INTERNATIONAL MONETARY FUND

HOW TO

NOTES

How to Forecast Corporate Income Tax Revenues

Sebastian Beer, Brian Erard, and Tibor Hanappi

NOTE 25/10

How to Forecast Corporate Income Tax Revenues

NOTE/2025/010

Sebastian Beer, Brian Erard, and Tibor Hanappi*

Cataloging-in-Publication Data

IMF Library

Names: Beer, Sebastian, author. | Erard, Brian, author. | Hanappi, Tibor, author. | International Monetary Fund, publisher.
Title: How to forecast corporate income tax revenues: / Sebastian Beer, Brian Erard, and Tibor Hanappi
Other titles: How to notes.
Description: Washington, DC : International Monetary Fund, 2025. | Nov. 2025. | NOTE/2025/010. | Includes bibliographical references.
Identifiers: ISBN:

9798229031257	(paper)
9798229031271	(ePub)
9798229031318	(WebPDF)

Subjects: LCSH: Tax revenue estimating. | Corporations--taxation.
Classification: LCC HJ2351.B4 2025

RECOMMENDED CITATION: Beer, Sebastian, Brian Erard, and Tibor Hanappi. 2025. "How to Forecast Corporate Income Tax Revenues." IMF How to Note 2025/010, International Monetary Fund, Washington, DC.

Publication orders may be placed online, by fax, or through the mail:

International Monetary Fund, Publications Services
P.O. Box 92780, Washington, DC 20090, USA
Tel.: (202) 623-7430 Fax: (202) 623-7201
E-mail: publications@imf.org
bookstore.IMF.org
elibrary.IMF.org

*The authors would like to thank Alexander Klemm and Kevin Shoom, participants of an IMF seminar, staff of several tax policy units worldwide, and colleagues who provided guidance of the IMF's departmental review process for providing helpful comments and suggestions.

Contents

Boxes

Figures

Tables

How to Forecast Corporate Income Tax Revenues

Sebastian Beer, Brian Erard, and Tibor Hanappi

November 2025

Corporate income tax (CIT) collections are among the most difficult revenues to forecast—even with adequate staffing, comprehensive data, and a stable tax design. In practice, forecasting units typically operate under less ideal conditions. As institutional constraints take time to ease, this Note sets out a practical toolkit of methods to strengthen forecasting capacity across a wide range of country contexts. It outlines techniques that provide unbiased forecasts even when the impact of past reforms is only partially known, introduces approaches to account for ongoing and prospective policy changes to leverage time-series approaches, and highlights the potential efficiency gains achievable through structural modeling. A simple empirical assessment of forecasting specifications shows that parsimonious regression models, when backed by sufficient data, can improve prediction accuracy, even though the benchmark of assuming CIT revenues grow in line with GDP remains difficult to beat.

Introduction

Accurate revenue forecasting is the basis for effective fiscal planning. Baseline estimates inform stakeholders about the available budget envelope, and the potential need for reforms to implement the government's spending priorities. Forecasting errors can result in funding gaps that may require additional public debt issuance or supplementary budgets, or may lead to the revocation of approved projects, damaging the credibility of the budget process. Although governments are keenly aware of the need for accurate fiscal projections, the development of such forecasts can be a daunting task even in the best of circumstances, where granular revenue data and reliable macroeconomic forecasts are available, and tax regimes have been stable.

CIT revenues are characterized by extreme year-over-year volatility. During the pandemic, CIT-to-GDP ratios fell by an average of 24 percentage points year over year (Figure 1, panel 1). Although CIT collections have recovered since and now account for sizable shares of government revenues worldwide, their volatility remains elevated compared to other taxes: across a sample of 134 advanced economies, emerging markets, and low-income countries, CIT revenues have been 36 percent more volatile than overall tax revenue (Figure 1, panel 3).[1] When combined with the relative importance of those collections

[1] Volatility is measured using concepts from the capital asset pricing literature. The data includes information from 2000 to 2022. For details on the derivation, see Annex 1.

for total revenues, this implies that the CIT explains about one quarter of overall revenue volatility worldwide. Although CIT revenues are relatively more volatile in advanced economies, their greater contribution to total revenues in low-income and emerging market economies—where they account for 16–21 percent of tax collections—means that CIT fluctuations have a larger impact on overall revenue volatility in these countries.

Volatility by itself would not be a problem if it could be explained with variation in external predictors. Panel 2 of Figure 1 explores predictability, decomposing annual revenue growth across tax types and income groups into a portion that can be explained by GDP (yellow) and he residual (blue). Although GDP is a crude predictor for any tax, it is often the only one available to forecasting units. Moreover, the disparity of residual variation across taxes illustrates the challenge of forecasting the CIT: GDP growth explains about 45 percent of the variation in other tax revenues, but it accounts for only about 21 percent in the case of CIT. Conditioning revenue outcomes on general economic activity thus increases the relative volatility of the CIT.

Several factors contribute to the excessive volatility of CIT revenues. These include the higher sensitivity of income-based taxes rather than consumption-based taxes to economic fluctuations; the concentration of the corporate tax base among a relatively small number of leading firms and industries; the presence of loss-offset provisions that drive a wedge between current economic performance and tax liability; or the discretionary nature of tax prepayment requirements, which makes it difficult to anticipate tax revenue flows. Frequent changes in CIT provisions can exacerbate this volatility as tax policy changes not only affect the revenue generated from a given tax base, but they can also directly and indirectly (through behavioral responses) affect the size of the base itself. When corporate tax collections move with the business cycle, the CIT can help stabilize the economy by reducing the effective tax burden during downturns and increasing it during upswings. But pronounced CIT volatility complicates revenue forecasting.

In practice, both the approaches to and the accuracy of revenue forecasting are heavily dependent on institutional factors. Tight budget cycles or insufficient collaboration across tax policy units, legal and budget departments, tax administration agencies, and statistical offices can limit the availability of accurate and timely information. And where the independence of the forecasting process is not sufficiently protected, political aspirations and revenue targets can result in biased predictions, regardless of the underlying data quality.[2] Beyond securing adequate resources, collaboration between divisions can greatly enhance forecast quality, not least to ensure consistency in the assumptions underpinning distinct predictions that factor into the final revenue forecast.[3] Partnerships with research institutes and academic institutions can improve analytical capabilities, bring specialized expertise, and supplement internal resources. Some countries have established dedicated forecasting committees, comprising government officials, external experts, industry representatives, and think tanks, which serve to evaluate studies,

---

[2] Tax administrations commonly set internal revenue targets to guide their activities. In countries where revenue predictions are used as de facto revenue targets, and where meeting these targets is linked to staff's compensation programs, revenue forecasts tend to be overly cautious and therefore biased. By contrast, political considerations in other countries (such as a motivation to justify a tax rate reduction or a large spending package) might lead to pressures on the forecasting unit to produce optimistic revenue projections.

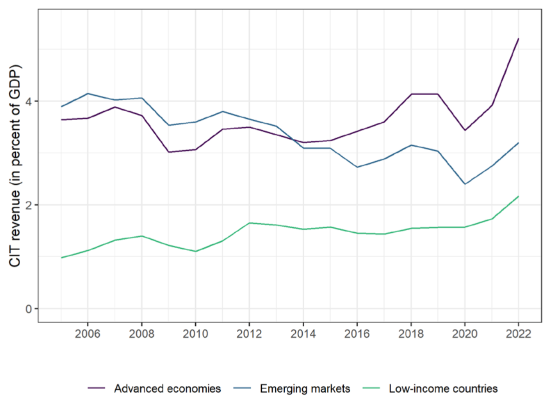
[3] In cases where multiple official research bodies, within and beyond the Ministry of Finance, are tasked with forecasting economic activities, a requirement for forecasts to be consistent with each other may be warranted. For example, economic profit estimates underlying the corporate income tax (CIT) revenue forecasts should be consistent with the official economic outlook.

consolidate diverse perspectives, and produce consensus-based forecasts. Improving collaborative structures, expanding access to timely and reliable data, and investing in the recruitment and training of qualified personnel is essential for the accuracy of tax revenue forecasts.

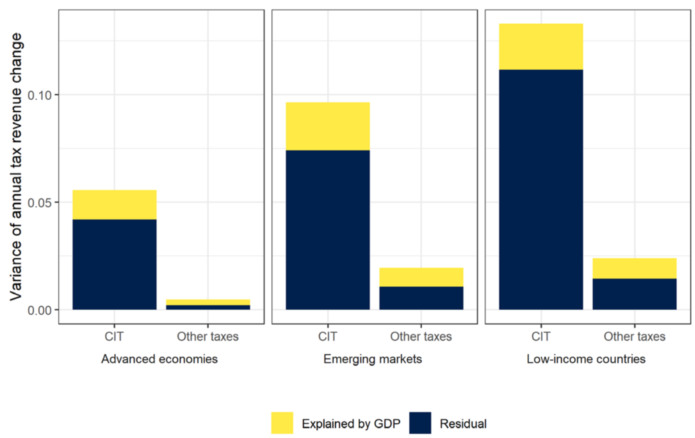
Figure 1. CIT Revenues: Contribution to Total and Relative Volatility

1. CIT Revenues

(In percent of GDP)



2. (Residual) Variation in Tax Revenue Growth



3. CIT Volatility by Country Category

Country Category	Relative Volatility of the CIT	Contribution to Total Tax Revenue	Contribution to Overall Volatility
Advanced economies	1.65	0.12	0.20
Emerging markets	1.36	0.21	0.29
Low-income countries	1.32	0.16	0.21
All countries	1.36	0.17	0.24

Source: WEO Database; WoRLD, and author computations.

Notes: Panel 2 presents variation of annual growth rates of the CIT and the aggregate of other tax revenues. Yellow areas represent the portion that can be explained by variation in GDP growth; dark areas represent the residual, unexplained variation. The table in panel 3 describes CIT volatility using two measures: the volatility of CIT revenues relative to the average volatility of other tax instruments; and the share of overall tax revenue volatility for which the CIT is responsible. All statistics are computed based on general tax revenues from all levels of government using the IMF World Revenue Longitudinal Database. CIT = corporate income tax.

This Note focuses on technical aspects that can support forecast accuracy of CIT revenue estimates for a wide range of countries. Resources, data availability, and, relatedly, the sophistication of modeling approaches vary widely across countries. In the best-case scenario, forecasters can rely on detailed information drawn directly from corporate tax returns, high-frequency economic indicators, and comprehensive costings of past policy reforms. With such inputs, statistical models can be devised that closely track the dynamics of the corporate tax base. In many countries, however, aggregate cash receipts are the only data series available, the effects of past reforms cannot be disentangled from economic shocks, and limited staff and analytical capacity constrain the use of sophisticated econometric techniques. Since institutional constraints take time to resolve, this Note reviews a spectrum of forecasting approaches that can be deployed depending on data availability and institutional capacity. The discussion of available methods assumes a basic understanding of econometrics.[4]

Although quantitative methods are an important starting point, they are insufficient on their own. Even the most sophisticated frameworks cannot capture all relevant information. Qualitative insights—such as periodic reports from industry associations or the guidance provided in investor communications of large companies—can play a valuable role in testing the plausibility of quantitative forecasts. In addition, CIT revenues represent only one component of total government revenue. Since overall revenue performance is the central metric in budget planning, assessments of the CIT forecast cannot be made in isolation. Overly optimistic projections in other revenue streams may justify a more cautious stance on the CIT outlook, ensuring that aggregate revenue forecasts remain credible and balanced.

The remainder of this Note is structured as follows. First, the "Forecasting Architecture" section outlines high-level design choices in the modeling architecture, focusing on which dependent variable should be used and how many models are ideally combined. Second, the "Accounting for Policy Change" section emphasizes the need to quantify the revenue effects of past policy changes and sets out methods to incorporate such information in the baseline estimate. Third, the "Robust Forecasting Techniques" section reviews robust forecasting methods that can be used when the impact of past policy changes is unknown. Fourth, the "Time-Series Methods" section reviews time-series methods. Fifth, the "Empirical Evaluation of Selected Methods" section evaluates forecasting performance of selected forecasting approaches. Finally, the "Documentation and Improving Forecast Performance" section highlights the importance of model documentation and improvement.

Forecasting Architecture

Budget preparation typically requires forecasts of cash receipts for the remainder of the current and for the subsequent fiscal year.[5] Although many countries develop CIT revenue forecasts that extend a number of years into the future, this Note focuses on projecting revenue for the current and subsequent fiscal years. Extension of the methods discussed to additional forecast years is, in most cases, reasonably straightforward. Forecasts are typically derived in the following two steps:

1. **The baseline estimate** provides a projection of future total cash receipts under current law. It is typically computed using aggregate information on tax liabilities or receipts, informed by official macroeconomic projections, and computed following a standard schedule that is determined by the budget calendar. Initial baseline estimates are frequently updated later in the year when new information, such as in-year revenue collections, is obtained. The difference between revenue

[4] Free online training on the fundamentals of Revenue Forecasting and Analysis is available at www.imf.org/en/Capacity-Development/Training/ICDTC/Courses/RFAx.

[5] Some countries, such as New Zealand, have adopted accrual accounting for use in government. Although this practice facilitates the estimation of tax liabilities, because economic predictors and recorded revenues both reflect current conditions, it does not remove the need to convert liabilities into cash receipts, because cash management remains an important component for fiscal stability.

collections today and the baseline are largely a function of economic change, holding tax design constant.[6]

2. **Costing estimates** quantify the likely budgetary effect of reform options that are being contemplated. They often leverage taxpayer-level information, sometimes using microsimulation models. Such information is necessary to focus on specific subsets of the population affected by various corporate tax reforms. The sum of individual costing estimates is added to the baseline estimate to derive a prediction of the available budget envelope. In contrast to the baseline estimate, costing estimates are often computed throughout the year, when new reform ideas emerge.

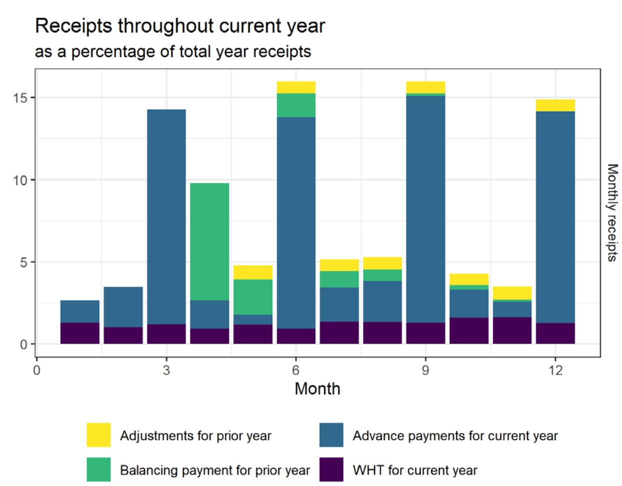
Baseline estimates build on a combination of models, each involving distinct data inputs and modeling approaches. For instance, revenue projections for the current fiscal year typically build on information on how much revenue has been collected year-to-date, while estimates for the subsequent year require more complex methods, as they need to be consistent with the official macroeconomic outlook and account for the impact of any proposed policy changes. In some countries, tax policy units produce CIT forecasts disaggregated by sector, region, or specific corporate taxpayers, when these groups contribute significantly to overall revenue collections or if policymakers have an intrinsic interest in these subpopulations.

This section highlights high-level design choices, including (1) the choice of the dependent variable, and (2) whether that variable should be estimated directly (reduced-form specification) or if it should be further broken down into subcomponent estimates (structural decomposition).

What Should Be the Primary Dependent Variable?

Revenue-forecasting units often face a trade-off in whether to rely on more recent or informative data. Revenue administrations typically share information on aggregate corporate cash collections on a timely basis. However, access to recent information on tax liabilities is often more restricted, because of either data privacy and confidentiality restrictions, or a lack of data sharing protocols.

Cash receipts collected from corporations consist of several subcomponents, some of which are unresponsive to current economic conditions. In most countries, corporate taxes are paid in installments throughout the year, with last year's liability determining the size of quarterly or semi-annual payments. Collections throughout the year typically also include withholding payments, final self-assessment payments, as well as penalties and interest charges, and only a small portion of these subcomponents respond instantaneously to changing economic conditions or to a changed



[6] When current legislation includes planned changes to the tax system parameters, the baseline forecast reflects the anticipated impact of those changes, and proposed tax system reforms are always forward-looking—and thus a function of future economic conditions.

tax design (Box 1). Accurate modeling of cash collections would thus require incorporating several years of economic and policy variables as explanatory variables. Since tax collections are only observed over a limited number of years, this strategy would quickly strain statistical power.

By contrast, tax liabilities are closely linked to the current economic and policy environment, and therefore more straightforward to explain with statistical models. For instance, while a prediction of tax collections for year t would, at least, require accounting for the effective tax rate (ETR) in year $t - 1$ and t, liabilities are largely determined by the ETR applicable in that year. As a result, liabilities can be modeled with greater accuracy. Receipts can then be computed based on projected tax liabilities. For example, advance payments received in the current year depend both on the tax liabilities of the previous year—because of tax system rules that link required prepayments to prior-year liabilities—and on the anticipated liabilities for the current year, as taxpayers are often allowed some discretion to adjust prepayments when they expect significant changes in profits. Although the precise relationship between advance payments and current and past liabilities is not known, it can be estimated using a regression analysis, which explains current-year cash receipts as a linear function of current and past year estimated tax liabilities. Similar regressions can be estimated for other receipt components.

Box 1. Components of Corporate Income Tax Receipts

While their relative importance differs across countries, the following components typically contribute to overall CIT collections:

Advance payments (or "estimated tax payments") of this year's tax liability are remitted periodically throughout the calendar year, often accounting for 60 to 70 percent of total CIT collections. Such payments are typically required on a monthly or quarterly basis, although alternative institutional arrangements, such as two or three required payments per year, are not uncommon. Reported CIT liabilities for the prior tax year typically determine the required advance payments for the current year. While most countries permit taxpayers to reduce payments if profits for the current tax year are expected to fall below the level for the prior year (in some cases only after applying for and receiving tax agency approval), the incentive to adjust prepayments differs depending on tax design aspects (such as interest and penalty charges if the prepayments turn out to be insufficient).

Withholding tax payments on certain forms of corporate income, such as dividends, interest, or royalties, are collected at higher frequency throughout the year. These prepayments are creditable against the final CIT liability if the taxpayer is a domestic resident. The importance of withholding tax payments differs across countries, but they can contribute up to 15 percent of aggregate CIT liabilities.

Final self-assessment (or balancing) payments reconcile the difference between the actual tax liability, and the advance payments and withholding tax payments that have been made previously. Filing requirements generally allow taxpayers to pay any outstanding amounts well into the subsequent tax year, which means that balancing payments made in the current year often depend on the tax system design and economic conditions that were present two years earlier (through the backward-looking determination of required advance payments). Balancing payments are typically small, contributing about 10 percent of total CIT collections.

Adjustments, including late payments, interest, and penalties for underreporting of prior year taxes, typically account for less than 5 percent of total annual CIT liability. The portion of overall adjustments attributable to enforcement activities often pertains to returns filed several years prior.

If recent information on aggregate tax liabilities is not available to the forecasting unit, an important first step is to obtain information on disaggregated cash receipts from the revenue administration. Disaggregated data improves forecasting in two ways. First, it enables direct modeling of specific cash receipt components. For instance, cross-border withholding tax collections depend primarily on current economic activity. As past policy changes can be easily controlled for (with revenues being the product of a rate and a base), a separate model for cross-border withholding tax collections can improve overall accuracy. Second, it allows for the approximation of aggregate liabilities as the sum of this year's balancing payments, prior year's withholding tax collections, and prior year's advance payments. This constructed liability series can then be used to strengthen forecasts of future cash receipts.

How Many Models Are Needed?

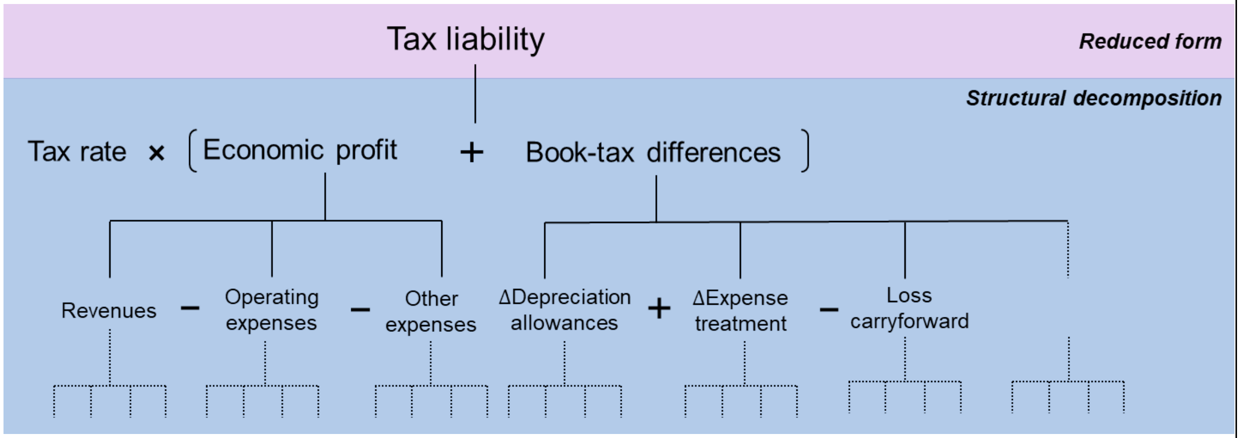
For any given year, there are multiple ways that distinct estimators can be combined to derive an estimate of total tax liabilities. At one end of the spectrum, a single estimate of aggregate liabilities could be derived. At the other end, there are two dimensions how this aggregate could be sliced:

1. **Disaggregation across tax liabilities.** Aggregate tax liabilities can be derived as the sum of projected tax liability subcomponents. For instance, the US Congressional Budget Office derives separate forecasts of key book-tax differences to convert projected aggregate economic profits, based on the historical profit series developed by the Bureau of Economic Analysis, into a forecast of aggregate taxable profit.[7] An ETR is then applied to produce a forecast of aggregate CIT liability.[8] Figure 2 illustrates this structural estimation approach, showing a decomposition of the tax base as the sum of economic profit and a residual (book-tax differences), which reflects the combined effect of tax policies—such as depreciation allowances, the treatment of expenses (including research and development credits or interest deductibility limitations), and the generosity of loss carryforwards—in driving a wedge between economic profit and taxable income. Each component of the forecast is a national aggregate measure, thus summing across all firms.
2. **Disaggregation across firms.** Alternatively, aggregate tax liabilities can be derived as the sum of tax liabilities projected at the firm-, sector-, or industry-level. Although taxpayer-level projections are relatively uncommon—with France, for example, being a notable exception—several advanced economies derive separate sector-specific estimates. For instance, resource-rich economies, such as the UK or Canada, often employ separate models to forecast CIT revenues from the natural resource sector, while countries with large financial sectors, such as the UK, sometimes estimate separate models for banks and insurance companies. Sector-specific models themselves can vary in their complexity: they can be derived using simple elasticity-based predictions or they can be based on a detailed structural framework that builds on a granular breakdown of a sector's overall tax liability.

[7] This first step already comes with the formidable challenge of splitting aggregate economic profit into positive profit and negative losses.

[8] Notably, while the US Congressional Budget Office derives baseline revenue estimates, revenue projections that include anticipated policy changes are then prepared by the Joint Committee on Taxation.

Figure 2. Reduced Form versus Structural Approaches



Note: Illustrative decomposition of aggregate tax liability, excluding capital gains and withholding taxes.

Naturally, these two dimensions can be combined. For instance, the UK's Office of Budget Responsibility differentiates between three sectors (industrial and commercial companies, life insurance companies, and banks), and projects for each of these components of income (such as profits and capital gains) and deductions (loss carryforward, capital allowances, and so on) using appropriate economic determinants. Predictions are derived for liabilities, which are then converted into a monthly profile of cash collections—crucial for accurate cash management of the government—using historical patterns between liabilities and cash collections.

In general, fewer estimators reduce complexity and data requirements. When there is only one dependent variable, model testing, selection, and data organization is much easier. Restricting the focus to fewer variables also increases the likelihood that more recent information can be leveraged, so that short-term forecasts are based on more complete information for the current and recent prior years. By contrast, disaggregated forecasts require more granular information, ideally leveraging tax returns.[9] Although data privacy laws and administrative policies often restrict the sharing of tax records even between government units, they should not present a significant barrier for forecasting: in most cases, forecasts can be more readily computed based on aggregated information. For example, Korea has strict confidentiality laws that bar the tax revenue-forecasting unit from directly leveraging tax returns. However, the National Tax Service compiles highly granular information on the components of tax liabilities across various dimensions, including company size or sector. These can be leveraged to predict components of tax liabilities. Nevertheless, the preparation of aggregate tax return information can be time-consuming,[10] especially where data processing or data sharing routines have not been established.

[9] Although financial reporting information can help gauge the approximate size of some tax base components, such information is typically only available for large public firms, and differences in financial accounting and tax reporting requirements can make publicly available information only a crude proxy for the relevant tax components.

[10] For example, as tax liabilities only accrue for profitable entities, the aggregation procedure should differentiate between companies with and without positive taxable income. Overall CIT liability is then predicted largely from the aggregate taxable income of profitable companies.

However, disaggregated revenue forecasts—by sector, source, or component of CIT liability—can improve forecast accuracy and may provide policy-relevant insights. For instance, information on projected shortfalls of specific revenue subcomponents can inform targeted policy adjustments. But even when there is no intrinsic interest in the more granular revenue estimate, disaggregation can lead to a more accurate estimate of the total for two distinct reasons:

1. **Better-tailored predictive variables.** Disaggregated forecasts enable the selection of more targeted explanatory variables, improving the accuracy of baseline projections. Box 2 illustrates this insight using a simple conceptual example, where knowledge on the drivers of tax liability components improves the accuracy of coefficient estimates by allowing the incorporation of exclusion restrictions. Given that CIT collections are often concentrated among a small subset of large firms, any economic shock, industry-specific downturn, or other event affecting these firms' profitability can disproportionately influence overall revenues. Tailored models can incorporate drivers of profitability among such key firms, using firm-specific information such as significant developments reported in the media or quarterly financial statements. In Germany, for example, state-level revenue forecasts incorporate information on developments affecting key regional taxpayers, such as car manufacturers.
2. **Increased options to control for policy change.** More granular modeling also expands the ability to control for past policy changes, thus reducing the potential of an omitted variable bias for baseline projections, and it increases possibilities for the incorporation of costing estimates that are fully consistent with the baseline. For example, if corporate investments and depreciation allowances are modeled separately, both past and future changes in the (weighted average) depreciation rate could be readily incorporated. By contrast, the policy change would be more difficult to control using aggregate information on tax liabilities, because changes in depreciation rates affect total liabilities in a nonlinear manner.[11]

[11] The reduced-form estimate could also be extended to account for these policy changes. However, the reduced form equation would now need to include the depreciation rate as well as its interaction with the explanatory variable of investments, thus losing more degrees of freedom.

Box 2. Efficiency Gains from Structural Modeling

Consider a simplified computation of aggregate tax liability, where in year *t* a tax at a rate of τ is levied on companies' aggregate gross sales, *S*, net of depreciation allowances, which are a share δ of aggregate new investment, *I*. Aggregate tax revenue in year *t* (T_t) is then described by the following expression:

$$T_t = \tau(S_t - \delta_t I_t).$$

We assume aggregate sales and investments follow the stochastic processes $S_t = \theta_1{}'x_{1t} + \varepsilon_{1t}$ and $I_t = \theta' x_{2t} + \varepsilon_{2t}$, respectively, where the x terms represent sets of observable economic predictors, and the epsilons are idiosyncratic errors. Suppose that policy remained constant in the past.

A reduced-form estimate can be obtained by simply regressing overall tax liability against the two sets of economic predictors.[1] Denoting estimated coefficients by hats, such an estimate would be constructed as follows:

$$\widehat{T_t}^{rf} = \hat{\beta}_1{}'x_{1t} + \hat{\beta}_2{}'x_{2t}.$$

A structural estimator, by contrast, would involve the following two steps. First, separate regressions are respectively used to construct conditional predictions of sales ($\hat{S}_t$) and investments ($\hat{I}_t$).Second, the tax parameters are applied to these estimates to project overall tax liability as follows:

$$\widehat{T_t}^{st} = \tau\left(\hat{S}_t - \delta_t \hat{I}_t\right)$$

Whether the structural approach provides efficiency gains depends on what information is used to predict sales and investments. If the full set of explanatory variables $\{x_1, x_2\}$ is used for estimating both outcome variables, then the result of the structural approach is identical to direct estimation. However, if the correct exclusion restrictions are used, so that sales are modeled as a function of x_1 while investments are modeled as a function of x_2, and if the structural errors ε_{2t} and ε_{1t} are assumed to be uncorrelated, the structural and reduced form predictors of aggregate tax liability will both be unbiased, but the former will be more efficient (that is, more precise).

The intuition behind this result is that the exclusion of irrelevant predictors increases the precision of each sub-component estimate. Since total revenue is a linear combination of these uncorrelated subcomponents, the combined estimate is also more precise. Notably, if the structural errors are correlated, the superiority of the structural approach, which estimates the equations separately, is no longer assured. When the errors are negatively correlated, tax revenue is less volatile than the subcomponents and the reduced form approach may produce more efficient estimates.

[1] If the sets x_1 and x_2 include overlapping explanatory variables, the duplicates would be dropped from one of the sets.

Whether these forecast improvements can be realized in practice is uncertain. It depends on the extent to which the data generating processes for specific revenue components can be adequately captured using available data and modeling strategies. The best predictors and functional forms for each component are typically not known. Rather, they are selected through a process of model fitting and testing. In such cases, it is less clear whether a disaggregated modeling approach will lead to more or less precise forecasts of overall tax liability.[12] Moreover, in many countries, legal requirements (for example, mandatory aggregation across certain dimensions) imply that disaggregated models often rely on more outdated information. This, in turn, can require forecasting over longer horizons to produce annual point estimates, thereby compounding estimation uncertainty. Thus, although model disaggregation may offer potential benefits, it also introduces greater complexity into the forecasting process, with no guarantee of improved accuracy.

The forecasting architecture should thus favor simplicity and only incorporate additional estimation layers where these provide clear benefits. When building a structural, or disaggregated model, efforts should first focus on those components that can be most productively modeled, that are of specific interest for policy or administrative reasons, or on those that are critical for the determination of aggregate tax collections. For instance, when aggregated components of tax liabilities can be accessed, depreciation allowances, loss carryforwards, and tax incentives often have an outsized impact on total CIT liabilities.

Accounting for Policy Change

The robustness of an empirical baseline estimate depends largely on the extent that past policy changes can be identified and controlled for in the estimation process. Revenue forecasts for future years need to be informed by "external" predictors—such as GDP, gross operating surplus, or imports—to ensure consistency with the government's macroeconomic framework. Statistical forecasting methods work by replicating relationships that were observed in the past. Where periodic tax reforms have altered the structural relationship between tax receipts and external predictors over time, and when these reforms are not adequately captured through statistical models, the resulting prediction will fail to provide an accurate estimate of baseline revenues; in statistical terms, this is so as the model suffers from an omitted variable bias (see Box 3).

This section discusses methods of how the baseline estimate can account for past policy reforms and highlights issues for incorporating future policy measures. Two widely used methods for controlling for past policy changes are data-preprocessing and regression-based modeling. While data preprocessing is particularly relevant for reduced-form estimates of aggregate tax liabilities, regression-based modeling is particularly useful when more disaggregated information is available. Importantly, the choice of method for controlling past policy changes also shapes the way in which future policy changes can be incorporated into the forecasting framework.

[12] In practice, the error terms of the specifications for various sub-components of aggregate tax liability will tend to be correlated, in which case the estimation of separate models for each component will fall short of achieving fully efficient predictions of future liabilities.

Box 3. Omitted Variable Bias from Neglecting Policy Change

The neglect of policy changes in revenue forecasting is a classic case of omitted variable bias. Suppose logarithmic tax revenues follow the process:

$$T_t = \tau_t + b_t + \varepsilon_t$$

where τ_t is the logarithmic average effective tax rate (ETR) that was levied in year *t*, b_t is the logarithmic tax base, and ε_t is an idiosyncratic error. For simplicity, assume that the tax base is perfectly proportional to GDP, such that $b_t = \delta + \mathrm{GDP}_t$, where again the variables are expressed in logs. This implies that the elasticity of CIT revenues with respect to GDP is unity: if GDP rises by 10 percent, tax revenues should also increase by 10 percent.

Now consider a regression that omits the ETR and estimates the elasticity in first differences as

$$\Delta T_t = \beta \Delta \mathrm{GDP}_t + u_t$$

where Δ is the first-difference operator. The estimated coefficient on GDP would then be

$$\beta = \frac{\mathrm{Cov}(\Delta\tau + \Delta b, \Delta\mathrm{GDP})}{\mathrm{Var}(\Delta\mathrm{GDP})} = 1 + \frac{\mathrm{Cov}(\Delta\tau, \Delta\mathrm{GDP})}{\mathrm{Var}(\Delta\mathrm{GDP})}$$

Accordingly, the coefficient is equal to 1, the actual impact of changes in GDP, plus a term that depends on the covariance between changes in the ETR and GDP. When the ETR remains constant, this covariance is zero, and the estimate is correct: $\beta = 1$. In other cases, however, the second term is non-zero, leading to biased estimates.

For example, if the ETR has been declining over time—so that $\mathrm{Cov}(\Delta\tau, \Delta\mathrm{GDP}) < 0$—the estimated elasticity would be less than one. In this case, the empirical model systematically underestimates revenue growth under the baseline, because it implicitly projects the effects of past policy changes forward into the future. Conversely, if the ETR has been increased in some year that is represented in the data, then the covariance will be positive, and baseline revenues will be overestimated.

How to Control Past Policy Changes?

Data-preprocessing methods aim to create a current-law series of dependent variables by adjusting for historical tax policy reforms through algebraic manipulations. For instance, the impact of changing CIT rates on aggregate tax liabilities can be (at least partially) controlled by dividing the aggregate recorded tax liability each year by the CIT rate that was levied in that year and then multiplying the result by the current CIT rate. A generalization of this approach is the proportional adjustment method, which uses annual costing estimates to create a counterfactual series of aggregate tax liabilities that would have been observed if the current tax system had been in place during the previous years in the series (see Box 1 for a description of this method). The adjusted series can then be used to estimate the sensitivity of the tax base to economic changes at varying levels of sophistication. For instance, the adjusted tax liability series can be used to compute an elasticity of aggregate tax liability with respect to GDP,[13] or it

[13] Estimates of tax buoyancy rely on the time-series relationship between unadjusted aggregate tax liability and the tax base. Elasticity-based measures rely instead on the relationship between the adjusted tax liability series and the tax base.

can be used in a regression analysis that predicts aggregate CIT liability as a function of multiple economic predictors.

The proportional adjustment method is straightforward to implement, but unlikely to fully remove the effect of past policy changes (see Box 4).[14] For instance, a higher CIT rate will tend to reduce economic activity (unless the CIT is levied on economic rents, which is seldom the case in practice). An adjusted series obtained by dividing each year's aggregate CIT liability by the statutory rate for that year (before multiplying the result by the current rate) will tend to produce relatively small values in high-tax periods, irrespective of variations in economic conditions. Therefore, the mechanical adjustment would not fully remove the behavioral responses to past policy changes from the data, and a forecast based on the adjusted series might continue to provide biased results if the tax rate change is correlated with predictors, such as GDP.

Notably, because firms respond to policy changes, tax reforms can also affect GDP and other macroeconomic aggregates, creating forecasting challenges that cannot be easily resolved through mechanical revenue adjustments. For instance, Hebous, Klemm, and Wu (2021) document that profit shifting—the relocation of accounting profits from high- to low-tax jurisdictions by multinational enterprises to minimize tax liabilities—can lead to substantial revisions in measured GDP growth and other macroeconomic aggregates. When such aggregates are used as predictors of tax revenues, reverse causality can bias elasticities even if the direct revenue effect of the reform is controlled for. Although instrumental-variable methods can address such endogeneity in principle, suitable instruments are scarce, and estimates are often imprecise. A more practical solution is to construct a "policy-clean" predictor. For instance, if the response of multinational enterprises to a statutory tax rate change affected GDP, a reweighted measure of economic activity that is less exposed to the reform (such as gross operating surplus of domestic firms or the value added of nonfinancial firms) could provide a more robust predictor variable.

An alternative to data preprocessing is estimation in a regression framework with tax parameters included as additional explanatory variables. For example, the effect of past tax rate changes can be controlled for by including the statutory rate alongside other economic predictors of unadjusted CIT liability. The estimated partial effect that a change in economic conditions exerts on aggregate tax liability will then be independent from any direct or indirect effects of the CIT rate change. Direct modeling in a regression framework is primarily constrained by statistical power. The number of distinct partial effects (coefficients) that can be estimated is limited by the number of observations in the estimation sample,[15] which in turn restricts the number of tax policy parameters that can be included along with economic predictors, particularly when one relies on annual aggregate time-series data. This restriction becomes more binding where the relationship between different tax parameters is nonlinear: for instance, changes in the tax base are mediated by the tax rate, requiring the inclusion of interaction terms when both the rate and some base parameters have changed. Although such effects can be accommodated through the creation of appropriately defined new variables, doing so further reduces statistical power. A practical response is

[14] In addition, the impact of tax policy changes will not always be equal to a fixed proportion of existing liabilities. For instance, an increase in the statutory tax rate may lead to a disproportionate liability increase owing to interactions with the foreign tax credit.

[15] The minimum sample size is not a fixed number, but the quantity of observations necessary to make relevant *t*- or *F*-statistics large enough, given the analysis choice of statistical significance and type-1 error. As a rule of thumb, the inclusion of an additional quantitative predictor requires about 10 observations.

to design composite policy variables that capture multiple features of the tax system in a single measure. This approach helps control for policy changes without exhausting degrees of freedom. Examples include the following:

- **Weighted average parameters.** Tax parameters often differ across subgroups of the corporate population. For instance, some CIT systems impose a progressive rate structure as a function of the size of the business. Similarly, depreciation allowances typically differ for different asset classes. Rather than including each of these parameters as a separate predictor, (historic) micro-data can be used to compute a weighted average tax parameter that reflects the population. Although the weights would ideally change over time, constant weights would permit a more parsimonious specification and may perform reasonably well in practice.

- **Theoretical tax parameters.** Another option is to compute theoretical tax measures, such as forward-looking marginal or average ETRs (see, for instance, Devereux and Griffith 1999). Such tax measures combine various aspects of the tax system, including the generosity of depreciation allowances, the availability of credits, and the statutory rate; which implies variation in the computed measure when any of the underlying tax system parameter changes. As both measures are sensitive to assumptions and capture the tax treatment of firms over multiple years, their ability to explain current changes in revenue collections will vary and depend on country-specific contexts.

Box 4. Proportional Adjustment Method

The proportional adjustment method uses external estimates of the aggregate fiscal impact of tax policy reforms implemented in one year to construct an adjusted series of aggregate corporate income tax (CIT) liabilities that would have been collected under current policy. This is calculated as follows:

1. Denote actual aggregate tax liability in year *t* by R_t and costing estimates of policy changes implemented in that year by C_t. It follows ${R_t}^c = R_t + C_t$, indicating the estimated counterfactual liability that would have been collected under unchanged policy. If the actual aggregate liability is less than the counterfactual amount, the costs of policy reforms are positive $(C > 0)$.

2. Denote the ratio of actual to counterfactual aggregate CIT liability in year *t* by $\theta_t = \frac{R_t}{R_t^c}$. The proportional adjustment method corrects aggregate tax liabilities using these ratios as follows:
   - Year *T*: $A_T = R_T$.
   - Year *T* − 1: $A_{T-1} = R_{T-1}\theta_T$.
   - Year *T* − 2: $A_{T-2} = R_{T-2}\theta_{T-1}\theta_T$
   - Year *T* − *k*: $A_{T-k} = R_{T-k}\prod_{t=T-k+1}^{T}\theta_t$

To better understand the mechanics of the approach, suppose aggregate tax liability is the product of a base and a tax rate $R_t = B_t\tau_t$. If the only discretionary policy change over the period covered by the historical series was a change in the tax rate from τ to τ^{new} in year *T*, the adjustment ratio for year *T* would be specified as $\theta_T = \frac{\tau^{\text{new}}}{\tau}$, while the adjustment ratios for all prior periods would be set equal to one. As a result, the adjusted series would simply take the form $A_t = B_t\tau^{\text{new}}$ for all periods.

By contrast, discarding outlier observations rarely improves prediction quality. In cases involving a substantive past change in tax policy, some tax administrations introduce an additional dummy explanatory variable into their forecast to allow for a one-time shift in the intercept of the prediction equation at the time of the policy change. However, this is a rather crude approach, and the use of such dummies can significantly deplete the degrees of freedom in estimation if separate dummies are used to account for several separate past policy changes. Better alternatives include the use of robust estimation methods, such as reweighted least squares, that automatically reduce the impact of outlier observations, or to test for structural change in the series using formal tests (such as the Chow test).

How to Incorporate Future Tax Policy Change?

Costing estimates require taxpayer-level data. This is so as tax policy reforms typically affect only a subset of the taxpayer population, such as through changes in the treatment of specific expenses or adjustments to tax rates in selected sectors. Using administrative records, the historical revenue impact of some reforms can be readily calculated using aggregated information. For instance, the revenue effect of increasing the CIT rate on agricultural companies could be approximated using information on the aggregate liabilities of firms in that sector.

Corporate tax systems typically include several nonlinearities that can only be accurately captured with microsimulation models. For instance, in tax systems that provide investment or employment credits, a change in the tax rate by 10 percent will not affect final liabilities by the same proportion. Similarly, the revenue effect of restricting the deductibility of interest payments (such as through Earnings Before Interest Depreciation and Amortization [EBITDA] rules) depends on the interaction of firm-specific profits and interest payments. Aggregate information on these variables will not help to obtain an accurate assessment of revenue effects from changing the limitation rule. In those cases, microsimulation models that replicate the full structure of the tax system provide an effective solution that allows modeling several nonlinearities easily. Irrespective of whether aggregated data from tax returns is used or whether microsimulation models are employed, taxpayer data are frequently available only with some delay. And since costing estimates must reflect future revenue effects, it is necessary to adjust the underlying tax bases for expected changes in economic conditions.

Ensuring consistency between baseline forecasts and costing estimates requires close communication between units responsible for these tasks. In many countries, costings are prepared by dedicated tax policy divisions, while baseline forecasts are produced separately by forecasting units. Without systematic interaction, the assumptions underpinning the two sets of estimates risk diverging, undermining the consistency of the overall revenue forecast. In many advanced economies, forecasts are thus derived in an iterative process, where draft forecasts and costings are exchanged, compared, and revised until they align along key dimensions. This back-and-forth not only improves technical accuracy but also strengthens institutional coherence and shared ownership of the results. At a minimum, consistency checks should verify that the growth rate of the tax base implicit in the costing aligns with the projected growth of the aggregate CIT base in the baseline forecast. For example, if baseline CIT revenues are expected to increase by 10 percent between this year and next, then the impact of a tax rate change—estimated on current data—should also be scaled by roughly 10 percent, unless there is evidence that the affected base will evolve differently from the overall CIT base.

Although regression models would provide the benefit of automatically capturing behavioral effects, they are difficult to use in practice. Many reforms are unique events that cannot be represented as continuous variables, and regression-based estimation can only generate costing estimates if similar reforms have occurred in the past. The use of summary measures, such as forward-looking ETRs, can expand the scope of past variation that is exploitable, but many reforms remain idiosyncratic events that cannot be captured in a quantitative analysis. In practice, regressions are best suited to capturing broad changes in tax rates or bases, while micro-data analysis is generally more effective for assessing narrower reforms targeted at specific taxpayer groups.

Robust Forecasting Techniques

Forecasting units often operate under conditions that limit the benefit of time-series methods. When aggregate cash receipts are the only data available, and the effects of past tax policy changes are not comprehensively recorded, time-series predictions rely on the strong (and untestable) assumption that the effects of past reforms are uncorrelated with key explanatory variables. If they are not, the baseline prediction will suffer from an omitted variable bias (see Box 3).

This section summarizes methods that provide robust forecasts even if past policy reforms cannot be accounted for. Estimates for the current year can draw on current cash collections and they can leverage the inertia built into the tax system through prepayments. Current-year receipts are thus relatively straightforward to predict with high confidence. The subsequent section discusses forecasting techniques for future years, which could also be applied to obtain current-year estimates.

Methods for the Current Year

Forecasting units often leverage the inertia built into corporate tax systems for forecasting current-year revenues. In Austria, for example, prepayments are calculated as a fixed mark-up on a taxpayer's latest tax liability and estimates of this year's advance payments can leverage this mechanical relationship. The estimated advance payments, combined with the balancing payments made in the current fiscal year (for last year's taxes), will account for the lion's share of this year's cash receipts.

In-year revenues provide another critical information source to predict full-year collections. Revenues collected throughout the year are typically subject to the same tax rules, which allows constructing a variable that is independent of policy design. The central input for in-year revenue projections is the progress ratio—defined as the ratio of full-year revenue to cumulative collections observed up to a given month. Since both the numerator and the denominator are determined by the same tax system design, the ETR cancels out, allowing for a meaningful comparison of these ratios across years.[16] This principle is illustrated in Table 1, which simulates a scenario where the tax rate doubles in the second year, resulting in a doubling of tax revenues. Despite that doubling, quarter-specific progress ratios remain unchanged over the two years.

[16] Balancing payments are determined by past policy design and should thus be disregarded in the construction of progress ratios.

Table 1. Hypothetical In-Year Revenue Collections

Year	Year 1 (tax rate of 10 percent)				Year 2 (tax rate of 20 percent)			
Quarter	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4
Cumulative collections	10	20	30	40	20	40	60	80
Share collected	0.25	0.5	0.75	1	0.25	0.5	0.75	1
Progress ratio	4	2	1.33	1	4	2	1.33	1

Source: Author's calculations

Using the progress ratio, full-year revenues are predicted by multiplying cumulative receipts up through a specified month by that month's (estimated) progress ratio. For example, if historical data indicate that annual revenue is typically 1.3 times the collections through September, then the full-year forecast can be derived by multiplying this year's cumulative revenue up to September by 1.3. The share of collections varies by month, and each month therefore has a distinct progress ratio. For instance, by November, almost all taxes will have been collected, and the progress ratio for November will be very close to 1. Since the monthly progress ratio reflects seasonal patterns, there is no need for an additional correction or adjustment of collections for seasonality. The econometric task boils down to finding out what the representative ratio is for a given month.[17]

The optimal approach to estimate this year's value or the ratio depends on the underlying stochastic dynamics of revenue collections. Two basic and commonly used approaches, that can be optimal in corner scenarios, include the following:

1. **Simple average.** One option is to use the simple average of historic progress ratios of a given month. This method is optimal when the progress ratio follows a white-noise process, where outcomes hover around some constant mean.[18] For such processes, the expected value for all future outcomes is the (unknown) mean of the process. Note that the mean places equal weight on all prior realizations.

2. **Last observed value.** Alternatively, one might rely just on the most recently observed value for this ratio rather than the average of past realizations.[19] This approach to forecasting is optimal, in a statistical sense, when the progress ratio equals last year's value plus some random noise.[20] In other

[17] Note that simple predictions based on the inverse of the progress ratio—cumulative collections relative to total year revenues—would yield biased estimates of full year revenues. The reason is that the prediction would be computed as cumulative collections up to some month, divided by the estimated (inverse) ratio. Since the inverse of an expectation is different from the expected inverse, the resulting forecast would consistently underestimate full-year revenues. Formally, suppose the inverse progress ratio (cumulative collections relative to full year revenues) can be represented as a constant plus an error term: $r = c + \varepsilon$. Denoting cumulative in-year collections by *M*, the full year estimate using the inverse progress ratio would be computed as $R = \frac{M}{r}$. Taking expectations gives $E[R] = ME\left[\frac{1}{r}\right]$. By contrast, the prediction using the estimated inverse progress ratio is $\hat{R} = \frac{M}{E[r]}$. Noting that the division is a convex function, Jensen's inequality implies that $E[R] \geq \hat{R}$. To avoid such bias, the progress ratio can simply be defined in reverse, as full year revenues relative to cumulative collection to date.

[18] The white-noise process follows $y_t = c + \varepsilon_t$, where **c** is a constant and the period-specific error terms ε_t are independent and identically distributed (i.i.d) with zero mean and constant variance.

[19] This alternative approach is equivalent to recording the percentage change in cumulative in-year tax collections as of the specified forecast date between last year and the current year, and assuming that annual total liabilities will change by the same percentage. For instance, if the year-over-year percentage change in cumulative CIT receipts by the specified date is found to be 20 percent, total annual tax liability for the current year would be predicted to be 20 percent higher than the amount observed for the prior year.

[20] The random walk process follows $y_t = y_{t-1} + \varepsilon_t$, where the period-specific error terms ε_t are i.i.d. with zero mean and constant variance. A constant term can be added to this model in which case the process represents a random walk with "drift" $y_t = c + y_{t-1} + \varepsilon_t$, where the constant c is the drift term. Depending on whether this term is positive or negative, the series will tend to drift either upwards or downwards over time. Observe that the first difference of this series $(y_t - y_{t-1}) = c + \varepsilon_t$ is a random walk

words, the year-over-year change in the outcome variable is white noise. When the target variable follows a random walk, its future outcomes are unpredictable in the sense that future movements in the series are independent of past movements. Rather, the statistical expectation of next year's outcome will simply be equal to this year's outcome. Accordingly, the optimal forecast for such a process is the most recently observed value in the series. Unlike a white-noise process, no weight whatsoever is given to earlier outcomes.

Exponential smoothing is an alternative approach that accommodates a wide range of intermediate dynamics by placing less weight on more distant observation. In practice, the dynamic process underlying the time path of the progress ratio is likely to fall somewhere between the extremes of white noise and a random walk. Typically, the current outcome depends on some combination of prior realizations of the outcome variable and random shocks to the series.[21] For these processes, the statistical expectation of next year's outcome is generally a weighted average of all past observations, where the weights are exponentially decaying over time from the most recent to the oldest observation in the series. This means that the optimal forecast of the target variable in the next period will fall somewhere between the most recently observed value and the long-term mean. Exponential smoothing places most weight on recent realizations of the outcome variable, but all past realizations play at least some role in determining the forecast value.

Although the anticipation of policy change can temporarily affect measured progress ratios, they will typically revert to their long-term pattern. For example, the announcement of more generous depreciation rules may incentivize firms to hold back on investments until the new policy is in place. Such anticipatory behavior could alter the within-year distribution of collections in the year preceding—and sometimes following—the reform. Over time, however, intra-year investment and tax payment patterns should return to some baseline distribution that is determined by nontax factors. To avoid an outsized impact of unusual observations, the progress ratio can be estimated using robust techniques. For instance, a widely used approach is iteratively reweighted least squares, which begins with an initial OLS fit, then repeatedly updates observation weights based on residual size using weighting functions such as Huber's or Tukey's bisquare, until convergence.

In-year receipts can also be used to update initial revenue estimates or macroeconomic projections. For instance, the Dutch forecasting unit, a part of the Bureau of Economic Policy Analysis, prepares two revenue projections annually (February/March and September). The installment payments are primarily used for the September forecast. Unexpectedly weak or strong installment payments up to that point can influence the full-year revenue forecast, and this information is also shared with the macro-forecast team for consideration when updating the macroeconomic projections. Similarly, in the UK's Office of Budget Responsibility, in-year deviations of actual from anticipated prepayments are seen an indication of unanticipated strength or weakness in economic profits. These deviations are often "pushed through" to the future-year forecasts.

---

process, so that the best predictor of *the change* between this year's outcome and next year's outcome is the average of the previous year-over-year changes in the series.

[21] These dynamics might be characterized, for instance, by $y_t = \alpha y_{t-1} + \varepsilon_t$, where $\alpha \in (0,1)$ and ε_t follows a moving average, such as: $\varepsilon_t = u_t + \theta_1 u_{t-1}$. Alternatively, the orders of these processes might involve multiple lags.

Methods for Future Years

Predictions for the future need to build on external predictors, such as official forecasts of GDP or gross operating surplus. The incorporation of such external predictors—or proxy tax bases—is necessary to ensure consistency between revenue projections and the government's macroeconomic framework. When key parameters in the relationship between external predictors and revenues cannot be reliably estimated, they must instead be inferred from theory. A common assumption, for example, is that CIT (and other) revenues grow in line with GDP, implying a unitary elasticity.[22] This assumption builds on the fact that GDP is the sum of wages and economic profits. If the wage share of total economic activity remains broadly constant in the long term—as is commonly assumed—then profits, and by extension CIT revenues, should expand proportionately with GDP under unchanged policy. Accordingly, in the absence of additional information, a theoretically justified approach to predicting next year's aggregate CIT liability is to assume that liabilities will grow in line with GDP.

A unit elasticity of CIT liabilities with respect to GDP is often a useful benchmark—and one that is difficult to improve on empirically (see the "Documentation and Improving Forecast Performance" section)—but several factors could explain deviations from it. The unit elasticity would be observed when economic profits remain a constant share of GDP, when aggregate taxable income changes one-for-one with economic profits captured in the national accounts, and when the average tax rate does not respond to changes in economic profits. Each of these assumptions, however, could fail (Box 5):

1. **Cyclical effects.** Empirical studies typically find that corporate profits rise disproportionately during the early stages of economic expansion. This heightened response occurs because wages are often rigid in the short term, so a greater share of additional revenues accrues to capital owners, boosting corporate profits. Over time, as labor markets tighten and wages adjust upward, profit margins are gradually compressed. As a result, the elasticity of CIT liabilities with respect to GDP can exceed or fall below unity depending on the phase of the business cycle.

2. **Tax base and profit responses**. The elasticity of the tax base with respect to profits can deviate from unity because of structural features of the tax system. Fixed deductions or credits that shield a portion of profits from taxation can cause the tax base to expand more than proportionally in response to an increase in financial profits. Another reason is the asymmetric treatment of profits and losses. Profits are taxed immediately, whereas losses can generally only be deducted against future profits through carryforward provisions.[23] During downturns, tax revenues do not fall as sharply as profits, because companies do not receive refunds for losses. Conversely, during expansions, loss carryforwards dampen the growth in tax revenues. On balance, this dynamic tends to reduce the elasticity of tax revenue with respect to profits below unity.

---

[22] This follows from noting that $\frac{\partial P}{\partial \mathrm{GDP}} \frac{\mathrm{GDP}}{P} = 1$ when $P = c * \mathrm{GDP}$, where *P* is the profit and *c* is some constant.

[23] In Canada, losses can be offset against past profits. Because this provision allows firms to receive an immediate cash refund when loss making, rather than waiting to offset losses against future profits as under carryforwards—the provision is highly valued, especially by firms facing cash-flow constraints. For revenue forecasters, however, loss carrybacks increase the volatility of CIT receipts relative to a system with only loss carryforwards.

3. **The sensitivity of the tax rate with respect to changes in profit**. Some tax systems apply reduced rates to small businesses. As GDP grows, more firms may transition to the standard (higher) tax rate, which would imply an elasticity of tax revenue with respect to aggregate taxable income that exceeds one.[24] However, if the threshold for the lower rate is set relatively low, it is unlikely that this progression alone would justify an elasticity above unity.

Box 5. Causes for Deviations from Unit Elasticity

Suppose tax revenue is given by $T = \tau B$, where τ is the national average tax rate (which may differ from the marginal rate, for instance, when progressive rates apply) and B is aggregate corporate taxable income. A small change in tax revenue can be traced back to changes in the tax rate and the tax base as: $dT = Bd\tau + \tau dB$.

Letting P denote financial profit and G aggregate output, the elasticity of tax revenue with respect to GDP, denoted by $\varepsilon_{T,G} = \frac{\partial T}{\partial G}\frac{G}{T}$, can be decomposed as

$$\varepsilon_{T,G} = (\varepsilon_{\tau,P} + \varepsilon_{B,P})\varepsilon_{P,G}$$

The equation offers three explanations why the elasticity may deviate from unity:

- First, the average tax rate may respond to increases in financial profit, meaning that $\varepsilon_{\tau,P} > 0$.
- Second, the tax base may not increase one for one with changes in financial profit because of tax design features, meaning that $\varepsilon_{B,P} \neq 1$.
- Third, financial profit may not increase one for one with changes in GDP, because of structural or cyclical shifts in the capital share of output, meaning $\varepsilon_{P,G} \neq 1$.

Calibrating these deviations requires information on key design aspects of the tax system and parameter estimates, drawn either from literature or from other sources. For instance, the elasticity of the national average tax rate in response to a percentage increase in economic profit could be estimated using a simple microsimulation model. At its core, this exercise involves increasing operating profits for all taxpayers in the dataset by one percent and recording the resulting percentage change in the average ETR (measured as total tax liability divided by aggregate taxable income). Such an exercise, however, cannot be purely mechanical: it requires knowledge of the tax system's key deductions and credits, and whether their value adjusts with changes in operating profits. Another example is the incorporation of cyclical effects. The elasticity of CIT revenues with respect to the output gap—the difference between actual and potential output—is often estimated at about 1.5 (see Girouard and Andre 2005). When an estimate of a country's potential output is available, the forecasting unit could use such estimates from the literature to forecast revenues over the business cycle (see Box 6).

[24] Typically, the small business threshold is not indexed, so nominal income growth leads some small businesses to transition to the standard CIT rate.

Box 6. Incorporating Cyclical Elasticities into Forecasting

An elasticity-based prediction can be complemented with information on the current output gap to incorporate cyclical effects. Starting from simple assumptions that revenues include cyclical and structural components, which are also at the core of estimations of structural deficits (see Annex 2), a cyclical-sensitive elasticity can be defined as

$$\varepsilon_{T,G} = 1 + (\beta - 1)\frac{g_{t+1} - s_{t+1}}{g_{t+1}}$$

where $g_{t+1} = \log\left(\frac{G_{t+1}}{G_t}\right)$ is the predicted growth rate of GDP and $s_{t+1} = \log\left(\frac{S_{t+1}}{S_t}\right)$ is the predicted growth rate of potential growth.

Accordingly, when the output gap is at a turning point, so that structural and actual GDP grow at the same rate, the elasticity of revenue with respect to GDP is unity. When the economy is in an upswing, revenue responds more than proportionally to GDP for $1 < \beta < 2$, whereas in downswings, revenue responds less than proportionally.

Simulation-based predictions of tax revenues, common in the natural resource sector, are another viable option. For instance, the IMF's fiscal analysis of resource industries tool is an Excel-based model that projects the government's share of a resource project's total pre-tax net cash flows. It does so by combining information on the project's expected production volumes, costs that reflect the stage of the project (exploration, development, production, closure) with economic variables, such as prices, inflation, interest and discount rates. Starting from the calculation of the project net cash flows before any fiscal impositions, the model calculates each fiscal payment according to the fiscal regime parameters. The results are then used to estimate different indicators that allow for the evaluation of the fiscal regime along relevant criteria.

A project- or company-level modeling approach is particularly appealing where revenues are concentrated around a few large-scale projects. Fewer companies facilitate data collection and model updating. Moreover, project-level information does not only enable revenue projections but also provides data that can be used to inform the government's macroeconomic framework, including on exports, prices, portfolio and foreign direct investment, interest and debt amortization, and production volumes.

Time-Series Methods

Ideally, revenue predictions should be fully empirical to capture country-specific tax base dynamics. When an adequate time-series data sample is available for the target variable (aggregate tax liabilities or cash receipts) and the target variable has been adjusted to account for any relevant policy changes over the estimation period (or when uncosted policy changes had no meaningful impact on total collections), empirical estimation approaches will improve forecast accuracy.

Two simple and widely used approaches to forecasting liabilities, which can be implemented using basic spreadsheet tools, include the following:

- **Elasticity-based predictions** were previously introduced as a theory-driven approach to forecast revenues. However, the same forecasting procedure, which involves updating the latest CIT liability estimate using information on the elasticity along with the growth in a proxy tax base, can also be implemented using an estimated elasticity.[25] The elasticity-based approach is similar to a regression specification that estimates the relationship between the target variable and a single predictor in (logarithmic) differences.

- **ETR-based prediction** relies on the ratio of tax liabilities relative to the proxy tax base, which is commonly referred to as an ETR. Under this approach, the ETR is first projected into the future based on its past trajectory. For instance, a simple approach is to use the average of recent ETR realizations as the future predicted value. More elaborate techniques can use exponential smoothing or ARIMA models. The baseline forecast is then obtained by applying this ETR estimate to the projected value of the proxy tax base. The ETR approach is comparable to a regression that explains CIT liability as a linear function of a single predictor (that is, the proxy tax base) with the constant term omitted.

Regressions provide a more flexible approach to forecasting future aggregate tax liabilities, because they can account for the effect of multiple predictors. For instance, incorporating trade indicators such as import and export values can help ensure that long-term trends and structural shifts in the economy are accurately reflected in revenue projections. Rather than choosing between an elasticity-, an ETR-, and a regression-based prediction, a more fundamental question is therefore what type of regression equation best captures the structural relationship between liabilities and one or several predictors.

This section discusses conceptual issues and formal tests that can guide the specification of reduced-form regression models to project CIT revenues.[26] It emphasizes that a key decision concerns whether the relationship between the dependent variable and the proxy tax base should be estimated in levels or in differences. Among other factors, this choice depends on the forecasting horizon and the reliability of the proxy tax base. Using the logarithm of the dependent variable can often improve model fit, although it requires an adjustment when projecting nominal amounts. The section concludes by highlighting the importance of out-of-sample testing procedures to evaluate model performance.

Which Explanatory Variables Can Be Used?

There are two types of variables that help explain long-term developments of CIT collections:

1. **Proxy tax bases.** The main explanatory variables in reduced-form models are typically drawn from national accounts. They should be highly correlated with the corporate tax base and robust forecasts of the variable should be available. In practice, these requirements often involve a trade-off. For example, gross operating surplus is conceptually closely aligned with the CIT base. However, it is typically

[25] Since elasticities measure the sensitivity of liabilities to changes in the proxy tax base, holding other factors constant, they need to be measured using an aggregate tax liability series that has been corrected for tax past policy changes. By contrast, measurement based on an unadjusted tax liability series would capture the combined effects of changes in the proxy tax base and changes in tax policy. This combined percentage impact measure is known as *buoyancy*. Because a buoyancy measure conflates these two effects, it is not very well suited for producing a baseline forecast of next year's aggregate tax liability.

[26] The section highlights selected issues that are important for forecasting aggregate CIT revenues using historical time-series data. Introductory treatments of time-series estimation issues are provided in Mills (2015) and Hyndman and Athanasopoulos (2025). A more advanced treatment is provided in Box and others (2016).

estimated as a residual,[27] making it vulnerable to sizable measurement errors, and forecasts are often not available. On the other hand, broader macroeconomic variables like GDP are more consistently forecast, but they may not track corporate taxable profits as accurately, especially in the short term. It is helpful to view the proxy tax base—whether a single variable or a combination of explanatory variables—as a signal of the underlying true tax base that is measured with noise.

2. **Statistical controls.** Time trends or indicator variables capturing structural shifts or outlier years can sometimes help capture gradual or sudden changes in the dependent variable that are not reflected in long-term predictors. Although such variables can help improve the fit of models on historic data, the extrapolation into the future is often difficult. For instance, a time trend may pick up gradual improvements in administrative capacity. But whether, and at which speed, such improvements would continue is difficult to gauge. Similarly, dummy variables that capture outlier years can improve fit, but they also remove variation that could help estimate the impact of other explanatory variables more accurately.

Short-term (or high-frequency) variables provide insights into more immediate fluctuations in aggregate CIT liabilities. These indicators can be derived from publicly available sources, such as quarterly financial statements of companies or production indices published by the statistical office. They have the advantage of being independent of tax policies, meaning they require little or no adjustment prior to their inclusion in forecasting models. Administrative tax information can also provide a powerful source for improving short-term forecasts. As previously discussed, cumulative CIT receipts over a portion of the fiscal year can provide a good indication of aggregate current-year tax liability, especially when taxpayers have flexibility in adjusting their CIT prepayments based on current economic conditions. This information can be included in a regression model that predicts current-year liabilities using different structural indicators. Information on revenue collections from other sources can also help predict aggregate CIT liabilities. For instance, monthly receipts from value-added taxes or customs duties can indicate current economic conditions. Ideally, receipts from other taxes should be adjusted to account for past policy changes. For example, monthly value-added tax revenue can be standardized by dividing it by the applicable standard value-added tax rate, thereby creating a more consistent measure of underlying economic activity.

When building models incorporating short-term indicators, a common challenge is missing information for future years. Whereas long-term indicators often have predicted values available several years into the future, short-term indicators may only be available for the current period. This challenge can be addressed either by imputing missing short-term indicator values for future periods—for instance, assuming that these variables will revert toward their long-term averages—or by employing such indicators only in specifications used to predict aggregate liability for the current tax year (and using specifications that exclude such variables to forecast future values).

Which Model Specification?

An important concept in the context of forecasting liabilities is the notion of a long-term equilibrium relationship. The true tax base maintains a stable, long-term relationship with tax liabilities under

[27] Total industry output minus intermediate consumption, employee compensation, and net taxes on production and imports.

unchanged policy. When both the tax base and aggregate tax liability are nonstationary series,[28] that move together over time such that a linear combination of the two series is stationary, the series are said to be co-integrated; a concept that shapes the efficiency of estimation techniques. Moreover, when the proxy is a noisy signal of the true tax base, there is also a co-integrating relationship between liabilities and the proxy. Two features then become central for forecasting: (1) how accurate the proxy is in levels (the volatility in the wedge between the true tax base and its proxy) and (2) whether the wedge between the proxy and the true tax base is stable over time (see Annex 3).

Regressions in levels and differences capture distinct aspects of the structural relationship between aggregate CIT liability and the proxy tax base:

- **Levels**. A regression in levels identifies the long-term relationship between aggregate tax liability and the proxy tax base. To avoid spurious estimation results,[29] regression in levels should only be estimated after statistical tests have confirmed that the proxy tax base is indeed co-integrated with the target variable. In practical forecasting, levels tend to perform best when the proxy is a clean measure in levels. With longer samples, co-integration implies that the long-term relationship can be estimated very precisely even in the presence of classical measurement error. Predictions from a regression in levels (or an ETR-based forecast) are especially informative for longer-horizon projections.
- **Differences**. A regression in differences identifies the short-term response of aggregate liability to changes in the proxy tax base. This approach reduces the risk of capturing a spurious relationship and can identify short-term responses more reliably under less restrictive assumptions, particularly when the wedge between the proxy and the true base is persistent. In this case, differencing strips out much of the noise while preserving genuine short-term signals that can help predict revenues. By contrast, if changes in the wedge between the proxy and the true tax base are volatile and close to white noise, differencing increases the noise and estimated short-term elasticities can be attenuated; in such cases a levels specification will often dominate. Predictions based on regressions in differences (or elasticity-based predictions) rely heavily on the most recently observed level of liabilities and are particularly useful for short-term forecasting.

More generally, however, the sole focus on either of the two specifications risks neglecting some component of the actual dynamics. The equations in levels and in differences both capture unique and complementary information about the process. Rather than focusing on one of these, a better option is to estimate the relationship in dynamic form when sufficiently many observations are available. For instance, this could be done by adding lagged levels of both the dependent and the explanatory variables to the differenced equation. Alternatively, when the presence of a co-integrating relationship is established through appropriate statistical tests,[30] that relationship can be leveraged to develop both short- and long-

---

[28] A nonstationary time series is characterized by variations in its statistical properties, such as its mean, variance, and covariance, over different time periods There are formal statistical "unit root" tests that can be used to assess whether a series is nonstationary. Two commonly used tests are the Augmented Dickey–Fuller Test and the Phillips–Perron Test.

[29] When regressing one trending variable against another, the results often suggest a strong causal relationship, with high goodness-of-fit and statistical significance. However, these results could be spurious, merely reflecting shared trends rather than a true underlying relationship. A levels regression would also be appropriate if the variables are found to be trend-stationary and detrended before estimation or in the unlikely event that they are found through testing to be stationary.

[30] Co-integration among a set of time-series variables is defined through two properties. First, the variables must be integrated of the same order, meaning they individually only become stationary after the same degree of differencing. Second, there exists some set of weights for which the weighted sum of the variables is stationary. Intuitively, this means that variables may deviate from each other in the short term, but they tend to revert to a stable relationship over time.

term forecasts for the outcome variable of interest using an error-correction model. With dynamic regressions, near-term projections are more heavily influenced by the most recent level of overall tax liability (and hence a prediction based on differences), while more distant realizations are more heavily dependent on the long-term equilibrium (and hence by the prediction implied from the equation in levels). In that sense, a dynamic regression resembles exponential smoothing of a single target variable.

As the number of historic aggregate liability observations is often limited, revenue forecasts need to rely on parsimonious regression specifications. In principle, the minimum number of required observations depends on the volatility of the process, desired targets for statistical power and type-1 error, as well as the number of included regressors. As a rule of thumb, the estimation of a regression requires about 10 observations per predictor. In revenue-forecasting settings, this benchmark will be difficult to achieve for some target variables. Model building should thus start with a parsimonious specification and only add complexity (by adding more variables) if it meaningfully reduces out-of-sample forecasting performance (discussed in the following section). Moreover, models that include economic priors (such as Bayesian estimation) or shrinkage (such as ridge regressions) can help deal with limited information.

Log Specifications

Time-series models that involve macroeconomic aggregates are commonly specified in natural logarithms (logs) rather than levels for several reasons. One is that such series are often multiplicative in the sense that they tend to exhibit exponential rather than linear growth. Consequently, the log of the series tends to follow a linear trend, which is convenient for estimation and prediction. Another reason why logarithmic transformations of the variables in a time-series regression specification are incorporated is to improve the symmetry of the error distribution, something which can be checked, for instance, by using normal quantile plots of the residual. Although the symmetry of the residual distribution is not a prerequisite for consistent model estimation, it is required (and typically assumed) for standard inference procedures. Using a log transformation can also be attractive from a conceptual perspective: as the logarithm of a product is the sum of the logarithms of the underlying factors, the logarithm of a revenue series additively separates the tax base from the rate. Past policy changes can therefore be conveniently modeled by including (the log of) some tax parameters as well as the logarithmic proxy tax base, which can simplify estimation and interpretation of results.[31]

An aspect that is often overlooked in applied revenue forecasting is that the exponential of a predicted logarithm will typically under-predict the level of the underlying variable—the quantity that is at the core interest of revenue forecasting. The reason is that the logarithm is a concave function of the original variable. As a result, the average of the transformed series is consistently smaller than the logarithm of the original series' average.[32] When transforming the logarithmic series back, this difference needs to be accounted for by multiplying the exponentiated prediction by a "smearing" term—a factor that is always larger than one and sometimes by enough to make a meaningful difference in the predicted value. This

---

[31] Suppose revenue is $R = B\tau$, where *B* is the base and τ is the rate. The natural logarithm of revenue can then be expressed as $\ln(R) = \ln(B) + \ln(\tau)$, which provides a good starting point for deciding on the functional form in a regression framework.

[32] Another way to understand this point is to consider the model $\ln y_t = c + \beta_1 \ln x_t + \varepsilon_t$. For this model, we can express the level of the outcome variable as $y_t = \exp(c + \beta_1 \ln x_t) * \exp(\varepsilon_t)$. Although the expectation of the error term ε_t is equal to zero, observe that the expectation of $\exp(\varepsilon_t)$ is a positive constant. For instance, if the error term is normal and independently distributed over time with mean zero and variance σ^2, then $E[\exp(\varepsilon_t)] = e^{\sigma^2/2}$. Observe that, since this expression is greater than one, the expected value of *y* will exceed $\exp(c + \beta_1 \ln x_t)$.

correction is also necessary when the dependent variable is the difference of a logarithmic series (as that difference is the same as the logarithm of a ratio).

Testing

Standard test statistics that are routinely provided as part of regression summaries can provide initial guidance for selecting between competing models. For instance, eliminating predictors that are found to be (jointly) statistically insignificant often leads to improved forecast precision with little (if any) loss in forecast accuracy; and a higher adjusted R-square statistic, which quantifies the relative variation in the outcome variable explained by included regressors, often indicates better model forecasting performance. However, these statistics are insufficient for evaluating predictive power, especially across different specifications. In particular, the usefulness of the adjusted R-squared statistic for model selection is limited, as it critically depends on the specified functional form.[33] For instance, a specification that describes a trending variable in levels, using a linear trend, would generally provide a higher adjusted R-squared than an alternative specification that models the year-over-year change of the same variable as a constant, even when the predictive performance of the two specifications is comparable.

Testing the relative accuracy of different models requires computing forecast error statistics. The bases for such statistics are the difference between the predicted and the realized values of the nontransformed dependent variable (for example, level rather than growth). The predictive performance of alternative specifications is commonly evaluated using the root mean-squared error (MSE) or the mean absolute error, with smaller values in both statistics reflecting better performance. When performance statistics are computed based on the same data sample that was used for estimation, measured performance will generally be higher among specifications that include more explanatory variables, thereby favoring specifications that tend to “overfit” the estimation sample.

Therefore, model selection should be based on an evaluation of out-of-sample forecast performance. The simplest procedure is to split the sample into two subsets, where one is used for model estimation and the other is used for computing some test statistics. Although this method is straightforward to implement, the resulting statistics are highly sensitive to how the data was split. One generalization of this approach is a rolling or expanding window. Here, the dataset is split into an estimation and testing sample while preserving its temporal order: estimation is done on earlier data and testing on later data. Rather than estimating the model only once, models are estimated multiple times on different windows. Forecast accuracy can be assessed using one- or multiple periods ahead forecasts after each window. Two variants of this approach are the rolling window procedure, where the estimation set size is fixed and moves forward in time as new data is included, and the expanding window, where the estimation set grows over time as new data is added.[34]

---

[33] Moreover, it is computed differently depending on whether the specification includes an intercept.

[34] An alternative generalization of the train-test-split approach is *k*-fold cross validation. Here, the dataset can be divided into k subsets (or folds). Each fold contains one or more years of data. The regressions are then sequentially estimated using $k - 1$ folds and tested on the remaining fold with a different fold excluded each time. This process is repeated *k* times, and the average test statistic across all folds is calculated. A special case of this approach is the leave-one-out cross validation, where *k* equals the number of data points.

Empirical Evaluation of Selected Methods

This section tests the relative performance of specific regression specifications, drawing on a large sample of annual country-level data. The dependent variable is CIT cash collections recorded in the IMF's WoRLD revenue database. For simplicity and lack of comprehensive alternatives, GDP recorded in the IMF's World Economic Outlook Database serves as the sole predictor.

Previous evidence on the forecast performance of different CIT models is scarce. An early assessment by Basu, Emmerson, and Frayne (2003) of the Institute for Fiscal Studies' forecasting record found that simple models based on macroeconomic drivers perform reasonably well but are often outperformed by forecasts that incorporate expert judgment—particularly for short horizons when recent receipts provide valuable real-time information. For longer horizons, regression-based approaches linking CIT to profits or GDP tended to yield smaller systematic errors, provided the relationships were re-estimated regularly and adjusted for policy changes. The work by Shahnazarian, Solberger, and Spånberg (2017) applied a Bayesian vector autoregression framework to Swedish data, decomposing CIT revenues into macro-linked components such as operating surplus and business income. Their results show that integrating macroeconomic dynamics improves forecast accuracy relative to trend-based specifications.

The findings on forecast precision should be interpreted with caution. They represent a lower-bound estimate of empirically informed accuracy as the same crude predictor is used for each country. At the same time, forecast errors overstate accuracy to the extent that observed, rather than predicted, GDP is used. In practice, differences between predicted and realized proxy tax bases account for sizable shares of forecast uncertainty. For instance, Basu and others (2003) show that errors induced by uncertainty in the proxy tax base (for example, mispredicted economic profit growth) amount to 50–150 percent of the residual error—where the residual is the component that would remain even if the proxy were predicted perfectly. At short horizons, the residual accounts for a larger share of the total forecast error.

Evaluation Design

The importance of policy changes is tested indirectly, using a set of theoretical and empirical forecasting approaches. Because no data are available on historic policy costings, the predictions cannot distinguish between baseline and policy reform scenarios. Moreover, many countries implemented gradual tax reforms over time, and time series–based predictions for these countries may suffer from an omitted variable bias. The analysis addresses this issue indirectly by applying two theoretical and eight empirical forecasting models (Table 2). In cases where policy changes affected historic revenue data, theoretically informed forecasts are expected to provide improved forecasts.

Table 2. Tested Models

	Model No.	Prediction Name	Prediction Equation for *t* + 1	Notes
Theoretical	1	Unit elasticity	$(1+g_{t+1})T_t$	Prediction assumes CIT will grow in line with GDP.
	2	Adjusted unit elasticity	$(1+\varepsilon_{t+1}\,g_{t+1})T_t$	Elasticity is adjusted conditional on the output gap, based on Box 6. The first stage output gap is estimated by assuming potential GDP is a second order polynomial in time.
Empirical	3	Constant growth	$(1+\alpha)T_t$	Estimating equation is growth of CIT revenue on a constant, α.
	4	Elasticity	$(1+\beta g_{t+1})T_t$	Estimating equation is growth of CIT revenue on GDP growth. No constant is included.
	5	Growth on growth	$[1+(\ \alpha+\beta_1 g_{t+1})]T_t$	Estimating equation is growth of CIT revenue on a constant and GDP growth.
	6	Growth on growth with rate	$\left[1+\left(\ \alpha+\beta_1 g_{t+1}+\beta_2\frac{\mathrm{d}\tau}{\tau}\right)\right]T_t$	This is same as Model No. 5 but with percentage change in statutory CIT rate included.
	7	ETR	$\exp[\beta_1\ln(G_{t+1}))]*s$	Estimating equation is log CIT on log GDP. No constant included. Prediction includes a smearing term *s*.
	8	Log-level	$\exp[\alpha+\beta_1\ln(G_{t+1}))]*s$	Estimating equation is log CIT on log GDP and a constant. Prediction includes a smearing term *s*.
	9	Log-level with rate	$\exp[\alpha+\beta_1\ln(G_{t+1})+\beta_2\ln(\tau)]*s$	This is the same as Model No. 6 but with the log tax rate included.
	10	Dynamic	$[1+(\ \alpha+\beta_1 g_{t+1}+\beta_2\ln(T_t)+\beta_3\ln(G_t))]T_t$	Expands growth rate prediction by including lagged CIT revenues and lagged GDP as additional explanatory variables.

Notes: Variables are defined as follows: g is GDP growth, ε is the computed elasticity of CIT revenue with respect to GDP (see Box 6 for computation), *G* is GDP, CIT revenue is *T*, and *s* is the smearing term $\frac{1}{N}\sum_i \exp(u_i)$. CIT = corporate income tax; ETR = effective tax rate.

The models are estimated on an unbalanced panel covering 148 countries, between 1999 and 2017, and each model is used to produce about 3500 predictions of CIT collections. Countries were excluded if annual CIT collections fell by more than 75 percent or rose by more than 150 percent in any year, or if fewer than six years of CIT data were available. Each model is estimated using a rolling window of varying length. Specifically, 12 country-specific subsets are used, differing both in the last year of estimation (2014–17) and in the number of preceding observations included (*N* = 5, 10, or 15). Forecasts are generated for both the following year (*t* + 1) and for the subsequent year (*t* + 2), giving a total of 24 forecasts per country and model. To allow a comparison of forecast performance across countries, the forecast error is then divided by the actual revenue outcome to obtain the percentage deviation.

Performance is evaluated using two complementary measures, both of which are based on the mean absolute percentage error (MAPE):

- **Pooled MAPE**. The pooled MAPE is the mean absolute forecast error as a share of actual revenues, averaging across countries, horizons, and estimation window lengths. It measures overall forecast accuracy, treating each forecast error equally. A lower pooled MAPE indicates smaller average errors. However, this measure can mask systematic differences across countries, motivating the use of a second statistic.

- **Best-in-Class Share (BICS)**. The BICS records the share of cases in which a given model achieves the lowest MAPE among competing models. A higher BICS implies greater reliability, as the model delivers the most accurate forecast more often. BICS is computed only when all competing models produce valid forecasts for a given observation. Unlike pooled MAPE, BICS is less sensitive to outliers, since it records the frequency of wins rather than magnitudes of errors.

Findings

Overall, the results show that the unit-based elasticity prediction is difficult to beat when costings are not available to account for past tax policy changes. Table 3 summarizes forecast performance statistics for a restricted set of forecasting models, focusing on predictions that do not control for tax rate changes. On average, a prediction that assumes CIT collections will grow in line with GDP produces forecasts that deviate from actual CIT outcomes by an average of 14 percent. This average error is lower than those of any other tested model. For instance, the absolute error of a dynamic model is, on average, 60 percent larger (22.4 percent of actual CIT collections). The pooled MAPE statistics indicate that forecast accuracy generally worsens as more variables in general, and level information in particular, are included. When comparing relative forecast performance, the unit-based elasticity continues to outperform empirical models: in almost 25 percent, the unit-based elasticity provides a lower forecast error than competing models. If all models performed equally well, the BICS would be 14 percent (= 1/7).

The pooled MAPE averages across countries, estimation samples, and forecast horizons, masking important variation in forecast accuracy at those lower levels. This becomes apparent when contrasting evaluation metrics for empirical models, where the relative ranking of models shifts depending on whether MAPE or BICS is used. For instance, based on BICS, the dynamic model outperforms the constant growth model, even though the constant growth model produces significantly smaller average errors. This seemingly contradictory finding reflects differences in the conditional forecast accuracy: in cases where the dynamic model works, it produces more accurate forecasts. But where it does not, it can lead to outsized forecast errors that affect the pooled MAPE. The best performer among empirical models in terms of BICS is the dynamic model, which provides the most accurate prediction in 15.5 percent of evaluated cases.

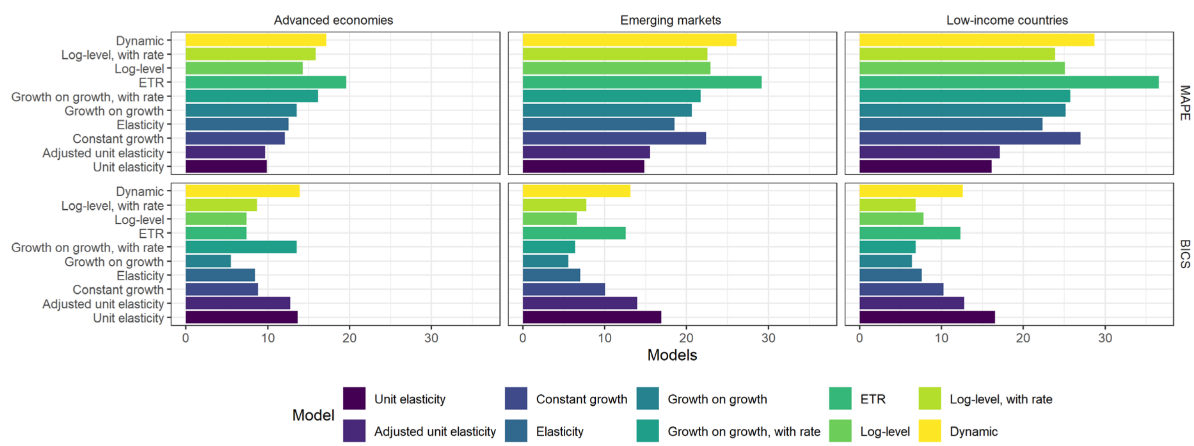
Table 3. Aggregate Forecast Performance

Model Class	Theoretical	Empirical					
Statistic	Unit Elasticity	Constant Growth	Elasticity	Growth on Growth	ETR	Log-Level	Dynamic
MAPE	0.140	0.209	0.180	0.200	0.288	0.213	0.244
BICS	0.232	0.136	0.103	0.122	0.123	0.130	0.155

Note: BICS = best-in-class share; ETR = effective tax rate; MAPE = mean absolute percentage error.

A more nuanced picture emerges when contrasting forecast performance across countries with different income levels (Figure 3). The breakdown shows that, irrespective of which model is used—theoretical or empirical, conditional on GDP or not—MAPE is inversely related to the income level within a given country: higher volatility of tax revenues and GDP makes forecasting more challenging in lower-income countries. In these more volatile environments, the relative superiority of the simple unit-elasticity model becomes more apparent. The BICS shows that the simple unit-elasticity approaches, and the extension incorporating variation in response to output gap dynamics produce the most accurate results in nearly 15 percent of cases in emerging and low-income countries.[35] The flipside of this observation is that empirical models work reasonably well—and outperform the unit-based elasticity in some cases—in advanced economies. In terms of MAPE, the most accurate prediction method in this subgroup is the adjusted elasticity approach. However, both the dynamic and the growth-on-growth regression that controls for tax rates provide a higher relative performance, measured by BICS, than the elasticity-based approaches. This implies that when empirical models fail—because of data gaps or unaccounted-for policy changes—they produce much larger errors than the unit-elasticity approach. Where they are appropriate, however, they deliver more accurate results.

Figure 3. Forecast Performance Conditional on Income Group



Notes: Forecast accuracy is higher when MAPE is low and when BICS is high. BICS = best-in-class share; ETR = effective tax rate; MAPE = mean absolute percentage error.

More broadly, the results underscore that the performance of time-series models depends on the forecasting horizon and data availability. Table 4 summarizes forecast statistics conditional on estimation sample size and forecast horizon. Overall, the unit-based elasticity continues to outperform other models in most cases. Moreover,

- **The MAPE of empirical models tends to decrease with larger estimation samples.** The exception are models that use information in levels (ETR, log-level, log-level with rate), which become less

[35] Note that the graph includes more models than Table 3 and therefore scales BICS differently. If all models performed equally well, the BICS would be 10 percent (= 1/10) in the graph, suggesting that the relative superiority of the elasticity approach is 150 percent (= 0.15/0.1). In Table 3, which does not differentiate between income groups and includes fewer models, the relative superiority is about 160 percent (= 0.23/0.14).

accurate as more distant information is included. The reason for lower forecast performance is likely the presence of structural breaks or continuous tax system change that make distant information less relevant. Among models that use growth rates as the dependent variable, the dynamic model benefits most from a larger sample size, with MAPEs reduced by almost 50 percent when the estimation sample includes 15 rather than 5 observations.

- **Parsimonious specifications tend to perform better with smaller estimation samples.** Among models that use CIT growth as the dependent variable, the elasticity-based model provides the most accurate forecast (evaluated in terms of BICS and MAPE) when only five observations are available to estimate the model. Although the ETR model performs badly in terms of MAPE, also compared to other models that use levels information, it outperforms other level-based models in terms of BICS when only five observations are available. However, with 15 observations in the estimation sample, the performance of more flexible models improves. In particular, all models that incorporate tax rate information outperform their otherwise identical counterparts that exclude it with 15 observations.

- **Information in levels becomes more important for two-year-ahead predictions.** Growth-based regressions tend to outperform level-based regressions for one-year-ahead forecasts. For example, with 15 observations, the elasticity model records a MAPE of 0.136, compared with 0.199 for the log-level model that includes rate information. However, for two-year-ahead forecasts, the gap in MAPEs narrows and the level-based regressions controlling for tax rate changes performs marginally better in terms of BICS than other growth-based regressions.

- **Dynamic models tend to outperform other empirical specifications in terms of BICS.** Compared to other empirical specifications, the dynamic model always obtains the highest or second-highest BICS irrespective of the forecast horizon or the sample estimation size. Like other specifications that use growth as the dependent variable, the MAPE decreases with sample size. Although its MAPE remains higher than the unit-based elasticity on average, the dynamic model matches the unit-based elasticity on BICS for two-year-ahead predictions when 15 observations are available (both 0.158).

Table 4. MAPE and BICS Conditional on Forecast Horizon and Estimation Sample Size

	Statistic: MAPE					
Forecast Horizon	*T* + 1			*T* + 2		
Observations	5	10	15	5	10	15
Unit elasticity	0.122*	0.122*	0.122	0.158*	0.158*	0.158*
Adjusted unit elasticity	0.126	0.127	0.119*	0.168	0.168	0.164
Constant growth	0.161	0.156	0.152	0.265	0.261	0.259
Elasticity	0.149	0.135	0.136	0.233	0.211	0.214
Growth on growth	0.172	0.145	0.145	0.266	0.234	0.236
Growth on growth, with rate	0.181	0.147	0.151	0.295	0.237	0.241
ETR	0.215	0.273	0.339	0.256	0.302	0.367
Log-level	0.168	0.180	0.219	0.238	0.222	0.256
Log-level with rate	0.173	0.179	0.199	0.247	0.223	0.243
Dynamic	0.276	0.169	0.157	0.387	0.229	0.218
	Statistic: BICS					
Unit elasticity	0.150*	0.143	0.165*	0.179*	0.154*	0.158*
Adjusted unit elasticity	0.142	0.150*	0.108	0.167	0.116	0.101
Constant growth	0.079	0.119	0.138	0.079	0.085	0.093
Elasticity	0.091	0.058	0.079	0.079	0.054	0.095
Growth on growth	0.061	0.062	0.042	0.065	0.069	0.038
Growth on growth, with rate	0.075	0.091	0.093	0.079	0.089	0.095
ETR	0.122	0.107	0.070	0.126	0.130	0.087
Log-level	0.072	0.075	0.073	0.059	0.079	0.071
Log-level with rate	0.084	0.070	0.095	0.055	0.070	0.104
Dynamic	0.122	0.125	0.136	0.110	0.152	0.158*

Notes: The top panel summarizes MAPE, where lower numbers imply more accurate forecasts. The lowest MAPE in each column indicated with an asterisk. The bottom panel summarizes BICS, where higher numbers imply more reliable forecasts. When all models worked similarly well, the BICS would be 10 percent. The highest BICS in each column indicated with an asterisk. BICS = best-in-class share; ETR = effective tax rate; MAPE = mean absolute percentage error.

The empirical findings highlight that model testing, using out-of-sample forecast statistics, should be a fundamental component of model design. Ideally, such tests are performed using that were available at the time the forecast was produced, so that both model and predictor uncertainty are factored into the overall assessment of forecast performance.

Documentation and Improving Forecast Performance

Appropriate procedures should be in place for documenting and archiving data, statistical programs, and methodological approaches so that relevant information is available to future forecasting teams and other stakeholders. In addition, it is important to evaluate past forecasting performance, diagnose shortcomings, and explore potential data and modeling improvements with an eye toward improving the accuracy and robustness of future forecasts. The procedures discussed in the following are broadly appropriate not just for CIT forecasting but also for all revenue forecasts.

Documentation

Documentation of existing models used for revenue forecasting serves a number of important purposes, including

- Ensuring that the model developers have carefully thought through all the steps required to make a robust, accurate, and useful model.
- Allowing other contributors (for example, other staff in the Ministry or from other ministries, outside advisors) to understand the model and advise on its construction and use.
- Providing guidance to new model developers when there is staff turnover.
- Supplying enough operational guidance so a new modeler will know how to operate the model if the model developer is unavailable.

Models should be documented in two ways. First, a standalone text document should describe the methodology and associated data used in the model. The level of detail provided should increase with the complexity of the model. The second form of documentation is inside the model itself. Brief operational instructions and explanations should be provided automatically when the model file is opened. These instructions should be easily bypassed by those routinely using or updating the model. Furthermore, all data and calculations inside the model should be clearly labeled. Labels should be long enough to avoid confusion between related data points, but short enough to avoid being cumbersome.

Following a modeling standard can help to ensure clear internal documentation. There are several available commercial and open-source financial modeling (spreadsheet) standards. They generally follow similar principles of good practice, such as the separation of inputs, calculation and outputs; consistency of structure and layout; and clarity of labeling. It is important for a modeling unit to pick a single standard and apply it consistently. Which standard is picked is less important. It is often useful to customize an existing standard to better fit the revenue-forecasting context so that the procedures are both effective and straightforward to implement. This customized standard should nonetheless be consistently followed by everyone in the modeling unit. Available modeling standards include Institute of Chartered Accountants in England and Wales Modeling Code; FAST Standard (Flexible, Appropriate, Structured, Transparent); Smart/Mazars Standard, Operis, and Modano BPM.

Version Control and Archiving

Revenue-forecasting models and data used in their estimation are constantly changing, as model developers gain new insights, receive new data, face new legislative requests, or fix old bugs. It is essential that a robust system of version control and archiving be implemented. Careful file naming and storage is all that is necessary. Operational files should always be stored in a shared, backed-up folder. Filenames should be clear, unique, and should include a version number.

In general, linkages between files should be avoided, as they create too many opportunities for errors. The one exception to this rule is when a model uses a large and relatively unchanging dataset. In that case, it is useful to separate the data and the model into separate files. In this way, the same data file can be used with multiple versions of the model, without the need to create and store multiple copies of the same data.

When a dataset and model have been used to produce published results, that file (or files) should be locked and archived. Records must be kept of which dataset and model version produced which published result. Published results can then be reproduced at any time in the future. Tracking procedures should be incorporated into the models so that new results (using new data or methodology) can easily be compared with prior results.

Model Verification and Evaluation

Model verification involves two separate issues, namely correctness and accuracy. The first requirement is that the model correctly calculates what it is supposed to calculate. Verifying that a model's calculations are correct requires breaking those calculations into small discrete pieces that can easily be checked for logic and including cross-checks in model calculations. Use of a small test dataset, explicitly designed to trigger all model calculations, should be used frequently and repeatedly when building or modifying a model to ensure that each upgrade or new calculation is correct. Full-scale revenue-forecasting models are sometimes enormously complex. In such cases, it impossible to verify whether calculations are correct without following these procedures.

How accurately a model forecasts tax revenues can, of course, only be determined after the forecast period is over. This is typically more than a year after the model was built and executed. Procedures should be established to compare published model results (from the archive) with actual revenue performance and to record and store the findings. Significant biases and errors identified from past forecasts should be investigated, and insights gained from such investigations should be used to explore potential data and modeling improvements for future forecasts. Forecast errors along with a summary of the analysis should ideally be published to inform stakeholders about revenue-forecasting performance.

References

Basu, Suman, Carl Emmerson, and Christine Frayne. 2003. “An Examination of the IFS Corporation Tax Forecasting Record.” IFS Working Papers No. 03/21, Institute for Fiscal Studies, London.

Box, George E. P., Gwilym M. Jenkins, Gregory C. Reinsel, and Greta M. Ljung. 2016. *Time Series Analysis: Forecasting and Control*. Hoboken, NJ: John Wiley & Sons, Inc.

Devereux, Michael P., and Rachel Griffith. 1999. “The Taxation of Discrete Investment Choices.” Institute for Fiscal Studies Working Paper Series No. W98/17. https://taxation-customs.ec.europa.eu/document/download/6eff89b7-96cf-4256-ad06-fea820f27d15_en

Girouard, Nathalie, and Christophe Andre. 2005. “Measuring Cyclically-Adjusted Budget Balances for OECD Countries.” Economics Department Working Papers No. 434, OECD.

Hebous, Shafik, Alexander D. Klemm, and Yuou Wu. 2021. “How Does Profit Shifting Affect the Balance of Payments?” IMF Working Paper WP/21/41, International Monetary Fund.

Hyndman, Rob J., and George Athanasopoulos. 2025. *Forecasting: Principles and Practice*, 3rd edition. Melbourne: OTexts. https://otexts.com/fpp3/

Shahnazarian, Hovick, Martin Solberger, and Erik Spånberg. 2017. “Forecasting and Analysing Corporate Tax Revenues in Sweden Using Bayesian VAR Models.” *Finnish Economic Papers* 28 (1): 50–74.

Mills, Terence C. 2015. *Time Series Econometrics: A Concise Introduction.* New York: Palgrave McMillan.

Annex 1. Corporate Income Tax Volatility across Countries

We measure corporate income tax (CIT) volatility, borrowing insights from the capital asset pricing model literature on the measurement of volatility in securities prices. Just as that framework relates the volatility in the price of an individual security to the overall volatility of the market portfolio of securities using the beta measure, we relate the volatility of CIT revenues to the overall volatility of a country's "portfolio" of tax instruments using this measure.

In this context, the CIT beta for a given country is the estimated slope coefficient obtained by regressing the growth rate in its aggregate CIT revenue (approximated by the year-over-year change in the natural log of aggregate CIT revenue) against the growth rate of its overall tax revenue:

$$\left[\ln\left(R_t^{\mathrm{CIT}}\right) - \ln\left(R_{t-1}^{\mathrm{CIT}}\right)\right] = \alpha + \beta\left[\ln\left(R_t^{\mathrm{Total}}\right) - \ln\left(R_{t-1}^{\mathrm{Total}}\right)\right] + \varepsilon_t.$$

Letting the share contributed by tax *i* given by $\omega_i = \frac{R^{\mathrm{Total}}}{R^i}$. The estimated slope coefficient can be expressed as follows:

$$\hat{\beta} = \frac{\mathrm{cov}(g_{\mathrm{CIT}},\ g_{\mathrm{TOT}})}{\mathrm{var}(g_{\mathrm{TOT}})} = \frac{\mathrm{cov}(g_{\mathrm{CIT}}, g_{\mathrm{TOT}})}{\sum \omega_i \mathrm{cov}(g_i, g_{\mathrm{TOT}})}.$$

Here, the numerator quantifies the extent to which CIT revenue growth co-varies with overall tax revenue growth, reflecting how changes in CIT translate into aggregate volatility. The denominator standardizes this covariance by the overall volatility of tax revenue. The second equality provides an alternative interpretation, showing that total revenue volatility is essentially a weighted average of the covariances between the growth of each individual tax instrument and total revenue, where the weights (ω_i) reflect the shares of total tax revenue accounted for by each instrument. Consequently, a beta greater than 1 implies that CIT revenue is more volatile relative to total revenue than the average tax instrument. When $\widehat{\beta}$ is multiplied by the share of overall tax revenue accounted for by the CIT, ω_{CIT}, one obtains the weighted-beta measure, which represents the share of overall tax revenue volatility for which the CIT is responsible.[36]

[36] The CIT revenue share is based on the most recent year for which tax revenue data is available for a given country. A modest adjustment is made to the CIT weighted beta measure to ensure that the weighted beta measures for CIT and non-CIT revenues sum to one (as the share of CIT revenue volatility accounted for by all tax instruments combined must be 100 percent).

Annex 2. Incorporating Cyclical Elasticities

The model builds on three basic equations that describe cyclical fluctuations of GDP and their interaction with revenues:

- First, the actual output is the product of potential output and a cyclical component, $G = SC$, where *G* is actual GDP, *S* is potential GDP, and *C* is the cyclical factor.
- Second, revenues, given by $R = R^s R^c$ include a component that responds to output gap variation, given by $R^c = e^{\beta \log\left(\frac{G}{S}\right)}$, where β is the elasticity of the cyclical factor with respect to the output gap, $\log\left(\frac{G}{S}\right)$.
- Third, the structural component of revenue responds to potential output $R^s = \tau S$ with τ being the ETR. This third equation imposes a long-term stable relationship between revenue and GDP.

Combining these equations shows that the sensitivity of revenues to changes in the output gap can be estimated using the following equation:

$$\log\left(\frac{R}{S}\right) = \log(\tau) + \beta \log\left(\frac{G}{S}\right)$$

Studies focusing on corporate income taxes find estimates of the coefficient β are typically about 1.5.

This framework has a straightforward and intuitive implication for the elasticity of revenue. Taking first differences and rearranging the equation shows the elasticity of revenue with respect to GDP is given by

$$\varepsilon_{R,G} = 1 + (\beta - 1)\frac{d\log(G) - d\log(S)}{d\log(G)}$$

Annex 3. Proxy Dynamics and the Accuracy of Levels and Differences

Whether levels or differences gives a more accurate prediction depends largely on the wedge between the true and the proxy tax base, and how this wedge evolves over time. To see this, consider a simplified data generating process, where

- the (log) tax revenues are given by $T_t = \tau + b_t + \varepsilon_t$, with *b* denoting the true tax base;
- the actual tax base follows as a random walk with drift: $b_t = b_{t-1} + \mu + v_t$; and
- the proxy tax base is a noisy measurement of the true: $g_t = b_t + u_t$, meaning that the proxy tax base is co-integrated with tax liabilities.

We assume that the errors are mutually independent but allow for serial correlation in the wedge between the true and the proxy tax base.

MSE of the levels regression. A regression in levels, of *T* on *g*, identifies the coefficient on *g* as follows:

$$\beta^{\mathrm{Lev}} = \frac{\mathrm{Cov}(T, g)}{\mathrm{Var}(g)} = \frac{\mathrm{Var}(b_t)}{\mathrm{Var}(b_t) + \sigma_u^2}$$

Because the tax base has a unit root, its variance increases over time and the coefficient estimate converges toward 1 (the true partial effect), in spite of the measurement error. In the limit, the level regression will produce the following error $e^{\mathrm{Lev}} = T_{t+1} - \alpha - \beta^{\mathrm{Lev}} g_{t+1} = \varepsilon_{t+1} - u_{t+1}$, meaning that the MSE is

$$\mathrm{MSE}^{\mathrm{Lev}} = \sigma_\varepsilon^2 + \sigma_u^2$$

MSE of the differenced regression. A regression of ΔT on Δg, identifies the coefficient on Δg as follows:

$$\beta^{\mathrm{Lev}} = \frac{\mathrm{Cov}(\Delta T, \Delta g)}{\mathrm{Var}(\Delta g)} = \frac{\sigma_v^2}{A + \sigma_v^2}$$

where $A = 2\sigma_u^2(1-\rho)$, σ_x^2 denotes the variance of *x*, and ρ is the first-order auto correlation coefficient in *u*. Unless the proxy noise is has a first-order auto correlation coefficient of 1, the estimated coefficient is thus downward biased because of noise in the measurement of the true tax base.

The differenced regression will produce the following error $e^{\mathrm{Dif}} = T_{t+1} - T_t - \beta^{\Delta}\Delta g_{t+1} = v_t + \varepsilon_{t+1} - \varepsilon_t - \beta^{\Delta}(v_{t+1} + u_{t+1} - u_t)$, meaning that the MSE is

$$\mathrm{MSE}^{\Delta} = 2\sigma_\varepsilon^2 + \frac{\sigma_v^2 A}{\sigma_v^2 + A}$$

Which one is better? The levels regression will provide a more accurate prediction when

$$\mathrm{MSE}^{\mathrm{Lev}} < \mathrm{MSE}^{\Delta} \rightarrow \; \sigma_v^2[A - (\sigma_u^2 - \sigma_\varepsilon^2)] > A(\sigma_u^2 - \sigma_\varepsilon^2).$$

Where we assume that tax noise is smaller than variation in the proxy tax wedge: $\sigma_\varepsilon^2 < \sigma_u^2$.

A few boundary cases are useful to consider:

- **Perfect serial correlation in the proxy tax wedge.** In this case, the tax wedge does not change from one period to the next, implying that $A = 0,$ and the condition would require $\sigma_v^2 < 0$ for levels to beat differences. As the variance of the tax base cannot become negative, the difference equation is more accurate for all combinations of other errors.

- **No serial correlation in the proxy tax wedge and not tax noise.** When errors in the proxy are white noise, differencing doubles the noise, meaning that $A = 2\sigma_u^2$. Furthermore, without tax noise $\sigma_\varepsilon^2 = 0$ the condition simplifies to the requirement that $\sigma_v^2(\sigma_u^2)^{-1} > 2$. The left-hand side in this condition is the signal-to-noise ratio of the levels prediction. When there is either a lot of stochastic movement in the true tax base, which is picked up by the proxy, and when the proxy tax base is measured with great accuracy, the signal-to-noise ratio is large and the levels regression will give a more accurate prediction.