

Modeling Real GDP in Kazakhstan Using Regression-Based Differential Equations

Yernur Galym, Abdizhan Zhapan

Supervisor: Dr. Piotr Sebastian Skrzypacz

Second Reader: Dongming Wei

April 2026

Abstract

This capstone project studies the dynamic behavior of real GDP in Kazakhstan using regression-based differential equation models. Rather than identifying causal effects between variables, the project focuses on describing and estimating how log real GDP evolves over time within a continuous-time framework. The resulting extrapolations are therefore interpreted as model-based extensions of the fitted trajectory rather than structural macroeconomic forecasts.

This approach is appropriate for three reasons. First, differential equation models directly describe rates of change, which makes them useful for studying adjustment dynamics and long-run behavior. Second, many economic theories, especially growth and accumulation models, are naturally formulated in continuous time, so this framework is consistent with their underlying structure. Third, once estimated, the model can be used to generate model-based extrapolations and compare short-run and long-run implications of the fitted trajectory.

Overall, regression-based differential equation modeling provides a mathematically coherent and empirically interpretable framework for analyzing real GDP dynamics in Kazakhstan.

1 Introduction

Real GDP evolves over time through continuous adjustment processes rather than isolated changes between observed dates. This makes continuous-time modeling a natural

framework for studying macroeconomic dynamics. In this project, we use regression-based differential equation models to describe and estimate the evolution of real GDP in Kazakhstan.

The goal of the project is not to identify causal relationships between macroeconomic variables. Instead, we aim to construct a mathematically coherent and empirically interpretable dynamic model that captures how log real GDP changes over time, how it adjusts toward long-run patterns, and how the fitted trajectory can be extended beyond the observed sample.

Our approach is motivated by two main considerations. First, many economic theories, especially those related to growth, accumulation, and adjustment, are naturally formulated in continuous time [1, 2, 3]. Second, differential equation models provide a direct way to represent the law of motion of a variable, which is more informative for dynamic analysis than a purely static regression framework.

To implement this idea, we model the rate of change of log real GDP through a first-order differential equation with a deterministic polynomial component. This specification allows us to combine analytical tractability with enough flexibility to capture trend behavior in the data. Once the model is derived and estimated, it can be used to interpret adjustment speed, study long-run behavior, and generate medium-term model-based extrapolations implied by the fitted path.

The remainder of the project focuses on establishing the mathematical correctness of the model, deriving its solution, fitting it to real GDP data, and evaluating whether the resulting specification is interpretable and useful for empirical analysis.

2 Related Literature

This project is related to three strands of literature. The first is the literature on GDP forecasting and macroeconomic time-series models. Recent studies emphasize that GDP forecasting usually requires comparing several model classes rather than relying on a single specification. For example, Perone and Zambrano-Monserrate [4] compare SARIMA, ETS, NNAR, TBATS, and hybrid models for GDP growth forecasting in South Caucasian economies. For Kazakhstan specifically, Akhmet [5] develops a dynamic factor model with regularization and exponential decay of observations, highlighting the importance of unstable and time-varying macroeconomic relationships in GDP forecasting.

The second strand is the literature on differential-equation-based economic growth modeling. Cheow et al. [6] use integer-order and fractional-order differential models combined with regression analysis to model Malaysia's GDP and compare model performance using criteria such as AIC, RMSE, and R^2 . This is close to the present project because we also combine a differential equation structure with regression-based estimation and information criteria.

The third strand concerns time-varying parameter and state-space approaches. Awe et

al. [7] use a time-varying parameter state-space model to study the relationship between GDP and macroeconomic indicators, showing that allowing coefficients to evolve over time can help capture changing economic relationships. Christensen, Neri, and Parra-Alvarez [8] also emphasize the methodological importance of estimating continuous-time macroeconomic models from discretely observed data. This supports our interpretation of annual GDP observations as discrete realizations of an underlying continuous-time trajectory.

Compared with these studies, our contribution is narrower but more focused. We do not build a structural macroeconomic forecasting system. Instead, we develop a parsimonious regression-based differential equation model for Kazakhstan's log real GDP, compare polynomial forcing terms, and examine whether a time-varying adjustment coefficient improves the baseline specification.

3 Mathematical Correctness of the Model

The purpose of this project is not to identify causal effects between macroeconomic variables, but to describe and estimate how Kazakhstan real GDP evolves over time under a continuous-time dynamic structure.

Let $u(t)$ denote the logarithm of real GDP at time t :

$$u(t) = \log(\text{Real GDP}(t)).$$

We model the evolution of log real GDP rather than the level of real GDP because the logarithmic transformation smooths the scale of the series and makes the dynamic path easier to interpret. We assume that its evolution is governed by a first-order linear ordinary differential equation of the form

$$\frac{du}{dt} + au(t) = P_m(t), \quad (1)$$

where

$$P_m(t) = b_0 + b_1t + b_2t^2 + \dots + b_mt^m \quad (2)$$

is a polynomial of degree m , and $a, b_0, b_1, \dots, b_m \in \mathbb{R}$ are constant parameters.

In this formulation, t denotes continuous time, while $u(t)$ represents the underlying path of log real GDP. The term $au(t)$ captures the dependence of the rate of change on the current level of log real GDP, and the polynomial term $P_m(t)$ represents a deterministic component that allows for flexible long-run behavior. Thus, the model directly describes the law of motion of log real GDP rather than a static contemporaneous relationship.

To ensure that the model is mathematically well defined, we impose the following assumptions. Time is taken on an interval $I \subseteq \mathbb{R}$. The function $u(t)$ is assumed to be differentiable on I . The coefficients $a, b_0, b_1, \dots, b_m$ are real constants. Since $P_m(t)$ is a polynomial, it is continuous on $\mathbb{R}$ and therefore continuous on the interval of interest.

Under these assumptions, equation (1) belongs to the standard class of first-order linear ordinary differential equations with continuous coefficients. Hence, for any prescribed initial condition

$$u(t_0) = u_0, \quad (3)$$

where $t_0 \in I$ and $u_0 \in \mathbb{R}$, there exists a unique solution on the interval of interest. Therefore, the model is mathematically well posed [9].

This point is essential for the logic of the project. Since the observed real GDP data are recorded at discrete dates, the model does not treat them as isolated observations. Instead, it interprets them as realizations of an underlying continuous-time trajectory generated by (1). This is consistent with many economic theories, especially growth and accumulation frameworks, in which the relevant variables evolve continuously over time.

Moreover, the structure of the model is internally coherent with the objectives of the capstone project. The differential equation describes how log real GDP adjusts over time, allows for long-run trend behavior through the polynomial component, and remains analytically tractable. As a result, the model provides a formal basis for subsequent estimation, interpretation, and extrapolation.

Hence, the proposed specification is mathematically correct in the following sense: the variables and parameters are clearly defined, the assumptions are explicit, the model is well posed, and the differential equation generates a unique continuous-time trajectory for every admissible initial condition.

In the baseline specification, the adjustment coefficient a is treated as constant over the full sample period. This assumption is useful because it keeps the model analytically tractable and allows the solution to be derived in closed form. However, from an economic perspective, the constant value of a should be interpreted as an average adjustment coefficient over the sample rather than as a literal structural constant.

A natural extension is to allow the adjustment coefficient to vary over time:

$$\frac{du}{dt} + a(t)u(t) = P_m(t). \quad (4)$$

In this more general formulation, $a(t)$ represents a time-varying speed of adjustment. This is similar in spirit to time-varying parameter models in stochastic process and time-series analysis, where coefficients are allowed to change across periods rather than remaining fixed over the entire sample [10, 11, 12].

In the empirical extension, we use the simple specification

$$a(t) = \exp(\gamma_0 + \gamma_1 z_t), \quad (5)$$

where z_t is normalized time. The exponential form ensures that $a(t) > 0$ for all years, preserving the stabilizing adjustment interpretation.

4 Derivation of the Solution

In this section, we derive the explicit solution of equation (1) for the baseline case in which the adjustment coefficient a is constant. Recall that the model is given by

$$\frac{du}{dt} + au(t) = P_m(t), \quad (6)$$

where

$$P_m(t) = b_0 + b_1t + b_2t^2 + \cdots + b_mt^m \quad (7)$$

is a polynomial of degree m .

To solve this equation, we use the integrating factor method [9]. Since the coefficient of $u(t)$ is the constant a , an integrating factor is

$$\mu(t) = e^{at}. \quad (8)$$

Multiplying both sides of equation (1) by e^{at} gives

$$e^{at} \frac{du}{dt} + ae^{at}u(t) = e^{at}P_m(t). \quad (9)$$

By the product rule, the left hand side can be written as

$$\frac{d}{dt} (e^{at}u(t)) = e^{at}P_m(t). \quad (10)$$

Integrating both sides with respect to t , we obtain

$$e^{at}u(t) = \int e^{at}P_m(t) dt + C, \quad (11)$$

where $C \in \mathbb{R}$ is a constant of integration. Therefore, the general solution is

$$u(t) = e^{-at} \left(C + \int e^{at}P_m(t) dt \right). \quad (12)$$

If we impose the initial condition

$$u(t_0) = u_0, \quad (13)$$

then it is more convenient to integrate over the interval from t_0 to t . Doing so gives

$$e^{at}u(t) - e^{at_0}u_0 = \int_{t_0}^t e^{as}P_m(s) ds. \quad (14)$$

Hence, the unique solution satisfying the initial condition is

$$u(t) = e^{-a(t-t_0)}u_0 + e^{-at} \int_{t_0}^t e^{as}P_m(s) ds. \quad (15)$$

This formula gives the continuous-time trajectory generated by the model for any admissible initial condition. It shows that the solution consists of two parts. The first part,

$$e^{-a(t-t_0)}u_0, \quad (16)$$

depends on the initial condition and describes the transient component of the dynamics. The second part,

$$e^{-at} \int_{t_0}^t e^{as} P_m(s) ds, \quad (17)$$

is determined by the polynomial forcing term and captures the systematic evolution implied by the model.

Moreover, since $P_m(t)$ is a polynomial, the integral term is well defined for all t . When $a > 0$, the transient term decays over time, which means that the influence of the initial condition becomes weaker as t increases. When $a < 0$, the transient term grows over time. Thus, the parameter a plays an important role in governing the speed and direction of adjustment.

It is also useful to note the special case $a = 0$. In that case, equation (1) reduces to

$$\frac{du}{dt} = P_m(t), \quad (18)$$

and the solution with initial condition $u(t_0) = u_0$ becomes

$$u(t) = u_0 + \int_{t_0}^t P_m(s) ds. \quad (19)$$

Thus, the model admits a closed-form solution in all cases. This is important for the empirical part of the project, since it allows us to connect the theoretical continuous-time specification with observed real GDP data and to estimate the unknown parameters using the explicit solution formula.

If the adjustment coefficient is allowed to vary over time as $a(t)$, the integrating factor becomes

$$\mu(t) = \exp \left(\int a(t) dt \right). \quad (20)$$

This shows that the time-varying version remains mathematically interpretable, although it is less convenient for empirical estimation. For this reason, the present capstone estimates the constant- a version as a parsimonious baseline and treats the time-varying version as a robustness check.

5 Fitting Procedure

After deriving the explicit solution of the model, the next step is to fit it to observed real GDP data. The purpose of this section is to explain how the theoretical continuous-time trajectory is connected to the available discrete observations and how the unknown parameters are estimated.

Suppose that we observe real GDP at discrete times

$$t_1, t_2, \dots, t_n, \quad (21)$$

with corresponding observed log values

$$u_1, u_2, \dots, u_n. \quad (22)$$

In our context, these observations represent annual values of Kazakhstan log real GDP. Although the data are observed only at discrete dates, the model assumes that they arise from an underlying continuous-time process.

Recall that the model is given by

$$\frac{du}{dt} + au(t) = P_m(t), \quad (23)$$

where

$$P_m(t) = b_0 + b_1 t + b_2 t^2 + \dots + b_m t^m. \quad (24)$$

Let

$$\theta = (a, b_0, b_1, \dots, b_m, C) \quad (25)$$

denote the vector of unknown parameters in the general solution, or equivalently the parameters together with the initial condition when the initial value formulation is used.

Using the explicit solution derived in the previous section, we define the fitted trajectory

$$\hat{u}(t; \theta), \quad (26)$$

which depends on the unknown parameter vector θ . The idea is to choose parameter values so that the theoretical path $\hat{u}(t; \theta)$ is as close as possible to the observed log real GDP data.

To measure the discrepancy between the model and the data, we define the residual at time t_i by

$$\varepsilon_i = u_i - \hat{u}(t_i; \theta), \quad i = 1, 2, \dots, n. \quad (27)$$

The parameters are then estimated by minimizing the sum of squared residuals:

$$S(\theta) = \sum_{i=1}^n (u_i - \hat{u}(t_i; \theta))^2. \quad (28)$$

Thus, the estimation problem is

$$\hat{\theta} = \arg \min_{\theta} S(\theta). \quad (29)$$

This is a nonlinear least squares problem, because the solution $\hat{u}(t; \theta)$ depends on the parameters in a nonlinear way [13]. In particular, the parameter a enters the solution

through exponential terms, while the coefficients of the polynomial component determine the deterministic part of the trajectory.

The fitting procedure consists of the following conceptual steps. First, a degree m is chosen for the polynomial term $P_m(t)$. Second, the corresponding closed-form solution is written down. Third, the objective function in equation (28) is constructed using the observed data. Finally, numerical optimization is used to obtain the parameter estimates that minimize the sum of squared residuals.

This approach is appropriate for the present project because the model is derived in continuous time, while the data are available only at discrete dates. Rather than approximating derivatives directly from the data, we estimate the parameters through the explicit solution itself. This makes the fitting procedure consistent with the theoretical structure of the model.

Once the parameter estimates are obtained, the fitted solution can be compared to the observed series in order to evaluate how well the model captures the overall trajectory of the data. The estimated parameters can then be interpreted in economic terms. In particular, the parameter a governs the adjustment behavior of the system, while the polynomial coefficients determine the long-run deterministic component.

In practice, several candidate specifications may be considered by varying the degree m of the polynomial term. These models can then be compared using goodness of fit, visual agreement with the data, and interpretability of the estimated dynamics. In this way, model selection is based not only on numerical fit but also on whether the resulting specification remains economically meaningful and analytically coherent.

Therefore, the fitting procedure provides the empirical link between the theoretical differential equation model and the observed real GDP data. It allows the continuous-time specification to be estimated, evaluated, and later used for interpretation and model-based extrapolation.

6 Interpretation of Parameters and Economic Meaning

After estimating the parameters of the model, it is important to interpret them in a way that is both mathematically precise and economically meaningful. The purpose of this section is to explain how the coefficients of the regression-based differential equation determine the dynamic behavior of log real GDP under study.

Recall that the baseline model is given by

$$\frac{du}{dt} + au(t) = P_m(t), \quad (30)$$

where

$$P_m(t) = b_0 + b_1t + b_2t^2 + \cdots + b_mt^m. \quad (31)$$

The parameter a plays a central role in the dynamic structure of the baseline model. For $a \neq 0$, the model can be rewritten as

$$\frac{du}{dt} = -a \left(u(t) - \frac{P_m(t)}{a} \right). \quad (32)$$

This representation shows that the system adjusts toward the time-dependent target path

$$u^*(t) = \frac{P_m(t)}{a}. \quad (33)$$

Therefore, a measures how quickly deviations of log real GDP from this target path disappear.

When $a > 0$, the homogeneous part of the solution contains the factor e^{-at} , which decays over time. This means that the effect of the initial condition becomes weaker as time passes. In this case, the system exhibits a form of stabilizing adjustment, since deviations from the long-run path tend to diminish over time.

When $a < 0$, the factor e^{-at} grows over time. This means that the influence of the initial condition increases rather than disappears. In such a case, the system may display unstable behavior, since deviations from the underlying path can become larger over time.

When $a = 0$, the model reduces to

$$\frac{du}{dt} = P_m(t), \quad (34)$$

so the rate of change is determined entirely by the polynomial component. In this case, there is no adjustment term linked to the current level of the variable.

Therefore, the sign and magnitude of a provide information about the adjustment speed and dynamic stability of the log real GDP process. A larger positive value of a implies faster convergence of the transient component, while a smaller positive value implies slower adjustment.

The coefficients $b_0, b_1, \dots, b_m$ determine the polynomial forcing term $P_m(t)$, which represents the deterministic component of the model. These coefficients shape the long-run movement of log real GDP over time.

The constant term b_0 captures the baseline component of the rate of change. The coefficient b_1 determines the linear trend in the rate of change. Higher-order coefficients such as b_2, b_3 , and so on allow for curvature and more flexible patterns in the dynamic trajectory. As a result, the polynomial term makes it possible to represent growth paths that are not purely linear and that may accelerate or decelerate over time.

From an economic point of view, this means that the polynomial component can reflect gradual structural tendencies in real GDP. For example, it may capture persistent trend behavior associated with long-run growth or broader changes in the economic environment. However, these coefficients should be interpreted as describing the shape of the trajectory rather than as isolated causal effects.

The initial condition also has an important interpretation. It determines the starting point of the continuous-time trajectory and affects the transient part of the solution. In empirical applications, the initial value anchors the model to the observed data and helps determine how the theoretical path begins relative to the sample period.

Taken together, the parameter a , the polynomial coefficients, and the initial condition describe different aspects of the system. The parameter a governs adjustment behavior, the polynomial coefficients determine the systematic trend component, and the initial condition determines the starting position of the trajectory.

This interpretation is important for the goals of the capstone project. The purpose of the model is not only to fit the observed data well, but also to provide an interpretable dynamic representation of real GDP evolution in Kazakhstan. A specification that fits the data numerically but produces implausible adjustment behavior or economically unreasonable long-run dynamics would be less useful than a slightly simpler model with clearer interpretation.

For this reason, the evaluation of the estimated model should consider both statistical fit and economic meaning. In the following part of the project, the fitted trajectories and estimated parameter values can be examined in order to assess how well the model captures the observed data and whether the implied dynamics are reasonable in the real GDP context.

At the same time, the assumption that a is constant is restrictive in macroeconomic applications. Kazakhstan's economy passed through several different phases during the sample period, including transition, commodity-driven growth, crisis episodes, and post-pandemic recovery. Therefore, the true speed of adjustment may plausibly vary over time.

For this reason, the estimated value of a in the baseline model should be interpreted as an average adjustment speed over the sample. A more flexible version of the model would replace a with a time-varying coefficient $a(t)$:

$$\frac{du}{dt} = -a(t) (u(t) - u^*(t)). \quad (35)$$

Under this formulation, $a(t)$ measures the local adjustment speed at time t . This connects the model to stochastic process and time-varying parameter approaches, where the dynamic response of the system may change across different economic regimes.

7 Empirical Results for Kazakhstan Real GDP

In this section, we apply the regression-based differential equation model to annual real GDP data for Kazakhstan over the period from 1993 to 2025, using the World Development Indicators data source [14]. Let

$$u(t) = \log(\text{Real GDP}(t)),$$

where real GDP is measured in constant million tenge and $t = 0, 1, 2, \dots, n$ denotes annual time, with $t = 0$ corresponding to 1993. For readability, figures and extrapolation tables may report GDP levels in trillion tenge after converting from the original million-tenge units.

The purpose of this section is fourfold. First, we assess whether the continuous-time model can reproduce the historical trajectory of Kazakhstan real GDP. Second, we compare several specifications of the deterministic forcing term. Third, we interpret the estimated parameters in economic terms. Finally, we determine which specification is the most appropriate for the subsequent analysis.

7.1 Model Specifications

We estimate three nested specifications of the model

$$\frac{du}{dt} + au(t) = P_m(t), \quad (36)$$

where $P_m(t)$ is a polynomial forcing term.

The first specification uses a constant forcing term:

$$\frac{du}{dt} + au(t) = b_0. \quad (37)$$

The second specification allows for a linear trend:

$$\frac{du}{dt} + au(t) = b_0 + b_1 t. \quad (38)$$

The third specification allows for quadratic curvature:

$$\frac{du}{dt} + au(t) = b_0 + b_1 t + b_2 t^2. \quad (39)$$

These three models provide a natural progression from the simplest dynamic structure to a more flexible specification. This allows us to evaluate whether increasing the complexity of the forcing term materially improves the empirical fit while preserving interpretability.

7.2 Estimation Procedure

For each specification, the explicit solution derived in the previous section is fitted to the observed annual log real GDP data using nonlinear least squares. More precisely, for each degree m , the unknown parameter vector is chosen to minimize the sum of squared residuals

$$S(\theta) = \sum_{i=1}^n (u_i - \hat{u}(t_i; \theta))^2, \quad (40)$$

where u_i denotes observed log real GDP and $\hat{u}(t_i; \theta)$ is the fitted value implied by the differential equation solution.

The comparison across models is based on the sum of squared errors, root mean squared error, the coefficient of determination R^2 , and information criteria such as AIC and BIC. The information criteria are useful because they penalize additional model complexity, which is important when comparing constant, linear, and quadratic forcing terms.

7.3 Estimation Results

Table 1 reports the estimation results for the three candidate specifications.

Table 1: Comparison of candidate real GDP models

Model	$\hat{a}$	RMSE	R^2	AIC	BIC
Constant forcing term	0.0611	0.4159	0.9530	41.7480	46.2375
Linear forcing term	1.1959	0.1860	0.9906	-9.3692	-3.3832
Quadratic forcing term	1.8232	0.1075	0.9969	-43.5351	-36.0526

Table 2: Estimated parameters of the preferred quadratic model

Parameter	Estimate	Interpretation
$\hat{a}$	1.8232	Average adjustment speed over the sample
$\hat{C}$ or $\hat{u}_0$	10.2835	Initial condition or integration constant
$\hat{b}_0$	24.4226	Baseline component of the forcing term
$\hat{b}_1$	0.4478	Linear trend component
$\hat{b}_2$	-0.0043	Curvature of the forcing term

The results show a clear improvement in fit as the polynomial forcing term becomes richer. The constant forcing specification already captures the broad upward trend in real GDP, but it leaves a substantial amount of variation unexplained. Allowing the forcing term to depend linearly on time improves the fit considerably. The quadratic specification provides the strongest fit among the three, with the lowest root mean squared error and the highest coefficient of determination.

Thus, the empirical evidence suggests that the real GDP trajectory over the sample period is not adequately represented by a purely constant trend component. Instead, the data exhibit curvature that is better captured by a quadratic forcing term.

7.4 Graphical Fit

Figure 1 compares the observed log real GDP series with the fitted trajectories from the three models. Figure 2 then presents the observed real GDP series together with the fitted path implied by the preferred quadratic specification.

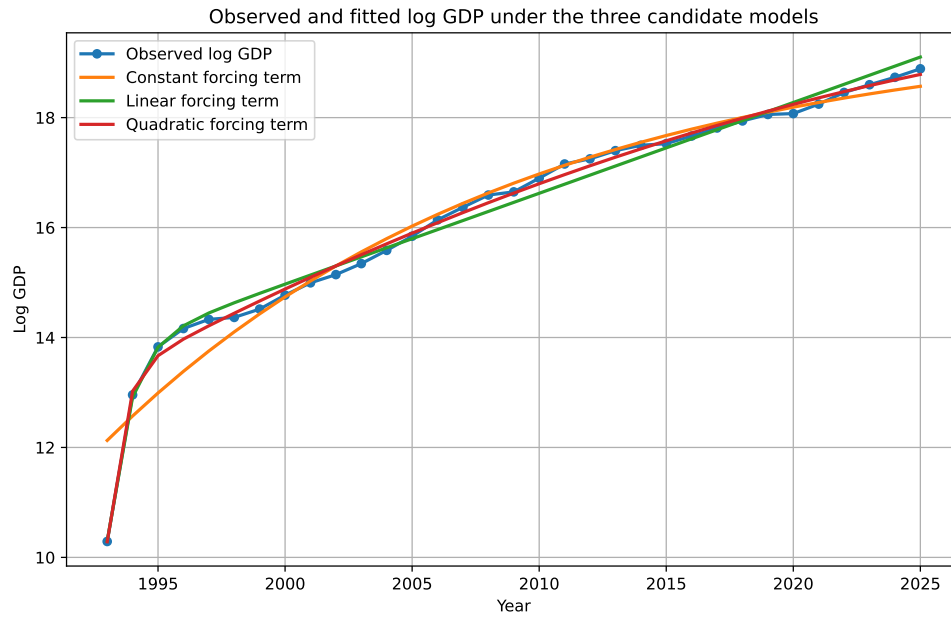


Figure 1: Observed and fitted log real GDP under the three candidate models

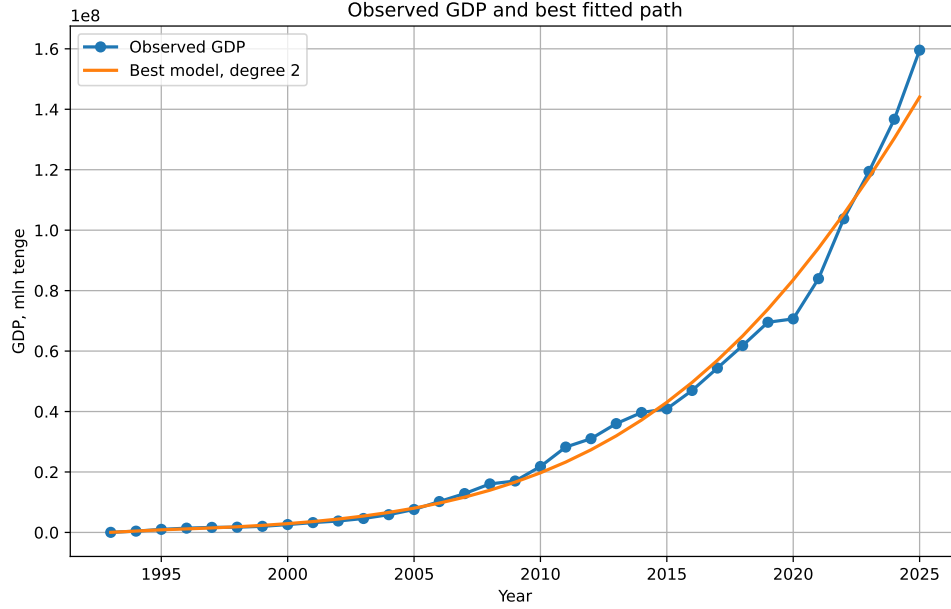


Figure 2: Observed real GDP and fitted path under the preferred quadratic model

The graphical evidence is consistent with the numerical results. The constant forcing model is too rigid to track the changing slope of the real GDP path over time. The linear forcing model captures the general direction much better, but still misses some curvature. The quadratic specification tracks the observed series most closely over the full sample.

7.5 Residual Diagnostics

In addition to comparing fitted trajectories visually, it is important to examine whether the preferred model leaves systematic patterns in the residuals. Let the residuals of the preferred quadratic model be defined as

$$\hat{\varepsilon}_i = u_i - \hat{u}(t_i; \hat{\theta}). \quad (41)$$

A well-fitting reduced-form model should leave residuals that do not display a strong trend or persistent structure over time.

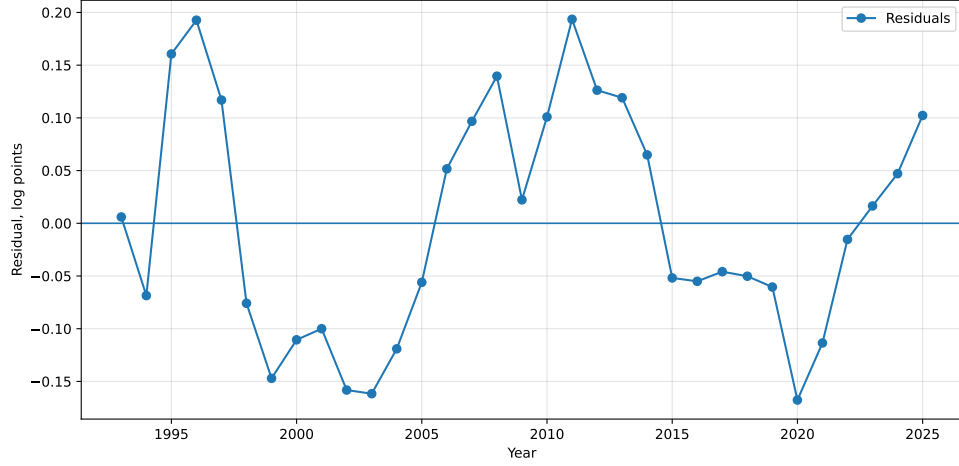


Figure 3: Residuals from the preferred quadratic model

The residual plot shows that the preferred quadratic model captures the broad deterministic trajectory, but some systematic patterns remain. To examine this more directly, we also inspect the residual autocorrelation function.

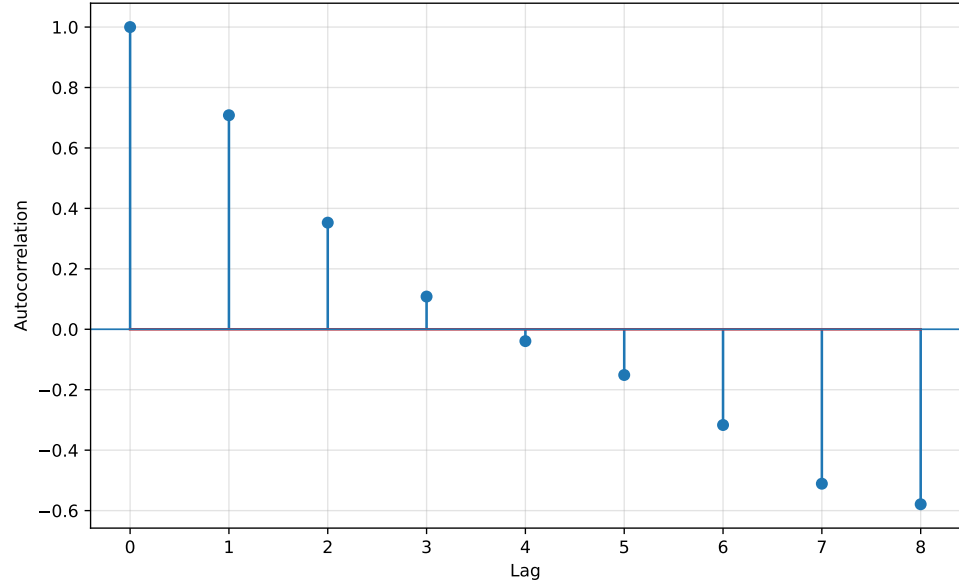


Figure 4: Autocorrelation function of residuals from the preferred quadratic model

The residual autocorrelation function indicates that some short-run dependence remains, especially at the first few lags. This suggests that the deterministic differential-

equation model captures the main trend component of real GDP, but does not fully capture short-run time-series dynamics. A natural extension would be to add a stochastic error component or compare the model with ARIMA or state-space benchmarks.

7.6 Time-Varying Adjustment Coefficient as a Robustness Check

In response to the possibility that the adjustment speed may not be constant over the full sample, we also estimate an extended specification with a time-varying adjustment coefficient:

$$\frac{du}{dt} + a(t)u(t) = b_0 + b_1t + b_2t^2. \quad (42)$$

To keep the extension parsimonious, we specify the adjustment coefficient as

$$a(t) = \exp(\gamma_0 + \gamma_1 z_t), \quad (43)$$

where z_t denotes normalized time. This functional form guarantees that $a(t)$ remains positive, which preserves the stabilizing adjustment interpretation.

The estimated time-varying model gives

$$\hat{\gamma}_0 = 0.4196, \quad \hat{\gamma}_1 = -0.0850.$$

The negative value of $\hat{\gamma}_1$ implies that the estimated adjustment coefficient declines gradually over time. This suggests that the speed of adjustment may have become lower in later years of the sample.

However, the improvement in fit is very small. The RMSE decreases only from 0.1075 in the constant- a quadratic model to 0.1064 in the time-varying model, while AIC and BIC worsen after penalizing the additional parameter. Therefore, the time-varying model is useful as a robustness check, but it does not provide enough evidence to replace the simpler constant- a quadratic specification as the preferred baseline.

Table 3: Comparison of constant- a and time-varying $a(t)$ quadratic models

Model	RMSE	R^2	AIC	BIC
Constant- a quadratic model	0.1075	0.9969	-43.5351	-36.0526
Time-varying $a(t)$ quadratic model	0.1064	0.9969	-42.2329	-33.2538

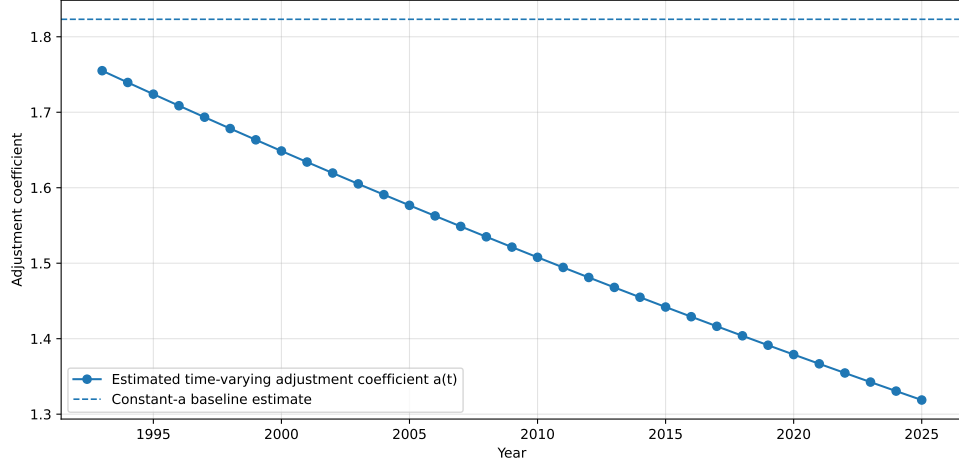


Figure 5: Estimated time-varying adjustment coefficient $a(t)$

7.7 Interpretation of the Preferred Model

Among the three candidate specifications, the quadratic model provides the best balance between fit, parsimony, and interpretability. Its estimated adjustment parameter is positive, with

$$\hat{a} = 1.8232.$$

In the baseline model, this value should be interpreted as the average adjustment speed over the full sample period. Since $\hat{a} > 0$, the transient component of the solution decays over time, which is consistent with a stabilizing adjustment mechanism. In other words, deviations from the fitted deterministic path do not persist indefinitely, but gradually become less important.

At the same time, the constant- a assumption should not be interpreted as saying that Kazakhstan's economy adjusted at exactly the same speed in every year. The time-varying robustness check suggests that the adjustment coefficient may decline gradually over time. However, because the improvement in fit is very small and information criteria worsen, the constant- a quadratic model remains the preferred specification.

7.8 Model-Based Real GDP Extrapolations for 2026 to 2030

In this subsection, we use the preferred quadratic specification to generate out-of-sample extrapolations for Kazakhstan real GDP over the period from 2026 to 2030. Since the quadratic model provided the best balance between empirical fit and interpretability within the observed sample, it is the most appropriate specification for model-based extrapolation in the context of this capstone.

These extrapolated values should not be interpreted as structural forecasts driven by fully specified macroeconomic mechanisms. Rather, they should be understood as model-based extrapolations implied by the estimated continuous-time trajectory. Their main purpose is to illustrate how the fitted regression-based differential equation can be used to extend the historical path beyond the observed period and to examine the medium-term implications of the estimated dynamics.

Figure 6 presents the observed real GDP series together with the fitted historical path and the extrapolated values for 2026 to 2030. Table 4 reports the corresponding extrapolated real GDP levels.

The extrapolations suggest that, under the continuation of the estimated quadratic trend, Kazakhstan real GDP declines slightly from the 2025 observed level to 158.37 trillion tenge in 2026, after which it resumes an upward path and reaches 220.73 trillion tenge by 2030. This pattern is consistent with the shape implied by the preferred model. At the same time, these values should be interpreted with caution, since extrapolation uncertainty increases outside the observed sample and the model does not explicitly incorporate structural breaks, policy changes, or external shocks.

The extrapolated path is generated from the fitted model rather than forced to pass through the observed 2025 value. Therefore, the visual path should be interpreted as the continuation of the fitted trajectory, while the table reports the corresponding model-implied values.

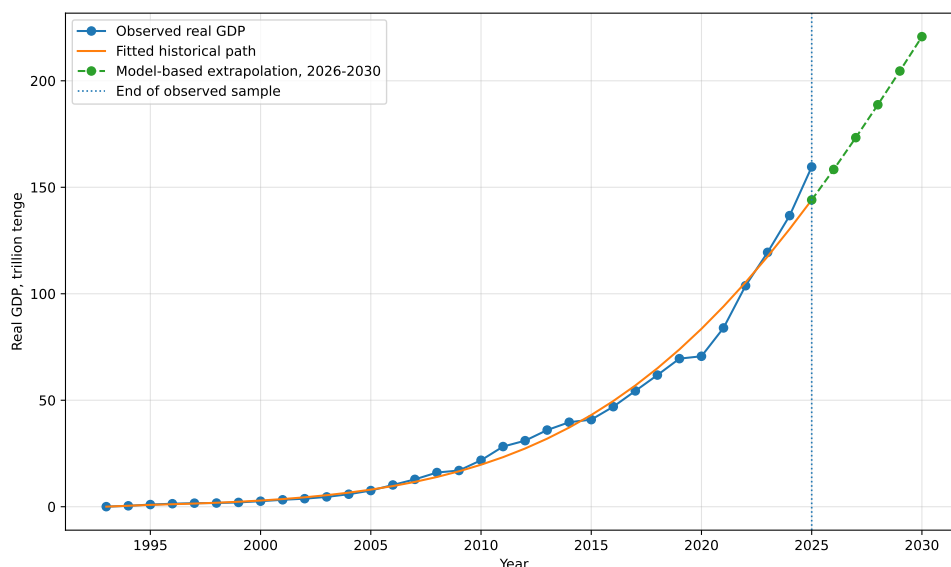


Figure 6: Observed, fitted, and extrapolated real GDP under the preferred quadratic model

Table 4: Model-based extrapolated real GDP for Kazakhstan, 2026 to 2030 under the preferred quadratic model

Year	Extrapolated real GDP (trillion tenge)
2026	158.37
2027	173.30
2028	188.74
2029	204.59
2030	220.73

7.9 Remarks on Extrapolation Limits

Although the preferred quadratic specification provides a useful extension of the historical real GDP path, the resulting extrapolations should be interpreted with caution. The model is designed to describe and estimate the dynamic trajectory of real GDP within a continuous-time framework, but it does not explicitly incorporate structural macroeconomic mechanisms, policy interventions, external demand shocks, or unexpected regime changes.

For this reason, the extrapolated values for 2026 to 2030 should be understood as extrapolations implied by the fitted model rather than as fully structural forecasts. Their main value lies in showing how the estimated continuous-time trajectory behaves beyond the observed sample and in illustrating the medium-term implications of the preferred specification.

As the extrapolation horizon becomes longer, uncertainty naturally increases. This is especially important in macroeconomic applications, where future outcomes may be affected by factors not captured by a deterministic polynomial forcing term. Therefore, the extrapolations reported in this capstone should be viewed as model-based scenarios conditional on the continuation of the historical dynamic pattern represented by the fitted quadratic model.

7.10 Summary of Empirical Findings

The empirical application yields four main conclusions. First, the regression-based differential equation model is capable of reproducing the historical path of Kazakhstan real GDP with a high degree of accuracy. Second, the flexibility of the deterministic forcing term matters substantially for model performance. Third, the quadratic specification emerges as the preferred model because it combines strong empirical fit with a coherent interpretation of the adjustment mechanism. Fourth, when the preferred specification is extended beyond the observed sample, it generates a model-based extrapolation in which real GDP declines slightly in 2026 and then resumes growth through 2030, reaching 220.73 trillion tenge by the end of the extrapolation horizon.

These findings support the broader argument of the capstone project: continuous-time differential equation models can provide a mathematically coherent, empirically interpretable, and practically useful framework for analyzing Kazakhstan real GDP. They can describe the historical trajectory well and also generate medium-term model-based extrapolations, provided that these are interpreted with appropriate caution.

8 Limitations and Possible Extensions

Several limitations should be noted. First, the model is reduced-form and descriptive rather than causal. It describes the trajectory of real GDP but does not identify the effects of specific macroeconomic variables such as investment, government spending, oil prices, exchange rates, or external demand.

Second, the deterministic polynomial forcing term captures smooth trend behavior but does not explicitly account for structural breaks, policy changes, or external shocks. This is important for Kazakhstan, where real GDP dynamics may be affected by commodity cycles, exchange-rate movements, global crises, and domestic policy changes.

Third, the baseline model treats the adjustment coefficient a as constant over the full sample period. This provides a parsimonious estimate of average adjustment speed, but it may hide variation across different macroeconomic regimes. The time-varying extension suggests that $a(t)$ may decline over time, although the improvement in model fit is not large enough to justify replacing the baseline model.

Fourth, the residual diagnostics indicate that some short-run autocorrelation remains. This suggests that the model captures the main deterministic trend but does not fully describe all short-run time-series dynamics.

Finally, the extrapolated values for 2026 to 2030 should not be interpreted as structural forecasts. They are conditional scenarios implied by the fitted trajectory under the assumption that the historical dynamic pattern continues. Future work could extend the model by adding stochastic error terms, estimating richer time-varying parameters, comparing against ARIMA or state-space benchmarks, and constructing confidence intervals around the extrapolated path.

9 Conclusion

This capstone project develops a continuous-time framework for modeling Kazakhstan real GDP using regression-based differential equations. The main goal was not to identify causal relationships, but to construct a mathematically correct, empirically interpretable, and practically useful dynamic model for real GDP.

The analysis showed that the proposed first-order differential equation with polynomial forcing is mathematically well posed and admits a closed-form solution. This makes

it suitable for connecting theoretical dynamic structure with discrete macroeconomic observations through nonlinear least squares estimation. The framework also provides interpretable parameters, especially through the adjustment coefficient and the deterministic forcing term. The estimated adjustment coefficient should be understood as an average adjustment speed in the baseline model, while the time-varying extension illustrates how this coefficient could change across the sample period.

In the empirical application to Kazakhstan real GDP, the model was able to reproduce the observed historical trajectory with high accuracy. Among the candidate specifications, the quadratic forcing model provided the best balance between fit and interpretability, suggesting that Kazakhstan real GDP follows a dynamic path with meaningful curvature over time. When extended beyond the observed sample, this preferred specification also generated medium-term model-based extrapolations for 2026 to 2030, illustrating how the fitted continuous-time model can be used not only for description but also for cautious forward-looking analysis. This result supports the use of flexible but tractable continuous-time models in macroeconomic analysis.

Overall, the project demonstrates that regression-based differential equation models can serve as a useful tool for studying Kazakhstan real GDP. They offer a coherent way to describe adjustment behavior, long-run patterns, and model-based extrapolations within a unified mathematical structure.

References

- [1] F. P. Ramsey, “A Mathematical Theory of Saving,” *The Economic Journal*, Vol. 38, No. 152, 1928, pp. 543–559.
- [2] R. M. Solow, “A Contribution to the Theory of Growth,” *The Quarterly Journal of Economics*, Vol. 70, No. 1, 1956, pp. 65–94.
- [3] R. J. Barro and X. Sala-i-Martin, *Economic Growth*, 2nd ed., MIT Press, Cambridge, MA, 2004.
- [4] G. Perone and M. A. Zambrano-Monserrate, “Forecasting of GDP Growth in the South Caucasian Countries Using Hybrid Ensemble Models,” *Econometrics*, Vol. 13, No. 3, 2025, Article 35.
- [5] A. Akhmet, “Forecasting Kazakhstan’s GDP Based on a Dynamic Factor Model with Regularization,” *National Bank of the Republic of Kazakhstan Economic Study*, No. 2026-01, 2026.
- [6] Y. H. Cheow, K. H. Ng, C. Phang, and K. H. Ng, “The Application of Fractional Calculus in Economic Growth Modelling: An Approach Based on Regression Analysis,” *Heliyon*, Vol. 10, 2024, e35379.

- [7] O. O. Awe, I. Crandell, A. A. Adepoju, and S. Leman, “A Time Varying Parameter State-Space Model for Analyzing Money Supply-Economic Growth Nexus,” *Journal of Statistical and Econometric Methods*, Vol. 4, No. 1, 2015, pp. 73–95.
- [8] B. J. Christensen, L. Neri, and J. C. Parra-Alvarez, “Estimation of Continuous-Time Linear DSGE Models from Discrete-Time Measurements,” Working paper, 2024.
- [9] W. E. Boyce, R. C. DiPrima, and D. B. Meade, *Elementary Differential Equations and Boundary Value Problems*, 11th ed., Wiley, Hoboken, NJ, 2020.
- [10] J. D. Hamilton, *Time Series Analysis*, Princeton University Press, Princeton, NJ, 1994.
- [11] A. C. Harvey, *Forecasting, Structural Time Series Models and the Kalman Filter*, Cambridge University Press, Cambridge, 1989.
- [12] J. Durbin and S. J. Koopman, *Time Series Analysis by State Space Methods*, 2nd ed., Oxford University Press, Oxford, 2012.
- [13] D. M. Bates and D. G. Watts, *Nonlinear Regression Analysis and Its Applications*, Wiley, New York, 1988.
- [14] World Bank, *World Development Indicators*, Kazakhstan real GDP, constant LCU, World Bank Data, accessed April 2026.