MethodsX 7 (2020) 101160

Contents lists available at ScienceDirect

MethodsX

journal homepage: www.elsevier.com/locate/mex

Method Article

How to apply the novel dynamic ARDL simulations (dynardl) and Kernel-based regularized least squares (krls)

Samuel Asumadu Sarkodie*, Phebe Asantewaa Owusu

Nord University Business School (HHN), Post Box 1490, 8049 Bodø, Norway

ABSTRACT

The application of dynamic Autoregressive Distributed Lag (dynardl) simulations and Kernel-based Regularized Least Squares (krls) to time series data is gradually gaining recognition in energy, environmental and health economics. The Kernel-based Regularized Least Squares technique is a simplified machine learning-based algorithm with strength in its interpretation and accounting for heterogeneity, additivity and nonlinear effects. The novel dynamic ARDL Simulations algorithm is useful for testing cointegration, long and short-run equilibrium relationships in both levels and differences. Advantageously, the novel dynamic ARDL Simulations has visualization interface to examine the possible counterfactual change in the desired variable based on the notion of *ceteris paribus*. Thus, the novel dynamic ARDL Simulations and Kernel-based Regularized Least Squares techniques are useful and improved time series techniques for policy formulation.

- We customize ARDL and dynamic simulated ARDL by adding plot estimates with confidence intervals.
- A step-by-step procedure of applying ARDL, dynamic ARDL Simulations and Kernel-based Regularized Least Squares is provided.
- All techniques are applied to examine the economic effect of denuclearization in Switzerland by 2034.

ARTICLE INFO

Method name: Novel dynamic ARDL Simulations and Kernel-based Regularized Least Squares

Keywords: Dynamic autoregressive distributed lag simulations, Kernel-based regularized least squares, Response surface regressions, Average marginal effects, Pointwise derivatives, time series techniques, Counterfactual change, Impulse-Response, Dynardl, Krls

Article history: Received 1 October 2020; Accepted 20 November 2020; Available online 27 November 2020

* Corresponding author.
E-mail address: asumadusarkodiesamuel@yahoo.com (S.A. Sarkodie).

https://doi.org/10.1016/j.mex.2020.101160

Specifications table

Subject Area:	Environmental Science
More specific subject area:	*Energy, Environmental and Health econometrics*
Method name:	*Novel dynamic ARDL Simulations and Kernel-based Regularized Least Squares*
Name and reference of original method:	S. Kripfganz, D.C. Schneider, *Response Surface Regressions for Critical Value Bounds and Approximate p-values in Equilibrium Correction Models1*, Oxf. Bull. Econ. Stat. 82 (2021) 1456–1481, Jordan, Soren, and Andrew Q. Philips. *"Cointegration testing and dynamic simulations of autoregressive distributed lag models."* The Stata Journal 18.4 (2018): 902–923.
Resource availability:	*Dataset attached as supplementary material*

Introduction

Though nuclear power is a clean source of energy yet, has several long-term environmental (management of radioactive waste) and health costs [1,2]. The short-range characteristic of emitted particles from nuclear reactors and electromagnetic interactions of atoms in solid matter has serious health consequences in living organisms [3]. Following the nuclear accidents that occurred in Ukraine (Chernobyl) and Japan (Fukushima Daiichi), several countries including Switzerland are phasing out nuclear power plants [1,3]. In this regard, we assess the possible economic effect of phasing out nuclear power plants in Switzerland for 20 years using novel estimation techniques. We employ four data series from 1970 to 2018 namely GDP, gross fixed capital formation, exportation of goods, and services (obtained from the World Bank[1]), labor[2], and consumption of nuclear energy[3].

Method details

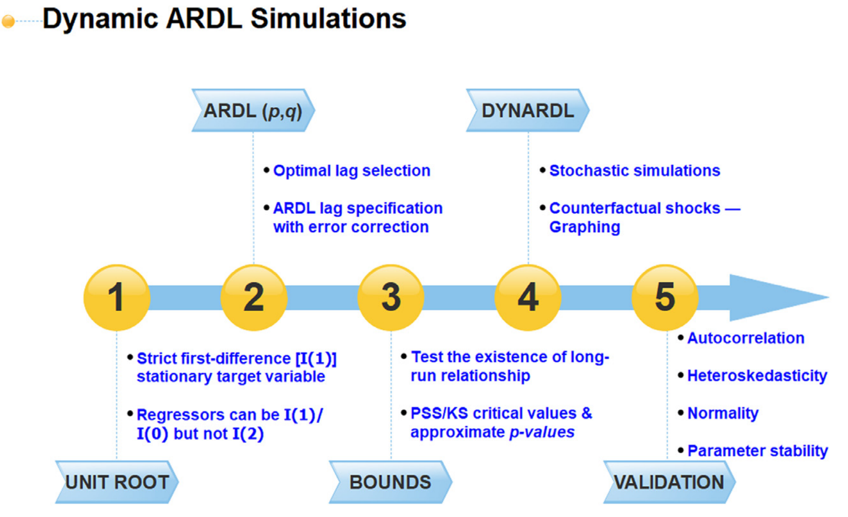
The application of the novel dynamic ARDL Simulations follows simple but technical guidelines presented in this method (Scheme 1). The ARDL bounds testing procedure used in the novel dynamic ARDL simulations requires a strict first-difference stationary, $I(1)$ dependent variable [4]. This implies that the only possible entrant for cointegration is a dependent variable that is non-stationary at level, $I(0)$. In contrast, bounds testing procedure with a dependent variable violating the initial conditions can be tested using the standard but modified ARDL bounds test with surface regression [5]. To test this conditional requirement, several unit root tests can be employed such as augmented Dickey-Fuller (ADF), Phillips-Perron (PP), Kwiatkowski-Phillips-Schmidt-Shin (KPSS), Dickey-Fuller Generalized Least Squares (DF-GLS), among others. Second, all sampled independent variables can either be $I(0)$ or integrated of order one, $I(1)$ but not greater than $I(1)$ devoid of a structural break, autocorrelation, and heteroskedasticity. We generate the variables in natural logarithms to control for potential heteroskedasticity [6]. After importing the data into STATA, we declare the dataset as time series using: *tsset Years, yearly*

Step 1: unit root test

To control for potential spurious regression, we examine the stationarity properties of the variables using PP and ADF tests. To do this, we run PP and ADF unit root tests in both level and first difference as: *pperron lnGDP*; *pperron d.lnGDP*; *dfuller lnGDP*; *dfuller d.lnGDP*

Options such as nocons, trend, lags(#) can be included. The results of PP and ADF tests are reported in Table 1. While we fail to reject (except for lnNUKE) the null hypothesis of unit root at level in Table 1, we strongly reject the null hypothesis at first-difference based on *p-value<0.01*.

[1] https://buff.ly/2ShVBtP.
[2] https://buff.ly/2GoaOa6.
[3] https://buff.ly/3ni9zuf.



Scheme 1. Salient steps in applying the dynamic ARDL simulations.

Table 1
Unit root tests.

Variable	Level.PP	Δ.PP	Level.ADF	Δ.ADF
lnGDP	0.350	−4.972***	0.407	−5.097***
lnNUKE	−6.367***	−5.876***	−4.613***	−5.689***
lnGFCF	−0.366	−4.267***	−0.159	−4.289***
lnLABOR	0.363	−3.833***	0.785	−3.833***
lnEXPORTS	−0.753	−8.831***	−0.697	−8.044***

Notes: Where Level.PP and Δ.PP denote the level and first-difference of Phillips-Perron unit root test; Level.ADF and Δ.ADF denote the level and first-difference of augmented-Dickey Fuller unit root test; ***denotes rejection of the null hypothesis of no unit root at 1% significance level.

Step 2: ARDL estimation

After meeting the condition of strict first-difference stationary dependent variable (lnGDP), we determine the optimal lag for the proposed model using *varsoc lnGDP lnNUKE lnGFCF lnLABOR lnEXPORTS, maxlag(2)*. Using the optimal lag selected, we test for cointegration using Pesaran, Shin, and Smith (PSS) bounds test with novel Kripfganz & Schneider (KS) critical values and approximate *p*-values. Before running the customized ARDL model, the following packages [*parmest, eclplot, dynardl, krls*] must be installed using:

ssc install parmest; *ssc install eclplot*; *ssc install dynardl*; *ssc install krls*

We modify the original model specification of the ARDL to express the estimated parameters in a plot expressed as:

parmby "xi:ardl lnGDP lnNUKE lnGFCF lnLABOR lnEXPORT, maxlag(2 2 2 2 2) nocons ec1 regstore(res)", label norestore
sencode parm, gene(parmid)
eclplot estimate min95 max95 parmid

Where *nocons* suppresses the constant term, *ec1* estimates the long-run parameter in time, t-1; *regstore* saves the estimated regression for validation through diagnostic tests. The resulting parameters based on ARDL(1,2,2,0,0) are presented in Fig. 1 with empirics repeated in Table 2 for clarity.

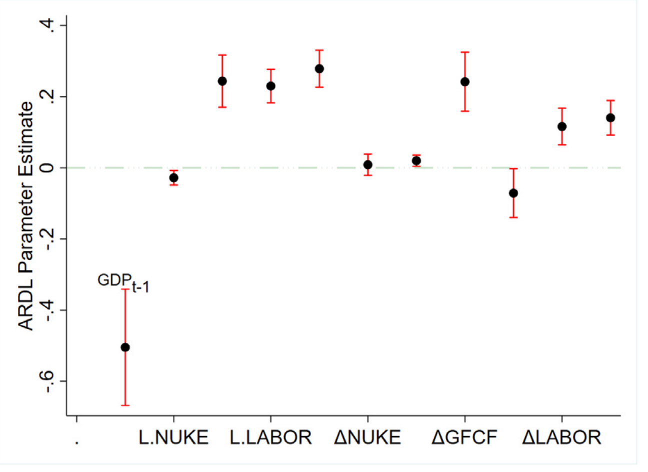


Fig. 1. Parameter estimates of the ARDL model. *Notes:* black (●) is the estimate in a log-log model, olive teal long-dash 3-dots is the reference line, red-spike denotes lower 95% and upper 95% confidence limit. **Legend:** GFCF represents Gross Fixed Capital Formation, LABOR represents labor, EXPORTS denotes exportation of goods and services from Switzerland, and NUKE means consumption of nuclear energy. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 2
ARDL estimation model.

EQN	Parm	Estimate	SE	*P*-value	Min 95	Max 95
ECT	$\ln GDP_{t-1}$	−0.505	0.081	0.000***	−0.668	−0.341
Long-Run	$\ln NUKE_{t-1}$	−0.028	0.010	0.008***	−0.048	−0.008
	$\ln GFCF_{t-1}$	0.244	0.036	0.000***	0.171	0.317
	$\ln LABOR_{t-1}$	0.230	0.023	0.000***	0.183	0.277
	$\ln EXPORTS_{t-1}$	0.279	0.026	0.000***	0.227	0.331
Short-Run	$\Delta\ln NUKE$	0.009	0.015	0.564	−0.021	0.039
	$\Delta\ln NUKE_{t-1}$	0.020	0.008	0.014**	0.004	0.036
	$\Delta\ln GFCF$	0.242	0.041	0.000***	0.159	0.325
	$\Delta\ln GFCF_{t-1}$	−0.071	0.034	0.042**	−0.139	−0.003
	$\Delta\ln LABOR$	0.116	0.025	0.000***	0.065	0.168
	$\Delta\ln EXPORTS$	0.141	0.024	0.000***	0.092	0.189
ARDL(1,2,2,0,0)	Obs	47	R^2	0.916	Root MSE	0.008

Notes: Where SE is the standard error; ***, ** denote statistical significance at 1, 5% level. **Legend:** GFCF represents Gross Fixed Capital Formation, LABOR represents labor, EXPORTS denotes exportation of goods and services from Switzerland, and NUKE means consumption of nuclear energy.

After testing the unit root properties of sampled variables, we proceed to examine cointegration using the modified PSS bounds test with KS critical values and approximate *p*-values. Based on ARDL(1,2,2,0,0), we run the long-run relationship using: *estat ectest*

The subsequent results of the bounds test are reported in Table 3. The estimated F-statistic based on a finite sample of 4 variables, 47 observations, 4 short-run coefficients is 18.563 whereas t-statistic is −6.245 — which is above the upper bound critical values (3.832, −3.625) at 5% significance level and above the critical values of all I(1) variables in 10 and 1% level. This is further validated by Kripfganz & Schneider approximate *p*-values [*p-value*<*0.01*], hence, rejecting the null hypothesis of no level relationship. Thus, both PSS bounds test and Kripfganz-Schneider critical values with approximate *p*-values confirm the presence of cointegration.

Table 3
Pesaran, Shin, and Smith bounds testing.

		10%		5%		1%		*p-value*	
	K	I(0)	I(1)	I(0)	I(1)	I(0)	I(1)	I(0)	I(1)
F	18.563	2.021	3.227	2.476	3.832	3.547	5.225	0.000***	0.000***
t	−6.245	−1.608	−3.231	−1.962	−3.625	−2.660	−4.398	0.000***	0.000***

Notes: Where *I(0)* and *I(1)* denote the lower and upper band critical values at 10%, 5% and 1% significance level of Pesaran, Shin, and Smith bounds test; *P*-value is Kripfganz & Schneider critical values and approximate *p-values*; ***denotes rejection of the null hypothesis of no level relationship at 1% significance level.

Table 4
Breusch-Godfrey LM test for autocorrelation.

lags(p)	F	df	Prob > F
1	0.068	1, 37	0.796
2	0.275	2, 36	0.761
3	0.611	3, 35	0.612
4	0.567	4, 34	0.689

Table 5
Cameron & Trivedi's decomposition of IM-test.

Source	chi^2	df	*p*-value
Heteroskedasticity	47.00	46	0.4313
Skewness	15.78	9	0.0717
Kurtosis	0.63	1	0.4274
Total	63.41	56	0.2316

Table 6
Skewness/Kurtosis tests for normality.

Variable	Obs	Pr(Skewness)	Pr(Kurtosis)	joint adj $chi^2(2)$	Prob>chi^2
Residuals	47	0.2155	0.7297	1.74	0.4187

Step 3: diagnostics of ARDL estimation

As part of the initial conditions of the dynamic ARDL simulations, we perform several tests to get rid of serial correlation, heteroskedasticity, violation of normality, and structural breaks. First, we restore the saved estimated regression using: *estimates restore res*

Second, we examine the residuals of the estimated model for autocorrelation using Breusch-Godfrey LM test by running: *estat bgodfrey, lags(1/4) small*

The resulting estimates of Breusch-Godfrey LM test with four lags are presented in Table 4. We fail to reject the null hypothesis of no serial correlation based on 5% significance level — confirming the residuals of the estimated ARDL(1,2,2,0,0) model are free from autocorrelation.

Third, we test for heteroskedasticity in the residuals using Cameron & Trivedi's decomposition of IM-test by running: *estat imtest, white*

It can be observed from Table 5 that the null hypothesis of homoskedasticity cannot be rejected at 5% significance level — confirming the residuals are homoskedastic.

Next, we assess the independence of the residuals by testing for normality using Skewness/Kurtosis tests by running: *predict res1, residuals; sktest res1*

The results in Table 6 reveal that the null hypothesis of normal distribution cannot be rejected at 5% significance level.

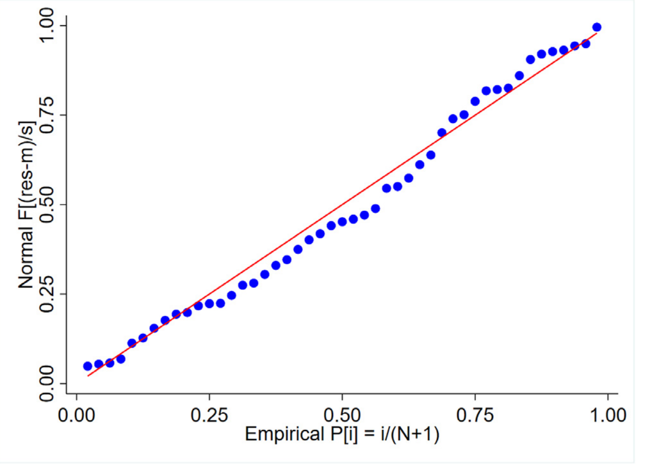


Fig. 2. Standardized normal probability plot.

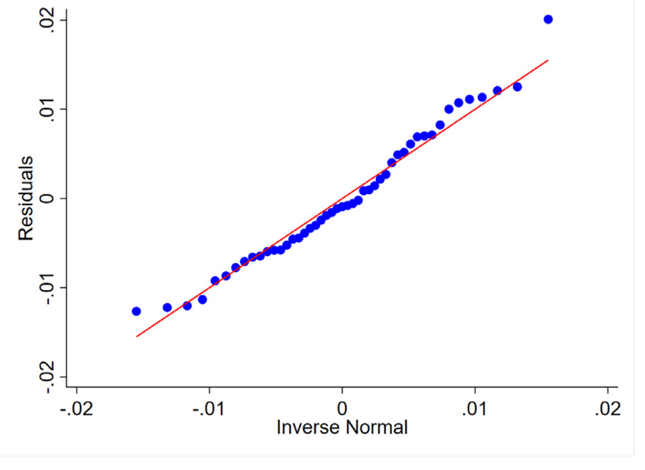


Fig. 3. Quantiles of residuals against quantiles of normal distribution.

We further validate the distribution using both standardized normal probability plot (Fig. 2) and quantiles of residuals against quantiles of normal distribution estimates (Fig. 3) by running: *pnorm res1*; *qnorm res1*

The resulting plots (Figs. 2 and 3) confirm the residuals based on the estimated ARDL(1,2,2,0,0) are normally distributed.

Finally, we investigate potential structural breaks using cumulative sum test for parameter stability by running: *estat sbcusum, ols*

Evidence from Fig. 4 reveals that the estimated test statistic is within the 95% confidence band, hence, confirming that stability of the estimated coefficients over time.

Step 4: applying dynamic ARDL simulations

The novel dynamic ARDL simulations technique has been utilized in several studies to capture future shocks in socioeconomic and climatic indicators [7,8]. In contrast, we present policy-based

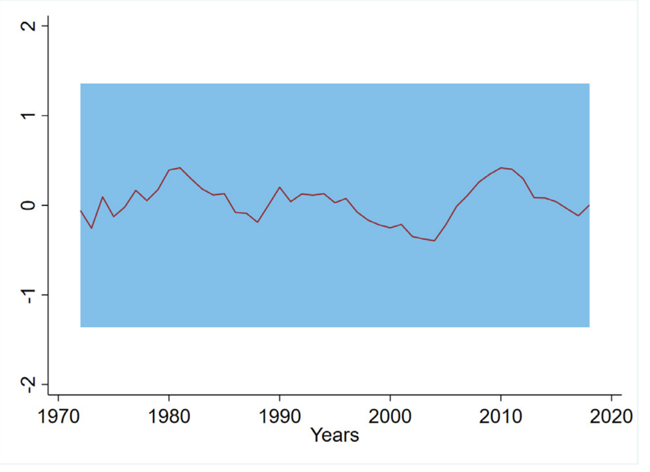


Fig. 4. Cumulative sum test using OLS CUSUM plot for parameter stability.

specific inputs to account for potential shocks due to the recent phasing out of nuclear plants in Switzerland [1]. The dynamic ARDL simulation is based on ~21% (2018 estimate from BP [9]) contribution of nuclear to the energy mix used as counterfactual shock over 20 years from 2018 to 2038. The model specification of the proposed dynamic ARDL simulations can be expressed as [4,10]:

$$\begin{aligned} ln\ (GDP)_t = {} & \beta_0 ln\ (GDP)_{t-1} + \beta_1 ln\ (NUKE)_t + \beta_2 ln\ (NUKE)_{t-1} + \beta_3 ln\ (GFCF)_t + \beta_4 ln\ (GFCF)_{t-2} \\ & + \beta_5 ln\ (LABOR)_t + \beta_6 ln\ (LABOR)_{t-} + \beta_7 ln\ (EXPORTS)_t + \beta_8 ln\ (EXPORTS)_{t-1} + \epsilon_t \end{aligned} \tag{1}$$

Where GDP denotes economic growth, *GFCF* is Gross Fixed Capital Formation, *LABOR* represents Labor, *EXPORTS* means exports of goods and services, and *NUKE* denotes nuclear energy consumption. ϵ is the error term in time t.

Thus, the dynamic ARDL simulations technique is applied by running:

parmby "xi:dynardl lnGDP lnNUKE lnGFCF lnLABOR lnEXPORT, lags(1, 1, 2, 1, 1) diffs(., ., 1, 1, 1) shockvar(lnNUKE) nocons ec shockval(−21) time(10) range(30) graph change sims(5000)", label norestore

Afterward, we run:

sencode parm, gene(parmid)
eclplot estimate min95 max95 parmid

Here, *shockvar* is the variable to examine potential shocks whereas *shockval* is the amount of shock to be applied to the target variable. It is noteworthy that the length of scenario (*range*) should always be greater than the scenario *time*. The parameter plot of the dynamic simulated ARDL is depicted in Fig. 5 whereas the expounded empirics are presented in Table 7. Like the ARDL estimates, long-term nuclear energy consumption has depreciating effects on economic development. This may perhaps be linked to environmental and health costs of radioactive waste management, decommissioning, and health hazards in Switzerland [1]. In contrast, increasing level of labor, gross fixed capital formation, exportation of goods and services have economic expansion effect in both short and long -run (i.e. in both ARDL and dynARDL).

To account for the effect of decreasing marginal returns of nuclear energy on sustained economic growth, we assess the counterfactual shocks via the dynamic ARDL simulations by incorporating the

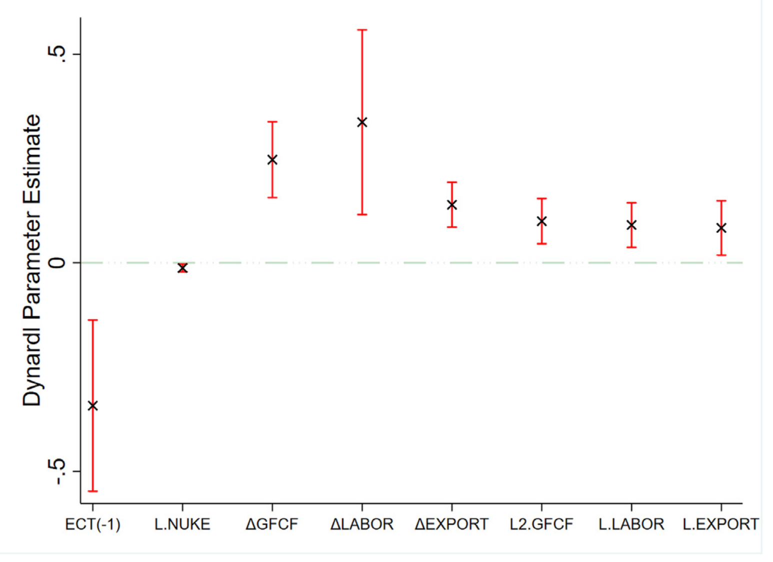


Fig. 5. Parameter estimates of dynamic ARDL Simulations. *Notes:* black (×) is the estimate in a log-log model, olive teal long-dash 3-dots is the reference line, red-spike denotes lower 95% and upper 95% confidence limit. **Legend:** GFCF represents Gross Fixed Capital Formation, LABOR represents labor, EXPORTS denotes exportation of goods and services from Switzerland, and NUKE means consumption of nuclear energy. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 7
Estimates of dynamic simulated ARDL model.

Parm	Estimate	SE	*P*-value	Min 95	Max 95
$\ln GDP_{t-1}$	−0.343	0.101	0.002***	−0.548	−0.137
$\ln NUKE_{t-1}$	−0.012	0.005	0.010***	−0.022	−0.003
$\Delta \ln GFCF$	0.247	0.045	0.000***	0.157	0.338
$\Delta \ln LABOR$	0.337	0.110	0.004***	0.116	0.559
$\Delta \ln EXPORT$	0.139	0.027	0.000***	0.085	0.193
$\ln GFCF_{t-2}$	0.100	0.027	0.001***	0.046	0.154
$\ln LABOR_{t-1}$	0.091	0.026	0.001***	0.037	0.144
$\ln EXPORT_{t-1}$	0.084	0.032	0.013**	0.018	0.149
Prob > *F*	0.000***	R^2	0.906	Root MSE	0.009

Notes: Where SE is the standard error; ***, ** denote statistical significance at 1, 5% level. **Legend:** GFCF represents Gross Fixed Capital Formation, LABOR represents labor, EXPORTS denotes exportation of goods and services from Switzerland, and NUKE means consumption of nuclear energy.

share of nuclear energy in the energy portfolio (~21% [9]), and period estimated for denuclearization (2018–2038). The plot showing dynamic ARDL simulations reveals that −21% shock in predicted nuclear energy consumption may affect economic growth in the first period but growth accelerates thereafter (Fig. 6). Thus, denuclearizing the economy will have no lasting impact on sustained economic growth.

Step 5: applying Kernel-based regularized least squares

We subsequently apply Kernel-based Regularized Least Squares (KRLS), a machine learning algorithm that implements the pointwise derivatives to examine the causal-effect relationship. The

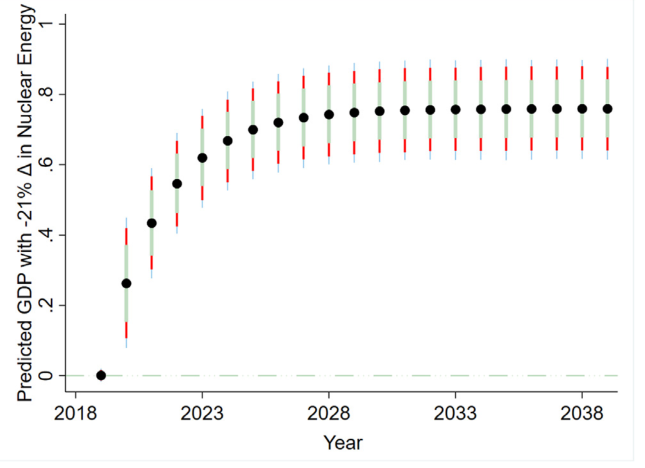


Fig. 6. Representation of counterfactual shock in predicted nuclear energy using dynamic ARDL simulations. *Notes:* black dot (•) is the predicted GDP by −21% shock in nuclear energy in a log-log model; olive teal, red and light-blue spikes denote 75, 90, and 95% confidence interval. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 8
Pointwise derivatives using KRLS.

lnGDP	Avg.	SE	t	P>t	P-25	P-50	P-75
lnNUKE	0.023	0.009	2.463	0.018	0.001	0.036	0.057
lnGFCF	0.223	0.020	10.876	0.000	0.163	0.249	0.289
lnLABOR	0.421	0.035	12.084	0.000	0.238	0.472	0.607
lnEXPORTS	0.093	0.011	8.848	0.000	0.036	0.098	0.156
Diagnostics							
Lambda	0.091	Sigma	4.000	R^2	0.998	obs	40.000
Tolerance	0.049	Eff. Df	11.220	Looloss	0.059	***F***-test	5.886

Notes: Where Avg. is the average marginal effect; SE is the standard error; P-25, P-50, and P-75 represent 25th, 50th· and 75th percentile. **Legend:** GFCF represents Gross Fixed Capital Formation, LABOR represents labor, EXPORTS denotes exportation of goods and services from Switzerland, and NUKE means consumption of nuclear energy.

mathematical elaborations of the technique can be found in Hainmueller and Hazlett [11]. To account for the 2034 plan to denuclearize the economy, we examine the structural adjustments in economic growth using empirical estimation via pointwise marginal effect. We re-run the economic function with KRLS as: *krls lnGDP lnNUKE lnGFCF lnLABOR lnEXPORT, graph*

The pointwise derivatives of the estimated KRLS model are presented in Table 8. The model is statistically significant at 1% level, with a predictive power of 0.998. Meaning that the regressors explain 99.8% variation in economic development. An assessment of heterogeneous marginal effects using derivatives of regressors is reported as 25th, 50th· and 75th percentiles in Table 8. We observe no evidence of heterogeneous marginal effects across sampled variables, thus, confirming the robustness of the pointwise derivatives. It can be observed that the mean pointwise marginal effect of nuclear energy consumption, gross fixed capital formation, labor, and exports of goods and services are 0.02%, 0.22%, 0.42%, and 0.09%, respectively. This underscores the importance of nuclear energy, gross fixed capital formation, labor, and exports of goods and services in sustaining economic development in Switzerland. The question still persists on how phasing out of nuclear energy will affect future economic development. Going further, we examine the long-term variation in nuclear

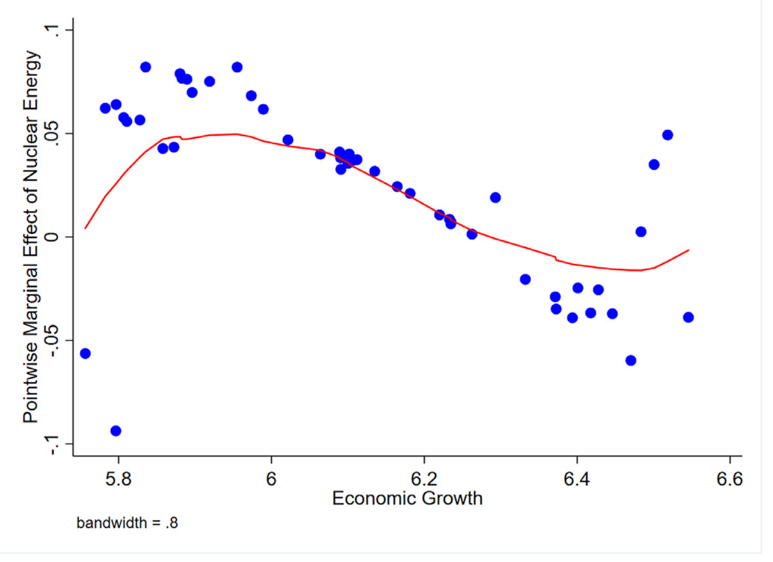


Fig. 7. Representation of Pointwise marginal effect of nuclear energy.

energy consumption and how it affects economic growth and vice versa. To do this, we plot the pointwise derivative of nuclear energy consumption against GDP to capture varying marginal effects. We run *lowess deriv_ lnNUKE lnGDP*

It can be observed in Fig. 7 that higher levels of nuclear energy consumption increase economic growth at lower levels to a threshold where increasing marginal returns occur, however, declines nuclear energy consumption thereafter with increasing economic growth. Thus, nuclear energy consumption has decreasing marginal returns with increasing economic growth. This infers potential energy technological obsolescence with increasing growth.

Conclusion

Decoupling nuclear energy consumption from economic growth has several structural implications but advantageous to reducing environmental risk and nuclear weapon proliferation. Here, we investigated the relationship between nuclear energy consumption and economic growth in Switzerland over the period 1970–2018. With Switzerland's energy policy of phasing out nuclear energy production by 2034, we examined the long-term economic structural impact by utilizing novel estimation techniques such as Kernel-based Regularized Least Squares (krls) and dynamic ARDL simulations (dynardl) to capture counterfactual shocks in denuclearizing the economy. We find that decoupling nuclear energy from the economy will affect economic growth in the first year but has a rebound effect afterward. Our customized ARDL and dynamic simulated ARDL are useful in producing plot estimates with confidence intervals — useful for policy modeling in environment, health, and energy economics.

Declaration of Competing Interest

The Authors confirm that there are no conflicts of interest.

Acknowledgment

Open Access funding provided by Nord University.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.mex.2020.101160.

References

[1] WNA. Nuclear power in Switzerland. 2020; Available from: https://buff.ly/2GwIUbR.
[2] J.P. Christodouleas, et al., Short-term and long-term health risks of nuclear-power-plant accidents, N. Engl. J. Med. 364 (24) (2011) 2334–2341.
[3] J.W. Tester, et al., Sustainable Energy: Choosing Among Options, MIT Press, 2012.
[4] S. Jordan, A.Q. Philips, Cointegration testing and dynamic simulations of autoregressive distributed lag models, Stata J. 18 (4) (2018) 902–923.
[5] S. Kripfganz, D.C. Schneider, Response Surface Regressions for Critical Value Bounds and Approximate p-values in Equilibrium Correction Models1, Oxf. Bull. Econ. Stat. 82 (2021) 1456–1481, doi:10.1111/obes.12377.
[6] S.A. Sarkodie, P.A. Owusu, How to apply dynamic panel bootstrap-corrected fixed-effects (xtbcfe) and heterogeneous dynamics (panelhetero), MethodsX 7 (2020) 101045.
[7] S.A. Sarkodie, et al., Energy–climate–economy–population nexus: an empirical analysis in Kenya, Senegal, and Eswatini, Sustainability 12 (15) (2020) 6202.
[8] A.H. Shabbir, et al., Predicting the influence of climate on grassland area burned in Xilingol, China with dynamic simulations of autoregressive distributed lag models, PLoS ONE 15 (4) (2020) e0229894.
[9] B.P. BP statistical review of world energy 2019. 2020 December 14th, 2019]; Available from: http://www.bp.com/statisticalreview.
[10] S.A. Sarkodie, et al., Environmental sustainability assessment using dynamic autoregressive-distributed lag simulations—nexus between greenhouse gas emissions, biomass energy, food and economic growth, Sci. Total Environ. 668 (2019) 318–332.
[11] J. Hainmueller, C. Hazlett, Kernel regularized least squares: reducing misspecification bias with a flexible and interpretable machine learning approach, Polit. Anal. 22 (2) (2014) 143–168.