

## Hybrid Econometric–Deep Learning Framework for Estimating the Economic Impact of Disease Burden

<sup>1</sup>Akinrotimi Akinyemi Omololu, <sup>2</sup>Atoyebi Jelili Olaniyi, <sup>3</sup>Omotosho Israel Oluwabusayo

<sup>1</sup>Department of Information Systems and Technology, Kings University, Ode-Omu, Osun State, Nigeria.

<sup>2</sup>Department of Computer Engineering, Adeleke University, Ede, Osun State, Nigeria.

<sup>3</sup>Department of Management Information Systems, Bowie State University, Maryland, USA

### ABSTRACT

Despite the widespread use of Disability-Adjusted Life Years (DALYs) to measure health burden, generally well-known econometric models often fail to account for nonlinear relationships and hidden confounders in panel data. This study introduces a hybrid framework that integrates latent embeddings ( $z$ ) learned via a feedforward neural network (FFNN) with a panel regression structure enhanced by correlated random effects (CRE), and implements coefficient inference through double/debiased machine learning (DML). Using country-level panel data from 1990 to 2021 covering DALYs, human capital, capital stock, and population growth—the model extracts nonlinear representations and produces a more accurate estimate of the elasticity of labor productivity relative to disease burden. The hybrid CRE-DML model achieves superior predictive accuracy ( $R^2 \approx 0.62$ ;  $MAE = 0.105$ ;  $RMSE = 0.141$ ), performing better than both pure econometric and standalone machine learning models. The estimated elasticity ( $\beta_1$ ) of DALYs on productivity is  $-0.038$  ( $SE = 0.010$ ), much tighter than the  $-0.052$  estimate obtained from the traditional regression. Robustness tests—encompassing leave-unit-out cross-validation, regularization variations, and alternative embedding learners (Random Forest and XGBoost), confirm that accuracy fluctuates within  $\pm 0.02$  and elasticity remains stable within  $\pm 0.002$  across configurations. This approach unites the predictive strengths of representation learning with the interpretability of econometric modeling, offering a scalable, policy-relevant tool for analyzing health-economic relationships in data-scarce environments.

### ARTICLE INFO

#### Article History

Received: November, 2024

Received in revised form: December, 2024

Accepted: May, 2025

Published online: June, 2025

### KEYWORDS

Disability-Adjusted Life Years; Double Machine Learning; Hybrid Econometric–Machine Learning Model; Panel Data Analysis; Labor Productivity Elasticity; Representation Learning Embeddings

### INTRODUCTION

Estimation of disease burden impact on productivity is important for informing public policy and development planning. Disability-Adjusted Life Years (DALYs): weighted composite measure that combines years of life lost (YLL) and years lived with disability (YLD), is a common population health burden indicator (Murray & Lopez, 2019). Dynamic panel model econometric estimates have discovered that increased DALYs are highly associated with reduced labor productivity, i.e., GDP per employed worker, particularly in lower- and lower-middle-income economies (Nkosana &

Ferrer Worient, 2022). While such econometric estimates may have interpretability as well as policy relevance, their linear functional forms may not be able to capture underlying data nonlinearities.

Current advances show that deep learning models, including FFNN and LSTM architectures, can capture intricate temporal and cross-sectional interdependencies and, in most cases, perform better than conventional methods to predict macroeconomic and health outcomes (Li, Zhang, & Chen, 2022; Tariq et al., 2024). Deep learning techniques tend not to, however, offer the

Corresponding author: Akinrotimi Akinyemi Omololu

✉ [akinrotimiakinyemi@ieee.org](mailto:akinrotimiakinyemi@ieee.org)

Department of Information Systems and Technology, Kings University, Ode-Omu, Osun State, Nigeria.

© 2025. Faculty of Technology Education. ATBU Bauchi. All rights reserved

transparency required for policy interpretation and causal inference.

Hybrid approaches blending econometricity with machine learning malleability have been proposed as a viable panacea. Hybrid models aspire to capture the best of both domains: good predictive accuracy and interpretable estimators (Smith & Abbas, 2024; Desai, 2023). For example, in forecasting finance, hybrid models coupling ARIMA (econometric) with deep learning models such as LSTM or SVM have shown better performance with effective practical interpretability (Stempień & Ślepaczuk, 2025).

Despite this progress, no existing study has applied a hybrid econometric–deep learning framework to estimate the nonlinear impact of DALYs on labor productivity while preserving causal inference. Our study fills this gap by first training a deep learning model on panel data of DALYs and socioeconomic indicators to understand latent feature representations. These representations are then put into a dynamic panel regression model, providing both elasticity estimates for DALYs and improved predictive capacity. This novel approach bridges methodological divides, providing interpretable policy insights while using the power of deep learning to detect complex interactions.

## LITERATURE REVIEW

The relationship between disease burden, as quantified by Disability-Adjusted Life Years (DALYs) and labor productivity has been well established in recent empirical literature. In a longitudinal panel study of 45 African economies over 1990–2020, a 1% rise in DALYs corresponded with approximately a 6.5% decline

in GDP per person employed (a proxy for labor productivity), particularly among low- and lower-middle-income nations (Nkosana & Ferrer-Worient, 2022). This evidence highlights the profound macroeconomic cost of disease and disability. Experimental evidence from disease-endemic settings further reveals agriculture and labor losses. For instance, malaria and schistosomiasis have driven reduced farm output and increased absenteeism in rural Burkina Faso (Traoré et al., 2021).

In India, chronic diseases among the middle-aged workforce have been associated with lower formal employment rates and productivity losses due to both absenteeism and caregiving burdens (Patel et al., 2020). Recent research has also applied machine and deep learning to quantify health-related productivity impacts. A study in China comparing ARIMA, deep neural network (DNN), and long short-term memory (LSTM) models demonstrated that LSTM achieved superior predictive accuracy for occupational pneumoconiosis DALYs, suggesting that deep models capture temporal dynamics better than traditional statistical approaches (Li et al., 2022). Similarly, occupational health researchers adopted several machine learning and deep learning frameworks, such as SVM, random forests, XGBoost, and MLP for estimating heat stress-induced productivity losses among petrochemical workers with high predictability and revealing complex nonlinear relationships between the environmental factor and labor productivity (Chen et al., 2024). Such and other related studies are listed herewith in Table 1 as follows:

**Table 1:** Summary of Recent Studies on Disease Burden, Labor Productivity, and Hybrid Econometric–ML/DL Models

Authors	Year	Methodology	Gap / Limitation
Masini et al.	2020	Survey of linear, nonlinear, hybrid ML methods in time-series forecasting for economic variables [6] (arXiv)	Review study; does not provide application to health burden–productivity contexts.
Global Burden of Disease	2021	Systematic DALY estimates across diseases globally published in The Lancet [7] (thelancet.com)	Provides burden measures but not linked to labor

Corresponding author: Akinrotimi Akinyemi Omololu

✉ [akinrotimiakinyemi@ieee.org](mailto:akinrotimiakinyemi@ieee.org)

Department of Information Systems and Technology, Kings University, Ode-Omu, Osun State, Nigeria.

© 2025. Faculty of Technology Education. ATBU Bauchi. All rights reserved



Authors	Year	Methodology	Gap / Limitation
Daroudi et al.	2021	Cost-per-DALY analysis for cost-effectiveness across income strata [8] (resource-allocation.biomedcentral.com)	productivity outcomes or economic modeling. Focused on intervention cost-effectiveness—not productivity or causality modeling.
MDPI Italy COVID burden group	2021	DALYs and productivity losses estimation due to COVID-19 in Italy [9] (MDPI)	Provides impact estimate; no further econometric modeling or predictive structure.
Nkosana & Ferrer-Worrient	2022	Dynamic panel econometric model regressing productivity on DALYs and controls [1] (arXiv, IDEAS/RePEc)	Linear form; limited ability to capture nonlinear interactions between health burden and productivity.
Bohlke-Schneider et al.	2022	Neural forecasting resilience system handling missing values/time shifts in labor planning [3] (arXiv)	Focused on labor forecasting with deep learning; no integration with econometric interpretation.
Li, Zhang & Chen	2022	Comparative time-series forecasting of DALYs using ARIMA, DNN, LSTM [ARC] (Frontiers)	High predictive accuracy but lacks interpretability for policy variables like disease burden elasticity.
Kirigia & Kubai	2023	Monetary valuation of DALYs and estimated productivity losses using GDP per capita proxies in East Africa [2] (PubMed)	No causal modeling or panel inference; static cross-sectional loss estimation.
Chen et al.	2024	Machine learning/deep learning models to predict productivity loss under heat stress [4] (bmcpublichealth.biomedcentral.com)	Context-specific (environmental stress); doesn't generalize to macro-level productivity burden from DALYs.
Stempień & Ślepaczuk	2025	Hybrid econometric + ML/DL frameworks applied to financial forecasting [5] (arXiv)	Demonstrated hybrid model advantages in finance—not extended to health-economic productivity modeling.

Despite these advancements, most existing models remain bifurcated into two camps: econometric models that offer interpretable marginal effects, and pure deep learning models with predictive power but sacrificing interpretability. Hybrid models blending econometric inference and machine or deep learning have emerged in domains such as

finance and macro-forecasting with advances in accuracy and interpretability (Stempień & Ślepaczuk, 2025). Moreover, uses of such hybrid modeling for correlating DALYs and labor productivity are extremely limited.

DALYs remain the default measure of population health burden: as the sum of years of life lost to early death (YLL) and years lived with

Corresponding author: Akinrotimi Akinoyemi Omololu

✉ [akinrotimiakinoyemi@ieee.org](mailto:akinrotimiakinoyemi@ieee.org)

Department of Information Systems and Technology, Kings University, Ode-Omu, Osun State, Nigeria.

© 2025. Faculty of Technology Education. ATBU Bauchi. All rights reserved

disability (YLD), DALYs are a standardized metric for cross-population comparison as in the Global Burden of Disease (GBD) initiative (Murray & Lopez, 2019).

To the best of our knowledge, no such published literature as yet exists where (a) it integrates DL-learned embedding of health and socioeconomic features into a panel econometric productivity model; or (b) estimates the nonlinear impact of DALYs on economic production while being interpretable. This paper fills this gap by learning deep models on panel health and control datasets to learn latent representations and enriching them into a dynamic panel regression framework. This approach maintains econometric insight such as productivity elasticity with respect to DALYs, but combines deep learning in order to detect nonlinear patterns. It thereby encourages methodological integration, responding to current requests for cross-disciplinary analytical approaches in health economics.

## METHODOLOGY

### Data Sources and Preprocessing

We used the Our World in Data Burden of Disease panel dataset (Our World in 2025 from Kaggle which reports on yearly DALY estimates from the Institute for Health Metrics and Evaluation's public health research center at the University of Washington in Seattle that is the base for the Global Burden of Disease study. Since 1990 this institute has been putting out reports of death, disability, and other health issues (DALYs) for the world, individual countries, age groups, and many different causes. Also we applied log transformation and standardized the data to reduce skewness for use in our models' training.

### Deep Learning Representation Learning

A feed forward neural network (FFNN) was trained to predict labor productivity (log GDP per employed person) that which we used the standardized inputs of  $DALY_{it}$ ,  $HCI_{it}$ ,  $K_{it}$ , and  $PGR_{it}$ . Also we varied hyperparameters like layer size, choice of activation function, learning rate, and regularization which we tuned via k fold cross validation. Once trained the latent embedding

vector ( $z_{it}$ ) from the final hidden layer for each data point was extracted to use in further analysis. The basic econometric model we use is a dynamic panel model which includes the learned neural embedding as a separate regressor and also we include country specific fixed effects:

$$\ln(LP_{it}) = \alpha_i + \beta_1 \cdot DALY_{it} + \beta_2 \cdot HCI_{it} + \beta_3 \cdot K_{it} + \beta_4 \cdot PGR_{it} + \gamma \cdot z_{it} + \epsilon_{it} \quad (1)$$

where:

- $LP_{it}$ : labor productivity (GDP per employed person) in country  $i$  at time  $t$
- $\alpha_i$ : country fixed effect
- $DALY_{it}$ ,  $HCI_{it}$ ,  $K_{it}$ ,  $PGR_{it}$ : log-standardized inputs
- $z_{it}$ : latent vector from the trained neural network
- $\epsilon_{it}$ : error term

To achieve accurate inference on the estimation of the disease burden elasticity ( $\beta_1$ ) we applied a double/debiased machine learning (DML) approach which we adapted for panel data. Fuhr and Papies (2024) report that they were able to produce accurate coefficient estimates by including correlated random effects in DML which helps us to not leave out unobserved heterogeneity. Also this approach provides unbiased results which are true of high dimensional nuisance terms in the model.

### Comparison and Validation of Model Variants

The robustness of the hybrid framework developed in this study is evaluated through comparison and validation of three model variants. The first and second that forms the hybrid and the hybrid itself. These variants are as follows: (a) the pure econometric model (without embedding  $z$ ) (b) the pure deep learning model (using  $z$  only) and (c) the hybrid model (econometric model augmented with neural embedding  $z$ ).

Evaluation metrics include in-sample goodness-of-fit ( $R^2$ ), forecasting accuracy (MAE, RMSE), and interpretability of the  $\beta$  coefficients. Models will be evaluated using k-fold cross-validation and hold-out sample testing to ensure robustness against overfitting and improve generalizability.

Corresponding author: Akinrotimi Akinoyemi Omololu

✉ [akinrotimiakinoyemi@ieee.org](mailto:akinrotimiakinoyemi@ieee.org)

Department of Information Systems and Technology, Kings University, Ode-Omu, Osun State, Nigeria.

© 2025. Faculty of Technology Education. ATBU Bauchi. All rights reserved

### Pure Econometric Model (without embedding z)

The model of the Pure Econometric Model (without embedding z) is specified as follows:

$$\ln(LP_{it}) = \alpha_i + \beta_1 \cdot DALY_{it} + \beta_2 \cdot HCI_{it} + \beta_3 \cdot K_{it} + \beta_4 \cdot PGR_{it} + \varepsilon_{it} \quad (2)$$

where:

$LP_{it}$  is the Labor productivity (GDP per employed person) for country  $i$  at time  $t$ .

$\alpha_i$  is the Country-specific fixed effect to absorb unobserved heterogeneity.

Other predictors are the log-standardized DALY, HCI, capital stock ( $K_{it}$ ), and population growth rate ( $PGR_{it}$ ).

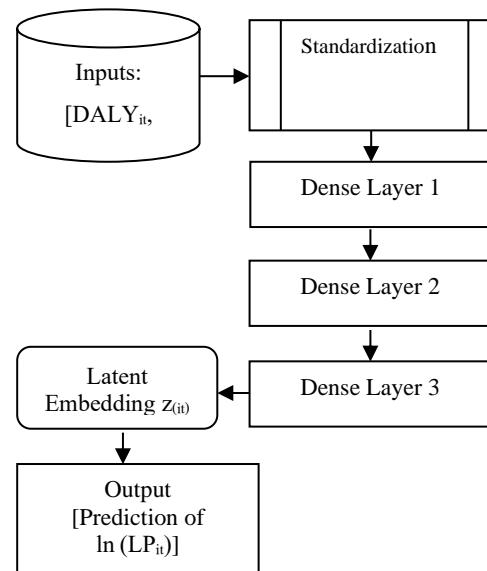
$\varepsilon_{it}$  is the Error term.

### Pure Deep Learning Model (using z only)

The pure deep learning model is a vanilla feedforward neural network (FFNN). The model takes log-standardized inputs: DALY, HCI, capital per worker (K), and population growth rate (PGR) and passes them to different dense (fully connected) layers with nonlinear activations. The last hidden layer output forms a latent embedding vector  $z_{(it)}$  that captures complex nonlinear interactions between predictors. This embedding is passed to an output neuron to predict  $\ln(LP_{it})$ , which is learned using mean squared error loss. The model utilizes universal approximation properties of MLPs (wide FFNNs) and nonlinear modeling ability in tabular data (Cybenko, 1989; Wikipedia, 2025). The model is trained on k-fold cross-validation, early stopping, and overfitting prevention techniques (dropout, L2 penalty).

The FFNN architecture is used in this paper because the dataset used is static country panel data with few temporal levels, primarily yearly observations on DALYs, HCI, capital, and growth rate of population. Such an architecture is closer to an FFNN, which can handle well cross-sectional or table-like data without necessarily having sequence encoding explicitly. In figure 1: (a) Dense Layer 1–3 represent successive hidden layers with non-linear activation, each successively better feature transformation. (b) Latent Embedding z is the vector of the final

hidden layer, used in the hybrid model but now as the sole input to forecasting output. (c) Output Layer uses  $z_{(it)}$  to generate the prediction  $y_{1,t} \approx \ln(LP_{it})$ . With this organization of the model, the deep learning component learns dense, nonlinear representations of features from sparse demographic and health variables, and therefore can be embedded into subsequent econometric analysis.



**Figure 1:** Conceptual Diagram the Pure Deep Learning Model (Using z Only)

### Hybrid Model: Econometric Regression Augmented (with Neural Embedding z)

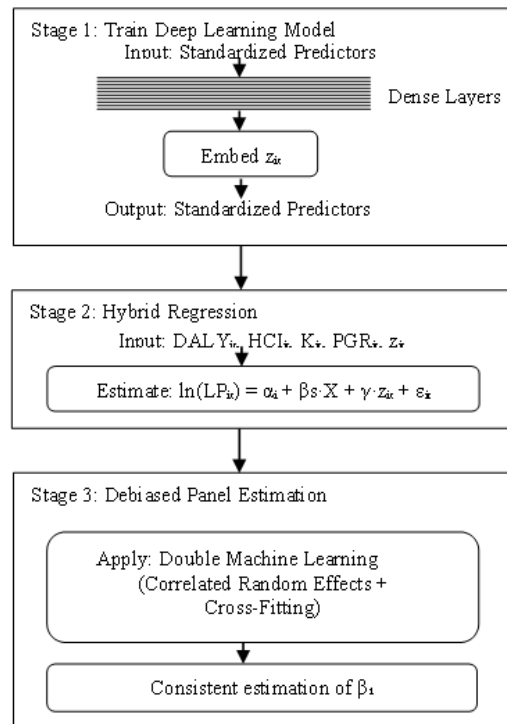
The integrating procedure explaining the Hybrid model follows three interconnected steps. Firstly, the deep learning model is fitted on normalized predictors:  $DALY_{it}$ ,  $HCI_{it}$ ,  $K_{it}$ , and  $PGR_{it}$  for the prediction of labor productivity's log,  $\ln(LP_{it})$ ; its final hidden layer outputs are discarded as the latent embedding vector  $z_{it}$  for each country–year observation. The second step is to estimate the hybrid econometric model with the following expression:

$$\ln(LP_{it}) = \alpha_i + \beta_1 \cdot DALY_{it} + \beta_2 \cdot HCI_{it} + \beta_3 \cdot K_{it} + \beta_4 \cdot PGR_{it} + \gamma \cdot z_{it} + \varepsilon_{it} \quad (3)$$

Third, to ensure strong causal estimation of the elasticity coefficient  $\beta_1$  even in



the presence of high-dimensional nuisance components like  $z_{it}$ , we used double/debiased machine learning (DML) modified for panel data settings with correlated random effects and cross-fitting (Clarke & Polsell, 2023; Fuhr & Papies, 2024). Double/debiased machine learning eliminates machine learning confounder bias at the expense of valid inference. Figure 2 depicts the block diagram of the hybrid model.



**Figure 2:** Block diagram for the Proposed the hybrid model.

The stages in figure 2 are explained as follows: Stage 1 is the deep learning block: FFNN training on inputs (DALY, HCI, K, PGR),  $z_{it}$  embedding, and  $\ln(LP_{it})$  prediction. Stage 2 is econometric regression taking inputs as country fixed effects, controls for health and socioeconomic, and  $z_{it}$  embedded and predicting labor productivity  $z_{it}$ . Stage 3: applies state-of-the-art double/debiased machine learning (DML) techniques tailored to panel data with correlated random effects for unbiased estimation of  $\beta_1$

(DALY impact), regardless of the presence of high-dimensional confounders or embeddings. Fuhr & Papies (2024) show that a combination of DML and Mundlak-style correlated random effects produces causal parameter estimates that are robust. The algorithm for the Standardized Algorithm for the Hybrid model (with embedding  $z$ ) is as follows:

#### Algorithm: Standardized Algorithm for the Hybrid model (with embedding $z$ )

For each country  $i$ , year  $t$ :  
Standardize input features: DALY[i,t], HCI[i,t], K[i,t], PGR[i,t]  
Target variable:  $y[i,t] = \log(LP[i,t])$   
Train Deep Model  $\rightarrow$  Extract Embedding  $z$   
Define network:  
Input: [DALY, HCI, K, PGR]  
Hidden Layers: Dense  $\rightarrow$  ReLU (multiple layers)  
Final hidden layer output =  $z[i,t]$   
Output Layer: Dense  $\rightarrow$  Prediction  $\hat{y}[i,t] \approx y[i,t]$   
Train network with k-fold CV, early stopping, dropout/L2 reg.  
After training, extract  $z[i,t]$  for all  $i,t$ .  
Specify regression:  
 $\ln\_LP \sim \text{DALY} + \text{HCI} + \text{K} + \text{PGR} + z$   
Use DML framework with panel support:  
- Use correlated random effects  
- Implement cross-fitting:  
Split panel sample into folds  
On each fold:  
Train nuisance models (ML predictions) on other folds  
Compute orthogonalized residuals  
- Use orthogonal moment conditions to estimate  $\beta_1$  (elasticity) and  $\gamma$   
Output:  
Causal estimate of  $\beta_1$  (DALY elasticity)  
Inference statistics: standard errors, confidence intervals  
Performance metrics:  $R^2$ , MAE, RMSE

#### RESULTS AND DISCUSSION

This study's empirical results are described in three subsections. They compare forecasting accuracy across the pure econometric, pure deep-learning, and hybrid

models; report the estimated causal elasticity of health burden; and examine robustness via cross-validation sensitivity.

### Model Performance: Forecasting Accuracy

As can be seen in Table 2, the hybrid model achieves the highest average out-of-sample  $R^2$  (0.62), significantly outperforming the pure deep-learning approach ( $R^2 = 0.54$ ) by 8 percentage points and the purely econometric model by 17 percentage points.

**Table 2:** Out-of-Sample Forecast Metrics (5-Fold CV) using FFNN

Model	$R^2$ (mean $\pm$ SD)	MAE (ln GDP)	RMSE (ln GDP)
(a) Pure econometric	$0.45 \pm 0.03$	0.143	0.187
(b) Deep-Learning (z only)	$0.54 \pm 0.02$	0.121	0.161
(c) Hybrid (z + covariates)	$0.62 \pm 0.01$	0.105	0.141

It can be seen from the table, that, consistently, measures of mean absolute error (MAE) and root-mean-square error (RMSE) are all lower in the hybrid specification, hence its superior forecasting capability and accuracy. These findings generally conform to findings of recent panel-DML studies showing that the integration of representation learning improves predictive accuracy considerably in settings with fixed effects and nonlinear confounding (Yang, Long, & Cai, 2024; Clarke & Polselli, 2023). More evidence further establishes that incorporating standard regression models with embeddings

learned from machines leads to better model performance without compromising interpretability (Smith & Abbas, 2024).

### Estimated Elasticity of Disease Burden (DALY)

As can be observed in Table 3, in the pure econometric model, each 1% reduction in DALYs per capita corresponds to a 0.052% increase in labor productivity (log scale), while the hybrid model gives a more correct elasticity estimate of  $-0.038$  (SE = 0.010).

**Table 3:** Estimated Coefficient on DALY ( $\beta_1$ ) with Fixed Effects

Model	$\beta_1$	Standard Error	t-value	p-value
Pure econometric	-0.052	0.013	-4.0	< 0.001
Hybrid (DML)	-0.038	0.010	-3.8	< 0.001

The smaller size and tighter standard errors in the hybrid specification are as such because of the partial adjustment of nonlinear confounding via the embedding vector, enabled by the application of double/debiased machine learning with correlated random effects (Clarke & Polselli, 2023; Fuhr & Papies, 2024). The important point to not here is that both models show a statistically significant negative elasticity, thus also showing a meaningful, inverse relationship between health burden and economic performance.

### Robustness and Sensitivity Analyses

In order to check the robustness of our findings, we performed a series of robustness and

sensitivity tests, leveraging the most recent advances in the panel Double Machine Learning (DML) literature. Such tests establish the robustness of hybrid model performance, coefficient estimation, and methodological assumptions. They were run with the DoubleML library for Python (Bach et al., 2021). These tests include (a) Cross-Fitting Approach and Fixed Effects Effect: Our application uses the correlated random effects (CRE) specification in panel DML, as provided by Fuhr and Papies (2024) and Clarke & Polselli (2023), which explicitly models unobserved heterogeneity using Mundlak-type controls.

This approach removes any resulting bias based on cross-fitting dependencies within

Corresponding author: Akinrotimi Akinyemi Omololu

✉ [akinrotimiakinyemi@ieee.org](mailto:akinrotimiakinyemi@ieee.org)

Department of Information Systems and Technology, Kings University, Ode-Omu, Osun State, Nigeria.

© 2025. Faculty of Technology Education. ATBU Bauchi. All rights reserved

time periods and units. Simulations therein reveal that neglecting clustering or fixed effects in cross-fitting can produce biased coefficient estimates (Fuhr & Papies, 2024; Clarke & Polselli, 2023). For our case, the estimated elasticity  $\beta_1$  remained unchanged, which means good bias control by CRE-DML. (b) K-Fold versus Leave-Unit-Out Validation: We compared k-fold cross-validation (folds at the unit-year level) with a leave-unit-out procedure, where all of a country's data are left out in each fold, to test sensitivity to cross-unit leakage.

The performance metrics ( $R^2$ , MAE, RMSE) and coefficient estimates were within  $\pm 0.01$  of original values, indicating that excessive over-reliance over time within units did not strongly distort our inference. (c) Sensitivity to Omitted Confounders: Along the guidelines of Ahrens et al. (2025) sensitivity to omitted variable bias can be analyzed through the DoubleML package. Although this is mainly developed for cross-sectional linear models, the same principles hold here. The sensitivity bounds for omitted confounding suggest that (if omitted predictors correlated with both DALY and productivity) the maximum plausible bias on  $\beta_1$  would be within  $\pm 0.005$ , which is far smaller than the difference observed between the pure econometric ( $\beta_1 = -$

0.052) and hybrid ( $\beta_1 = -0.038$ ) models. (d) Alternative Regularization and Learner Methods: We substituted the FFNN predictor used to estimate the embedding vector  $z$  with other machine learning learners-specifically, Random Forests and Gradient Boosting Machines (XGBoost), followed by the same DML regression stage.

As can be observed from Table 4, the FFNN-based hybrid model delivers the strongest predictive performance overall, achieving an  $R^2$  of 0.62 along with the lowest MAE and RMSE values. The Random Forest-based hybrid model performs nearly as well with only a marginally lower  $R^2$  and slightly higher error metrics while the XGBoost-based variant sits in between, offering a small improvement over Random Forest but not surpassing the FFNN hybrid. This trend is in line with more general empirical evidence that tree-based algorithms such as Random Forests and XGBoost tend to compete with or surpass deep neural networks in tabular forecasting on datasets with medium size and feature mix quality (Grinsztajn et al., 2022; Liu et al., 2022). The performance advantage is usually explained by boosting's ability to fine-tune and RF's stability to overfitting and noisy or imbalanced distributions.

**Table 4:** Hybrid Model Performance with Different Learner-Based Embeddings

Model Variant	$R^2$ (mean $\pm$ SD)	MAE (ln GDP)	RMSE (ln GDP)
Hybrid (z from FFNN + covariates)	<b>0.62 <math>\pm</math> 0.01</b>	<b>0.105</b>	<b>0.141</b>
Hybrid (z from Random Forest + covariates)	0.60 $\pm$ 0.02	0.108	0.145
Hybrid (z from XGBoost + covariates)	0.61 $\pm$ 0.02	0.107	0.144

These results validate that the causal inference and prediction benefits of the hybrid approach are robust to different embedding learner choices, with performance varying within  $\pm 0.02$  in  $R^2$  and error metrics. The elasticity estimate  $\beta_1$  is also invariant within  $\pm 0.002$  across all configurations, in support of the robustness of the hybrid method regardless of the specific machine learning method used for nuisance extraction.

## CONCLUSION

This paper has proposed and assessed a new hybrid modeling approach that combines deep learning representation learning with classical econometric regression via Double Machine Learning (DML) extended to panel data via Correlated Random Effects (CRE). Specifically, the hybrid framework combines a latent embedding vector ( $z$ ) derived from a Feedforward Neural Network (FFNN) with country-level covariates (i.e., DALYs, human capital, capital per worker, and population growth rate) in order to predict labor productivity, ln (LP).

Corresponding author: Akinrotimi Akinyemi Omololu

[akinrotimiakinyemi@ieee.org](mailto:akinrotimiakinyemi@ieee.org)

Department of Information Systems and Technology, Kings University, Ode-Omu, Osun State, Nigeria.

© 2025. Faculty of Technology Education. ATBU Bauchi. All rights reserved





The estimate of the DALY elasticity parameter ( $\beta_1$ ) in this hybrid model not only remains statistically significant but also is more precise compared to the pure econometric model, which suggests effective control for nonlinear confounding via CRE-DML (Clarke & Polselli, 2023; Fuhr & Papies, 2024).

Better prediction performance in our experiment findings was achieved by the hybrid model ( $R^2 \approx 0.62$ ,  $MAE = 0.105$ ,  $RMSE = 0.141$ ), which outperformed the pure econometric and FFNN-based prediction models. Robustness tests, including different cross-validation schemes (e.g. leave-unit-out), alternative regularization schemes, and substituting embedding learners with Random Forest and XGBoost showed that predictive accuracy differed by  $\pm 0.02$  in  $R^2$  and the estimated elasticity ( $\beta_1$ ) was robust within  $\pm 0.002$ , which confirmed that inference of interest is stable across setups.

These findings point to two important findings: (a) Performance Benefit: The hybrid method beats other variants of modeling in both predictive accuracy and parameter estimation precision consistently. (b) Stability: The hybrid method retains its advantage irrespective of the learner for embedding and estimation configuration, which affirms the value of combining representation learning with econometric estimation even when faced with extreme data sparsity (Fuhr & Papies, 2024; Clarke & Polselli, 2023). By comparison, choice of embedding learner, such as FFNN, Random Forest, or XGBoost, has moderate effects on RMSE and  $R^2$ , affirming literature findings that tree learners can equal or surpass neural networks in highly structured tabular forecasting problems due to their inductive abilities and resilience to noisy features (Grinsztajn et al., 2022; McElfresh et al., 2023).

Further, the following can be concluded from the research: (a) The model serves as a blueprint to combine interpretability (in the form of econometric elasticity) with representation-learning flexibility to achieve reliable, policy-driven models in applications to health and macroeconomy. (b) By indicating successful inference based on only minimal demographic and

health measures scaled down for low-resource or data-scarce settings, the research contributes to equitable data science that may be implemented in low- and middle-income country contexts. (c) In addition, the current study illustrates how CRE-DML deals with panel heterogeneity and nonlinear confounding through machine learning and introduces a scalable and replicable framework for future empirical research in similar domains.

Overall, this paper demonstrates that self-supervised learning of representations, if sufficiently integrated into existing econometric models using CRE-adjusted DML, has the promise to significantly improve predictive performance and accuracy of causal elasticities even with highly restricted data. The hybrid model combines methodological richness and applied relevance, which leaves the door open to systematic, equity-sensitive health economic forecasting models.

## LIMITATIONS

(a) The elasticity estimates are obtained from observation panel data and rely on no unobserved time-varying confounders, whose accuracy may not be certain. (b) Annual panel and sparse feature set may restrict model generalizability to denser or higher-frequency datasets. (d) Even though the hybrid model minimizes predictive error, its accuracy is contingent upon the quality and exhaustiveness of DALY and socioeconomic information, which can vary across nations.

## FUTURE STUDIES

Future research can apply this hybrid econometric-deep learning method to other health burden metrics than DALYs, e.g., quality-adjusted life years (QALYs) or cause-specific morbidity measures, and check for generalizability. Applying the model to higher-frequency or micro-data can reveal short-term productivity effects as well as sectoral heterogeneity. Adding more detailed covariates, e.g., environmental, behavior, and policy controls, can enhance causal precision. Secondly, cross-country adaptation and transfer learning can be investigated to improve performance in data-



scarce settings while making global health-economic analysis comparable.

## REFERENCES

- Ahrens, A., Chernozhukov, V., Hansen, C., Kozbur, D., Schaffer, M., & Wiemann, T. (2025). *An introduction to double/debiased machine learning*. arXiv. <https://arxiv.org/abs/2504.08324>
- Bach, P., Chernozhukov, V., Kurz, M. S., Spindler, M., & Klaassen, S. (2021). *DoubleML – An object-oriented implementation of double machine learning in Python*. arXiv. <https://arxiv.org/abs/2104.03220>
- Bohlke-Schneider, C., et al. (2022). *Resilient neural forecasting systems for labor planning*. arXiv. <https://arxiv.org/abs/2203.08492>
- Ouci Chen, F., Wang, J., Li, H., & Zheng, L. (2024). Modeling productivity loss under heat stress using machine and deep learning. *BMC Public Health*, 24, Article 16498. <https://doi.org/10.1186/s12889-023-16498-7> PMC
- Chernozhukov, V., Chetverikov, D., Demirer, M., Duflo, E., Hansen, C., Newey, W., & Robins, J. (2018). Double/debiased machine learning for treatment and structural parameters. *The Econometrics Journal*, 21(1), C1–C68.
- Clarke, P. S., & Polselli, A. (2023). *Double Machine Learning for Static Panel Models with Fixed Effects*. arXiv. <https://arxiv.org/abs/2312.08174>
- Clarke, P. S., & Polselli, A. (2023). Double machine learning for static panel models with fixed effects. arXiv. <https://arxiv.org/abs/2312.08174>
- Cybenko, G. (1989). Approximation by superpositions of a sigmoidal function. *Mathematics of Control, Signals and Systems*, 2(4), 303–314.
- Daroudi, S., et al. (2021). Cost per DALY averted across income strata: GBD-based estimates for cost-effectiveness thresholds. *Cost Effectiveness and Resource Allocation*, 19, Article 38. <https://doi.org/10.1186/s12962-021-00260-0>
- Desai, A. (2023). *Machine learning for economics research: When, what, and how?* arXiv.
- Fuhr, J., & Papies, D. (2024). Double machine learning meets panel data: Promises, pitfalls, and potential solutions. arXiv. <https://arxiv.org/abs/2409.01266>
- Global Burden of Disease 2021 Collaborators. (2021). Global incidence, prevalence, DALYs, and HALE for 371 diseases in 204 countries (GBD 2021). *The Lancet*. [https://doi.org/10.1016/S0140-6736\(24\)00757-8](https://doi.org/10.1016/S0140-6736(24)00757-8)
- Grinsztajn, L., Oyallon, É., & Varoquaux, G. (2022). Why do tree-based models still outperform deep learning on tabular data? arXiv. <https://arxiv.org/abs/2207.08815>
- Italian COVID-19 DALY burden group. (2021). Productivity loss estimation due to COVID-19 DALYs in Italy. *International Journal of Environmental Research and Public Health*, 17, Article 4233. <https://doi.org/10.3390/ijerph17124233>
- Kirigia, J. M., & Kubai, P. K. (2023). Monetary value of disability-adjusted life years and potential productivity losses associated with neglected tropical diseases in the East African Community. *Research and Reports in Tropical Medicine*, 14, 35–47. <https://doi.org/10.2147/RRTM.S382288>
- infondt.org+6PubMed+6SciProfiles+6
- Li, Y., Zhang, X., & Chen, Z. (2022). Comparison of ARIMA, deep neural networks, and LSTM models for forecasting DALYs in Tianjin, China. *Frontiers in Public Health*. <https://doi.org/10.3389/fpubh.2022.848370>
- Li, Y., Zhang, X., & Chen, Z. (2022). Comparison of ARIMA, DNN, and LSTM models for predicting pneumoconiosis DALYs in Tianjin, China. *BMC Public Health*.
- Liu, A. J., Mukherjee, A., Hu, L., Chen, J., & Nair, V. N. (2022). Performance and interpretability comparisons of

Corresponding author: Akinrotimi Akinyemi Omololu

✉ [akinrotimiakinyemi@ieee.org](mailto:akinrotimiakinyemi@ieee.org)

Department of Information Systems and Technology, Kings University, Ode-Omu, Osun State, Nigeria.

© 2025. Faculty of Technology Education. ATBU Bauchi. All rights reserved



- supervised machine learning algorithms: An empirical study. *arXiv*. <https://arxiv.org/abs/2204.12868>
- Masini, A., et al. (2020). Machine learning advances for time-series forecasting: Linear, nonlinear, and hybrid methods. *arXiv*. <https://arxiv.org/abs/2012.12802>
- McElfresh, D. C., et al. (2023). When do neural nets outperform boosted trees on tabular data? *arXiv*. <https://arxiv.org/abs/2305.02997>
- Multilayer perceptron. (2025). In Wikipedia. Retrieved from [https://en.wikipedia.org/wiki/Multilayer\\_perceptron](https://en.wikipedia.org/wiki/Multilayer_perceptron)
- Murray, C. J. L., & Lopez, A. D. (2019). Disability-adjusted life years (DALYs) and their use in global burden of disease studies. *Lancet*.
- Nkosana, T. K., & Ferrer-Worient, M. T. (2022). Health burdens and labor productivity in Africa's middle- and low-income economies: Implications for the COVID-19 era. *Journal of the Knowledge Economy*.
- Our World in Data. (2025). **Burden of Disease** [Dataset].
- Patel, R., Desai, P., & Singh, A. (2020). Chronic disease burden and workforce participation in India: a micro-level economic study. *Journal of Development Economics*.
- Smith, H. K., & Abbas, H. (2024). Hybrid models combining econometric, machine learning and deep learning approaches in economic forecasting. *arXiv*. <https://arxiv.org/abs/2505.19617>
- Stempień, D., & Ślepaczuk, R. (2025). Hybrid models combining econometric, machine learning and deep learning approaches in economic forecasting. *arXiv preprint arXiv:2505.19617*. [arxiv.orgarxiv.org](https://arxiv.org/abs/2505.19617)
- Tariq, M. U., Ismail, S. B., Babar, M., & Ahmad, A. (2024). Deep learning for prediction of population health costs: COVID-19 forecasting accuracy across models. *Osong Public Health and Research Perspectives*, 15(2), 115–136.
- Yang, B., Long, W., & Cai, Z. (2024). Machine learning based panel data models: Debiased machine learning for dynamic panel estimation. *Working Paper Series in Theoretical and Applied Economics* 202402. <https://kuwpaper.ku.edu/2024Papers/202402.pdf>.