

Development of an Integrated Macroeconomic Database and Automated Modelling Framework for Enhanced Policy Analysis Concept Note

Chirume Admire Tarisirayi

2025-06-15

Contents

1	Introduction	3
2	Objectives of the Integrated System	3
3	Milestones Achieved	4
4	The Need for Centralized Hosting on a Server	7
5	Conclusion	8

1 Introduction

Economic policy-making relies heavily on robust data management and macroeconomic modelling. On numerous occasions in developing countries, the approach is fragmented, leading to inefficiencies, inconsistent data handling, and limited collaboration. This concept note proposes the development of an Integrated Macroeconomic Database and Automated Modelling Framework that will centralize data management, automate model execution, and streamline policy analysis. By integrating macroeconomic datasets into a structured SQL-based database and linking them with predefined economic models in R and Python, this initiative aims to enhance data accessibility, improve forecasting accuracy, and allow economists to focus on policy interpretation rather than the technical execution of models.

When economic model building involves repetitive data handling, manual execution of diagnostic tests, and interpretation of results, the process consumes valuable time that should be devoted to economic analysis and policy formulation. The proposed system eliminates these inefficiencies by storing predefined models in a centralized database, allowing users to retrieve model outputs, forecasts, and diagnostics through an intuitive interface. By automating model execution, the system ensures consistency, reduces human error, and enhances collaboration across departments. Furthermore, integrating machine learning techniques and AI-driven analytics will enhance the predictive power of macroeconomic models and provide real-time insights for economic decision-making.

2 Objectives of the Integrated System

The core objective of this initiative is to develop a centralized database that enhances data integrity, efficiency, and policy analysis. This system will allow seamless retrieval and analysis of macroeconomic indicators, including GDP, monetary aggregates, fiscal data, external sector statistics, and exchange rate dynamics. Another key objective is to integrate predefined macroeconomic models within the system, enabling automated execution of diagnostics, forecasting, and policy simulations. By shifting the focus from model execution to economic interpretation, economists will be able to concentrate on evaluating economic conditions, identifying risks, and formulating policy recommendations. The system also aims to improve collaboration by providing multi-user access, where different departments can leverage the same database and models, ensuring a unified approach to economic research.

Additionally, this initiative will incorporate AI-powered forecasting techniques to complement traditional econometric models. Machine learning methods such as ARIMA, VAR, LSTM, and XGBoost will be deployed to improve macroeconomic projections. Furthermore, a natural language processing (NLP)-based chatbot will be integrated to facilitate real-time data queries, automating routine data retrieval tasks and generating instant policy reports. The implementation of this database-driven system will thus optimize decision-making processes, improve forecasting accuracy, and create an efficient workflow.

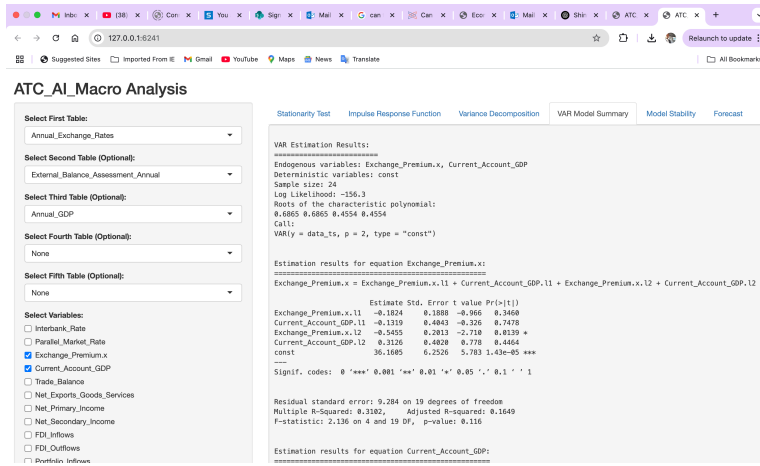
3 Milestones Achieved

Significant progress has already been made in the development of this system. A structured SQL database has been designed to store key macroeconomic variables, ensuring consistency and reliability in data management. The database includes well-defined tables for GDP, monetary and fiscal indicators, exchange rates, and trade statistics, with data frequencies standardized to accommodate monthly, quarterly, and annual time series. Below is a screenshot of the structure of the database and the proposed schema.

The screenshot shows the DB Browser for SQLite interface. On the left, a tree view lists the database tables. The main window displays the 'Net_Exports_Goods_Services' table, which contains macroeconomic data. The table has five columns: 'Net_Exports_Goods_Services', 'Net_Primary_Income', 'Net_Secondary_Income', and 'FDI_inflow'. The data is organized by time period, with rows for various years and quarters.

	Net_Exports_Goods_Services	Net_Primary_Income	Net_Secondary_Income	FDI_inflow
1856	205.260765245446	161.801586320997	1.84921883040627	168.102420046
1857	-89.980549783498	174.331827554862	-92.3870206487809	283.853823246
1858	-178.2399222915176	195.087346185196	38.3101352896146	289.847742679
1859	-118.522576941177	49.4570638984442	95.7435461003333	492.080988333
1860	228.49405431442	-88.3452813141048	65.8726701512933	394.11761041
1861	-269.353247044384	-70.3931351192296	38.4528329551965	84.5192018903
1862	294.717137049884	120.527192391455	87.6282443292439	142.325004485
1863	-49.4544439483434	63.9400762805829	-52.0817779470235	123.338102176
1864	286.806843073656	91.4841733269485	-9.39258700236678	492.558013514
1865	150.286355545993	185.419074259698	81.053933722228	178.00632866
1866	-188.009780785059	119.854871979654	34.4391037710011	272.868653584
1867	-32.9537520606816	159.180927090347	42.7350750568661	443.703002439
1868	-187.493956508115	35.7039380818605	-49.4130345992744	280.054110065
1869	-16.1307475995272	-33.008238441825	39.7958116896854	293.371096628
1870	-155.791245074943	-36.7177566071183	-38.3829086651983	371.260018704
1871	171.599056525156	-121.605642206999	-65.7351349706885	124.686079475
1872	-93.5265127569437	9.79290520772336	-45.491726979888	431.311756931
1873	234.488951973617	30.2949201318641	-40.3191420349541	85.5150047875
1874	209.642458960528	21.6032300911644	3.92737821675837	423.147628976
1875	-255.956198740761	99.7433044016361	22.60870359611273	481.328289795
1876	-8.06231056176138	112.597613969589	-9.28259944722772	-54.1591981425881
1877	748496772480965	-65.7039823941886	99.9400339614851	-73.8356138594776
1878	-3.52746730670333	-189.445321541128	56.7988323498189	12.864575330059
1879	-0.70127691142261	-164.451172850356	-8.58389474451542	11.0398183576763

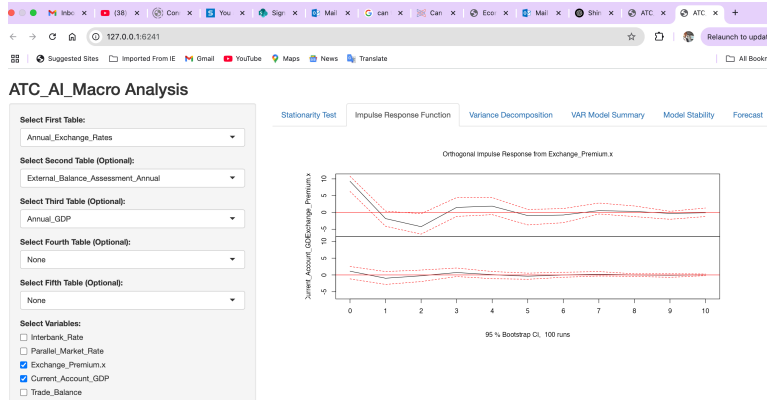
This database was developed in SQLite and will be migrated to MySQL to enhance scalability and allow multi-user access.



With the data infrastructure in place, the next milestone was integrating the database with economic modelling frameworks in R and Python. The system has been configured to execute predefined models in the background, automatically performing stationarity tests, cointegration analysis, and impulse response functions before presenting results to the user. This automation eliminates the need for manual model specification, allowing economists to directly access output diagnostics and forecasts through a user-friendly interface. Moreover, the system is equipped to store historical forecasts, enabling the evaluation of past model performance and refining predictive accuracy over time.

Below is a screenshot of the Graphic User Interface (GUI) which users across the departments with the Division can use to run the pre-defined AI and machine learning generated models

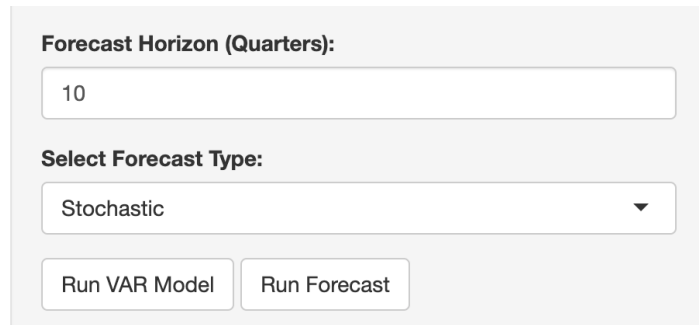
which are integrated in R and the Database.



The modelling framework underlying this project is robust and consistent with standard classical econometrics, making it ideal and plausible for use by individuals with or without a strong econometrics and modelling background. The Framework is also capable of generating standard IRF for policy analysis. The screenshot underneath is one such example using dummy data in the database.

The framework also provides options for stochastic or deterministic simulation and provides an automated reporting framework for policy analysis. Below is a dialog box showing the option for running standard VAR models for policy analysis, using data mined from the

database, and for forecasting.



The image shows a web form with a light gray background. At the top, it says "Forecast Horizon (Quarters):" in bold. Below this is a text input field containing the number "10". Underneath that, it says "Select Forecast Type:" in bold. Below this is a dropdown menu with "Stochastic" selected and a downward arrow. At the bottom, there are two buttons: "Run VAR Model" and "Run Forecast".

4 The Need for Centralized Hosting on a Server

Hosting the database on a centralized MySQL server is critical for multi-user access, data security, and enhanced computational performance. A centralized system ensures that all economists within the Division can access updated macroeconomic data in real-time, eliminating redundancy and the risk of working with outdated information. It also enhances data security by implementing controlled access, encryption, and regular backups to prevent data loss or tampering.

5 Conclusion

The establishment of an integrated Macroeconomic Database and Automated Modelling Framework is paramount and may mark a significant advancement in economic research and policy analysis. By automating model execution and integrating AI-driven forecasting tools, this initiative will free up valuable time for economists to focus on interpretation and policy recommendations. Given its transformative potential, it is recommended that resources be allocated to support the full implementation of this system, ensuring that the Division remains at the forefront of modern macroeconomic research and policymaking.