Modeling Threshold Effects with Kinked Regression: Methods, Applications, and Replicable Code

Provash Kumer Sarker[a]
[a]Bangladesh Bank, provash.econ@gmail.com

Abstract

Kinked regression is a powerful econometric method for capturing structural changes or threshold effects in the relationships among variables. This technique allows for slope shifts at a specific threshold—the kink point—while maintaining continuity in the dependent variable. Unlike standard linear regression, kinked regression accommodates piecewise linear relationships, making it particularly valuable for analyzing behavioral or policy responses at critical thresholds. Applications span diverse fields including labor economics, environmental science, finance, and policy evaluation. This document provides a comprehensive introduction to kinked regression methodology, distinguishes it from threshold regression, and offers practical implementation guides with replicable STATA and R code.

Keywords: Kinked regression; Threshold effects; Nonlinearity; Piecewise regression; STATA; R

1. Introduction

Economic relationships often shift when policy variables cross critical thresholds, leading to abrupt rather than smooth changes in behavior, making kinked regression an essential tool for capturing these nonlinear, threshold-driven dynamics. We provide two examples below.

A country introduces a carbon tax and expects emissions to decline steadily as the tax rate increases. However, after two years, the emissions data look confusing: small tax increases produce almost no change, while slightly larger increases produce rapid reductions. Policymakers notice an unusual pattern. When the carbon tax rises from 10 to 14 USD per ton, firms make only minor adjustments to their production methods. However, when the tax crosses roughly 18–20 USD, many firms suddenly switch to cleaner technologies, invest in energy-efficient machinery, or shift to renewable energy sources.

A central bank is examining how changes in the policy rate affect household borrowing. It expects a smooth, linear relationship: lower interest rates lead to more borrowing. However, post-crisis data tells a different story. For moderate rate cuts—say, from 5% to 4% or 4% to

https://doi.org/10.33774/coe-2025-rd36d **ORCID:** https://orcid.org/0000-0002-2767-7252 Content not peer-reviewed by Cambridge University Press. **License:** All Rights Reserved

3.5%—household loan demand barely changes. Households appear cautious, still worried about economic uncertainty. However, when the policy rate falls below a certain level—around 2%, borrowing suddenly accelerates. Mortgage applications jump, consumer loans rise, and credit card spending increases sharply. Economists suspect that households exhibit nonlinear sensitivity to monetary policy, reacting strongly only when rates fall below a psychological or liquidity threshold.

In both examples, policymakers observe a nonlinear response: emissions only begin to fall once the carbon tax crosses a certain threshold, and household borrowing reacts strongly only when interest rates fall below a critical level. A simple linear model cannot capture these behavioral tipping points. Allowing the slope of the emissions response to change after a threshold tax rate, or letting the interest-rate elasticity of borrowing shift after a cutoff point, is precisely why kinked regression is the appropriate econometric tool.

In real-world cases, many economic and environmental variables exhibit nonlinear behavior, particularly when key thresholds are crossed. Traditional linear regression models assume constant relationships across the entire range of independent variables, making them inadequate for capturing threshold dynamics. Consider these motivating questions:

- How do household saving rates respond when interest rates exceed inflation thresholds?
- Are central bank interventions less effective once inflation rates surpass target levels?
- Do environmental impacts accelerate beyond critical temperature or pollution thresholds?
- How do workers respond to changes in unemployment benefits at eligibility thresholds?

These questions illustrate situations in which the relationships between variables shift in intensity or direction at specific thresholds. Kinked regression provides a robust framework for modeling these threshold dynamics while maintaining continuity of relationships across critical points.

Kinked regression is a form of piecewise regression that models the relationship between an independent variable X and a dependent variable Y, where the slope changes at a specific point—the "kink." Unlike standard linear regression with uniform relationships, kinked regression allows for structural changes in slope at a threshold while ensuring continuity in the dependent variable.

This method is instrumental when relationships change at a particular value of X but remain linear on either side of that point. The approach captures both continuous relationships and nonlinear dynamics through piecewise linear approximations.

The kinked regression model has several important features that make it well-suited for analyzing threshold-driven relationships. It divides the data into two slope segments—one before and one after the kink point—allowing each segment to have its own rate of change. Despite this shift in slope, the relationship remains continuous at the kink, ensuring that the predicted values do not jump abruptly at the threshold. The kink point itself represents the critical level at which behavioral or structural changes occur, making the model particularly useful for identifying such thresholds. Overall, kinked regression offers greater flexibility than simple linear models when capturing complex, nonlinear responses.

The paper proceeds as follows. Section 2 presents an empirical review of the kinked regression. Sections 3, 4, and 5 provide the mathematical overview of modeling threshold effects using kinked regression. Section 6 illustrates the key difference between threshold and kinked regression, while Section 7 provides replication codes for Stata and R.

2. Literature Review: Applications for Kinked Regression

Kinked regression methods have been applied across diverse research domains, particularly through the regression kink design (RKD) framework developed for causal inference. This section reviews key applications demonstrating the versatility and power of kinked regression in empirical research.

2.1 Labor Economics and Social Policy

Card et al. (2015) pioneered the generalized regression kink design to study how unemployment insurance (UI) benefits affect unemployment duration in Austria. They exploited a kink in the benefit schedule where the replacement rate changes discontinuously at certain earnings thresholds. Their analysis revealed that higher UI benefits significantly increase unemployment duration, with an elasticity of approximately 0.3-0.5. This seminal work established RKD as a credible identification strategy for causal inference when treatment assignment has a discontinuous derivative but a continuous level at a threshold.

The methodology addresses a fundamental challenge in policy evaluation: estimating causal effects when policy rules create smooth changes in treatment intensity rather than sharp discontinuities. Card et al. (2012) developed the theoretical foundations, showing that under certain assumptions—including smoothness of the conditional expectation function and density of the assignment variable—the RKD can identify local treatment effects at the kink point.

Manoli and Weber (2016) applied kinked regression to analyze how pension wealth affects retirement timing. Using administrative data from Austria, they exploited kinks in the pension accrual schedule created by policy reforms. Their findings indicate that individuals respond to financial incentives embedded in pension rules, with elasticities varying by demographic characteristics and labor market conditions.

2.2 Public Finance and Taxation

Saez (2010) and subsequent researchers have used kinked regression to estimate elasticities of taxable income by exploiting kinks in progressive tax schedules. When marginal tax rates change at specific income thresholds, individual responses reveal behavioral elasticities. These studies typically find that high-income earners exhibit greater responsiveness to tax incentives, with implications for optimal tax design and revenue forecasting.

Devereux et al. (2014) employed kinked regression to study how corporate tax schedules affect firm investment decisions. They found that firms bunch at tax kink points and that investment responds to marginal effective tax rates, with elasticities higher for smaller and younger firms facing greater financial constraints.

2.3 Environmental and Agricultural Economics

Moeltner et al. (2021) developed a Bayesian kinked regression framework to test the von Liebig hypothesis in agricultural production—the theory that crop yields increase linearly with nutrient inputs until a critical threshold, beyond which additional inputs provide diminishing returns. Using data on corn yields and nitrogen fertilizer application, they found strong evidence of kinked relationships, with kink points varying by soil type and weather conditions. This approach helps farmers optimize input use while minimizing environmental externalities from excess fertilizer application.

Schlenker and Roberts (2009) employed nonlinear methods, including kinked regression, to analyze the effects of temperature on crop yields. They identified critical temperature thresholds beyond which yields decline rapidly, with the kink points varying by crop type. Their findings suggest that climate change impacts on agriculture may be substantially more

https://doi.org/10.33774/coe-2025-rd36d **ORCID:** https://orcid.org/0000-0002-2767-7252

severe than linear models predict, particularly if temperatures frequently exceed critical thresholds.

Koundouri and Xepapadeas (2004) applied kinked regression to groundwater extraction, where pumping costs increase gradually until aquifer levels fall below critical thresholds, after which costs accelerate. This application demonstrates how kinked regression captures regime changes in natural resource economics.

2.4 Health Economics and Policy

Einav et al. (2013) used kinked regression designs to estimate the extent of moral hazard in health insurance. They exploited kinks in cost-sharing schedules (deductibles and out-of-pocket maximums) to identify how insurance generosity affects healthcare utilization. Their results show significant behavioral responses to nonlinear cost-sharing, with implications for insurance design.

Alpert et al. (2015) applied RKD to study prescription drug demand, exploiting kinks in Medicare Part D benefit schedules. They found substantial price sensitivity among elderly patients, particularly at the coverage gap threshold, with important implications for medication adherence and health outcomes.

2.5 Financial Economics and Corporate Finance

Karlan and Zinman (2019) employed kinked regression to analyze how interest rates affect borrowing and default behavior in consumer credit markets. Using data from a South African lender with kinked pricing schedules, they identified separate elasticities for credit demand and default risk, informing optimal lending policies.

Researchers have used kinked regression to study how stock option exercise prices create kinks in executive pay-performance relationships. These studies reveal how compensation structures affect managerial effort and risk-taking behavior.

2.6 Methodological Extensions

Chiang and Sasaki (2019) extended the RKD framework to quantile regression, allowing researchers to estimate treatment effects at different points of the outcome distribution. This innovation is particularly valuable when policy effects are heterogeneous across the population or when outcomes have thick tails. Their method has been applied to study how unemployment benefits affect the distribution of unemployment duration, revealing that effects are stronger for workers at higher quantiles.

Several researchers have developed fuzzy RKD methods for situations in which the treatment kink is not perfectly sharp—analogous to fuzzy regression discontinuity designs. These methods use instrumental variables approaches to handle measurement error or imperfect compliance with policy rules.

Recent work has addressed estimation with multiple kinks, common in progressive tax systems, tiered benefit schedules, and regulatory frameworks. These extensions allow researchers to estimate multiple local treatment effects and test for consistency across kink points.

The literature demonstrates that kinked regression methods are particularly valuable in three contexts: (1) evaluating policies with nonlinear treatment rules, (2) identifying behavioral responses to economic incentives, and (3) modeling threshold effects in natural and social systems. The method's strength lies in its ability to provide credible causal estimates while accommodating complex, nonlinear relationships that characterize many real-world phenomena.

3. Mathematical Framework

3.1 Model Specification

The kinked regression model can be represented as:

$$Y = \beta_0 + \beta_1 X + \beta_2 (X - Kink)_+ + \varepsilon \qquad (1)$$

where:

- Y is the dependent variable
- X is the independent variable
- β_0 is the intercept
- β_1 is the slope before the kink point
- β_2 is the additional slope change after the kink point
- $(X - Kink)_+$ is the positive part that activates when X exceeds the kink
- ε is the error term

3.2 The Positive Part Function

The positive part function is defined as:

$$(X - Kink)_+ = \max(X - Kink, 0) \qquad (2)$$

This function equals zero when $X \leq$ Kink and equals (X − Kink) when X > Kink, ensuring the model activates the slope change only after the threshold is crossed.

3.3 Interpretation

Before the kink point

When $X \leq$ Kink, the term $(X - \text{Kink})_+$ equals zero, reducing the model to.

$$Y = \beta_0 + \beta_1 X + \varepsilon$$

After the kink point

When X > Kink, the full equation becomes active.

$$Y = \beta_0 + \beta_1 X + \beta_2 (X - Kink) + \varepsilon \qquad (3)$$

Key Insight: Below the kink point, the slope is β_1. After the kink point, the slope becomes $\beta_1 + \beta_2$, indicating an accelerated (if $\beta_2 > 0$) or decelerated (if $\beta_2 < 0$) rate of change.

4. Pedagogical Examples

4.1 Example 1: Price Elasticity of Demand

Let us consider the relationship between coffee price and quantity demanded. Economic theory suggests demand decreases with price, but becomes more elastic (sensitive) above a specific price threshold.

Scenario:

- Below \$1.00 per cup: Demand decreases slowly as price increases
- Above \$1.00 per cup: Demand decreases rapidly as price increases

Model Specification:

$$Q\_Demanded = \beta_0 + \beta_1 \times Price + \beta_2 \times (Price - 1.00)_+ + \varepsilon$$

Here, β_1 represents the rate of demand decrease when price < \$1.00, while β_2 captures the additional sensitivity (steeper slope) when price > \$1.00.

4.2 Example 2: Income and Savings Behavior

Household savings behavior often exhibits threshold effects: as income rises, savings initially increase modestly, but once a threshold is reached, households save a much higher proportion of additional income.

https://doi.org/10.33774/coe-2025-rd36d **ORCID:** https://orcid.org/0000-0002-2767-7252 Content not peer-reviewed by Cambridge University Press. **License:** All Rights Reserved

Scenario:

- Below $50,000 annual income: Modest marginal propensity to save
- Above $50,000 annual income: Higher marginal propensity to save

Model Specification:

$$Savings = \beta_0 + \beta_1 \times Income + \beta_2 \times (Income - 50{,}000)_+ + \varepsilon$$

Below the kink ($50,000), the savings rate follows slope β_1. Above the kink, the effective slope becomes $\beta_1 + \beta_2$, reflecting increased savings behavior.

5. Applications in Climate Econometrics

Climate change economics often involves nonlinear relationships that shift dramatically once tipping points are reached. Kinked regression effectively models relationships between environmental variables (temperature, CO_2 concentration) and outcomes (sea level rise, agricultural yield, biodiversity loss) that change significantly beyond critical thresholds.

5.1 Global Temperature and Sea Level Rise

The relationship between global temperature and sea level rise is nonlinear, with accelerating impacts beyond critical warming thresholds.

Scenario:

- Below 2°C above pre-industrial levels: Gradual, predictable sea level rise
- Above 2°C threshold: Accelerating rise due to ice sheet melting and thermal expansion

Model Specification:

$$Sea\_Level\_Rise = \beta_0 + \beta_1 \times Temperature + \beta_2 \times (Temperature - 2)_+ + \varepsilon$$

This specification captures both gradual baseline warming effects (β_1) and accelerated impacts beyond the 2°C threshold (β_2).

6. Kinked Regression versus Threshold Regression

While both methods model relationships that change at specific points, they differ fundamentally in how changes are represented and the nature of the transition.

6.2 Key Differences

Feature	Kinked Regression	Threshold Regression
Change Type	Continuous slope change at kink point	Abrupt change with potential discontinuity
Mathematical Form	Continuous function with slope change	Different functional forms before/after threshold
Application	Rate of change shifts but remains continuous	Drastic behavioral change after crossing the threshold
Examples	Temperature & sea level; CO_2 & crop yield; progressive taxation	Policy activation at pollution levels; ecosystem regime shifts

7. Implementation: STATA and R Code

This section provides replicable code for implementing kinked regression in both STATA and R. The examples use a kink point at X = 5 for illustration.

7.1 STATA Implementation

```
* Set sample size
set obs 100

* Generate independent variable
gen x = runiform(0, 10)

* Generate dependent variable with kink at x = 5
gen y = cond(x <= 5, 2 + 1.5*x, 2 + 1.5*5 + 0.5*(x-5)) + rnormal(0,1)

* Visualize the data
scatter y x, msize(small) title("Data for Kinked Regression")

* Create kinked variable (activates when x > 5)
gen kink = (x > 5) * (x - 5)

* Estimate kinked regression model
regress y x kink

* Estimate kink point using nonlinear regression
nl (y = {b0} + {b1}*x + {b2}*cond(x > {psi}, x - {psi}, 0))

* Display estimated kink point
nlcom (Kink_Point = {psi})
```

7.2 R Implementation

```
# Load required package
install.packages("segmented")
library(segmented)

# Set seed for reproducibility
set.seed(123)
n <- 100

# Generate data with kink at x = 5
x <- runif(n, 0, 10)
y <- ifelse(x <= 5,
        2 + 1.5*x + rnorm(n, 0, 1),
        2 + 1.5*5 + 0.5*(x-5) + rnorm(n, 0, 1))

data <- data.frame(x, y)

# Fit initial linear model
lm_model <- lm(y ~ x, data = data)

# Apply segmented regression to estimate kink point
seg_model <- segmented(lm_model, seg.Z = ~x, psi = 5)

# Display results
summary(seg_model)

# Plot results
plot(x, y, main = "Kinked Regression", xlab = "X", ylab = "Y")
plot(seg_model, add = TRUE, col = "red", lwd = 2)
```

References

Alpert, A., Hsi, H. J., & Jacobson, M. (2015). Evaluating the role of payment policy in driving vertical integration in the oncology market. *Health Affairs*, 34(9), 1559-1566.

Card, D., Lee, D. S., Pei, Z., & Weber, A. (2015). Inference on Causal Effects in a Generalized Regression Kink Design. *Econometrica*, 83(6), 2453-2483. https://doi.org/10.3982/ECTA11224

Card, D., Lee, D., Pei, Z., & Weber, A. (2012). Nonlinear policy rules and the identification and estimation of causal effects in a generalized regression kink design (No. w18564). *National Bureau of Economic Research*.

Chiang, H. D., & Sasaki, Y. (2019). Causal inference by quantile regression kink designs. *Journal of Econometrics*, 210(2), 405-433. https://doi.org/10.1016/j.jeconom.2019.02.005

Devereux, M. P., Liu, L., & Loretz, S. (2014). The elasticity of corporate taxable income: New evidence from UK tax records. *American Economic Journal: Economic Policy*, 6(2), 19-53.

Einav, L., Finkelstein, A., Ryan, S. P., Schrimpf, P., & Cullen, M. R. (2013). Selection on moral hazard in health insurance. *American Economic Review*, 103(1), 178-219.

Karlan, D., & Zinman, J. (2019). Price and control elasticities of demand for savings. *Journal of Development Economics*, 137, 105-122.

Koundouri, P., & Xepapadeas, A. (2004). Estimating accounting prices for common pool natural resources: A distance function approach. *Water Resources Research*, 40(6).

Manoli, D., & Weber, A. (2016). Nonparametric evidence on the effects of financial incentives on retirement decisions. *American Economic Journal: Economic Policy*, 8(4), 160-182.

Moeltner, K., Ramsey, A. F., & Neill, C. L. (2021). Bayesian Kinked Regression with Unobserved Thresholds: An Application to the von Liebig Hypothesis. *American Journal of Agricultural Economics*, 103(5), 1832-1856. https://doi.org/10.1111/ajae.12185

Saez, E. (2010). Do taxpayers bunch at kink points? *American Economic Journal: Economic Policy*, 2(3), 180-212.

Schlenker, W., & Roberts, M. J. (2009). Nonlinear temperature effects indicate severe damages to US crop yields under climate change. *Proceedings of the National Academy of Sciences*, 106(37), 15594-15598