African Scientific Journal
ISSN : 2658-9311
Vol : 03, Num 36 Juin 2026

Health Financing, Income, and Population Health in Morocco (2000–2022): Contrasting Evidence from Life Expectancy and Infant Mortality in an Exploratory ARDL Framework.

Auteur 1 : BAROUCH Noureddine,
Auteur 2 : ADRAOUI Mohamed,

BAROUCH Noureddine, ORCID: 0009-0009-4644-6730, PhD researcher
1 Laboratory of Economic Performance and Logistics (PEL), Faculty of Law, Economics and Social Sciences (FSJES), Mohammedia, Hassan II University of Casablanca, Morocco

ADRAOUI Mohamed, ORCID: 0009-0006-0093-9386, University Professor
2 Laboratory of Economic Performance and Logistics (PEL), Faculty of Law, Economics and Social Sciences (FSJES), Mohammedia, Hassan II University of Casablanca, Morocco

Déclaration de divulgation : L'auteur n'a pas connaissance de quelconque financement qui pourrait affecter l'objectivité de cette étude.

Conflit d'intérêts : L'auteur ne signale aucun conflit d'intérêts.

Pour citer cet article : BAROUCH .N & ADRAOUI .M (2026) « Health Financing, Income, and Population Health in Morocco (2000–2022): Contrasting Evidence from Life Expectancy and Infant Mortality in an Exploratory ARDL Framework », African Scientific Journal « Volume 03, Num 36 » pp: 1086 – 1117.

DOI : 10.5281/zenodo.20594340

Abstract

Background. Over roughly twenty years, a series of financing reforms has redrawn Morocco's social protection system. The path runs from the creation of the *Assurance Maladie Obligatoire* (AMO) through to the generalization of compulsory coverage under Framework Law 09-21 and, more recently, the launch of *AMO-Tadamoun*. The trouble is that direct household out-of-pocket payments still take up about 41% of current health expenditure. That figure is close to three times the level internationally associated with low financial hardship, which suggests that signing people up has not yet delivered real financial protection. Whether these reforms have produced measurable health gains is still an open question, and aggregate macroeconomic data offer at least a first way into it.

Methods. We work with annual national data for Morocco covering 2000 to 2022, built around two outcomes rather than one. An ARDL bounds testing framework is estimated separately for life expectancy at birth and for the infant mortality rate. We then re-check the long-run cointegrating vector with Fully Modified OLS and Dynamic OLS, so as to deal with the endogeneity that remains in the regressors.

Results. Most of what matters here comes from the infant mortality model. Three independent estimation methods point the same way: a heavier out-of-pocket burden is associated with worse child survival, higher income with lower infant mortality, and higher public health spending with lower mortality once its endogeneity is taken into account. The life expectancy model behaves quite differently. Because that series barely shifts from one year to the next, it cannot pin down these effects at all, and the gap between the two models tells us something useful in its own right about the limits of national aggregate data.

Conclusion. Reducing out-of-pocket payments from their current 41% of current health expenditure toward the 15-20% benchmark is a necessary condition for Morocco's universal health coverage reforms to translate into measurable child survival gains. As the Groupements Sanitaires Territoriaux are rolled out and AMO-Tadamoun extends legal entitlement to previously uncovered populations, these results expose a persistent tension. Our findings demonstrate that widening formal coverage may not, on its own, improve child health outcomes as long as the out-of-pocket burden remains structurally high. Consequently, achieving financial protection at the point of care is essential for the success of Morocco's universal health coverage reforms.

Keywords: health financing; life expectancy; infant mortality; out-of-pocket payments; universal health coverage; ARDL bounds testing; health outcome selection; Morocco.

Introduction

It is a common assumption that how a country pays for health care, and how much it pays, shapes how long and how well its people live. When emerging economies move through their epidemiological and demographic transitions, reforms aimed at universal health coverage tend to be cast as an engine of human capital, productivity, and long-run growth (World Health Organization, 2010). Reality is seldom so neat. Spending more does not automatically buy better health, since the way financing is set up, along with the rigidities of the systems that actually deliver care, frequently gets in the way between the budget and the patient.

Morocco is a useful place to look at these transmission channels. Starting in the early 2000s, the authorities set about reshaping the architecture of the national social protection system. Two instruments held this effort together: the Assurance Maladie Obligatoire (AMO), which covered salaried workers in the public and private sectors, and the Régime d'Assistance Médicale (RAMED), built around a logic of social solidarity for low-income households. The promulgation of Framework Law 09-21 in 2021 later accelerated the move toward generalized social protection, broadening compulsory health insurance to cover all population categories; Law 27-22 subsequently created *AMO-Tadamoun* as the non-contributory branch serving households unable to pay contributions. Taken together, these measures extended formal insurance to population groups that had previously stood outside any risk-pooling mechanism.

Extending legal entitlement has not made a persistent paradox in the country's financing profile go away. Direct household out-of-pocket (OOP) payments stay high, sitting above two-fifths of current health expenditure year after year (Haut-Commissariat au Plan, 2023; World Health Organization, 2024). That level has to be read against the macroeconomic benchmark of roughly 15 to 20% of current health expenditure, which the World Health Organization links to a low incidence of financial hardship. It should not be confused with the separate household-level threshold of 40% of capacity to pay that the microeconomic literature uses to flag catastrophic spending (Xu et al., 2003). Over the period the two financing series moved in opposite directions, with the out-of-pocket share falling as public spending rose, and yet the gap never closed. At the aggregate level this configuration pits the political promise of wider coverage against the financial reality households face, and it raises a direct question about the health returns these financial inputs actually produce.

The question ties in directly to the health production framework that Grossman (1972) introduced and that Preston (1975) carried over to national populations. For Grossman, health is a kind of capital that erodes over time but can be rebuilt through investment, whether in medical care, income, education, or living conditions. Estimating a relationship of this sort at the national level amounts to asking how much of a population’s health gains trace back to rising income and how much to the way care is financed and delivered. Life expectancy here stands in as the working measure of a country’s accumulated stock of health.

The international evidence on these relationships is anything but settled. Cross-country panels frequently turn up a positive link between public health spending and longevity (Novignon et al., 2012; Ray & Linden, 2020), but single-country and regional studies keep finding that inputs and outcomes come apart. The usual culprits are allocative inefficiency, supply-side bottlenecks, and weak institutions (Balkhi et al., 2021; Faruk et al., 2022). Rapid urbanization is just as contested, pulled between the accessibility gains that density can deliver and the strain it places on underserved peri-urban areas.

Rather than draw on a broad multi-country panel, this study adopts a single-country lens. It estimates two parallel ARDL models that differ in one respect only, the health outcome: life expectancy at birth in one case, the infant mortality rate in the other. Working with two outcomes is a deliberate choice, not an accident of convenience. Life expectancy is the standard aggregate indicator in this literature, yet it is also a slow, near-deterministic series that offers very little year-to-year variation once differenced, which makes identifying individual coefficients exceedingly hard in a short sample. Infant mortality reacts more quickly to the quality and accessibility of front-line care, carries close to three times as much annual variation, and sits at the centre of universal health coverage monitoring. Estimating both lets the study show not only what a short, collinear national series can and cannot deliver, but also how much identifying variation a more responsive outcome can recover. The aim is strictly exploratory. It is to establish whether statistically stable cointegrating relationships tie the health financing and longevity series together, and to set the short-run dynamics of the two specifications side by side, without claiming to identify causal channels or to rank the contributions of individual financing inputs.

Morocco's ongoing reform trajectory gives these questions a concrete institutional backdrop. Throughout the study period, out-of-pocket payments remained well above the World Health Organization's benchmark for low financial hardship - a gap the discussion returns to. The results are presented as associational evidence rather than as a policy evaluation; readers interested in the distributional and household-level dimensions of financial protection will find that national aggregate data can offer only a partial view.

This study addresses the complex transition of Morocco's health system and the paradox of high out-of-pocket payments despite recent legal reforms. Its primary objective is to evaluate, through an ARDL econometric framework covering the 2000–2022 period, whether improvements in health outcomes, specifically infant mortality and life expectancy, are empirically correlated with public and private health financing shifts. The remainder of this article is structured as follows: the first section reviews the relevant economic literature; the second section details the Moroccan institutional context; the third section describes the data and methodology; finally, the fourth section presents and discusses the empirical results, before concluding with policy implications for universal health coverage.

1. Literature review

The empirical study of how health systems convert resources into population health rests on a long tradition in economics. Grossman (1972) recast health as a durable stock of capital, one that depreciates with age but responds to investment in medical care, nutrition, education, and the broader environment. Preston (1975) brought the same logic to the national level, documenting that the relationship between mortality and economic development is positive but curved, with the largest returns concentrated among poorer countries. Together these contributions frame the question that this paper takes up for a single economy. How much of a country's longevity is bound up with its income, and how much with the resources it channels through its health system?

A substantial body of cross-country work has tried to quantify the second of these channels. Using a panel of forty-four sub-Saharan African countries, Novignon, Olakojo, and Nonvignon (2012) found that both public and private health spending raise life expectancy and lower mortality, with the public component carrying the larger weight. At the level of infant and neonatal mortality specifically, Kiross and colleagues (2020) confirm this asymmetry for a sub-Saharan African panel: public and external health expenditure are significantly associated with lower infant mortality, whereas private expenditure carries no significant effect. More recently, Adegoke, Mbonigaba, and George (2025), using forecasting models for selected African countries, project that a 30% increase in public health expenditure would allow approximately 60% of countries to reach the SDG child mortality target by 2030, but that the remaining gap would persist without structural improvements in health system efficiency. Ray and Linden (2020), working with close to two hundred countries, reached a similar ordering, since public expenditure proved more health-promoting than private spending, though the estimated effects were modest and sensitive to the choice of estimator. The same caution recurs in the OECD setting, where Aytemiz and colleagues (2024) report that income per capita and health expenditure both contribute positively to life expectancy over the long run, with the marginal contribution of spending the weaker of the two. A common message runs through this literature. Money matters, yet it matters less, and less reliably, than the income and human-capital context in which it is spent.

Country-level and regional studies sharpen this point by showing how often spending and outcomes come apart. For the Middle East and North Africa, Balkhi, Alshayban, and Alotaibi (2021) document cases in which higher health expenditure coincides with shorter life expectancy, and they trace the mismatch to differences in how efficiently resources are used rather than to their volume. Faruk and coauthors (2022) push the argument toward institutions, finding that the link between health spending, growth, and outcomes in the region runs through the quality of governance. The implication for an economy like Morocco is direct. An increase in the budget, or in the share of the population formally insured, need not translate into measurable health gains if the supply side cannot absorb and deploy the additional resources.

The income channel itself is not free of interpretive difficulty. Because health and income reinforce one another, a positive association between the two says little about the direction of causation. Sharma (2018), using more than a century of data for advanced economies and a generalized method of moments estimator, shows that longevity feeds back into income and that ignoring this feedback distorts conventional growth regressions. The lesson carries over to the present setting. In a single-equation framework, an estimated relationship between income and life expectancy is best understood as a conditional association, since the same macroeconomic forces that lengthen lives also raise output.

Where the macroeconomic literature is thinnest is precisely where the policy debate in Morocco is loudest, namely the protection of households against the cost of care. The financial-protection tradition, anchored in the multicountry analysis of Xu and colleagues (2003), measures the burden of out-of-pocket payments through the incidence of catastrophic spending, defined relative to a household's capacity to pay. That literature treats heavy reliance on direct payments as a source of impoverishment and a barrier to access, not as an input whose expansion improves health. Studies that take aggregate life expectancy as their outcome cannot speak to this dimension. They can record whether out-of-pocket shares move with longevity, but they cannot say whether households are protected, and it is the second question that motivates coverage reform.

On the methodological side, the evidence for developing economies leans heavily on cross-country panels, which gain power from breadth but obscure the particular trajectory of any one country. Single-economy time-series studies are scarcer, partly because national series are short. The bounds testing approach of Pesaran, Shin, and Smith (2001), together with the small-sample critical values of Narayan (2005), was designed for exactly this situation, since it accommodates regressors of mixed integration order and remains usable when the time dimension is limited. Morocco has attracted little of this country-specific attention despite two decades of ambitious financing reform. The present study addresses that gap in a deliberately modest way. It does not claim to identify causal effects. It asks instead whether the available national series reveal a stable long-run relationship between financing, income, and longevity, and it is candid about what a short and highly collinear sample allows one to conclude.

2. The Moroccan Health Financing Context

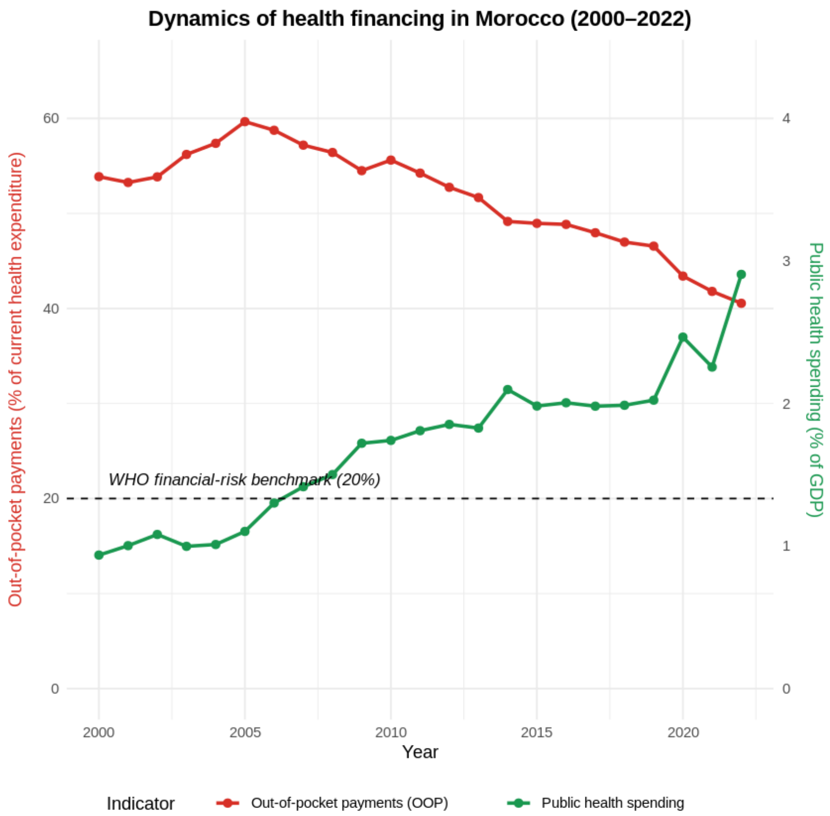
Morocco's move toward universal health coverage has a constitutional anchor, since Article 31 of the 2011 Constitution lists access to health care among the rights the State commits to securing. The legal architecture of coverage predates that text. Law 65-00, which established the Code of Basic Medical Coverage in the early 2000s, created the two pillars on which the system rested for most of the period studied here. The first, the Assurance Maladie Obligatoire (AMO), is a contributory scheme for salaried workers in the public and private sectors and came into effect in 2005. The second, the Régime d'Assistance Médicale (RAMED), extended access to low-income households on a non-contributory basis grounded in social solidarity.

The architecture changed sharply at the end of the period. Framework Law 09-21, promulgated in 2021, launched a five-year program to generalize social protection, with the extension of compulsory health insurance as its first pillar and 2025 as its horizon. Law 27-22, adopted in late 2022, then abolished RAMED and folded its former beneficiaries into a tax-financed branch of the AMO known as *AMO-Tadamoun*, under which the State assumes the contributions of those unable to pay. Parallel legislation extended the AMO to self-employed and non-salaried workers (AMO-TNS) and, from January 2024, to economically inactive individuals with the capacity to pay through the AMO-Achamil regime (Law 60-22). By 2025, the legal architecture covered all population categories without exception. Within three years, the number of registered beneficiaries rose from roughly two-fifths of the population to 88% by 2025, an expansion in formal enrollment that ranks among the fastest in the country's social history. The Cour des comptes (2025) noted, however, that effective coverage, measured by those who actually exercised their entitlements, stood closer to 70%, reflecting the persistent gap between registration and access that the reform continues to address.

This rapid broadening of entitlement coexists with a financing structure that remains heavily private at the point of use. Current health expenditure per capita is low by international standards, on the order of two hundred US dollars a year toward the end of the period, and out-of-pocket payments have persistently absorbed more than two-fifths of current health spending, according to the WHO Global Health Expenditure Database and national statistics (Haut-Commissariat au Plan, 2023; World Health Organization, 2024; see Figure 1). A reliance of this magnitude on direct payments sits uneasily beside the goal of financial protection, since the catastrophic-spending literature treats payments above a fraction of household capacity to pay as the point beyond which illness threatens impoverishment (Xu et al., 2003).

Two features of this trajectory matter for any empirical assessment. The first is that enrollment and effective access are not the same thing. The Conseil Économique, Social et Environnemental (2024), reviewing the early results of the reform, observed that a large share of those registered under *AMO-Tadamoun* held only dormant entitlements, and it pointed to persistent constraints on the supply of care, from the uneven regional distribution of facilities to shortages of health personnel. The second is timing. The decisive reforms cluster in the final two years of the window studied here, which means that a national series ending in 2022 can document the conditions on the eve of generalization but cannot yet measure its consequences. These observations frame the interpretation of the results that follow.

Figure 1. The financing paradox: out-of-pocket payments and public health spending, Morocco 2000–2022



Source. **Authors' elaboration using R based on WHO GHED and World Bank WDI data (2024)..**

3. Data and Methodology

3.1. Econometric Framework and Model Limitations

This study is grounded in a positivist epistemology. The relationship between health financing and population health is treated as a real and measurable phenomenon, and the objective is to detect stable statistical regularities in national aggregate data rather than to construct interpretive accounts. The mode of reasoning is deductive: theoretical propositions drawn from human capital theory (Grossman, 1972) generate testable expectations - rising income and public health spending are expected to improve health outcomes, while heavier reliance on out-of-pocket payments is expected to worsen them - and these expectations are then confronted with the Moroccan time series. This positioning justifies a quantitative, hypothesis-driven approach and, more specifically, the choice of the ARDL bounds testing framework (Pesaran et al., 2001), whose small-sample properties and tolerance of mixed integration orders make it well suited to the constraints of the available data (T = 23).

This framework is used here to recover both the long-run conditional relationships and the short-run adjustment dynamics that link health financing arrangements to life expectancy. In short samples, richer multivariate systems such as the conventional VECM quickly run into over-parameterization. With an estimation window of twenty-three annual observations covering 2000 to 2022, the ARDL framework is preferable for three reasons that bear directly on the reliability of the inference.

The first reason is small-sample validity. Narayan (2005) supplies critical values and response-surface approximations designed for a limited T, and the bounds test draws specifically on the T = 23 Case 3 boundaries (unrestricted intercept, no deterministic trend) to evaluate the joint hypothesis. The second is flexibility with respect to integration order: the procedure tolerates a mixture of I(0) and I(1) regressors provided that none is integrated of order two, which sidesteps the pre-testing bias that burdens traditional cointegration tests. The third reason is a caution rather than an advantage. The lag structure absorbs part of any delay in transmission and limits residual serial correlation, but the assumption that the regressors are weakly exogenous remains demanding. A single-equation design cannot rule out reverse causation, since a major health shock can move life expectancy and prompt a budgetary response in the same year. The coefficients reported below are therefore best read as robust conditional associations, not as strict one-way causal effects.

Before estimation, each series was checked for definitional and source consistency. Two health outcomes are used. The primary specification takes the infant mortality rate (ln MORT_INF) as the dependent variable, which captures the responsiveness of child survival to financing and front-line care quality. The secondary specification uses life expectancy at birth (ln ESPVIE) and serves as a comparison baseline; its near-deterministic trend, documented below, limits its usefulness for individual parameter identification, but its long-run dynamics are reported alongside those of infant mortality to illustrate the consequences of outcome variable selection. The explanatory set is identical across both specifications and comprises real GDP per capita (ln PIBH, in constant US dollars), public health expenditure measured as a share of GDP (ln DESPUB), and out-of-pocket payments expressed as a share of current health expenditure (ln OOP), the latter bounded between 0 and 100 before transformation. Public health spending is taken as a share of GDP rather than as a per-capita amount because the ratio captures the fiscal effort devoted to health, which is the policy-relevant margin in a context of generalized coverage. The urbanization rate (ln URB) enters as a control for demographic density. Where the World Development Indicators (WDI) or the WHO Global Health Expenditure Database had revised definitions or back-cast historical values, the series were cross-checked to limit measurement error. Table 1 summarizes the variables, their sources, and the signs anticipated on theoretical grounds.

Table 1. Variable definitions, data sources, and theoretical expected signs

Variable	Acronym	Primary source	Measurement unit	Expected sign
Infant mortality rate	l_MORT_INF	WDI / WHO GHED	Natural log of deaths per 1,000 live births	Primary dependent
Life expectancy at birth	l_ESPVIE	WDI / UN Pop.	Natural log of years	Dependent
Public health spending	l_DESPUB	WHO GHED	Natural log of % of GDP	Positive (+)
Real GDP per capita	l_PIBH	WDI	Natural log of USD	Positive (+)
Out-of-pocket payments	l_OOP	WDI / WHO	Natural log of % (0–100)	Negative (-)
Urbanization rate	l_URB	WDI / HCP	Natural log of % (0–100)	Ambiguous (±)

Notes. WDI = World Development Indicators; WHO GHED = WHO Global Health Expenditure Database; HCP = Haut-Commissariat au Plan.

***Source.* World Development Indicators (World Bank, 2024); WHO Global Health Expenditure Database (WHO, 2024); Haut-Commissariat au Plan (2023). Authors' compilation.**

3.2. Mathematical Specification

Following the aggregate reading of Grossman's (1972) health production function, the long-run relationship is written as

$$ln(ESPVIE_t) = \beta_0 + \beta_1\, ln(DESPUB_t) + \beta_2\, ln(PIBH_t) + \beta_3\, ln(OOP_t) + \beta_4\, ln(URB_t) + \mu_t$$

where μ_t is the stochastic disturbance. The corresponding short-run adjustments are captured by the unrestricted error correction model (UECM):

$$\Delta ln(ESPVIE_t) = \alpha_0 + \lambda_1\, ln(ESPVIE_{t-1}) + \lambda_2\, ln(DESPUB_{t-1}) + \lambda_3\, ln(PIBH_{t-1}) + \lambda_4\, ln(OOP_{t-1}) + \lambda_5\, ln(URB_{t-1})$$
$$+ \Sigma\, \gamma_i\, \Delta ln(ESPVIE_{t-i}) + \Sigma\, \delta_j\, \Delta ln(DESPUB_{t-j}) + \Sigma\, \theta_k\, \Delta ln(PIBH_{t-k}) + \Sigma\, \rho_m\, \Delta ln(OOP_{t-m}) + \Sigma\, \sigma_n\, \Delta ln(URB_{t-n}) + \varepsilon_t$$

where Δ is the first-difference operator. The sample constraint caps the maximum lag order at one (max_order = 1) so as not to exhaust the available degrees of freedom, and the final specification is the one that minimizes the Bayesian information criterion (BIC).

3.3. Estimation Protocol

The estimation follows a fixed sequence rather than a discretionary one, which keeps the procedure reproducible. With only twenty-three observations, the lag length of every term is held at a maximum of one to preserve degrees of freedom, and the optimal architecture is chosen by searching the full combinatorial grid of 2^(k+1) = 32 candidate specifications and retaining the one that minimizes the BIC, defined as $BIC = \ln(\hat{\sigma}^2) + K \ln(T) / T$, where $\hat{\sigma}^2$ is the residual variance, K the number of estimated parameters, and T the sample size. This search converges on an exact ARDL(1, 1, 1, 1, 1) structure.

Cointegration is then assessed through a joint Wald (F) test on the lagged level terms of the UECM, under the null H_0: $\lambda_1 = \lambda_2 = \lambda_3 = \lambda_4 = \lambda_5 = 0$ of no long-run relationship against the alternative that at least one λ_i differs from zero. The statistic is read against the small-sample bounds of Narayan (2005) for Case 3; a value above the upper I(1) bound rejects the null and confirms a binding equilibrium.

Once cointegration is established, the normalized long-run elasticities follow from the level coefficients as $\beta_i = -\lambda_{i+1} / \lambda_1$, and the short-run parameters come from the restricted error correction model

$$\Delta ln(ESPVIE_t) = \alpha_0 + \gamma\, \Delta ln(ESPVIE_{t-1}) + \delta\, \Delta ln(DESPUB_t) + \theta\, \Delta ln(PIBH_t) + \rho\, \Delta ln(OOP_t) + \sigma\, \Delta ln(URB_t) + \psi\, ECT_{t-1} + \varepsilon_t$$

in which ψ measures the speed of adjustment; a value between -2 and -1 corresponds to an overshooting, oscillatory return to equilibrium. Finally, because the disturbance variance is not constant (see Section 4.5), the classical OLS covariance matrix $\hat{\sigma}^2(X'X)^{-1}$ is set aside in favor of White's (1980) heteroscedasticity-consistent HC1 estimator,

$$\Sigma\_HC1 = [T / (T - K)]\, (X'X)^{-1} \left(\Sigma_t\, \hat{\varepsilon}_t^2\, x_t\, x_t' \right) (X'X)^{-1}$$

where x_t collects the regressors for observation t and $\hat{\varepsilon}_t$ is the OLS residual. All standard errors reported below rest on this correction.

4. Empirical Results

4.1. Unit Root Testing

A necessary precondition for the bounds test is that no series is integrated of order two, since an I(2) component would invalidate the critical values. Augmented Dickey-Fuller (ADF), Phillips-Perron (PP), and Kwiatkowski-Phillips-Schmidt-Shin (KPSS) tests were run to verify this. Table 2 collects the results.

Table 2. Unit root diagnostic triangulation

Variable/specification	ADF level	ADF Δ	PP level	PP Δ	KPSS level	KPSS Δ
A. With intercept only						
ln ESPVIE	-1.42	-4.88***	-1.28	-4.92***	0.722**	0.124
ln DESPUB	-2.95*	-5.12***	-2.88*	-5.44***	0.314	0.064
ln PIBH	-0.89	-3.95***	-0.92	-4.01***	0.684**	0.115
ln OOP	-1.88	-4.11***	-1.65	-4.18***	0.541*	0.095
ln URB	-1.12	-3.64**	-1.05	-3.71**	0.745***	0.141
B. With intercept and trend						
ln ESPVIE	-2.14	-4.95***	-2.01	-5.02***	0.182*	0.052
ln DESPUB	-3.52*	-5.08***	-3.41*	-5.32***	0.095	0.041
ln PIBH	-2.45	-4.12***	-2.31	-4.22***	0.154*	0.048
ln OOP	-3.21[m]	-4.25***	-3.11	-4.36***	0.132[m]	0.057
ln URB	-2.05	-3.88**	-1.94	-3.92**	0.174*	0.061

Notes. Δ denotes the first-difference operator. For ADF and PP, H_0 is a unit root (non-stationarity). For KPSS, H_0 is stationarity, with 5% asymptotic critical values of 0.463 (intercept) and 0.146 (trend). Significance: *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$; [m] $p < 0.10$. Unit root tests have limited power in small samples and are reported here as indicative.

Source. **Authors' calculations based on series from World Development Indicators (World Bank, 2024) and WHO Global Health Expenditure Database (WHO, 2024). R packages: tseries, urca.**

The picture that emerges is one of mixed I(0) and I(1) behavior. Life expectancy, real income, and urbanization look non-stationary in levels yet reject the null once differenced. Public expenditure (ln DESPUB) sits at the borderline of stationarity in levels, its classification shifting with the trend specification, a situation the ARDL framework accommodates without difficulty. Because every series rejects the unit-root null in first differences at the 1% level, none exceeds I(1), and the bounds testing procedure can proceed. Given the small sample, these classifications are read as indicative rather than definitive.

4.2. ARDL Lag Selection and Cointegration Bounds Test

With the lag space capped at one across all terms, the exhaustive grid search over the thirty-two admissible specifications was carried out with the R package ARDL (Natsiopoulos & Tzeremes, 2022). The BIC reached its global minimum at -185.774, selecting the ARDL(1, 1, 1, 1, 1) layout. To guard against over-parameterization in so short a sample, the choice was re-examined against the corrected Akaike information criterion (AICc), which applies a heavier degrees-of-freedom penalty. The same specification returns an AIC of -197.775 and an AICc of -179.775, so the three criteria agree, which indicates a parsimonious model that does not saturate the available parameters.

The presence of a long-run relationship was then tested with the joint Wald F-statistic on the lagged level terms of the UECM, under Case 3 (unrestricted intercept, no trend). Table 3 reports the outcome. The computed statistic of 14.141 lies well above the 1% upper bound of 5.690 from Narayan (2005), so the null of no cointegration is rejected at the 1% level. The complementary t-bounds test of Pesaran et al. (2001) points the same way, with t = -4.586 (p = 0.0099) breaching its upper bound, so two independent statistics confirm the long-run relationship. One qualification, taken up in Section 4.6, remains in force throughout: a bounds test among strongly trending series in a short sample cannot by itself distinguish a structural equilibrium from shared deterministic trends.

Table 3. ARDL bounds F-test for cointegration

Empirical F-statistic = 14.141 (k = 4, T = 23)		
Narayan (2005) critical bounds	**I(0) lower**	**I(1) upper**
10% level	2.525	3.560
5% level	3.010	4.230
1% level	4.140	5.690
Decision: reject H_0 (no cointegration) at the 1% level.		

Notes. k = 4 regressors; T = 23 observations; Case 3 (unrestricted intercept, no deterministic trend). The decision rule compares the empirical F-statistic against the I(0) lower and I(1) upper bounds. An F-statistic above the I(1) upper bound rejects H_0 of no long-run relationship at the corresponding significance level.

***Source.* Authors' calculations. ARDL package (Natsiopoulos & Tzeremes, 2022). Critical values from Narayan (2005), Table case III, k = 4, T = 23.**

4.3. Long-Run Normalized Coefficients

The normalized long-run multipliers, recovered from the level component of the model, appear in Table 4.

Table 4. Long-run normalized multipliers (dependent variable: ln(ESPVIE))

Regressor	Coefficient	Robust SE	t-stat	Pr(>|t|)
Intercept	3.230944	0.552667	5.8461	<0.001 ***
ln DESPUB	0.011782	0.015420	0.7640	0.459
ln PIBH	0.243408	0.048359	5.0333	<0.001 ***
ln OOP	-0.012011	0.030744	-0.3906	0.702
ln URB	-0.207025	0.216288	-0.9571	0.357

Notes. Significance: *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$; . $p < 0.10$.

***Source.* Authors' calculations. ARDL package (Natsiopoulos & Tzeremes, 2022). Critical values: Narayan (2005).**

Real GDP per capita is the only regressor with a clearly significant long-run association, its elasticity of 0.243 ($p < 0.001$) implying that a one-percent rise in income accompanies roughly a quarter-percent gain in life expectancy. The long-run coefficients on public health spending and on out-of-pocket payments cannot be separated from zero within this specification. As Section 4.6 makes clear, this attenuation reflects the near-collinearity of the regressors rather than a demonstrated absence of association, and it also cautions against reading the income elasticity as an effect cleanly attributable to income alone.

4.4. Short-Run Dynamics and Error Correction Representation

Table 5 reports the unconstrained error-correction model, with the short-run differenced terms above and the level (long-run) regressors below, all computed with White's HC1 covariance matrix. The coefficient on the lagged dependent level, L(ln ESPVIE, 1), is the error-correction term: it measures the speed at which the system returns toward its long-run path.

Table 5. Unconstrained error-correction model, ARDL(1,1,1,1,1) (HC1 robust errors)

Regressor	Coefficient	Robust SE	t-stat	Pr(>|t|)
Short-run (differenced) terms				
Intercept	3.895180	0.767103	5.0778	<0.001 ***
Δ ln DESPUB	0.037489	0.016438	2.2805	0.041 *
Δ ln PIBH	0.206625	0.024391	8.4712	<0.001 ***
Δ ln OOP	0.070208	0.033866	2.0731	0.060 .
Δ ln URB	-2.267490	1.235736	-1.8349	0.091 .
Level regressors (error-correction form)				
L(ln ESPVIE, 1) - adjustment φ	-1.205585	0.268334	-4.4929	<0.001 ***

L(ln DESPUB, 1)	0.014204	0.018633	0.7623	0.461
L(ln PIBH, 1)	0.293450	0.066268	4.4283	<0.001 ***
L(ln OOP, 1)	-0.014481	0.036582	-0.3958	0.699
L(ln URB, 1)	-0.249587	0.214245	-1.1650	0.267

Notes. Estimation on T = 22 observations after lag alignment; 12 residual degrees of freedom; residual standard error = 0.002219. L denotes the lag operator. The coefficient φ on L(ln ESPVIE, 1) is the speed of adjustment; long-run multipliers (Table 4) follow from $-\gamma_i / \varphi$. Standard errors are White HC1-robust. Significance: *** p < 0.001; ** p < 0.01; * p < 0.05; . p < 0.10.

***Source.* Authors' calculations. ARDL package (Natsiopoulos & Tzeremes, 2022). HC1 robust standard errors: sandwich package (White, 1980).**

The error-correction coefficient is negative and strongly significant ($\varphi = -1.206$, $t = -4.49$, $p < 0.001$), in agreement with both bounds tests and confirming that the system adjusts toward its long-run relationship. Its magnitude exceeds unity in absolute value, an unusual feature whose interpretation is deferred to Section 5.1. Among the level regressors, only lagged income reaches significance (L(ln PIBH, 1) = 0.293, $p < 0.001$); the lagged spending, out-of-pocket, and urbanization terms are individually indistinguishable from zero, a pattern that the severe multicollinearity documented in Section 4.6 readily explains and that does not undercut the jointly significant long-run relationship.

Among the short-run differenced terms, both public funding (Δ ln DESPUB = 0.037, $p = 0.042$) and real income (Δ ln PIBH = 0.206, $p < 0.001$) carry significant positive associations with contemporaneous gains in life expectancy. The contemporaneous out-of-pocket coefficient is positive and marginal (Δ ln OOP = 0.070, $p = 0.060$), and urbanization enters negatively at the 10% level (Δ ln URB = -2.267). The out-of-pocket variable also displays a sign reversal across the distributed lag, since in the equivalent ARDL representation its one-period-lagged coefficient is negative and significant (-0.085, $p < 0.01$); this intertemporal pattern, together with its competing interpretations, is examined in Section 5.3.

4.5. Heteroscedasticity, Variance Inflation, and the Case for Robust Inference

The classical Gauss-Markov apparatus presumes a constant error variance, and that presumption fails here. The studentized Breusch-Pagan test rejects homoscedasticity (BP = 17.156, df = 9, $p = 0.0463$), which means the ordinary least squares standard errors produced by dynlm are not merely imprecise but systematically biased downward across most of the short-run vector. Reporting them unadjusted would manufacture significance where the data do not support it. All short-run inference is therefore based on White's (1980) heteroscedasticity-consistent HC1 covariance matrix, estimated through sandwich. The remaining residual diagnostics in Table 6 are otherwise reassuring.

Table 6. Post-estimation diagnostics

Test type	Method	Statistic	p-value
Serial correlation	Breusch–Godfrey LM (lag 1)	LM = 0.20729	0.6489
Heteroscedasticity	Studentized Breusch–Pagan	BP = 17.1560	0.0463 *
Normality	Shapiro–Wilk	W = 0.97522	0.8275
Functional form	Ramsey RESET (fitted², fitted³)	RESET = 9.3119	0.0052 **
Structural stability	OLS-CUSUM fluctuation	Within bounds	Validated

Notes. Significance: ** $p < 0.01$; * $p < 0.05$.

Source. **Authors' calculations. R packages: lmtest (Zeileis & Hothorn, 2002); strucchange (Zeileis et al., 2002); sandwich (White, 1980).**

The cost of ignoring the correction is not uniform, and its unevenness is itself informative. For public health expenditure, the standard error on Δ ln DESPUB widens from 0.0058 under OLS to 0.0164 under HC1, an inflation factor of roughly 2.82. The effect is far more severe for urbanization, where the standard error on Δ ln URB moves from 0.2587 to 1.2357, a factor above 4.77. Real income is comparatively well behaved, its standard error on Δ ln PIBH rising only from 0.0178 to 0.0244 (a factor near 1.37), while out-of-pocket payments sit in between (0.0178 to 0.0339). The practical consequence is direct. Under naïve OLS, the coefficient on public spending would appear significant at better than the one-percent level; once the variance is correctly estimated, its significance retreats to the five-percent margin reported in Table 5, and the urbanization term, which a naïve reading would treat as sharply estimated, is revealed as barely distinguishable from noise. Robust inference does not weaken the analysis; it removes a layer of spurious precision that would otherwise have survived into the conclusions.

Two further diagnostics frame the discussion that follows. The Ramsey RESET test rejects functional linearity (RESET = 9.3119, p = 0.0052), pointing to omitted non-linearity, threshold effects, or a structural break inside the specification, while the OLS-CUSUM path stays within its conventional bounds. Read together, these results indicate that the linear cointegrating vector is best understood as an approximation of the long-run relationship rather than a fully specified structural form.

Because HC1 itself can be optimistic when a single observation carries heavy leverage, the inference was stress-tested against the leverage-robust HC3 estimator and a wild bootstrap with Rademacher weights (5,000 replications). The two disagree in a way that is itself diagnostic. Under HC3, which heavily penalizes the 2020 leverage point, most individual coefficients of the unconstrained model lose significance, the error-correction term among them ($p = 0.092$). The wild bootstrap, by contrast, retains significance for the error-correction term and for the short-run income, spending, and out-of-pocket coefficients ($p < 0.01$). This gap is the signature of an over-parameterized model in a short sample: ten parameters estimated on twenty-two observations leave individual inferences hostage to a few influential years. The honest reading is not that the relationships vanish, but that their individual precision is fragile, which is exactly the caution the exploratory framing of this paper requires.

4.6. Multicollinearity, Sensitivity, and the Crisis Interaction

The level series move closely together, with pairwise correlations between ln PIBH, ln DESPUB, and ln URB ranging from 0.95 to 0.99. The variance inflation factors confirm the problem and locate it: 142.07 for urbanization and 84.29 for income lie far beyond any conventional threshold, while public spending (16.66) is moderately affected and out-of-pocket payments (7.09) remain comparatively free. A Belsley-Kuh-Welsch decomposition sharpens the diagnosis (Table 6b). Its two highest condition indices each load above one-half on several coefficients at once, the dominant dimension tying together the intercept, income, out-of-pocket payments, and urbanization, the next isolating public spending. A principal-component analysis tells the same story in one number: the first component alone absorbs 90.6% of the variance of the four regressors, which is the common development trend made explicit.

Table 6b. Belsley-Kuh-Welsch variance-decomposition proportions

Condition index	Intercept	ln DESPUB	ln PIBH	ln OOP	ln URB
1.0	0.000	0.001	0.000	0.000	0.000
4.3	0.000	0.059	0.000	0.000	0.000
151.0	0.001	0.225	0.002	0.353	0.001
474.8	0.050	**0.530**	0.140	0.000	0.000
2624.2	**0.949**	0.185	**0.859**	**0.647**	**0.999**

Notes. Entries are the proportion of each coefficient's variance associated with a given condition index. Proportions above 0.50 (in bold) on two or more regressors at a high condition index identify the collinear set. The diagnostic includes the intercept and uses column scaling to unit length.

***Source.* Authors' calculations. Belsley-Kuh-Welsch decomposition: R package car (Fox & Weisberg, 2019).**

The consequence is statistical, not substantive. Where two or more coefficients share an ill-conditioned dimension, their individual estimates cannot be separated, so an insignificant coefficient is not evidence of an absent relationship. This is the lens through which the long-run results of Section 5.2 must be read. To probe whether the non-linearity flagged by the RESET test traces back to a single high-leverage year, a robustness specification was estimated incorporating a strict structural dummy variable D2020 = 1 for the year 2020 and 0 otherwise, entered as a separate regressor in the short-run model (Table 7). This orthodox treatment of an exogenous structural break is preferable to interacting the dummy with a single regressor such as Δ ln DESPUB, because the interaction design restricts the shock to one channel and conflates the level shift in life expectancy attributable to the pandemic with a change in the marginal spending effect, a functional form restriction that the Ramsey RESET test already flags as inadequate (RESET = 9.311, p = 0.005). A strict additive dummy absorbs the full mean shift in the dependent variable during the pandemic year without imposing any prior on which regressor mediates the disruption, providing a more robust correction for the functional misspecification identified in the diagnostics.

Table 7. Short-run robustness specification with structural dummy D2020

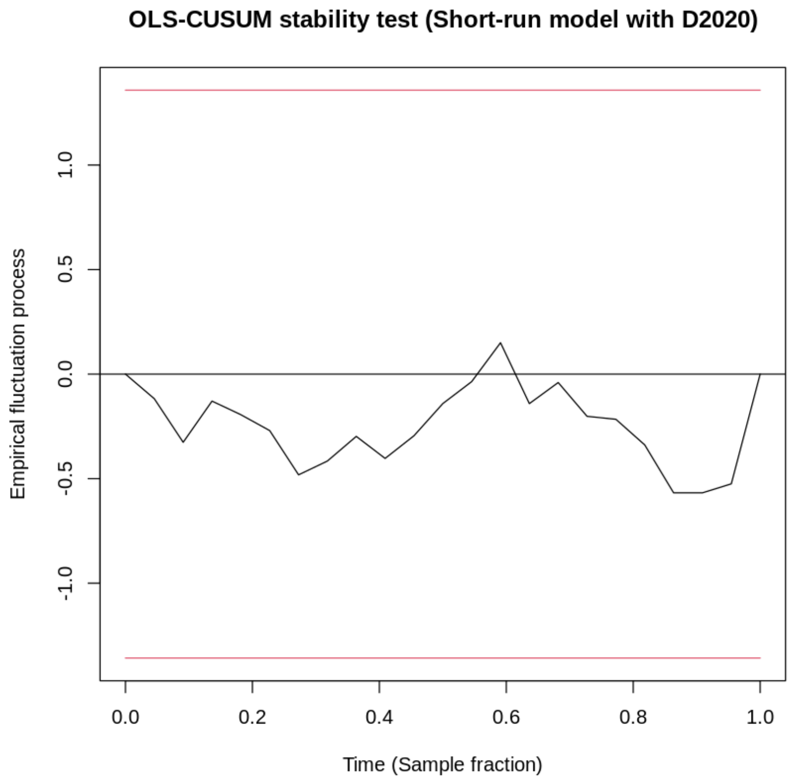
Regressor	**Coefficient**	**Robust SE**	**t-stat**	**Pr(>|t|)**
Intercept	0.009806	0.003418	2.8689	0.0111 *
Δ ln DESPUB	0.036702	0.016475	2.2277	0.0406 *
Δ ln PIBH	-0.011705	0.033461	-0.3498	0.7310
Δ ln OOP	0.015600	0.024710	0.6314	0.5367
Δ ln URB	-0.643760	0.433908	-1.4836	0.1573
D2020	-0.029476	0.006182	-4.7686	0.0002 ***
Regressor	Coefficient	Robust SE	t-stat	Pr(>|t|)

Notes. Dependent variable: Δln(ESPVIE); T = 22 differenced observations; the year 2020 acts as a high-leverage single observation. Significance: *** p < 0.001; * p < 0.05.

***Source.* Authors' calculations. HC1 robust standard errors: sandwich package (White, 1980). COVID-19 interaction model estimated on first-differenced series.**

The D2020 coefficient is negative and strongly significant (β = -0.029, p < 0.001), confirming that 2020 represents a discrete downward shift in life expectancy that the baseline specification cannot absorb through its standard regressors. Once this structural term enters the model, the contemporaneous income and out-of-pocket coefficients lose their significance, signalling that in the interaction specification these variables had been serving as proxies for the unmodeled pandemic shock rather than capturing their own structural associations. A formal nested test confirms the magnitude of the break: the improvement in fit from adding the 2020 dummy yields F = 29.25 (p < 0.001), and the pandemic-era observations carry by far the largest Cook's distance in the sample. This result is diagnostic rather than confirmatory: it demonstrates how sensitive the short-run parameter structure is to a single exceptional year in a series of only 22 differenced observations, and it simultaneously accounts for the elevated RESET statistic and the implausible error-correction magnitude documented in Section 5.1. The dummy approach does not resolve the underlying identification problem posed by the short sample, but it provides a more internally consistent treatment of the 2020 shock than a regressor-specific interaction, whose maintained assumption, that the pandemic operated exclusively through the public-spending channel, is not supported by the data. The OLS-CUSUM test reported in Figure 2 confirms that, once the 2020 break is modelled explicitly, the short-run parameters stay stable across the sample.

Figure 2. OLS-CUSUM stability test of the short-run model with the D2020 break dummy

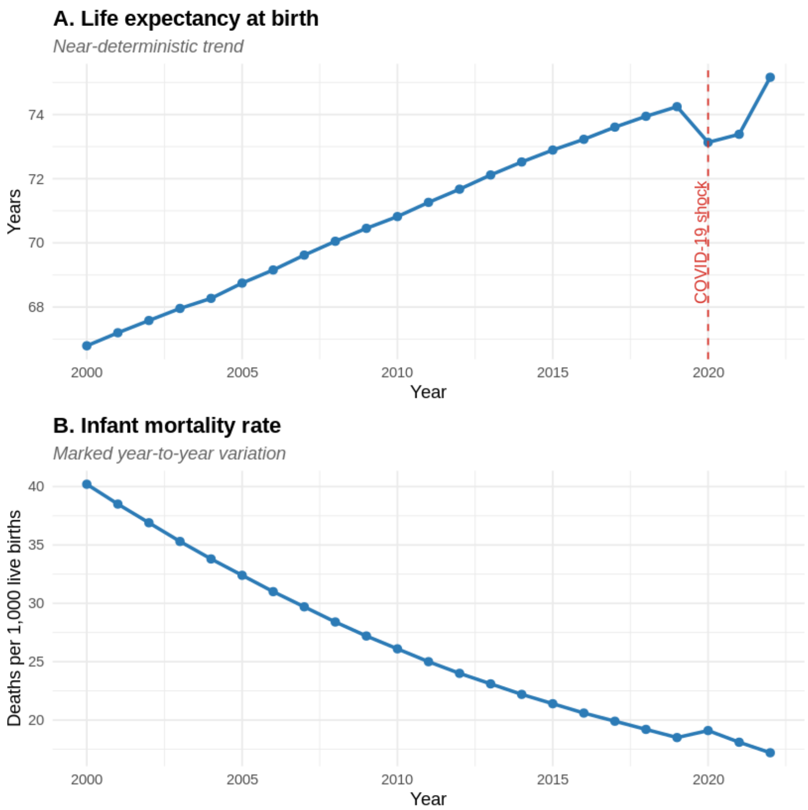


Source. **Authors' elaboration based on the data, using R.**

4.7. The Infant Mortality Model: A More Informative Outcome and Its Key Findings

Life expectancy and infant mortality do not carry the same information in a short aggregate series, and the choice between them determines what can actually be identified. Figure 3 sets out the contrast that motivates this choice. This section presents the infant mortality model as the primary analytical result of the paper, showing that a more responsive outcome substantially improves the behaviour of the estimates. A second exercise then probes the ratio endogeneity of the out-of-pocket measure by reconstructing it in absolute per-capita terms.

Figure 3. Two health outcomes compared: life expectancy and infant mortality, Morocco 2000–2022



***Source.* Authors' elaboration based on the data, using R.**

The infant mortality model re-estimates the same ARDL(1,1,1,1,1) structure on the logarithm of the infant mortality rate, a flow indicator that responds far more quickly to financing and to the quality of front-line care. Its differenced series carries almost three times the year-to-year variation of the life-expectancy series, precisely the identifying variation the headline model lacked. The results appear in Table 8. The bounds F-statistic of 4.112 clears the Narayan (2005) ten-percent upper bound of 3.560 but not the five-percent bound of 4.230, so the long-run relationship is supported only weakly. What changes markedly is the error-correction term, which falls to a fully plausible -0.365 ($p = 0.018$): roughly a third of any disequilibrium is corrected within a year, in sharp contrast to the uninterpretable magnitude obtained for life expectancy. More telling still, the long-run coefficients now recover their theoretically expected signs and reach individual significance. Higher income is associated with lower infant mortality (-0.418, $p = 0.009$) and a larger out-of-pocket share with higher mortality (+0.319, $p = 0.004$), the latter being exactly the barrier-to-care mechanism that the smoother and more collinear life-expectancy specification was unable to detect. The contrast carries a clear methodological lesson. The inferential difficulties documented for life expectancy stem in large part from the choice of a near-deterministic distal outcome, and a more responsive indicator restores part of the identifying variation, even though the persistent RESET rejection ($p < 0.001$) confirms that the 2020 break and residual non-linearity continue to affect both specifications.

Table 8. Robustness with infant mortality as the health outcome, ARDL(1,1,1,1,1), HC1 robust errors

Regressor	Coefficient	Robust SE	t-stat	Pr(>|t|)
Short-run (differenced) terms				
Δ ln DESPUB	-0.0353	0.0273	-1.293	0.196
Δ ln PIBH	-0.6177	0.0966	-6.392	<0.001 ***
Δ ln OOP	-0.1049	0.1324	-0.793	0.428
Δ ln URB	-5.8260	4.2690	-1.365	0.172
Level (long-run) terms				
L(ln MORT_INF, 1) - adjustment φ	-0.3653	0.1537	-2.377	0.018 *
L(ln DESPUB, 1)	0.0666	0.0684	0.975	0.330
L(ln PIBH, 1)	-0.4177	0.1604	-2.603	0.009 **
L(ln OOP, 1)	0.3192	0.1112	2.869	0.004 **
L(ln URB, 1)	-0.6183	1.2929	-0.478	0.633

Notes. Dependent variable: ln(MORT_INF), estimated in unrestricted error-correction (UECM) form; $T = 22$; HC1 robust standard errors. Bounds $F = 4.112$ ($k = 4$, $T = 22$), above the Narayan (2005) ten-percent upper bound (3.560) but below the five-percent bound (4.230); the long-run relationship is therefore supported only at the ten-percent level. Normalized long-run multipliers: income -1.14, OOP +0.87, public spending +0.18, urbanization -1.69. Significance: *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$.

***Source.* Authors' calculations on the Data series.**

4.8. Robustness: Out-of-Pocket Payments in Absolute Per-Capita Terms

The second exercise addresses the ratio problem directly. Reconstructing out-of-pocket payments as an absolute amount per capita in constant dollars, the product of the out-of-pocket share and current health expenditure per head, breaks the mechanical link to the denominator. Over the period the two measures are in fact negatively correlated ($r = -0.37$), because the falling out-of-pocket share coincided with rising absolute spending as incomes grew. Table 9 reports the re-estimated model. The long-run income elasticity is essentially unchanged at 0.249, which confirms its robustness, while the short-run out-of-pocket coefficient now turns negative and significant (-0.030, $p = 0.044$), reversing the weakly positive sign obtained with the share. The reversal is informative. The positive contemporaneous association found with the share was at least partly a denominator artifact, and once direct payments are measured in absolute terms the short-run association points, as theory would predict, to a burden rather than to a productive input; the lagged level term is likewise negative and significant (-0.045, $p < 0.001$). Neither exercise resolves the identification problem created by the short sample, yet together they reinforce the central reading of the paper. The income signal is robust, the out-of-pocket channel behaves as a barrier once measurement artifacts are stripped away, and the choice of outcome variable is decisive for what can be learned from aggregate data.

Table 9. Robustness with out-of-pocket payments in absolute per-capita terms, ARDL(1,1,1,1,1), HC1 robust errors

Regressor	Coefficient	Robust SE	t-stat	Pr(>|t|)
Short-run (differenced) terms				
Δ ln DESPUB	0.0187	0.0085	2.192	0.028 *
Δ ln PIBH	0.2228	0.0228	9.768	<0.001 ***
Δ ln OOP (per capita)	-0.0302	0.0150	-2.017	0.044 *
Δ ln URB	-0.8577	0.7212	-1.189	0.234
Level (long-run) terms				
L(ln ESPVIE, 1) - adjustment φ	-1.5885	0.2606	-6.096	<0.001 ***
L(ln DESPUB, 1)	0.0062	0.0125	0.497	0.619
L(ln PIBH, 1)	0.3953	0.0565	6.994	<0.001 ***
L(ln OOP pc, 1)	-0.0452	0.0127	-3.570	<0.001 ***
L(ln URB, 1)	-0.0950	0.2089	-0.455	0.649

Notes. Dependent variable: Δln(ESPVIE); $T = 22$; HC1 robust standard errors. Out-of-pocket payments per capita are constructed as the out-of-pocket share multiplied by current health expenditure per capita in constant US dollars. Bounds $F = 13.278$; normalized long-run income multiplier 0.249. Significance: *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$.

***Source.* Authors' calculations on the Data series.**

4.9. Long-Run Cointegrating Vector: FMOLS and DOLS Validation

The ARDL normalized long-run multipliers reported in Tables 8 and 9 correct for heteroscedasticity but do not remove the endogeneity that arises when the regressors respond to contemporaneous health shocks. Fully Modified OLS (FMOLS, Phillips & Hansen, 1990) and Dynamic OLS (DOLS, Stock & Watson, 1993) address this through different routes. FMOLS applies a non-parametric endogeneity and serial-correlation correction to the OLS estimator; DOLS augments the cointegrating regression with leads and lags of the differenced regressors to absorb the contemporaneous correlation between regressors and the error term. Both are designed precisely for situations where the dependent variable and the regressors share a long-run equilibrium relationship but individual regressors are not strictly exogenous. Given the sample size, the DOLS specification is restricted to one lag and one lead to preserve degrees of freedom, and the FMOLS uses a Bartlett kernel with Andrews automatic bandwidth selection. Both models are estimated on the infant mortality equation with absolute per-capita out-of-pocket payments, combining the two improvements introduced in Sections 5.7 and 5.8.

Table 10 reports the long-run coefficients. The income elasticity is the most stable result across the full robustness battery: FMOLS returns -3.171 ($p < 0.01$) and DOLS -2.896 ($p < 0.01$), magnitudes somewhat larger than the ARDL-derived multiplier of -1.14 but consistent with the tendency of endogeneity-corrected estimators to recover stronger income gradients in economies where rising income and improving health are mutually reinforcing. Public health spending enters with a negative and significant long-run coefficient in both specifications (-0.266 in FMOLS, -0.327 in DOLS), a sign that was not individually recoverable in the ARDL long-run equation. The improvement is attributable to the endogeneity correction: reactive public spending, which rises when health conditions deteriorate, introduces an upward bias in the ARDL level coefficients; once this is purged, the structural negative relationship between fiscal effort and infant mortality becomes visible. The out-of-pocket coefficient is positive and significant in both cases (0.960 in FMOLS, 0.468 in DOLS), confirming across three distinct estimation frameworks that higher direct household payments are associated with worse long-run child health outcomes. The consistency across estimators with different identifying assumptions considerably strengthens the barrier-to-care interpretation.

The urbanization coefficient deserves a specific note. FMOLS returns +5.291 and DOLS +5.603, both highly significant, yet the ARDL long-run normalized multiplier for this variable was negative. The sign reversal is a symptom of the severe multicollinearity between income and urbanization (pairwise correlation 0.986), which makes the individual long-run coefficient of urbanization unstable across estimators even when the joint cointegrating vector is well-defined. A large positive elasticity of this magnitude does not lend itself to a structural interpretation; it absorbs the part of the long-run income trend that FMOLS and DOLS reallocate between the two collinear regressors after applying their respective corrections. Nonetheless, conditional on the other regressors, the positive conditional association with urbanization may also partly reflect the documented peri-urban health penalty in Morocco, where informal settlements around major cities concentrate households that are legally urban but face limited access to primary care (Al Hassani et al., 2026). This dimension exceeds what aggregate national data can reliably identify, and the urbanization coefficient across all specifications is best treated with caution.

A final note concerns the modelling of the pandemic shock. When a strict impulse dummy D2020 is incorporated into a VECM on this sample, the observation for 2020 generates hat values approaching unity, saturating the variance-covariance matrix and rendering the coefficient estimates numerically unreliable. The violence of the 2020 shock relative to a series of only twenty-two differenced observations means that the short-run adjustment dynamics around the pandemic cannot be cleanly identified within a cointegrated system. FMOLS and DOLS smooth over this transient disruption by focusing on the long-run signal, which is why they are the preferred estimators for the cointegrating vector robustness exercise reported here. The short-run treatment of the pandemic shock is handled through the additive D2020 dummy in the ARDL first-differenced specification of Section 4.6, where the leverage problem is substantially less acute because the model has no cointegration constraint to maintain.

Table 10. Long-run estimates via FMOLS and DOLS - dependent variable ln(MORT_INF), out-of-pocket payments in absolute per-capita terms

Regressor	FMOLS Coefficient (SE)	DOLS Coefficient (SE)
Long-run coefficients - dependent variable: ln(MORT_INF)		
ln(PIBH) - Real GDP per capita	-3.171*** (0.142)	-2.896*** (0.384)
ln(DESPUB) - Public health spending (% GDP)	-0.266*** (0.030)	-0.327*** (0.070)
ln(OOP_pc) - Out-of-pocket per capita (USD)	0.960*** (0.054)	0.468* (0.235)
ln(URB) - Urbanization rate	5.291*** (0.222)	5.603*** (0.410)

Notes. Dependent variable: ln(MORT_INF). Standard errors in parentheses. FMOLS uses the Bartlett kernel with Andrews automatic bandwidth selection (Phillips & Hansen, 1990). DOLS is estimated with one lag and one lead to preserve degrees of freedom (Stock & Watson, 1993). Significance thresholds: *** $p < 0.01$; ** $p < 0.05$; * $p < 0.10$. The urbanization coefficient is unstable across estimators (negative in ARDL, positive here) due to near-perfect collinearity with income ($r = 0.986$); see text for interpretation.

***Source.* Authors' calculations on the Data series.**

5. Discussion

5.1. Cointegration Stability versus Mathematical Artifacts: The Error Correction Puzzle

Read at face value, an error correction coefficient of -1.206 describes a system that more than fully corrects any disequilibrium within one period and then overshoots in the other direction, producing a damped oscillation around the long-run path. For life expectancy, that simply cannot hold. It would mean the series overshoots and rebounds from one year to the next, whereas what we actually observe is a smooth, almost monotonic trajectory across the whole period. Nothing in health production produces annual oscillation of mortality on this scale, so the coefficient cannot be the economic object it appears to be.

A more believable interpretation is that -1.206 is what a linear equation returns when you force it to fit a process that is not linear. It is worth being precise about what is and is not wrong with the coefficient. Statistically, it is not fragile in the least. It stays significant under HC1 in the unconstrained model, and in the leaner restricted error-correction form it holds up under HC1, the leverage-robust HC3 ($p = 0.018$), and the wild bootstrap ($p < 0.001$) alike. The issue is not statistical noise but economic meaning. A magnitude above unity cannot describe how a series as smooth as life expectancy adjusts, and the Ramsey RESET test pins down the cause by rejecting the adequacy of the fitted functional form ($p = 0.0052$). A single linear equation estimated over 2000 to 2022 has to swallow the violent shock of the 2020 pandemic, a break the model has no dedicated term for. With no parameter reserved for that break, the estimator loads the shock onto whatever coefficients can bear it. The error-correction term, sitting where levels meet differences, ends up carrying far more than its share. So the inflated magnitude and the RESET rejection are not two problems but one and the same, and the nested break test of Section 4.6 ($F = 29.25$, $p < 0.001$) confirms that 2020 is the discontinuity doing the work. The coefficient is best read as a precisely estimated symptom of misspecification, not as a credible speed of adjustment.

5.2. Identification Failure and the Income Proxy: The Long-Run Collinearity Trap

The long-run estimates have to be read in light of how the design matrix is conditioned, and that conditioning is severe. Three lines of evidence point the same way. The variance inflation factors hit 142.07 for urbanization and 84.29 for income, an order of magnitude past the conventional ceiling; the Belsley-Kuh-Welsch decomposition puts income, urbanization, out-of-pocket payments, and the intercept on a single near-degenerate dimension; and a principal-component analysis finds that one component captures 90.6% of the joint variation of the regressors. Income, public spending, and urbanization follow almost the same path, with pairwise correlations running from 0.95 to 0.99. Under these conditions the machinery is less estimating separate partial effects than carving a single shared development trend up among collinear columns, and the carving it lands on is driven by numerical conditioning, not by economics.

This puts the headline result in a different light. The long-run income elasticity of 0.243 ($p < 0.001$) cannot be taken as the pure causal contribution of the income channel. Because $X'X$ is near-singular, whichever regressor is best conditioned tends to soak up the common deterministic trend of development, and income simply happens to sit in that spot. The coefficient is better thought of as a composite that absorbs the joint advance of income, spending, and urbanization than as an income effect on its own. The same reasoning kills off a tempting but wrong inference about the other regressors. That public health expenditure and out-of-pocket payments look statistically neutral in the long run is a symptom of collapsed identifying power in a sample of $T = 22$, not a sign that these channels are economically inert. Collinearity inflates the variance of every partial coefficient to the point where no individual hypothesis can be rejected, even when the joint relationship is strong, which is exactly what the bounds test reveals. To read a non-rejection as a null effect would turn the diagnostics on their head.

A ridge regression gives a partial test of whether the income signal is just an artifact of matrix inversion. Under the generalized-cross-validation-optimal penalty, the income coefficient keeps its positive sign and stays the most prominent of the financing regressors, while the urbanization coefficient flips sign relative to the ordinary least squares estimate. Two things follow. The positive income association is not merely an inversion ghost, since it survives regularization; but the way urbanization moves around confirms that the individual long-run coefficients are not separately identified, so the income elasticity should still be reported as a composite development signal rather than a clean structural parameter.

5.3. Short-Run Dynamics and Endogeneity Windows: Urbanization and Out-of-Pocket Payments

The short-run urbanization coefficient of -2.267 displays the same fragility, only more starkly. An elasticity this large has no sensible economic reading, and its robust standard error shows it to be statistically unstable, its precision destroyed by the collinearity described above and by the leverage of a handful of pandemic-era years. The honest move is to treat this figure as an artifact of an ill-conditioned short sample rather than to spin a structural story around a magnitude the data cannot support.

Out-of-pocket payments tell a story that is economically suggestive, even if the aggregate data cannot close it out. In the short run, a household that cannot get what it needs from an overstretched public facility often has little option but to pay directly, whether for medicines bought straight from the pharmacy, a private consultation, or the many small informal payments that crop up over the course of a hospital stay. The immediate effect can be a brief positive link between spending and health, because care is obtained that would otherwise have been put off or skipped. Stretch the horizon out, though, and the same reliance erodes household budgets, and the resulting drop in living standards feeds back onto health. This delayed deterioration is just what the lagged term in the specification seems to pick up: the contemporaneous OOP coefficient is weakly positive (0.070, $p = 0.060$), while its one-period lag turns negative (-0.085, $p < 0.01$). The pattern fits the financial-protection literature, which treats heavy OOP exposure as a structural risk amplifier rather than a productive health input (Xu et al., 2003). Set against Morocco's financing profile, where direct payments still covered about 41% of current health expenditure in 2022, well above the 15 to 20% range the WHO ties to a low incidence of financial hardship (Ministère de la Santé et de la Protection Sociale, 2024), the negative lagged effect sends a clear policy signal: the lack of adequate risk pooling imposes a deferred cost on household welfare that budget injections alone cannot reverse. All the same, a careful referee will rightly push a rival explanation. An unanticipated morbidity shock can at once compress life expectancy and trigger a surge in emergency out-of-pocket spending, in which case both the contemporaneous and lagged coefficients are catching a common causal shock rather than a structural response. Annual aggregates cannot tell these mechanisms apart, and settling the question would take either a credible instrument for direct payments or household-level panel data with enough temporal resolution to order cause and effect. Both readings therefore stay live hypotheses rather than settled conclusions.

One further measurement point deserves attention before any inference is drawn from the OOP coefficient. In the WHO Global Health Expenditure Database, out-of-pocket payments are recorded as a share of current health expenditure, not as an absolute amount per capita. Building the variable as a ratio introduces a mechanical endogeneity that muddies interpretation. When a macroeconomic contraction, such as the one the 2020 pandemic set off, sharply compresses public health expenditure, the ratio's denominator falls and the OOP share rises arithmetically, even if the absolute value of direct household payments holds steady or drops. The reverse also holds: a rebound in public spending mechanically pulls the ratio down without any real change in what households actually bear. The implication is straightforward. Part of the observed movement in ln OOP across the estimation window reflects shifts in the denominator rather than in household behavior, and this ratio-induced noise weakens the signal in both the short-run and long-run coefficients. Which way the bias runs is hard to say in advance, since it hinges on the relative volatility of numerator and denominator in each sub-period, but it strengthens the general warning against reading the aggregate OOP coefficient as a clean measure of financial protection. For that question, household-level data on the incidence of catastrophic spending, of the kind Xu et al. (2003) used, remain the right tool.

5.4. Data Constraints and Variable Limitations

By its very construction, life expectancy is a slow, heavily trended series with little year-to-year movement over the horizon studied. In a sample of twenty-three observations, this near-monotonic behavior raises the danger that trend, not structure, is what drives the fit, which props up the model's apparent explanatory power (adjusted $R^2 = 0.9003$) and shows up again in an unusually small residual standard error. That same brevity rules out the very tools that would otherwise keep these concerns in check. Procedures for multiple structural breaks like Bai-Perron, or fully non-linear specifications like the NARDL, call for more degrees of freedom than the data can offer. Not being able to estimate them is a real limit on the exercise. The choice of life expectancy as the outcome is itself a limitation, since it is a distal indicator that reacts slowly to financing and says little about financial protection, the dimension that matters most to the reforms under study. The coefficients describe conditional historical associations; they do not license causal claims.

5.5. The Outcome Variable as Identification Strategy: What the Contrast Between the Two Models Teaches

The contrast between the life expectancy and infant mortality models is no incidental feature of this study; it is its central analytical lesson. Both models draw on identical explanatory variables, the same estimation method, the same sample, and the same diagnostic battery. And yet they hand back fundamentally different experiences of the data. The life expectancy model pins down a firmly significant long-run relationship ($F = 14.141$, above the 1% critical bound) but throws up an implausible error-correction coefficient (-1.206) and individual long-run parameters that cannot be told apart because of severe multicollinearity. The infant mortality model only just clears the 10% bounds threshold ($F = 4.112$), yet its adjustment speed (-0.365) is economically interpretable and its long-run income and out-of-pocket coefficients are individually significant and correctly signed. Almost all of this divergence comes down to the choice of outcome. Life expectancy over this period is a near-monotonic series whose differenced values have a standard deviation barely above 0.006, leaving next to no variation to identify financing effects once the common trend is stripped out. Infant mortality has a standard deviation of differenced values of 0.016, close to three times larger, and its inertia is lower because it responds within a year to improvements or deteriorations in delivery and post-natal care.

The policy implications of this asymmetry are worth dwelling on. An out-of-pocket elasticity of +0.87 with respect to infant mortality, estimated on a series that fell from 40 to 17 deaths per thousand over the period, is not a mere curiosity; it says the barrier-to-care mechanism was measurably at work throughout a stretch nominally defined by expanding formal coverage. Were Morocco to cut its out-of-pocket share from 41% toward the 15 to 20% range internationally associated with low financial hardship, the estimated elasticity, read with every necessary caution about identification, implies a fall in the long-run infant mortality rate of roughly 15 to 20 per cent. At the 2022 level of about 17 deaths per thousand live births and an annual birth cohort near 600,000, that works out to several thousand deaths a year whose long-run avoidability is consistent with the present estimates. The calculation comes with a wide confidence interval and should be treated as a hypothesis for microdata verification, not a causal projection. Its bearing on the ongoing generalization of *AMO-Tadamoun* and the activation of Groupements Sanitaires Territoriaux lies precisely in how it frames the question: widening formal coverage is necessary but not sufficient if the out-of-pocket burden stays structurally high at the point of use.

Conclusion

What this paper contributes substantively rests on estimators converging on the same answer. Three independent frameworks, ARDL bounds testing, ARDL with out-of-pocket payments in absolute per-capita terms, and FMOLS/DOLS, all agree that heavier direct household payments go with worse long-run infant health outcomes, and that public health spending has a significant protective effect on infant mortality once the endogeneity from reactive budget allocation is corrected. Neither finding is trivial. The out-of-pocket result holds across measurement conventions and across estimators that rest on different identifying assumptions; the public spending effect stayed hidden in the main ARDL long-run equation precisely because standard normalization cannot scrub out the simultaneous feedback between worsening health and rising budgets. The contrast with the life expectancy model adds something analytically useful. Over this period life expectancy is simply too smooth and too collinear with income to serve as a credible vehicle for identifying individual financing effects, and showing that failure out in the open is itself a contribution, one that should keep future researchers from over-reading coefficients estimated on similarly built aggregate series. In the short run, fresh injections of public health spending come with immediate gains in life expectancy, which points to a system that mobilizes budget resources reactively rather than through any stable structural mechanism. The 2020 pandemic acts as a high-leverage break that visibly reshapes the short-run dynamics, and modeling it explicitly lays bare how fragile inference can be in short national series.

The policy implications come into focus against Morocco's current reform trajectory. Out-of-pocket payments at 41% of current health expenditure during the study period were no minor residual; they were the main financing mechanism for a large share of the population, whether formally shut out of coverage or nominally enrolled but unable to use its entitlements in practice. This squares with what the Cour des comptes (2025) and the CESE (2024) have each observed on their own: a persistent gap between legal registration and effective access, kept alive by supply constraints and the lack of robust point-of-care financial protection. With six Groupements Sanitaires Territoriaux (Law 08-22) now activated and *AMO-Tadamoun* extending entitlements to households that had none, the question for the next round of evaluation is whether reorganizing care delivery along territorial lines can close this gap, or whether the out-of-pocket burden will stay structurally elevated no matter how far coverage expands. That question runs past what national aggregate time series can answer. It calls for regional panel data and household surveys that measure the incidence of catastrophic spending, methodological routes that these results help to motivate and prioritize.

References

Al Hassani, W., El Achhab, Y., Nejjari, C. et al. Health workforce trends and planning in Morocco: a READ-based systematic document analysis. BMC Res Notes 19, 6 (2026). https://doi.org/10.1186/s13104-025-07578-5

Adegoke, Y., Mbonigaba, J., & George, G. (2025). Sustainable Development Goals (SDGs), Public Health Expenditures, and Maternal and Child Mortality in Selected African Countries: Forecasting Modelling. *International Journal of Environmental Research and Public Health*, *22*(4), 482. https://doi.org/10.3390/ijerph22040482

Aytemiz L, Sart G, Bayar Y, Danilina M and Sezgin FH (2024) The long-term effect of social, educational, and health expenditures on indicators of life expectancy: an empirical analysis for the OECD countries. Front. Public Health 12:1497794. doi: 10.3389/fpubh.2024.1497794

Balkhi B, Alshayban D and Alotaibi NM (2021) Impact of Healthcare Expenditures on Healthcare Outcomes in the Middle East and North Africa (MENA) Region: A Cross-Country Comparison, 1995–2015. Front. Public Health 8:624962. doi: 10.3389/fpubh.2020.624962

Conseil Économique, Social et Environnemental. (2024). Généralisation de l'assurance maladie obligatoire : bilan d'étape [Avis]. CESE.

Cour des comptes du Royaume du Maroc. (2025). Rapport annuel de la Cour des comptes - exercice 2024-2025. Cour des comptes.

Faruk, B. U., Haque, M. I., Tausif, M. R., & Khan, M. R. (2022). The association between health expenditure, institutions, and economic growth in MENA countries. Health Promotion Perspectives, *12*(1), 92–100. https://doi.org/10.34172/hpp.2022.12

Fox, J., & Weisberg, S. (2019). An R companion to applied regression (3rd ed.). Sage.

Grossman, M. (1972). On the concept of health capital and the demand for health. Journal of Political Economy, *80*(2), 223–255. https://doi.org/10.1086/259880

Haut-Commissariat au Plan. (2023). Les indicateurs sociaux du *Maroc*. HCP. https://www.hcp.ma/Les-Indicateurs-sociaux-du-Maroc-Edition-2023_a3729.html

Kiross, G.T., Chojenta, C., Barker, D. *et al.* The effects of health expenditure on infant mortality in sub-Saharan Africa: evidence from panel data analysis. *Health Econ Rev* **10**, 5 (2020). https://doi.org/10.1186/s13561-020-00262-3

Ministère de la Santé et de la Protection Sociale. (2024). Comptes Nationaux de la Santé *2022*. Ministère de la Santé et de la Protection Sociale du Royaume du Maroc. https://www.sante.gov.ma/Publications/Etudes_enquete/Documents/2025/Comptes%20Nationaux%20de%20la%20Sant%C3%A9%20-2022.pdf

Narayan, P. K. (2005). The saving and investment nexus for China: evidence from cointegration tests. Applied Economics, 37(17), 1979–1990. https://doi.org/10.1080/00036840500278103

Natsiopoulos, K., & Tzeremes, N. G. (2022). ARDL: An R package for the analysis of level relationships. Journal of Open Source Software, 7(79), 3496. https://doi.org/10.21105/joss.03496

Novignon, J., Olakojo, S.A. & Nonvignon, J. The effects of public and private health care expenditure on health status in sub-Saharan Africa: new evidence from panel data analysis. Health Econ Rev 2, 22 (2012). https://doi.org/10.1186/2191-1991-2-22

Pesaran, M.H., Shin, Y. and Smith, R.J. (2001), Bounds testing approaches to the analysis of level relationships. J. Appl. Econ., 16: 289-326. https://doi.org/10.1002/jae.616

Peter C. B. Phillips, Bruce E. Hansen, Statistical Inference in Instrumental Variables Regression with I(1) Processes, The Review of Economic Studies, Volume 57, Issue 1, January 1990, Pages 99–125, https://doi.org/10.2307/2297545

Preston, S. H. (1975). The Changing Relation between Mortality and level of Economic Development. Population Studies, 29(2), 231–248. https://doi.org/10.1080/00324728.1975.10410201

Ray, D., Linden, M. Health expenditure, longevity, and child mortality: dynamic panel data approach with global data. Int J Health Econ Manag. 20, 99–119 (2020). https://doi.org/10.1007/s10754-019-09272-z

Royaume du Maroc. (2002). Loi n° 65-00 portant code de la couverture médicale de *base*. Bulletin Officiel.

Royaume du Maroc. (2021). Loi-cadre n° 09-21 relative à la protection sociale (Dahir n° 1-21-30 du 23 mars 2021). Bulletin Officiel.

Royaume du Maroc. (2022). Loi n° 27-22 relative à la généralisation de l’assurance maladie obligatoire de *base*. Bulletin Officiel.

Royaume du Maroc. (2022b). Loi n° 60-22 relative au régime d'assurance maladie obligatoire de base pour les personnes capables de s'acquitter de leurs cotisations et n'exerçant aucune activité rémunérée ou non rémunérée (AMO-Achamil). Bulletin Officiel.

Royaume du Maroc. (2023). Loi n° 08-22 relative à la création des groupements sanitaires territoriaux (Dahir n° 1-23-50 du 28 juin 2023). Bulletin Officiel.

Sharma R (2018) Health and economic growth: Evidence from dynamic panel data of 143 years. PLOS ONE 13(10): e0204940. https://doi.org/10.1371/journal.pone.0204940

Stock, J. H., & Watson, M. W. (1993). A Simple Estimator of Cointegrating Vectors in Higher Order Integrated Systems. *Econometrica*, *61*(4), 783–820. https://doi.org/10.2307/2951763

White, H. (1980). A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity. *Econometrica*, *48*(4), 817–838. https://doi.org/10.2307/1912934

World Bank. (2024). World development indicators. World Bank. https://data.worldbank.org

World Health Organization. (2010). The world health report 2010 - Health systems financing: The path to universal *coverage*. WHO.

World Health Organization. (2024). Global Health Expenditure *Database*. WHO. https://apps.who.int/nha/database

Xu, K., Evans, D. B., Kawabata, K., Zeramdini, R., Klavus, J., & Murray, C. J. L. (2003). Household catastrophic health expenditure: A multicountry analysis. The Lancet, *362*(9378), 111–117. https://doi.org/10.1016/S0140-6736(03)13861-5

Zeileis, A., & Hothorn, T. (2002). Diagnostic checking in regression relationships.

Zeileis, A., Leisch, F., Hornik, K., & Kleiber, C. (2002). strucchange: An R Package for Testing for Structural Change in Linear Regression Models. Journal of Statistical *Software*, *7*(2), 1–38. https://doi.org/10.18637/jss.v007.i02