic Imputation Fisher Price Index is computed using the Hedonic Imputation Laspeyres Price Index and Hedonic Imputation Paasche Price Index. A regression equation is fitted for each t^{th} period considering log value of price *(P)* as the dependent variable and k number of selected price determining factors as explanatory variables *(X)*, where β represents their coefficients, similar to the previous method. Under this approach, only the matched price is imputed.

The Hedonic Imputation Laspeyres (HIL) Index

The index is calculated based on the imputed prices of base period properties at period t, compared to their base period actual prices.

$$\text{Hedonic Imputation Laspeyres Index} = \frac{Exp(\sum_{n=1}^{N_0} \sum_{k=0}^{k} \beta_k^t x_{nk}^0)}{\sum_{n=1}^{N_0} p_n^0} \times 100 \tag{9}$$

The Hedonic Imputation Paasche (HIP) Index

The index is calculated based on the actual prices of t[th] period properties at period t, compared to their imputed prices for base period.

$$\text{Hedonic Imputation Paasche Index} = \frac{\sum_{n=1}^{N_t} p_n^t}{Exp(\sum_{n=1}^{N_t} \Sigma_{k=0}^{k} \beta_k^0 x_{nk}^t)} \times 100 \qquad (10)$$

The Hedonic Imputation Fisher (HIF) Index

The price index is calculated by taking the geometric mean of HIL Index and HIP Index.

$$(P_{HIF}^{0t}) = [P_{HIL}^{0t}.\ P_{HIP}^{0t}]^{1/2} \qquad (11)$$

3.1.4 The Rolling Window Time Dummy (RWTD) Method

Considering the standard version of the RWTD method with a window length of k+1 periods and the 1[st] period in the window is period t, the 1[st] step to estimate a semi-log hedonic model is as follows:

$$\ln p_h = \sum_{c=0}^{C} \beta_c Z_{hc} + \sum_{s=t+1}^{t+k} \delta_s D_{hs} \qquad (12)$$

Where h indexes the property sales/advertised information data set, p_h the price and c indexes the set of selected characteristics of the property sold/advertised (eg: floor area, floor level, bedrooms, bathrooms, etc.). The explanatory variables are given by the Z_{hc} matrix, while D_{hs} is a matrix of dummy variables that equals 1 when s is the period in which the property is sold/advertised, and 0 otherwise.

The change in the price index from period t+k-1 to period t+k is then calculated as follows:

$$\frac{P_{t+k}}{P_{t+k-1}} = \frac{\exp(\delta_{t+k})}{\exp(\delta_{t+k-1})} \times 100 \qquad (13)$$

Initial price index at period T, where T=t+k:

$$P_T = \frac{\exp(\delta_T)}{\exp(\delta_{T-1})} \times P_{T-1} \times 100 \qquad (14)$$

3-month moving average index at period T:

$$P_{T'} = \frac{P_T + P_{T-1} + P_{T-2}}{3} \qquad (15)$$

Final index series at period T, where the months/quarters of the base period are denoted by n and the number of months or quarters for the year are denoted by m:

$$P_{T''} = \frac{P_{T'} \times \mathrm{m} \times 100}{\sum_{n=1}^{m} P_{Base\ year\ n'}} \qquad (16)$$

3.2 The index compilation process

This section discusses the index compilation process followed in developing the price indices. The data collected by scraping contains some erroneous or implausible data and need to be trimmed to form acceptable data sets. Therefore, data trimming was carried out monthly considering standard deviations of data distributions to remove extremely small or large properties from the data set and to retain the data which would reflect the commonly used types of real estate properties in the Colombo District.

Thereafter, variable transformations were identified in experimental work and bedroom/bathroom categories and location groups were assigned. In order to use the RWTD method, data pooling was done such that 12 months data were used for monthly series and 4 quarters data were used for quarterly series. In estimating regression models, the semi-log regression models were run in two stages: the first stage to identify outlier observations using Cook's Distance and the second stage to rerun the models without outliers. The coefficients obtained from the second regression output were used for compiling the RWTD Index. Accordingly, the initial index was compiled as per equation 14 and smoothing the initial index by taking the three months moving average to minimise the volatility of the monthly indices was done as per equation 15. Finally, the base period of the indices was set by converting the index such that the periodical average values for the selected base year equals to 100 as shown in equation 16.

4. Results and Discussion

Improvements made to the residential property price indices for Sri Lanka overtime, the process followed in compiling the current indices which adhere to international best practices and the movements of the price indices are discussed in this section.

4.1 Evaluation of different index compilation methodologies

Different index compilation methodologies were followed since 2016 in compiling real estate property price indices. Initially, the Hedonic Characteristic Laspeyres Price Index method was employed for condominium sales advertisements collected manually on a monthly basis from January 2016 to March 2017. Under its methodology, prices were imputed for a hypothetical unit which represent a property with average characteristics. However, it was challenging to select the perfect average property to track over time considering the heterogeneous nature of the market. This reduces the representative nature of wider market conditions in the index.

Secondly, the Hedonic Double Imputation Fisher Price Index method was employed during 2017 and 2018 using manually collected condominium advertisements. Although this satisfies the constant quality requirement, at certain times, the Hedonic Double Imputation Laspeyres Price Index and Hedonic Double Imputation Paasche Price Index moved in different directions resulting in the final index to be incompatible with the market conditions. Accordingly, it was observed that the Hedonic Double Imputation Fisher Price Index method was also not performing well with the manually collected data set. Further it was noted that the inadequate number of observations used was a constraint in following this methodology.

Since the results obtained using above methodologies for the manually collected data set failed to produce market representative residential property price indices in the Sri Lankan context, web scraping techniques were used since 2019 to expedite the data collection process while expanding the coverage.

Further, applying more statistical techniques for data pre-processing, experiments were carried out to compile the price indices following the Hedonic Double Imputation Fisher Price Index method, Hedonic Imputation Fisher Price Index method, and Hedonic Regression based Rolling Window Time Dummy method. Accordingly, property sales advertisement information was collected via web scraping monthly and this enabled collection of sufficiently large data sets required for index compilation. Thereafter, data sets were trimmed using trimming thresholds determined based on data distributions and summary statistics. Then, a descriptive analysis was conducted to identify the variables with suitable transformations which have an impact on the property prices considering their adequacy and significance, in order to be included in the regression model used for index compilation. Further, outliers were removed using Cook's Distance and the three months moving average technique was applied for smoothing purposes to minimise the statistical noise in the index, as followed by Patrick (2017). With these improvements in data pre-processing, testing was carried out for the three alternative index compilation methods mentioned above and the Hedonic Regression based Rolling Window Time Dummy method performed as the most suited method in compiling price indices for condominiums in Sri Lanka. Therefore, the index compilation process was expanded to cover the lands and housing markets as well, using the same methodology.

In the index compilation process following the Hedonic Regression based Rolling Window Time Dummy method, data manipulation is not carried out manually since the pre-defined methodologies are introduced for each stage. If any changes/instability is observed in coefficients which require a variable transformation different to the existing one, the change can be made without adjusting the past index series since the index follows a chain based method. If a new variable is identified as suitable to be included in the hedonic regression equation, that can be included once data are available for a time span of a window length, without changing the past index series.

4.2 Index development following the Hedonic Regression based Rolling Window Time Dummy Method

After selecting the most appropriate index compilation methodology, the steps followed in developing the price indices using the Hedonic Regression Based Rolling Window Time Dummy Method and the outcomes would be discussed in this section. A preliminary analysis was carried out for each index below to select the price determining characteristics to be used in the hedonic regression equations as explanatory variables. The variables or their transformations were selected based on their statistical significance and improvements to model specifications.

4.2.1 The Development of Asking Price Index for Condominiums

Data collection and transformation

Condominium sales advertisements published were collected through web scraping since January 2019 monthly using python codes developed to be run in Google Colab and Jupyter Notebook. The location of the property, number of bedrooms and bathrooms, floor area, availability of air conditioning, swimming pool, furniture, sea view, and price were the information that was collected. In addition, the distance to the city of Colombo from each property was found using Google Maps. In terms of data transformations, condominiums were classified into six predefined groups based on their location. The six location groups were determined with market experts' opinion considering the proximities and market conditions. When the size of the condominium was not given in square feet but in a different measure, formulas were used to convert them. Further, both bedrooms and bathrooms were classified based on their number. A descriptive analysis was conducted to identify the variables which impact the prices of condominiums, in order to be included in the regression model used for index compilation.

Data trimming

The distribution of the condominium size (sq ft) was positively skewed for each month. Therefore, the data set was trimmed by removing condominiums lying beyond three standard deviations from both the lower and upper ends of the distribution monthly. Further, condominiums with more than six bedrooms/bathrooms and below 100 sq ft in size were removed considering the nature of commonly used condominiums in the Colombo District. However, more than 97 per cent of the advertisements were retained in the data set for each month after trimming.

Index compilation

The index was calculated according to the RWTD method using pooled data sets of 12 months. Initially, a semi-log least square regression was fitted considering the log of price of the condominium as the dependent variable and the price determining factors as explanatory variables. In terms of using the number of bedrooms and bathrooms in the regression equation, some experimental work was carried out and it was identified that it was better to consider as a categorical variable by grouping bedrooms and bathrooms, rather than a numerical variable. Among the alternative grouping combinations that were compared, it was decided to consider as two groups based on containing one or more bedrooms and bathrooms. Therefore, floor area, bedroom category, bathroom category, availability of a swimming pool, distance to Colombo city and location group and month of the advertisement were retained in the regression equation as explanatory variables. Among the property characteristics collected, the availability of air conditioning, furniture, and sea view were excluded from the equation due to their lack of statistical significance towards the price based on the results of the preliminary analysis.

Thereafter, outliers of the data set were removed using regression based Cook's distance method and another regression was fitted for the selected variables. More than 95 per cent of the data was retained in each 12-month window after removing the outliers. The exponential value of the coefficients for months of this regression was used for initial index compilation. After the initial index was compiled, a 3-month moving average index was calculated, and then the final index series was computed considering the base period as 2019, by setting the average of 3 months moving average series for 2019 equals to 100 and converting the monthly index numbers accordingly.

4.2.2 The Development of a Price Index for New Condominiums

Data collection and transformation

Sales information of new condominiums were collected quarterly through a survey of condominium developers since the third quarter of 2017. The information obtained includes the location of the condominium, number of bedrooms and bathrooms, floor area, floor level, whether it is in a mixed development project, availability of sea view and furniture, and the price. In addition, the distance to the city of Colombo from each condominium property was recorded manually using Google Maps. Following the similar techniques used in the asking price index for condominiums in terms of variable grouping, new condominiums were then classified into six predefined groups based on their location and into two groups based on containing one or more bedrooms and bathrooms.

Data Trimming

The distribution of the condominium size (sq ft) was positively skewed for each quarter. Therefore, in order to remove the extremely small or large condominiums from the data set, condominiums lying beyond three standard deviations from both the lower and upper ends of the distribution on quarterly basis were removed by trimming. Further, condominiums with more than six bedrooms/bathrooms and below 100 sq ft in size were removed considering the nature of commonly used condominiums in the Colombo District. However, more than 97 per cent of the data were retained for each quarter after trimming.

Index compilation

The price index for new condominiums was calculated according to the Hedonic Regression based RWTD method using pooled data sets of four quarters. Initially, a semi-log least square regression was fitted considering the log of price of the condominium as the dependent variable and floor area, floor level, bedroom category, bathroom category, whether it is in a mixed development project, distance to the city of Colombo and location group and quarter in which the condominium was sold, as explanatory variables. Among the available property characteristics, availability of furniture and sea view were excluded from the equation since they become statistically insignificant based on the results of the preliminary analysis. Thereafter, outliers of the data set were removed using regression based Cook's distance method and another regression was fitted for the same variables. More than 92 per cent of the data were retained in each 4-quarter window after removing the outliers. The exponential value of the coefficients for quarters of this regression were used for initial index compilation. After the initial index was compiled, the final index series was computed considering the base period as 2019 by setting the average of initial index values of the four quarters of 2019 equals to 100 and converting the quarterly index numbers accordingly.

4.2.3 The Development of Asking Price Index for Lands

Data collection and transformation

Land sales advertisements published were collected through web scraping using python codes developed to be run in Google Colab and Jupyter notebook since January 2019 monthly. The location of the land, number of perches to be sold, and price per perch were the information that was collected. The total price of the land was used in the index compilation process as the dependent variable. Therefore, when the price was not displayed in total and when the land size was not displayed in perches in the advertisement, formulas were used to convert them into the required term. Further, lands were also classified into six predefined groups based on their location.

Data trimming

The distribution of the land size was highly positively skewed for each month. Therefore, in order to remove the extremely small or large lands from the data set, lands lying beyond three standard deviations from the lower end and 0.5 standard deviations from the upper end based on the land size (perches) distribution on a monthly basis and below four perches, were removed considering the nature of residential lands in the Colombo District. However, more than 99 per cent of the advertisements were retained in the data set for each month after trimming.

Index compilation

The asking price index for lands was calculated according to RWTD method using pooled data sets of 12 months. Initially, a semi-log least square regression was fitted considering the log of total price of the land as the dependent variable and the number of perches, location group and month of the advertisement as explanatory variables. Thereafter, outliers of the data set were removed using regression based Cook's distance method and another regression was fitted for the same variables. More than 95 per cent of the data was retained in each 12-month window after removing the outliers. The exponential value of the coefficients for months of this regression was used for initial index compilation. After the initial index was compiled, a 3-month moving average index was calculated, and the final index series was computed considering the base period as 2019, by setting the average of 3-month moving average series for 2019 equals to 100 and converting the monthly index numbers accordingly.

4.2.4 The Development of Asking Price Index for Houses

Data collection and transformation

House sales advertisements published were collected since October 2019 monthly using python codes developed to be run in Google Colab and Jupyter Notebook. The location of the house, land plot size (perches), house size (sq ft), number of bedrooms and bathrooms, and price of the property were the information that was collected. When the house size was not displayed in sq ft and land size was not displayed in perches in the advertisement, formulas were used to convert them. Further, houses were classified into six predefined groups based on their location as per market experts' opinion. In terms of bedroom and bathroom grouping, they were classified into three groups based on having one or two, three or four, or more than four bedrooms and bathrooms, based on experimental level analysis.

Data trimming

The distributions of the land size and house size were highly positively skewed for each month. Therefore, in order to remove the extremely small or large landed houses from the data set, properties lying beyond three standard deviations from the lower end and two standard deviations from the upper end based on the land size (perches) distribution first and then on house size (sq ft) distribution on a monthly basis were removed by trimming. Further, houses below 100 sq ft in size and built on lands below four perches were removed considering the nature of commonly used houses in the Colombo District. However, more than 95 per cent of the advertisements were retained in the data set for each month after trimming.

Index compilation

The index was calculated applying the RWTD method on pooled data sets of 12 months. Initially, a semi-log least square regression was fitted considering the log of price of the property as the dependent variable and the size of the land, size of the house, bedroom category, bathroom category, location group, and month of the advertisement as explanatory variables. Thereafter, outliers were removed using Cook's distance method and another regression was fitted for the same variables. More than 91 per cent of the data was retained in each 12-month window after removing the outliers. The exponential value of the coefficients for months of this regression were used for initial index compilation and then a 3-month moving average index was calculated. The final index series was computed considering the base period as 2019, by setting the average of 3-month moving average series for 2019 equals to 100 and converting the monthly index numbers accordingly.

Details of the four types of real estate property price indices compiled covering the Colombo District of Sri Lanka are summarised in the tables below.

Table 1. Summary of the Real Estate Property Price Indices

	Condominiums		Lands	Houses
Index Name	Price Index for New Condominiums	Asking Price Index for Condominiums	Asking Price Index for Lands	Asking Price Index for Houses
Frequency	Quarterly	Monthly	Monthly	Monthly
Data Source	Condominium Market Survey	Sales Advertisements	Sales Advertisements	Sales Advertisements
Coverage	Colombo District	Colombo District	Colombo District	Colombo District
Methodology	Hedonic Regression based Rolling Window Time Dummy	Hedonic Regression based Rolling Window Time Dummy	Hedonic Regression based Rolling Window Time Dummy	Hedonic Regression based Rolling Window Time Dummy
Initiation	3rd Quarter 2017	January 2019	January 2019	October 2019
Base Year	2019	2019	2019	2019
Window Length	4 Quarters	12 Months	12 Months	12 Months
Variables	Log (price), floor area, floor level, bedroom and bathroom category, whether it is in a mixed development project, distance to the Colombo city and location group	Log (price), floor area, bedroom and bathroom category, availability of a swimming pool, distance to the Colombo city and location group	Log (price), the number of perches and location group	Log (price), land plot size, house size, bedroom and bathroom category and location group
Data Trimming	Condominiums within 3 standard deviations from both lower and upper ends of the floor area distribution on monthly basis and those with less than 6 bedrooms/ bathrooms, and above 100 sq.ft. were considered	Condominiums within 3 standard deviations from both lower and upper ends of the floor area distribution on monthly basis and those with less than 6 bedrooms/ bathrooms, and above 100 sq.ft. were considered	Lands within 3 standard deviations from the lower end and 0.5 standard deviations from the upper end based on the land size (perches) distribution on monthly basis and those with a land size of more than 4 perches were considered	Houses between 3 standard deviations from the lower end and 2 standard deviations from the upper end based on the land size (perches) and house size (sq.ft.) distributions, and those above 100 sq.ft. built on lands more than 4 perches were considered

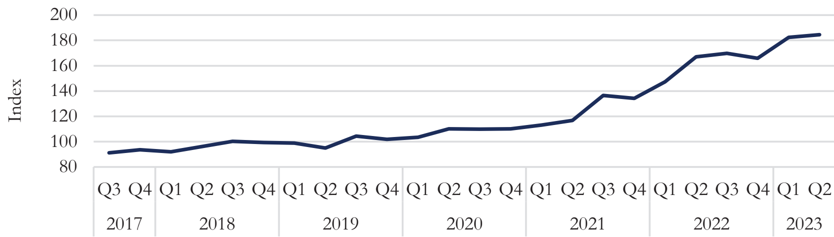
Table 2. Details of Data Used in Index Compilation

	Price Index for New Condominiums	*Asking Price Indices for* Condominiums	Lands	Houses
Average % of Data Retained after Trimming	97.9	97.9	99.6	95.4
Average % of Data Retained after Outlier Removal	92.3	95.5	95.0	91.1
Average no. of Data Points in a Window	981	7,215	30,576	38,500
Average R^2 of Hedonic Regression Models	0.95	0.74	0.71	0.78

4.3 Movements of real estate property price ındices compiled for Sri Lanka

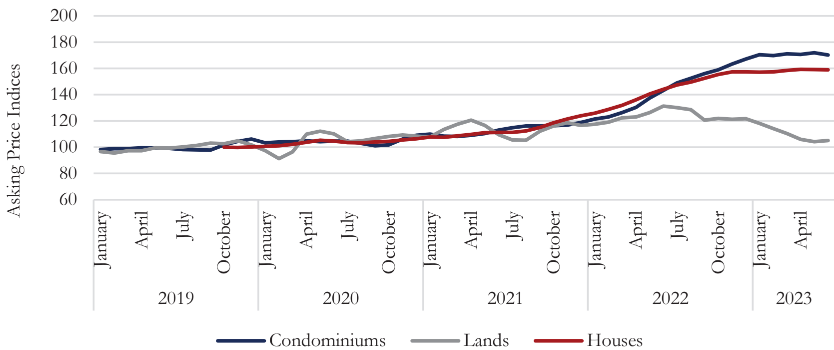
The latest available real estate property price indices compiled using the hedonic regression based RWTD method covering the Colombo Distrcit of Sri Lanka are presented below.

Figure 1. Price Index for New Condominiums (2019=100)



The index is compiled using sales information of new condominiums, which are obtained quarterly through the survey and thus reflects the primary condominium market.

Figure 2. Asking Price Indices for Condominiums, Lands and Houses (2019=100)



The asking indices for condominiums, lands, and houses are compiled using sales advertisements since 2019. Thus, the asking price indices for condominiums and houses represent both primary and secondary markets.

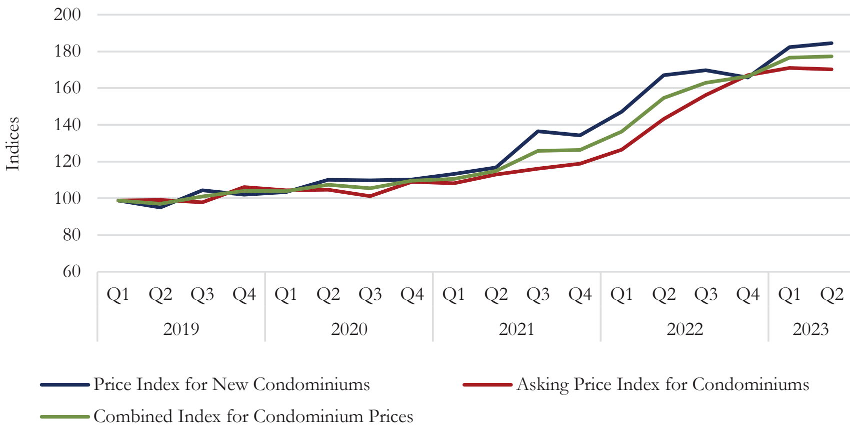
The four price indices reflect the trends observed in the data and the indices are more appropriate to observe long term trends, even if there are certain ups and downs in the short term. The regression equations used for index compilation explain more than 70 per cent[5] of the relevant data sets, which implies the level of precision of the indices. As per the index movements, new condominium prices are rapidly increasing while the asking prices of both condominiums and houses indicate stagnation at high price levels. Further, it is observed that the asking price indices for condominiums and houses follow similar trends with a strong positive correlation of 0.98. In the meantime, land prices have shown a declining trend during the last year.

In terms of the of the real estate market conditions in Sri Lanka, high construction costs that prevailed during the recent years owing to a depreciated currency, shortages of construction material due to

[5] Average R^2 of Hedonic Regression Models are shown in Table 2.

import restrictions, and high interest rates mainly led to an increase in prices of houses and condominiums. In addition, taxes imposed on condominium sales such as the 15 per cent value added tax and the 2.5 per cent social security contribution levy contributed further towards the price increases in condominiums. These factors created adverse market conditions which lowered the demand. Even against this background, property developers follow a wait and see approach and are not ready to reduce the prices to attract buyers due to their high-cost burden. As per real estate property experts', the real estate market is complex, and supply and demand are just one of many factors that impact property prices. Accordingly, the current market circumstances in Sri Lanka seems to be a cost push price increase rather than a demand push price increase. These market conditions are reflected in the price index for new condominiums. The price effect on new properties in the primary market is eventually reflected in the secondary market where prices of existing properties also remain stagnant at high levels. Further, the uncertainties in economic and political conditions affect the decisions of both buyers and sellers in purchasing properties for personal use and as an investment option. Consequently, the land market remains less active compared to the other types of properties as per market participants' views. Thus, the asking prices for condominiums, houses, and lands reflect the prevailing market conditions even at the recent turbulent environment.

Figure 3. Price Indices for Condominiums (2019=100)



In order to scrutinise the price indices for condominiums, a combined index is calculated by taking the geometric mean of the price index for new condominiums and the asking price index for condominiums[6]. As observed in Figure 3, the asking price index compiled based on the advertised prices and the price index compiled based on the actual selling prices follow similar trends over time and indicate a strong positive correlation of 0.97.

Similar studies are carried out by Shimizu *et al.* (2010) and Lyons (2019), demonstrating that the trends of prices collected at different stages of condominium buying and selling process align with each other even

[6] Since the asking price index is a monthly series which consists of three-month moving averages, the last month value of a quarter can be considered to match with the quarterly series of price index for new condominiums.

though there are differences in the price levels. Further, results of Lyons suggest that even in extreme market conditions, advertisement based price indices are a good measure of ultimate sale prices. Therefore, based on previous literature and the outcomes of this study, it is understood that the different price types collected at different stages of the property transaction process can be used in constructing property price indices.

5. Conclusion

In developing a property price index, a reliable data source, and a suitable price index compilation methodology are the principal components. The data source to be used for real estate property price indices can be collected at different stages of its buying and selling process and advertised prices can be used to develop asking price indices which performs as a good indicator to observe market price trends in the absence of official transaction data. Among the alternative residential property price index compilation methods, the Hedonic Regression based Rolling Window Time Dummy Method is identified as the best fitted method for thin real estate markets like Sri Lanka since it uses a pooled set of data for each window. The model specification improvement techniques, the index compilation method and smoothing techniques used are internationally accepted procedures followed by many countries. Thus, these indices are useful to observe trends in real estate property prices of Sri Lanka where no other reliable indices are available to monitor the sector.

References

De Haan, J. (2004). "Direct and Indirect Time Dummy Approaches to Hedonic Price Measurement". *Journal of Economic and Social Measurement*, vol. 29, no. 4, pp. 427-443.

Diewert, W.E. (2011). "Alternative Approaches to Measuring House Price Inflation". *Discussion Paper 10-10, Department of Economics, The University of British Columbia, Vancouver, Canada*.

Diewert, W.E., Heravi, S. and Silver, M. (2007). "Hedonic Imputation versus Time Dummy Hedonic Indexes". *IMF Working Paper. WP/07/234*.

Diewert, W.E. and Shimizu, C. (2014). "Residential Property Price Indexes for Tokyo". *Discussion Paper 13-07, School of Economics, The University of British Columbia, Vancouver, Canada*.

Eurostat, European Commition. (2013). *Handbook on Residential Property Price Indices (RPPIs).*

Hill, R. J., Scholz, M. and Shimizu, C. (2017). "Weekly Hedonic House Price Indices and the Rolling Time Dummy Method: An Application to Sydney and Tokyo". *Conference Proceedings at Hitotsubashi-RIETI International Workshop on Real Estate and the Macro Economy in Tokyo.*

Lyons, R. C. (2019). "Can list prices accurately capture housing price trends? Insights from extreme markets conditions". *Finance Research Letters,* vol. 30, pp. 228-232.

Patrick, G. (2017). "Redeveloping Ireland's Residential Property Price Index (RPPI)". *Conference Proceedings at Fifteenth Meeting of the Ottawa Group 2017 – International Working Group on Price Statistics.*

Shimizu, C., Nishimura, K. and Watanabe, T. (2010). "Housing Prices in Tokyo: A Comparison of Hedonic and Repeat Sales Measures". *Jahrbücher für Nationalökonomie und Statistik,* vol. 230, no.6, pp. 792-813.

Shimizu, C., Takatsuji, H., Ono, H. and Nishimura, K.G. (2010). "Structural and Temporal Changes in the Housing Market and Hedonic Housing Price Indices". *International Journal of Housing Markets and Analysis,* vol. 3, no. 4, pp. 351-368.

Shimizu, C., Nishimura, K.G. and Watanabe, T. (2011). "House Prices from Magazines, Realtors, and the Land Registry". *Property Market and Financial Stability, BIS Papers No.64, Bank of International Settlements,* pp. 29-38.

Song, H. S. and Wilhelmsson, M. (2010). "Improved price index for condominiums". *Journal of Property Research,* vol. 27, no.1, pp. 39-60.

ශ්‍රී ලංකා මහ බැංකුව
இலங்கை மத்திய வங்கி
CENTRAL BANK OF SRI LANKA