THE EVOLUTION OF THE NOMINAL EFFECTIVE EXCHANGE RATE

Companion Paper

Adrien Currat	Isaac Graber	Mathéo Bourgeois
HEC Lausanne	*HEC Lausanne*	*HEC Lausanne*
Msc. in Economics	*Msc. in Economics*	*Msc. in Economics*

January 12, 2025

1 Introduction

This companion paper provides all the mathematical proofs to support the code used for the project. It follows the same structure as the code, ensuring a clear and logical alignment between the theoretical and practical components. For each section, the paper provides detailed supplementary information, offering deeper insights into the theoretical foundations of the concepts and tools leveraged in the project.

2 Data Treatment

2.1 Denton-Cholette method

The Denton-Cholette method aims to minimize the following optimization problem:

$$\min_{\{x_t\}_{t=1}^T} \sum_{t=2}^{T} \left(\frac{x_t}{X_t} - \frac{x_{t-1}}{X_{t-1}} \right)^2 ,$$

subject to:

$$\frac{1}{3}\sum_{m=1}^{3} x_{q,m} = Q_q, \quad \forall q,$$

where:

- x_t: The interpolated monthly series to be estimated.
- X_t: The monthly indicator series (2010–2024), guiding the proportional movements of x_t.
- Q_q: The original quarterly series (1999–2023), providing the constraint for the quarterly averages.
- $x_{q,m}$: The interpolated monthly values within quarter q (e.g., $m = 1, 2, 3$ for months within the quarter).

Details

- **Objective Function:** The term $\left(\frac{x_t}{X_t} - \frac{x_{t-1}}{X_{t-1}}\right)^2$ minimizes the proportional differences between consecutive months, ensuring that the interpolated series x_t follows the proportional pattern of the indicator series X_t.
- **Constraint:** The constraint ensures that the average of the interpolated monthly values within each quarter matches the corresponding quarterly value Q_q:

$$\frac{1}{3}\sum_{m=1}^{3} x_{q,m} = Q_q, \quad \forall q.$$

2.2 The Hodrick-Prescott (HP) filter

Given the monthly industrial production index y_t, the HP filter solves this minimization problem:

$$\min_{\{\tau_t\}} \sum_{t=1}^{T} (y_t - \tau_t)^2 + 14400 \sum_{t=2}^{T-1} \left((\tau_{t+1} - \tau_t) - (\tau_t - \tau_{t-1})\right)^2,$$

where:

- y_t: The observed monthly industrial production index.
- τ_t: The estimated smooth trend component.
- T: The total number of monthly observations.
- $\lambda = 14400$: The smoothing parameter for monthly data.

This problem aims to find the optimal trend component in a minimization problem by balancing two forces: the fit of the trend (τ_t) to the observed time series and the smoothness of the trend component by minimizing changes in its slope. The parameter λ determines the extent to which the trend is allowed to be volatile. For example, a very high λ results in a very smooth trend.

3 Model estimation

3.1 Estimation of the model's coefficients

$$y = X\beta + e$$

where:

- y is an $n \times 1$ vector of the dependent variable (ΔS_{t+1}, size: 120×1)
- e is an $n \times 1$ vector of residuals (errors, size: 120×1)
- X is an $n \times k$ matrix of predictors (120×3) including:
  - A column of ones (intercept)
  - Inflation differential (Swiss inflation minus foreign inflation)
  - Output gap differential (Swiss output gap minus foreign output gap)
- β is a $k \times 1$ vector of coefficients (3×1) containing the intercept, $\beta_{\text{inflation}}$, and $\beta_{\text{output gap}}$.

The residuals are expressed as:

$$e = y - X\beta.$$

Sum of Squared Residuals (SSR)

The goal of Ordinary Least Squares (OLS) is to minimize the sum of squared residuals:

$$e'e = (y - X\beta)'(y - X\beta).$$

$$e'e = y'y - 2\beta'X'y + \beta'X'X\beta,$$

Using the transpose rule $(AB)' = B'A'$, the transpose of $-X\beta$ becomes $-\beta'X'$. Furthermore, since $y'X\beta$ is a scalar, we have $y'X\beta = \beta'X'y$, as the transpose of a scalar does not change its value. .

First Derivative of SSR

To minimize $e'e$, take the derivative with respect to β:

$$\frac{\partial e'e}{\partial \beta} = \frac{\partial}{\partial \beta}\big(y'y - 2\beta'X'y + \beta'X'X\beta\big).$$

The derivatives of each term are:

- $\frac{\partial y'y}{\partial \beta} = 0$ (constant term),
- $\frac{\partial(-2\beta'X'y)}{\partial \beta} = -2X'y$,
- $\frac{\partial(\beta'X'X\beta)}{\partial \beta} = 2X'X\beta$ (using $\frac{\partial(\beta'A\beta)}{\partial \beta} = 2A\beta$ for symmetric A).

Combining:

$$\frac{\partial e'e}{\partial \beta} = -2X'y + 2X'X\beta.$$

Setting this derivative to zero:

$$-2X'y + 2X'X\beta = 0 \implies X'y = X'X\beta.$$

To solve for β, the vector of OLS estimators, multiply both sides by $(X'X)^{-1}$ (assuming $X'X$ is invertible). You get the expression implemented in the code for the OLS estimators:

$$\beta = (X'X)^{-1}X'y.$$

Second Derivative and Minimum Check

To confirm this is a minimum, compute the second derivative:

$$\frac{\partial^2 e'e}{\partial \beta^2} = 2X'X.$$

Since $X'X$ is symmetric and positive definite (if X has full rank), the second derivative is positive, confirming a minimum.

The OLS estimator $\beta = (X'X)^{-1}X'y$ minimizes $e'e$ if X is full rank, which can be verified by checking $\det(X'X) \neq 0$. This ensures a unique solution for β.

4 Point Forecasting

4.1 Optimal forecasts

$$\Delta S_{t+1} = \alpha_1 + \beta_\pi^1 \pi_t + \beta_y^1 y_t + \eta_{t+1},$$

where:

$$\pi_t = \sum_{j=1}^{n} w_j(\pi_t - \tilde{\pi}_t^j), \quad y_t = \sum_{j=1}^{n} w_j(y_t - \tilde{y}_t^j),$$

and w_j are trade shares. Here, β_π^1 is associated with the differential of inflation in t, and β_y^1 is associated with the differential of the output gap in t minus its differential in $t-1$, as the output gap is first-differenced.

Conditional on the information set Ω_t, the optimal direct forecast for the one-month model is:

$$\mathbb{E}[\Delta S_{t+1}|\Omega_t] = \alpha_1 + \beta_\pi^1 \pi_t + \beta_y^1 y_t,$$

since $\eta_{t+1} \sim \mathcal{N}(0, \sigma^2)$ and $\mathbb{E}[\eta_{t+1}|\Omega_t] = 0$.

Thus:

$$\mathbb{E}[\Delta S_{t+1}|\Omega_t] = \alpha_1 + \beta_\pi^1 \pi_t + \beta_y^1 y_t.$$

The three-month model is given by:

$$\Delta S_{t+3} = \alpha_3 + \beta_\pi^3 \pi_t + \beta_y^3 y_t + \eta_{t+3}.$$

Conditional on Ω_t:

$$\mathbb{E}[\Delta S_{t+3}|\Omega_t] = \alpha_3 + \beta_\pi^3 \pi_t + \beta_y^3 y_t,$$

Same Process for 6, 12, and 18 Months

$$\mathbb{E}[\Delta S_{t+6}|\Omega_t] = \alpha_6 + \beta_\pi^6 \pi_t + \beta_y^6 y_t,$$
$$\mathbb{E}[\Delta S_{t+12}|\Omega_t] = \alpha_{12} + \beta_\pi^{12} \pi_t + \beta_y^{12} y_t,$$
$$\mathbb{E}[\Delta S_{t+18}|\Omega_t] = \alpha_{18} + \beta_\pi^{18} \pi_t + \beta_y^{18} y_t.$$

5 Forecast evaluation

5.1 Diebold-Mariano Test

The Diebold-Mariano (DM) test evaluates whether the forecast accuracy of two models differs significantly. In this implementation, the DM test is conducted as a one-sided test to determine if the model outperforms the benchmark.

Mathematical Formulation

Loss Differential:

$$d_t = \ell(e_{t,\text{model}}) - \ell(e_{t,\text{benchmark}}), \quad \ell(x) = x^2,$$

where $e_{t,\text{model}}$ and $e_{t,\text{benchmark}}$ are the forecast errors of the model and benchmark, respectively.

The mean loss differential is:

$$\bar{d} = \frac{1}{n} \sum_{t=1}^{n} d_t,$$

where n is the number of forecast periods.

Newey-West HAC Variance Estimator: The HAC-adjusted variance accounts for serial correlation:

$$\hat{\sigma}^2_{\text{HAC}} = \frac{1}{n} \sum_{t=1}^{n} (d_t - \bar{d})^2 + 2 \sum_{k=1}^{\ell} \left(1 - \frac{k}{\ell + 1}\right) \gamma_k,$$

where:

$$\gamma_k = \frac{1}{n} \sum_{t=k+1}^{n} (d_t - \bar{d})(d_{t-k} - \bar{d}),$$

and $\ell = \lfloor 1.5 n^{1/3} \rfloor$ is the optimal lag length.

DM Test Statistic:

$$DM = \frac{\bar{d}}{\sqrt{\frac{\hat{\sigma}^2_{\text{HAC}}}{n}}}, \quad DM \sim N(0,1).$$

Hypotheses

- H_0: $\mathbb{E}[d_t] \geq 0$ (benchmark performs at least as well as the model),
- H_1: $\mathbb{E}[d_t] < 0$ (model outperforms the benchmark).

The p-value for the one-sided DM test is:

$$p = 1 - \Phi(DM),$$

where $\Phi(\cdot)$ is the cumulative distribution function of the standard normal distribution.

Reject H_0 if $p < 0.05$, indicating that the model significantly outperforms the benchmark in terms of forecast accuracy.

5.2 Forecast Efficiency Test

The test evaluates whether the forecast errors (η_{t+1}) are uncorrelated with the predictors:

$$\sum_{j=1}^{n} w_j(\pi_t - \tilde{\pi}_t^j) \quad \text{and} \quad \sum_{j=1}^{n} w_j(y_t - \tilde{y}_t^j).$$

Regression Model

The regression is performed as:

$$\eta_{t+1} = \beta_0 + \beta_1 \sum_{j=1}^{n} w_j(\pi_t - \tilde{\pi}_t^j) + \beta_2 \sum_{j=1}^{n} w_j(y_t - \tilde{y}_t^j) + \epsilon_t,$$

where β_0 (intercept) is included in the regression.

Joint Hypothesis Test

The code tests:

$$H_0 : \beta_0 = \beta_1 = \beta_2 = 0$$

using the F-statistic:

$$F = \frac{\hat{\beta}^\top \hat{\Sigma}^{-1} \hat{\beta}}{k}, \quad p = 1 - F_F(F, k, T-k),$$

where:

- $\hat{\beta} = [\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2]$: Estimated coefficients (including the intercept).
- $\hat{\Sigma}$: HAC (Newey-West) covariance matrix.
- $k = 3$: Total number of coefficients tested (intercept + predictors).
- T: Number of observations.

Distribution

The test statistic follows an F-distribution with k and $T-k$ degrees of freedom.

Interpretation

- $p < 0.05$: Reject H_0; forecast errors are correlated with predictors (inefficient). - $p \geq 0.05$: Fail to reject H_0; forecast errors are uncorrelated with predictors (efficient).

5.3 Forecast Unbiasedness Test

The test evaluates whether forecasts are unbiased by testing if the mean forecast error (τ) is zero.

Hypothesis Test

Null Hypothesis (H_0): $\tau = 0$.

Mean Forecast Error ($\hat{\tau}$):

$$\hat{\tau} = \frac{1}{n} \sum_{t=1}^{n} \eta_{t+1},$$

where $\eta_{t+1} = Y_{t+1} - \hat{Y}_{t+1}$ represents the forecast errors.

Test Statistic (t):

$$t = \frac{\hat{\tau}}{\mathrm{SE}(\hat{\tau})}, \quad p = 2 \cdot (1 - F_t(|t|, \mathrm{df})),$$

with:

- $\hat{\tau}$: Sample mean forecast error.
- $\mathrm{SE}(\hat{\tau}) = \sqrt{\frac{1}{n^2}\sum_{t=1}^{n}(\eta_{t+1} - \hat{\tau})^2}$: Robust standard error of $\hat{\tau}$.
- n: Number of observations.
- $\mathrm{df} = n - 1$: Degrees of freedom.

Interpretation

$p < 0.05$: Reject H_0; forecasts are biased ($\hat{\tau} \neq 0$). $p \geq 0.05$: Fail to reject H_0; forecasts are unbiased ($\hat{\tau} = 0$).

6 Density Forecast

6.1 Prediction interval evaluation and Likelyhood ratio tests

Prediction intervals are computed as:

$$[\mathbb{E}[\Delta S_{t+h}|\Omega_t] - z_\alpha \cdot \sqrt{\mathrm{Var}(\Delta S_{t+h})}, \mathbb{E}[\Delta S_{t+h}|\Omega_t] + z_\alpha \cdot \sqrt{\mathrm{Var}(\Delta S_{t+h})}],$$

where z_α is the critical value (e.g., 80%, 90%, 95%), and according to section 6.2:

$$\mathrm{Var}(\Delta S_{t+h}) = \sigma^2 + \beta_\pi^2 \mathrm{Var}(\pi_t) + \beta_y^2 \mathrm{Var}(y_t) + 2\beta_\pi\beta_y \mathrm{Cov}(\pi_t, y_t).$$

Violation rates (α_{obs}) are calculated as:

$$\alpha_{\mathrm{obs}} = \frac{\text{Number of violations}}{\text{Total observations}},$$

where violations occur when ΔS_{t+h} lies outside the interval.

Then, we test the null hypothesis for each confidence interval, such as $\alpha_{\mathrm{obs}} = \alpha_{\mathrm{exp}} = 0.05$ for the 95% CI, using the likelihood ratio (LR) test:

$$\mathrm{LR} = 2\left(n_0 \ln\left(\frac{\alpha_{\mathrm{obs}}}{\alpha_{\mathrm{exp}}}\right) + n_1 \ln\left(\frac{1-\alpha_{\mathrm{obs}}}{1-\alpha_{\mathrm{exp}}}\right)\right),$$

where:

- n_0: Number of violations ($y_t \notin [\text{Lower}, \text{Upper}]$),
- n_1: Number of non-violations ($y_t \in [\text{Lower}, \text{Upper}]$).

The p-value is computed as:

$$p = P(\chi^2_1 > \text{LR}),$$

and we reject H_0 if $p < 0.05$.

6.2 Probability Integral Transform (PIT) Analysis

For the one-month model, for each forecast points:

$$\text{PIT}_{t+1} = \Phi\left(\frac{\Delta S_{t+1} - \mathbb{E}[\Delta S_{t+1}|\Omega_t]}{\sqrt{\text{Var}(\Delta S_{t+1})}}\right),$$

where Φ is the standard normal cumulative distribution function (`pnorm`).

Variance of the forecasted one-month change in the NEER

The variance of the one-month forecast is:

$$\text{Var}(\Delta S_{t+1}) = \text{Var}(\mathbb{E}[\Delta S_{t+1}|\Omega_t] + \eta_{t+1}),$$

where:

$$\text{Var}(\Delta S_{t+1}) = \text{Var}(\mathbb{E}[\Delta S_{t+1}|\Omega_t]) + \text{Var}(\eta_{t+1}),$$

since η_{t+1} and Ω_t are independent.

Expanding $\mathbb{E}[\Delta S_{t+1}|\Omega_t]$ according to section 4.1:

$$\mathbb{E}[\Delta S_{t+1}|\Omega_t] = \alpha_1 + \beta^1_\pi \pi_t + \beta^1_y y_t,$$

so:

$$\text{Var}(\mathbb{E}[\Delta S_{t+1}|\Omega_t]) = \text{Var}(\beta^1_\pi \pi_t + \beta^1_y y_t).$$

Using the variance properties:

$$\text{Var}(aX + bY) = a^2\text{Var}(X) + b^2\text{Var}(Y) + 2ab\text{Cov}(X, Y),$$

we get:

$$\text{Var}(\beta_\pi^1 \pi_t + \beta_y^1 y_t) = (\beta_\pi^1)^2 \text{Var}(\pi_t) + (\beta_y^1)^2 \text{Var}(y_t) + 2\beta_\pi^1 \beta_y^1 \text{Cov}(\pi_t, y_t).$$

Substituting back into the variance of ΔS_{t+1}:

$$\text{Var}(\Delta S_{t+1}) = \sigma^2 + (\beta_\pi^1)^2 \text{Var}(\pi_t) + (\beta_y^1)^2 \text{Var}(y_t) + 2\beta_\pi^1 \beta_y^1 \text{Cov}(\pi_t, y_t),$$

where $\sigma^2 = \text{Var}(\eta_{t+1})$.

Thus, the final expression for the variance is:

$$\text{Var}(\Delta S_{t+1}) = \sigma^2 + (\beta_\pi^1)^2 \text{Var}(\pi_t) + (\beta_y^1)^2 \text{Var}(y_t) + 2\beta_\pi^1 \beta_y^1 \text{Cov}(\pi_t, y_t).$$