MEMORANDUM

No 05/2026
October 2026

The Nominal Exchange Rate Regression – A critical appraisal of methodology and empirical results



Jennifer Castle and Ragnar Nymoen

ISSN: 0809-8832

Department of Economics
University of Oslo

This series is published by the
University of Oslo
Department of Economics

P. O.Box 1095 Blindern
N-0317 OSLO Norway
Telephone: + 47 22855127
Fax: + 47 22855035
Internet: http://www.sv.uio.no/econ
e-mail: post@econ.uio.no

In co-operation with
The Frisch Centre for Economic Research

Gaustadalleén 21
N-0371 OSLO Norway
Telephone: +47 22 95 88 20
Fax: +47 22 95 88 25
Internet: http://www.frisch.uio.no
e-mail: frisch@frisch.uio.no

Last 10 Memoranda

No 04/2026	Anders Harildstad, Ragnar Nymoen, and Olav Slettebø: From Manufacturing Productivity to National Wages: The Main-Course Variable and Norwegian Wage Formation, 1900–2025
No 03/2026	Andreas Kotsadam, Knut Løyland, Kjetil R. Madland, Oddbjørn Raaum, Gaute Torsvik and Arnstein Øvrum: The Effect of Punishment on Tax Compliance: Evidence from a Randomized Experiment
No 02/2026	Andreas Makoto Fukuda Andersen, Andreas Kotsadam, Knut Løyland, Gaute Torsvik, Oddbjørn Raaum and Arnstein Øvrum: The Limits of Low-Intensity VAT Enforcement: Experimental Evidence from Norway
No 01/2026	Ragnhild Schreiner, Gaute Torsvik and Gjermund Sterri Yndesdal: Who Bears the Burden of Payroll Taxes at the Top?
No 03/2025	Edwin Leuven and Ingrid Mikkelsen Semb: Entrepreneurial returns to college
No 02/2025	Zachary Grossman, Tony Hua, Jo Thori Lind and Karine Nyborg Unwillingly Informed: the Prosocial Impact of Third-Party Informers
No 01/2025	The mixed effects of making contribution behaviors observable Noel Z. Hoven og Karine Nyborg
No 03/2024	Effects of Compressed Work Schedules on Sickness Absence Elisabeth Fevang, Andreas Fidjeland, Karen Hauge, and Otto Lillebø
No 02/2024	Tore Adam Reiremo The Anatomy of Mismatch in the College Market
No 01/2024	Gunnar Bårdsen and Ragnar Nymoen U.S. wage-price dynamics, before, during and after COVID-19, through the lens of an empirical econometric model

Previous issues of the memo-series are available in a PDF® format at:
https://www.sv.uio.no/econ/english/research/publications/memoranda/

The Nominal Exchange Rate Regression – A critical appraisal of methodology and empirical results*

Jennifer Castle[a] and Ragnar Nymoen[b]

[a]Climate Econometrics and Magdalen College, Oxford University
[b]Department of Economics, University of Oslo

30 October 2026

Abstract

Nominal exchange rate regressions, in which the h-period change in the nominal exchange rate (NER) is regressed on the h-th lag of the real exchange rate (RER), are used to assess whether the RER predicts future NER adjustment. We show that these regressions are generally mis-specified relative to a VAR-type data generating process, and that the reliability of inference about the RER's predictive power depends critically on the horizon, on whether the RER is stationary, and on the sample size. Monte Carlo simulations show that, when the RER is stationary, the regressions estimate horizon-dependent projection coefficients rather than the adjustment coefficient, and cannot reveal shifts in the adjustment mechanism. When the RER is non-stationary, conventional tests reject the null of no predictive power far too often. At long horizons the regression also becomes spurious in the sense of Granger and Newbold, and non-standard critical values do not solve the problem. Empirically, we revisit the finding of Eichenbaum, Johannsen and Rebelo (2021) that, under inflation targeting, the RER predicts NER changes at horizons of three years and longer. NER-regressions for Norway and the UK reproduce this finding on the original sample period. However, the estimates shrink as the sample is extended. For the UK, for example, the lagged RER becomes insignificant once the regression is re-specified to have near white-noise residuals. The empirical regularity is therefore less robust than it appears, which matters for the assessment of monetary policy models that have been calibrated to match it.

Keywords: Nominal exchange rate regression; Exchange rate predictability; Relative PPP; Cointegration; Equilibrium correction; Spurious regression; Direct projections; Inflation targeting.
JEL: C22, C32, F31, E52.

*Thanks to Farooq Akram, Pål Boug, Steinar Holden and Genaro Sucarrat for comments on an earlier version of this paper, presented at the *Econometrics and Economic Policy Workshop*, arranged by Statistics Norway and Centre for Applied and Theoretical Econometrics (CATE, BI), Oslo, 24 September 2026. Norwegian Research Council project number 324472.

1 Introduction

In an influential article, Eichenbaum, Johannsen and Rebelo (2021), EJR hereafter, studied how monetary policy affects the relative importance of nominal exchange rates (*NER*s) and relative prices in shaping the response of real exchange rates (*RER*s) to shocks. A main conclusion is that the empirical relationship between *NER* and *RER* is regime dependent, but that it is strong and viable in inflation targeting regimes.

In EJR, it is assumed that the *RER* is stationary and hence that it adjusts back to its mean after a shock via changes in the *NER* or changes in relative prices. Under inflation targeting, shocks to the *RER* are not counteracted by movements in relative prices (it would be inconsistent with the regime) and therefore the equilibrating mechanism must logically be movements in the *NER*. In countries with fixed rates the situation is different, where changes in the relative price can restore the *RER* to its equilibrium after a shock.

With the aid of two types of regression models, labelled *Nominal exchange rate regressions* and *Relative-price regressions*, EJR document two facts about *RER*s and *NER*s for inflation targeting countries. The first fact comes from the *NER*-regressions estimated on data from inflation targeting countries. It is that the current *RER* is highly negatively correlated with future changes in the *NER* at horizons greater than two years. This correlation is stronger the longer the horizon. The second fact is that the *RER* is virtually uncorrelated with future relative inflation rates at all horizons, and is based on relative price regressions on data sets of fixed and quasi-fixed exchange rate regimes.

The results of the *NER*-regression imply that in well functioning inflation targeting regimes, "the *RER* adjusts to shocks in the medium and long run overwhelmingly through changes in the *NER*, not through differential inflation rates".[1] Or do they? In this paper we show that there are weaknesses in the econometric methodology of *NER*-regressions that call for a certain scepticism towards the results that EJR present as regularities and facts about the temporal properties of exchange rates under inflation targeting regimes.

Taken at face value, EJR's findings indicate that adjustments of *NER* over "long time distances" entail that weak relative PPP is a tenable assumption. This contradicts the widely held belief that under floating exchange rate regimes, *NER*s and *RER*s have similar features, and that non-stationarity is one of them. For example, one conclusion in a contemporary survey of the literature is that "the assumption of long-run mean reversion in *RER* is not generally justified" and that "*RER* is, in general, nonstationary and in fact integrated", Itskhoki (2021, p. 448).

Nevertheless, EJR maintain that in the *NER*-regressions, the significance of the *RER* as a predictor of changes in *NER* can be assessed without taking into consideration the possibility that the *RER* is non-stationary (by the use of unit-root tests for example). Instead, they provide supporting evidence for their assumption from a forecast comparison exercise.

A main objective of this paper is to show that the *NER*-regressions are in general inefficient for the purpose of estimating the predictive power that the *RER* may possess for future changes in the *NER* (of inflation targeting regimes), in particular for horizons like 3 years and longer. We identify one constellation as especially dubious if standard inference is used for testing the predictive power of *RER*: It arises when the dependent variable is the *NER*s change over time periods like two or three years and when the *RER* is non-stationary. This variant of *NER*-regressions can, with plausibility, be interpreted as a spurious regression, in a sense which is close to the original meaning in Granger and

[1] Eichenbaum et al. (2021, p. 193)

Newbold (1974).

We develop our analysis of the challenges to correct inference that face users of *NER*-regression with the aid of Monte Carlo simulations. We formulate data generating processes (DGPs) that encompass the different interpretations of the *NER*-regression. It is a VAR that generates time series for *NER* and the relative price, both in logs. Both series are generated as non-stationary in the DGP. In one version of the DGP, the two series are cointegrated in such a way that the *RER* is stationary and weak relative PPP holds (as variants we consider the DGP with and without exogeneity of the relative price). In the other version there is no cointegration and (specifically) the RER is non-stationary.

We show that with these DGPs, the (only) efficient way to test the hypothesis of no-relationship is the *NER*-regression where the first difference of *NER* (i.e., the depreciation rate) is regressed on the first lag of the *RER*. Using longer horizons induces residual misspecification and complicates inference, in a way that is not fixed by the use of standard errors that are consistent against residual heteroskedasticity and autocorrelation.

Although the issue we raise is relevant for both stationary and non-stationary DGPs, it is most consequential in the cases where the *RER* is non-stationary. The Monte Carlo analysis indicates that the probability of Type-I error is larger than the chosen significance level, and that the error probability may increase with the length of the horizon (differencing of the *NER*). Hence, unless care is taken, the nominal exchange rate regression may be a issue regression in hiding. The critique is different from the caveat that EJR first make, and then dismiss based on the forecast exercise, about over-rejection due to persistence of *RER* and short real-world sample sizes. The Monte Carlo analysis we show that, for long horizons, the bias in the *NER*-regression coefficient is large numerically for realistic sample sizes, and that the bias is reduced only very slowly as the sample length is increased. Estimation bias is only eliminated at unfeasible sample lengths.

The EJR paper has been cited in academic journal articles and working papers, as well as in the wider, policy oriented, literature.[2] The empirical regularity seems to have been accepted – at least it has not been seriously questioned. Any criticism has been dispersed and about the interpretation of the empirical regularity, rather than discussion of status of the econometric evidence. Itskhoki and Mukhin (2021) grant that inflation stabilisation can account for the predictive power of long lags in the *RER*, but objects that the mechanism "produces counterfactual predictions for alternative measures of the real exchange rate"[3]. However, at the same time, they endorse the empirical fact: the model they present reproduces EJR's finding that the current real exchange rate predicts future nominal changes. So the critique is of the explanation, not the regularity, see also Itskhoki (2021, p. 449).[4]

In a paper which is focused on forecasting, Engel and Wu (2023) note, like we do, that in multi-period forecasting equations care must be taken with regard to the potential integrated property of the predictor variable under investigation, as well as the induced persistence of the dependent variable as the forecast horizon is increased. However, the Engel and Wu article studies models where the lagged *NER* is the regressor, hence different from EJR's nominal exchange rate regressions.

[2]Papers from e.g., ECB, Federal Reserve System, BIS.

[3]Itskhoki and Mukhin (2021, Note 5)

[4]In particular Note 2, where EJR's result about *RER* as a predictor is recorded together with other examples of departures from an unpredictable random walk process for *NER*s, which is however modified by remarking that in many cases predictability exists only for specific currencies and during specific time periods, e.g., amid and after the 2008 global Financial crisis.

The *RER* is a central variable in macroeconomic models. Hence, EJR's *NER*-regressions results may be said to contradict the nominal exchange rate's macroeconomic disconnect property, Meese and Rogoff (1983). In order to avoid misunderstandings, it is worth mentioning that this paper's critique does not entail that *NER*'s cannot be explained by other empirical definitions of macroeconomic fundamentals than lagged *RER*.

In a recent paper, Stavrakeva and Tang (2026) regress daily *NER*s on lagged macroeconomic news, and claim to disprove the disconnection property, which is inconsistent with the incumbent full information rational expectations theory of exchange rates. Chahrour et al. (2024) estimate VARs and find a strong connection between exchange rates and other macroeconomic fundamentals. However, the finding relates current exchange rates to future fundamentals. Hence, the implied critique of EJR would concern the interpretation rather than the regularity.

The rest of the paper is organised as follows. Section 2 presents the main variable definitions and introduces the exact form of the nominal exchange rate regression that is the object of our study. Section 3 gives an econometric interpretation and characterisation of the regression, and sets the stage for the simulation based analysis in Section 4. Section 5 is the applied part of the paper, where we estimate empirical nominal exchange rate regressions. It is not meant as a replication study, but rather as a check on the usefulness of the theory developed, as an aid in interpretation of empirical *NER* regressions – and getting it right. Section 6 contains our conclusions.

2 The Nominal Exchange Rate Regression

2.1 Definitions

A nominal exchange rate, *NER*, represents the relative price of two currencies. In this paper, it is defined as the price of foreign currency in units of the domestic currency in a given time period, and is denoted by E_t.

The definition of the real exchange rate, *RER*, is:

$$Q_t = \frac{E_t P_t^f}{P_t} \tag{1}$$

where P_t^f denotes the consumer price index of the foreign country and P_t denotes the domestic consumer price index.

We use lower case letters to denote logs of these variables. Hence, q_t denotes $log(Q_t)$ and is defined by:

$$q_t = e_t + p_t^f - p_t \tag{2}$$

2.2 The *NER* regression

In terms of these definitions, the NER-regression of Eichenbaum et al. (2021), EJR hereafter, takes the form:

$$e_{t+h} - e_t = \alpha_h^e + \beta_h^e q_t + \epsilon_{h,t+h}, \qquad h = 1, 2, ..., H, \tag{3}$$

where α_h^e and β_h^e are parameters and $\epsilon_{h,t+h}$ denotes the error-term. The dependent variable is the (approximate) rate of nominal depreciation recorded between period $t+h$ and period t.

Different variants of nominal exchange rate regressions have been used in the literature. Engel and Wu (2023) replaced q_t by e_t on the right hand side, and examined whether a set of economic variables – in particular, measures of "global risk" which recent literature has focused on – add forecasting power to the lagged level of the log of the exchange rate itself. One of their conclusions was that persistent global risk variables did have additional forecasting power when the level *NER* is included as a predictor.

EJR assume that q_t is a stationary time series variable. Hence, estimates of β_h^e that are negative and significant are interpreted as consistent with the view that changes in *NER* (depreciations or appreciations) in part represent mean-reversion in *RER* after it has been hit by a shock. The hypothesis is that this kind of equilibrating mechanism is typical for floating exchange rate regimes.

A main finding was that the numerical and statistical significance of the β_h^e estimates increased with the horizon h. For inflation targeting countries EJR found that the estimated values of β_h^e were negative for all h, and statistically significant at three-year horizons or longer. Based on these findings, EJR suggested that the short horizons typical of *NER*-regressions in earlier studies explain why those regressions had failed to produce evidence in support of the weak relative PPP hypothesis.

In comparison, for fixed exchange rate regime data, the estimated β_h^es were close to zero for all h. Instead, it was in the relative-price regressions, with $(p_{t+h}^f - p_{t+h}) - (p_t^f - p_t)$ on the left hand side, that the adjustment parameter was found to be significant for fixed exchange rate regime data.

EJR judged these findings as strongly supportive of the view that in the established inflation targeting economies that constitute their benchmark data set, the real exchange rate was highly negatively correlated with changes in future nominal rates at horizons of three or more years. Moreover, regressions for countries that had changed monetary policy in the sample period (Norway was one) indicated strongly that the co-movement between nominal exchange rates, inflation and real rates in the new inflation targeting countries followed the same pattern as in the countries in the benchmark group. They viewed the results as supportive of their hypothesis that the monetary policy regime is a central determinant of the way that the real exchange rate adjusts to shocks.[5]

3 Econometric interpretation of *NER*-regressions

The *NER*-regression (3) can be interpreted as a multi-period forecasting equation, of e_{t+h} where $h = 1, 2, ..., H$ denotes the forecast horizons, Marcellino et al. (2006). However, as the name says, the starting point is that it is a regression equation that can be used to test the hypothesis that *RER* is a statistically significant predictor of changes in the NER. The subscripts to the coefficients, α_h^e and β_h^e indicate that in general, the answer depends on whether the dependent variable is measured as a quarterly change ($h = 1$), or a four quarter change ($h = 4$) – and so on. As noted, EJR estimated equations for 12, 20 and 28 quarters.

As the equation holds in all time periods, it is practical to use the period t *NER* equation:

$$\Delta_h e_t = \alpha_h^e + \beta_h^e q_{t-h} + \varepsilon_{h,t} \tag{4}$$

where Δ_h is the h period difference operator, hence $\Delta_h e_t = e_t - e_{t-h}$.

[5] Eichenbaum et al. (2021, p. 201).

As noted, EJR maintain the hypothesis that the *RER* q_t is stationary. In line with the general agreement of floating exchange rate regimes, the *NER* e_t is assumed to be well approximated by a random-walk process without drift. Internal consistency then requires that also $(p^f - p)_t$ is a random-walk.

Jointly, these assumptions can be interpreted as saying that while both e_t and $(p^f - p)_t$ are integrated series, denoted $I(1)$, they are cointegrated in such a way that $q_t \sim I(0)$. Cointegration implies equilibrium correction dynamics, Engle and Granger (1987). For example, assume that the data generation process is a VAR that generates e_t and $(p^f - p)_t$ as cointegrated $I(1)$-series. Cointegration then implies:

$$ec_t = b_1 e_t + b_2 (p^f - p)_t - m \sim I(0),$$

where b_1 and b_2 denote cointegration parameters, and m denotes the mean of the long-term relationship between e_t and $(p^f - p)_t$. Clearly, if the cointegration parameters are $b_1 = b_2 = 1$, then:

$$ec_t = q_t - m \sim I(0). \tag{5}$$

If the cointegrated VAR has first or second order dynamics, CI-VAR(1) or CI-VAR(2), the implied equilibrium correction model can be written as:

$$\Delta_1 e_t = -a_1[q_{t-1} - m] + c_{11}\Delta e_{t-1} + c_{12}\Delta(p^f - p)_{t-1} + u_{1,t},\ a_1 > 0, \tag{6}$$

$$\Delta_1 (p^f - p)_t = -a_2[q_{t-1} - m] + c_{21}\Delta e_{t-1} + c_{22}\Delta(p^f - p)_{t-1} + u_{2,t},\ a_2 \geq 0. \tag{7}$$

where the two error terms have a joint distribution, which can be assumed to be normal without loss of generality (for the representation). In the case of CI-VAR(1), the four c_{ij} coefficients are zero. If the DGP is CI-VAR(2), each c_{ij} coefficient is non-zero in general.

Cointegration implies equilibrium correction in *one* of the variables, not necessarily in both of them. Hence, to be consistent with adjustment in inflation targeting regimes (that the *NER* adjusts), we can set the equilibrium correction coefficients to $-a_1 < 0$ and $a_2 = 0$. The relative price $(p^f - p)_t$ is then weakly-exogenous with respect to the cointegration parameters and the remaining adjustment parameter $-a_1$, cf. e.g., Ericsson (1994, Ch.1) and Johansen (1995, Ch. 8), and Akram (2006) for an application to Norwegian exchange rate data. Hence, it is the adjustment of e_t that represents the equilibrium correction mechanism that produces the stationarity of q_t; price level adjustments do not contribute to the correction process.

It is convenient to re-write the DGP equation (6) as:

$$\Delta_1 e_t = \beta_0 + \beta_1 q_{t-1} + \beta_2 \Delta e_{t-1} + \beta_3 \Delta(p^f - p)_{t-1} + u_{1,t}, \tag{8}$$

where $\beta_0 \equiv a_1 m$, $\beta_1 \equiv -a_1$, $\beta_2 \equiv c_{11}$ and $\beta_3 \equiv c_{12}$.

If we first consider the CI-VAR(1) version of the DGP, we note that since $\beta_2 = \beta_3 = 0$ in (8), the *NER*-regression for $h = 1$:

$$\Delta_1 e_t = \alpha_1^e + \beta_1^e q_{t-1} + \varepsilon_{1,t} \tag{9}$$

is isomorphic with the DGP equation for $\Delta_1 e_t$, hence $\beta_1^e = -a_1$ and $\varepsilon_{h,t} \equiv u_{1,t}$. Hence, since *NER*-regression (9) emulates the DGP, the OLS estimator $\hat{\beta}_1^e$ is consistent (plim $\beta_1^e = -a_1$) and asymptotically efficient.

The correctly specified model of $\Delta_2 e_t$, again given that the DGP is CI-VAR(1), is:

$$\Delta_2 e_t = \beta_0 + \beta_1 q_{t-2} + (1 + \beta_1)\Delta_1 e_{t-1} + \beta_1 \Delta_1 (p^f - p)_{t-1} + u_{1,t}, \tag{10}$$

and we note that the OLS estimator of β_2^e in the *NER*-regression for $h = 2$:

$$\Delta_2 e_t = \alpha_2 + \beta_2^e q_{t-2} + \varepsilon_{2,t} \tag{11}$$

will not be consistent for the DGP adjustment coefficient a_1, since in general the two omitted variables $\Delta_1 e_{t-1}$ and $\Delta_1(p^f - p)_{t-1}$ are correlated with q_{t-2}. If we continue this line of reasoning and compare *NER*-regressions with increasing h with the implied multi-period forecasting equations, it emerges that the whole sequence shares the same issue: They do not yield consistent estimation of the true parameter in the relationship between the *NER* and lags of *RER*.

In the next section, in order to have a somewhat richer DGP we specify it as a cointegrated VAR with second order dynamics, CI-VAR(2). Given this DGP, OLS estimation of the NER-regression for $h = 1$ does not yield consistent estimation of a_1, since that *NER*-regression is not a valid simplification of the true conditional model (8). i.e., neither β_2 nor β_3 can be set to zero in this case. Hence when the DGP is CI-VAR(2), it appears that none of the *NER*-regressions give consistent and efficient estimation of the true adjustment coefficient between the nominal and real exchange rates.

What emerges from the above is that the relationship between *NER*-regressions and the (underlying) DGP is of importance. Also when the research purpose is 'only' to estimate the predictive power of *RER* across different horizons, rather than the adjustment coefficient of the DGP, this is true. Mis-specification of the regression model relative to the DGP may damage the statistical inference also about the predictability of *NER*.

Of course, there is nothing hindering that in practice, the DGP is well approximated by, for example, the *NER*-regression for $h = 4$. That would be in line with EJR's premise that the applied econometrician cannot expect to find predictive power of lagged *RER* over shorter horizons than a year. Within the VAR framework, such a DGP would entail fourth order dynamics and zero restrictions on the coefficients of the intervening differences of the *NER* and the relative price. However, it is difficult to see any good reason for not investigating whether those restrictions are credible and reasonable to impose on the empirical model. It can be done through mis-specification testing and by general to specific dynamic modelling, both standard.

As noted above, EJR use standard errors that are robust against residual autocorrelation and heteroskedasticity when they test the significance of q_{t-h} in the *NER*-regressions. It is, however, not certain that robust standard errors of β_h^e give valid inference when the source of the problem clearly is the omission of variables. Moreover, the statistical distributions used to make inference about β_h^e are affected if the stationarity assumption of $q_t \sim I(0)$ is not defendable, and is replaced by $q_t \sim I(1)$.

With reference to Table 1, the $\Delta_1 e_t$ *NER*-regression can be placed in both cell M1 and M3. In M1, both the dependent variable and the regressor are stationary variables. The reason why it makes sense to test H_0: $\beta_1^e = 0$ in M1, i.e., why there is no internal inconsistency under that null, is that stationarity of q_t could be maintained through adjustment of the relative price $(p^f - p)_t$. In the cointegrated system (6)-(7) the assumption of $q_t \sim I(0)$ is only inconsistent with $a_1 = 0 \Leftrightarrow \beta_h^e = 0$ if also $a_2 = 0$ in the relative-price equation.

In M3, the relevant inference theory is non-standard since q_t is non-stationary, but approximately correct critical values are available from the cointegration literature, e.g., Ericsson and MacKinnon (2002). As just noted, the purpose is then to test the null-hypothesis of $q_t \sim I(1)$ against the alternative of $I(0)$-stationarity. This case is, however, not considered by EJR.

Table 1: Types of *NER*-regressions, based on $e_t \sim I(1)$.

		q_t	
		I(0)	I(1)
$\Delta_h e_t$ ("small h")	I(0)	M1	M3
$\Delta_h e_t$ ("large h")	I(1)	M2	M4

We now make the following suggestion: As h is increased in the sequence of nominal exchange rate regressions, the temporal properties of the regressand are changed from stationarity towards non-stationarity. We indicate this by:

$$\Delta_h e_t \underset{h\to\infty}{\to} \doteq I(1) \tag{12}$$

If indeed $\Delta_h e_t$, for some realistic-to-use h behaves as if it was a random-walk, the implication is that, for large h, the nominal exchange rate regression can become either unbalanced (if $q_t \sim I(0)$) or spurious (if $q_t \sim I(1)$), in finite samples. In both these cases, care must be taken to avoid unreliable results for tests for the existence of a relationship.

Unbalanced regression denotes the case where $\Delta_h e_t$ is practically $I(1)$ while $q_t \sim I(0)$, i.e., M2 in Table 1. Logically, the adjustment mechanism that secures the stationarity of q_t is the relative price, while $\beta_h^e = 0$. However, the *NER*-regressions will be have autocorrelation residuals, which can lead to over-rejection of $\beta_h^e = 0$, also when consistent standard errors are used.

The constellation of $q_t \sim I(1)$ and $\Delta_h e_t$ practically $I(1)$, stands out as relevant, since the evidence for stationarity of (in particular) *RER*s is far from conclusive. M4 in Table 1, is therefore a case where *NER*-regressions can give yield spurious rejection of the null that $\beta_h^e = 0$, unless the relevant non-standard inference theory is used, cf. Granger and Newbold (1974), Phillips (1986).

An important question for applied work is therefore how large h needs to be for $\Delta_h e_t$ to become dominated by a low-frequency unit-root, since that is likely to have consequences for the true significance level (Type-I error probability) when the hypothesis of $\beta_h^e = 0$ is tested.

Apparently, inference problems are inherent in *NER*-regressions. It is only the combination of first order CI-VAR DGP and $h = 1$ *NER*-regression that gives trustworthy results (without further qualifications). This is also a reminder of a point made by Granger et al. (2001), namely that spurious regression occurs from a mis-specified model, under the economically relevant null hypothesis, also when the variables are stationary. In the simulations in the next section, we therefore simulate DGPs that generate both stationary and non-stationary *RER*.

4 Monte Carlo simulations

A range of simulations are undertaken to evaluate the properties of the above regressions.

4.1 DGP specification

The DGP is given by a bivariate VAR(2) where the first element $y_{1,t} = e_t$ and the second element is the relative price $y_{2,t} = (p^f - p)_t$ such that $y_t' = (y_{1,t}, y_{2,t})$:

$$y_t = A_0 y_t + A_1 y_{t-1} + A_2 y_{t-2} + \varepsilon_t \tag{13}$$

The VAR can be re-written as:

$$(I - A_0) y_t = I y_{t-1} + \Pi y_{t-1} - A_2 (y_{t-1} - y_{t-2}) + \varepsilon_t \tag{14}$$

where the long-run parameter matrix Π is defined as:

$$\Pi = (A_1 + A_2 - I), \tag{15}$$

and we define $A_0 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$. The Π-matrix is factorised as:

$$\Pi = \alpha \beta', \tag{16}$$

where β holds the cointegration parameters and α contains the adjustment coefficients, $\alpha' = \begin{pmatrix} \alpha_{11} & \alpha_{21} \end{pmatrix}$.

Using (15) and (16):

$$A_2 = (\alpha \beta' + I) - A_1. \tag{17}$$

The innovations are generated by:

$$\begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix} \sim \mathsf{IN} \left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0.04 & 0 \\ 0 & 0.02 \end{pmatrix} \right] \tag{18}$$

such that the errors are two independent normals with standard deviations of 0.2 and 0.14 to reflect the higher variance of the nominal exchange rate compared to the relative price.

Three alternative DGP specifications are considered, summarised in Table 2. In each DGP specification the A_1 matrix is identical, but A_2 is varied to give a specification with cointegration but no weak exogeneity (DGP#1), cointegration with weak exogeneity (DGP#2) and no cointegration (DGP#3). These DGPs correspond to (6) and (7), or nested versions thereof.

For DGP#1 and DGP#2, $\beta' = \begin{pmatrix} 1 & 1 \end{pmatrix}$ to reflect a stationary real exchange rate, q_t, where there is long-run adjustment in both the real exchange rate and the price differential. This corresponds to M1 and M2 in Table 1 and implies that the central bank responds to the price differential as well, in contrast to EJR:

$$\begin{aligned} \Delta y_{1t} &= -0.5 q_{t-1} + 0.4 \Delta y_{1,t-1} + 0.3 \Delta y_{2,t-1} + \varepsilon_{1t} \\ \Delta y_{2t} &= -0.25 q_{t-1} + 0.15 \Delta y_{1,t-1} - 0.15 \Delta y_{2,t-1} + \varepsilon_{2t} \end{aligned} \tag{19}$$

The second DGP maintains cointegration but imposes weak exogeneity of the price differential, matching the conclusion that EJR find, also in keeping with M1 and M2 in Table 1:

$$\begin{aligned} \Delta y_{1t} &= -0.5 q_{t-1} + 0.4 \Delta y_{1,t-1} + 0.3 \Delta y_{2,t-1} + \varepsilon_{1t} \\ \Delta y_{2t} &= -0.1 \Delta y_{1,t-1} - 0.4 \Delta y_{2,t-1} + \varepsilon_{2t} \end{aligned} \tag{20}$$

	DGP#1	DGP#2	DGP#3
A_2	$\begin{pmatrix} -0.4 & -0.3 \\ -0.15 & 0.15 \end{pmatrix}$	$\begin{pmatrix} -0.4 & -0.3 \\ 0.1 & 0.4 \end{pmatrix}$	$\begin{pmatrix} 0.1 & 0.2 \\ 0.1 & 0.4 \end{pmatrix}$
Π	$\begin{pmatrix} -0.5 & -0.5 \\ -0.25 & -0.25 \end{pmatrix}$	$\begin{pmatrix} -0.5 & -0.5 \\ 0 & 0 \end{pmatrix}$	0
Rank(Π)	1	1	0
α	$\begin{pmatrix} -0.5 \\ -0.25 \end{pmatrix}$	$\begin{pmatrix} -0.5 \\ 0 \end{pmatrix}$	None
β'	$\begin{pmatrix} 1 & 1 \end{pmatrix}$	$\begin{pmatrix} 1 & 1 \end{pmatrix}$	None
Weak exogeneity y_2	No	Yes	n/a

Table 2: Summary of DGP characteristics. All DGPs have $A_1 = \begin{pmatrix} 0.9 & -0.2 \\ -0.1 & 0.6 \end{pmatrix}$ and innovations generated by (18).

The third DGP specification corresponds to M3 and M4 in Table 1 , where there are two I(1) processes without cointegration to result in $q_t \sim I(1)$:

$$\begin{aligned} \Delta y_{1t} &= -0.1\Delta y_{1,t-1} - 0.2\Delta y_{2,t-1} + \varepsilon_{1t} \\ \Delta y_{2t} &= -0.1\Delta y_{1,t-1} - 0.4\Delta y_{2,t-1} + \varepsilon_{2t} \end{aligned} \tag{21}$$

The nesting conditional DGP for $y_{1,t}$ in VECM form is given in (8), i.e.,:

$$\Delta y_{1,t} = \beta_0 + \beta_1 q_{t-1} + \beta_2 \Delta y_{1,t-1} + \beta_3 \Delta y_{2,t-1} + \epsilon_{1,t}, \tag{22}$$

where $q_{t-1} = (y_{1,t-1} + y_{2,t-1})$. In DGPs #1 and #2, $\beta_1 \neq 0$. As the cointegrating relation is stationary, the OLS estimator of β_1 is standard and the approximate non-centrality of the $t-$statistic under the fixed alternative is:

$$t \approx \sqrt{T}\frac{-0.5}{\sigma_{\beta_1}}. \tag{23}$$

where σ_{β_1} is the asymptotic standard deviation of $\sqrt{T}\left(\hat{\beta}_1 - \beta_1\right)$. If the cointegrating relation is known, single equation OLS estimation is consistent for the first equation, and as the innovations are independent in the DGP design, the conditional model is sufficient to estimate as there is no efficiency gained from the joint innovation structure. As such, whether weak exogeneity holds or not is not relevant to the derivation below. For DGP#3, $\beta_1 = 0$. The concern is that as h increases, the induced non-stationarity properties of $\Delta_h y_{1,t}$ result in incorrect inference regarding β_1.

First, we consider the nesting DGP specification for $h = 2$. Using $\Delta_2 y_{1,t} = \Delta y_{1,t} + \Delta y_{1,t-1}$, we can insert (22) and the lag of (22) to obtain:

$$\begin{aligned} \Delta_2 y_{1,t} &= 2\beta_0 + \beta_1(q_{t-1} + q_{t-2}) + \beta_2(\Delta y_{1,t-1} + \Delta y_{1,t-2}) \\ &\quad + \beta_3(\Delta y_{2,t-1} + \Delta y_{2,t-2}) + \epsilon_{1,t} + \epsilon_{1,t-1}. \end{aligned} \tag{24}$$

As $\Delta q_{t-1} = q_{t-1} - q_{t-2} = \Delta y_{1,t-1} + \Delta y_{2,t-1}$, then:

$$q_{t-1} = q_{t-2} + \Delta y_{1,t-1} + \Delta y_{2,t-1}, \tag{25}$$

which results in:

$$\begin{aligned}\Delta_2 y_{1,t} &= 2\beta_0 + 2\beta_1 q_{t-2} + (\beta_1+\beta_2)\Delta y_{1,t-1} + (\beta_1+\beta_3)\Delta y_{2,t-1} \\ &\quad +\beta_2 \Delta y_{1,t-2} + \beta_3 \Delta y_{2,t-2} + \epsilon_{1,t} + \epsilon_{1,t-1},\end{aligned} \tag{26}$$

and for $h > 2$ the DGP representation generalises to:

$$\begin{aligned}\Delta_h y_{1,t} &= h\beta_0 + h\beta_1 q_{t-h} + \sum_{k=1}^{h-1} (k\beta_1+\beta_2)\,\Delta y_{1,t-k} + \sum_{k=1}^{h-1} (k\beta_1+\beta_3)\,\Delta y_{2,t-k} \\ &\quad +\beta_2 \Delta y_{1,t-h} + \beta_3 \Delta y_{2,t-h} + \sum_{j=1}^{h} \epsilon_{1,t-h+j},\end{aligned} \tag{27}$$

All short-run dynamic lags are included in the DGP representation. The coefficient on the cointegrating relation is a function of h and so will not converge. One could scale the dependent variable by defining:

$$\widetilde{\Delta_h y_{1,t}} = \frac{(y_{1,t} - y_{1,t-h})}{h} \tag{28}$$

as is done in the local projection literature, see, e.g., Jordà (2005).

The mapping in (27) is not a local projection, since it includes differenced regressors dated after the forecast origin. These regressors are generally correlated with the composite error term, so OLS estimation does not consistently recover the structural coefficients of the DGP. Equation (27) should be interpreted as an algebraic representation of the DGP process, whereas a valid direct projection instead conditions on information available at the forecast origin, developed below.

4.2 Model specification

The *NER* regression (4) is a direct projection (DP) which omits short-run dynamics:

$$\Delta y_{1,t} = \tilde{\beta}_0 + \tilde{\beta}_1 q_{t-1} + \epsilon_{1,t}. \tag{29}$$

We use $\tilde{\cdot}$ to distinguish the parameters on the DP model as $\tilde{\beta}_1$ no longer estimates the structural ECM loading, β_1. Instead, $\tilde{\beta}_1$ can be interpreted as the unconditional projection coefficient, leading to a sequence of projection coefficients over h:

$$\Delta_h y_{1,t} = \tilde{\beta}_{0(h)} + \tilde{\beta}_{1(h)} q_{t-h} + \epsilon_{1,t(h)}, \tag{30}$$

where $\epsilon_{1,t(h)}$ is a function of h.

As (29) omits relevant variables ($\Delta y_{1,t-1}, \Delta y_{2,t-1}$ for $h = 1$) we can calculate the omitted variable bias. Defining $x_t' = \begin{pmatrix} q_t & \Delta y_{1,t} & \Delta y_{2,t} \end{pmatrix}$, from (25) we obtain:

$$x_t = M x_{t-1} + G\epsilon_t \tag{31}$$

where:

$$M = \begin{pmatrix} 0.25 & 0.55 & 0.15 \\ -0.5 & 0.4 & 0.3 \\ -0.25 & 0.15 & -0.15 \end{pmatrix}, \quad G = \begin{pmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

From (18) the state innovation covariance matrix is:

$$\Omega = G\Sigma G' = \begin{pmatrix} 0.06 & 0.04 & 0.02 \\ 0.04 & 0.04 & 0 \\ 0.02 & 0 & 0.02 \end{pmatrix}.$$

Because x_t is stationary around the cointegrating equilibrium we can calculate:

$$P = \mathsf{E}(x_t x_t') = MPM' + G\Sigma G' \tag{32}$$

which solves to obtain:

$$P \approx \begin{pmatrix} 0.0971 & 0.0373 & 0.0131 \\ 0.0373 & 0.0591 & 0.0079 \\ 0.0131 & 0.0079 & 0.0258 \end{pmatrix}$$

where $P_{11} = Var(q_t) = 0.0971$ and $P_{21} = Cov(\Delta y_{1,t}, q_t) = 0.0373$.

We define $S_2' = \begin{pmatrix} 0 & 1 & 0 \end{pmatrix}$ which selects Δy_1 from the x vector and $S_1' = \begin{pmatrix} 1 & 0 & 0 \end{pmatrix}$ to select the cointegrating relation. Given $\Delta y_{1,t} = \begin{pmatrix} -0.5 & 0.4 & 0.3 \end{pmatrix} x_{t-1} + \epsilon_{1,t}$ and $\epsilon_{1,t} \perp x_{t-1}$, such that $Cov(\Delta y_{1,t}, q_{t-1}) = \begin{pmatrix} -0.5 & 0.4 & 0.3 \end{pmatrix} PS_1$. From (29) we can compute the theoretical direct projection estimate:

$$\tilde{\beta}_{1(1)} = \frac{Cov(\Delta y_{1,t}, q_{t-1})}{Var(q_{t-1})} = \frac{-0.0297}{0.0971} = -0.31. \tag{33}$$

Therefore, the DP coefficient for $h = 1$ will exhibit an omitted variable bias of magnitude 0.19 due to the omitted short run dynamics.

For $h > 1$ from (30):

$$\tilde{\beta}_{1(h)} = \frac{S_2' \left(\sum_{j=1}^{h} M^j \right) PS_1}{S_1' PS_1}, \tag{34}$$

such that for DGP#1 this implies $\tilde{\beta}_{1(2)} = -0.733$; $\tilde{\beta}_{1(3)} = -0.904$; $\tilde{\beta}_{1(5)} = -0.720$; $\tilde{\beta}_{1(10)} = -0.722$; and $\tilde{\beta}_{1(28)} = -0.709$. The non-monotonicity of the direct projection is discussed in the simulations below.

An alternative direct projection to that used by EJR augments the *NER* regression with short-run dynamics which are sufficiently lagged (using $\bar{\cdot}$ to distinguish the parameters of the augmented model):

$$\Delta_h y_{1,t} = \bar{\beta}_{0(h)} + \bar{\beta}_{1(h)} q_{t-h} + \sum_{i=0}^{h-1} \bar{\beta}_{2,i(h)} \Delta y_{1,t-h-i} + \sum_{i=0}^{h-1} \bar{\beta}_{3,i(h)} \Delta y_{2,t-h-i} + \epsilon_{1,t(h)}. \tag{35}$$

(35) is the most general augmented specification but it uses large degrees of freedom and includes irrelevant short run dynamics. A more parsimonious specification, denoted ADP for augmented direct projection, is:

$$\Delta_h y_{1,t} = \bar{\beta}_{0(h)} + \bar{\beta}_{1(h)} q_{t-h} + \bar{\beta}_{2(h)} \Delta y_{1,t-h} + \bar{\beta}_{3(h)} \Delta y_{2,t-h} + \epsilon_{1,t(h)} \tag{36}$$

which is correctly specified for DGP#1 and #2 at $h = 1$. We can calculate the theoretical values of $\bar{\beta}_{1(h)}$ along the lines of the DP model, but without the need for the covariance matrix adjustment, such that:

$$\bar{\beta}_{1(h)} = S_2' \left(\sum_{j=1}^{h} M^j \right) S_1. \tag{37}$$

For DGP#1 this implies $\bar{\beta}_{1(1)} = -0.5$ which gives the correct value of the β_1 parameter. Furthermore, $\bar{\beta}_{1(2)} = -0.9$; $\bar{\beta}_{1(3)} = -0.965$; $\bar{\beta}_{1(5)} = -0.713$; $\bar{\beta}_{1(10)} = -0.754$ and $\bar{\beta}_{1(28)} = -0.743$. Thus, the longer the horizon h, the less effect that the short run dynamics have on the direct projection coefficients and the unconditional coefficient from the DP and conditional coefficient from the ADP converge.

4.3 Stationarity of $\Delta_h y_{1,t}$

First, we investigate the stationarity properties of $\Delta_h y_{1,t}$. For all cases, $y_{1,t} \sim I(1)$ so $\Delta_1 y_{1,t} \sim I(0)$. For longer differences, we can use:

$$\Delta_h y_{1,t} = \sum_{i=0}^{h-1} \Delta_1 y_{1,t-i}$$

such that the long period difference is just the sum of a series of first differenced $I(0)$ processes so should be $I(0)$, but as $h \to \infty$ the long difference starts to exhibit spurious $I(1)$ behaviour due to the increasing serial correlation from overlapping shocks, hence $corr(\Delta_h y_{1,t}, \Delta_h y_{1,t+1}) \approx \frac{h-1}{h}$.

We run an ADF test on the $h-$differenced $y_{1,t}$, i.e.,:

$$\Delta\Delta_h y_{1,t} = \gamma_0 + \gamma \Delta_h y_{1,t-1} + \sum_{i=1}^{h} \kappa_i \Delta\Delta_h y_{1,t-i} + \eta_t \tag{38}$$

where all h lagged difference terms are included to account for serial correlation. The hypothesis test situation is: H_0:$\gamma = 0$ against H_1:$\gamma < 0$. We set $T = 200$ with 50 initial observations discarded to avoid dependence on initial conditions, and undertake $10,000$ replications. Critical values are obtained from MacKinnon's response surface approximations given as -3.463 at 1% and -2.876 at 5%, including an intercept but no deterministic trend.

Figure 1 records the ADF $\widehat{\mathtt{t}}_\gamma-$statistic across increasing h, along with critical values in the black solid lines on the top row, with each column corresponding to a DGP specification. The bottom row records the rejection frequency at 5% and 1%. $y_{1,t} \sim I(1)$ so we should reject the null hypothesis of a unit root for the hth difference. The null hypothesis is rejected for small h but as h increases the ADF test fails to reject as the induced serial correlation results in the data exhibiting non-stationary properties. In summary, longer differences behave as if non-stationary when the DGP implies the first difference is $I(0)$.

4.4 Distribution of projection coefficient

Simulations are undertaken to evaluate the properties of the *NER* regression. When $h = 1$ the projection coefficient is analogous to the adjustment coefficient of the equilibrium correction mechanism for the ADP, so can be interpreted as the speed of adjustment when the system deviates from long-run equilibrium. For all other cases, the interpretation of the projection coefficient is distinct from that of the DGP speed of adjustment. $M = 10{,}000$ replications of the DGPs given in (19)-(21) are generated and models (30) and (36) are estimated, varying $h = 1, \ldots, 28$ in steps of 1, with $T = 30, 50, 100, 200, 500, 1000$. Heteroskedastic and autocorrelation robust standard errors are used to compute the $\mathtt{t}-$statistics.

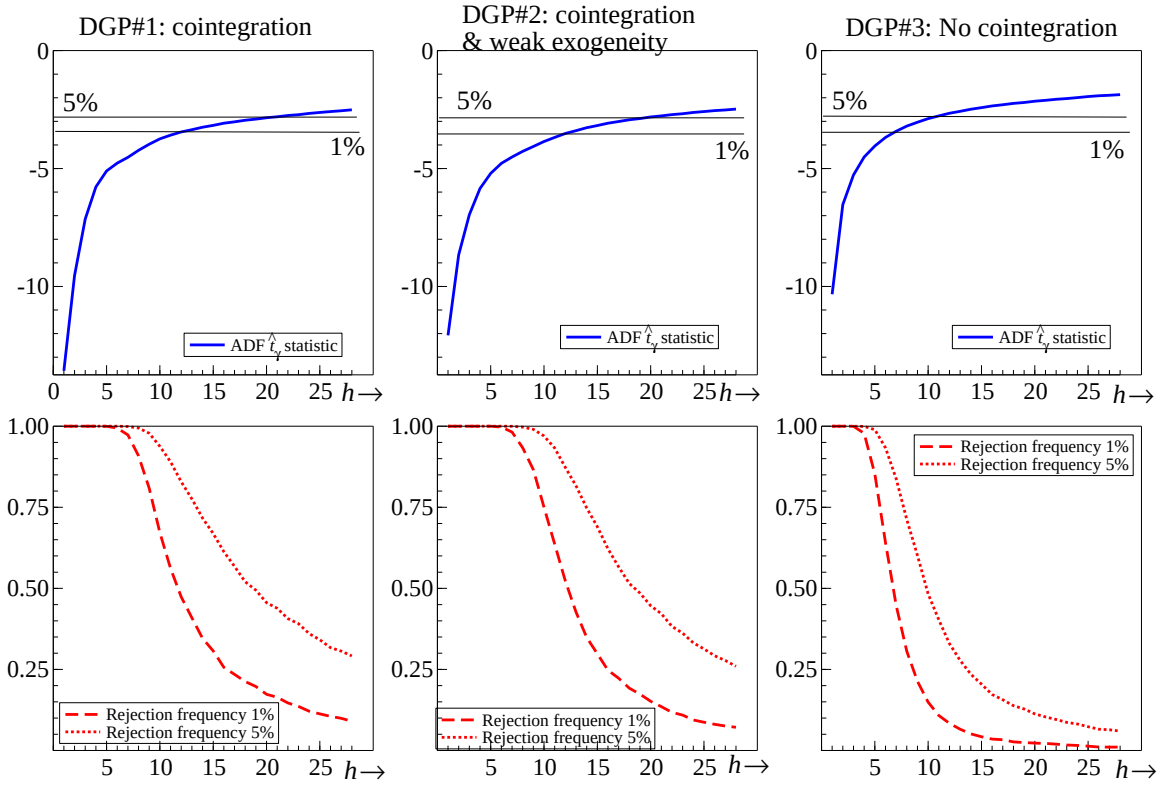


Figure 1: ADF tests for $\Delta_h e_t$ based on estimation of (38). Top panels report ADF$\hat{t_\gamma}$ statistics and bottom panels report rejection frequency at 1% and 5% using MacKinnon's response surface approximations given as $cv = -3.463$ for 1% and -2.876 at 5%.

Figure 2 records $\frac{1}{M}\sum_{i=1}^{M}\widehat{\tilde{\beta}}_{1(h),i}$ in the top row, $\frac{1}{M}\sum_{i=1}^{M}\widehat{t}_{\tilde{\beta}_{1(h),i}}$ in the middle row, and $\frac{1}{M}\sum_{i=1}^{M}\mathbf{1}_{\left\{|\widehat{t}_{\tilde{\beta}_{1(h),i}}|>cv_\alpha\right\}}$, where $\mathbf{1}$ is the indicator function and cv_α is the critical value at $\alpha = 5\%$ and 1% in the bottom row, for the DP (30). Figure 3 records the equivalent $\widehat{\tilde{\beta}}_{1(h)}$ statistics for ADP (36).

For DGP#1, the coefficient estimate on q_{t-h} matches the theoretical values in section 4.2, both for DP and ADP. The bias at $h = 1$ is evident in DP, where the mean estimate is -0.31, compared to ADP which is unbiased. The non-monotonicity of $\tilde{\beta}_{1(h)}$ and $\bar{\beta}_{1(h)}$ reflects the dynamic feedback inherent in the cointegrated system. The initial equilibrium correction is propagated through both y_1 and y_2 and the short-run dynamics, generating subsequent responses of alternating sign, so that the cumulative response need not increase monotonically with the projection horizon. The turning point is at $h = 3$ for this DGP specification. The sample size has little effect on the point estimate but is important for the $t-$statistic via the standard error. The $t-$statistic depends on $\widehat{\tilde{\beta}}_{1(h)}$ and $\widehat{\sigma}_{\tilde{\beta}_{1(h)}}$, which are both functions of h. The denominator, $\widehat{\sigma}_{\tilde{\beta}_{1(h)}}$ gets progressively worse as the sampling variation rises due to additional overlapping shocks, so coupling the non-monotonic numerator with the monotonically increasing denominator results in the $\mathtt{t}-$statistic paths in Figures 2 and 3.

For DGP#2 the results are similar in pattern but the magnitude differs due to weak exogeneity. When only one variable equilibrium-corrects there is less feedback leading to a smaller non-linearity, although the turning point is still at $h = 3$.

For both DGPs with cointegration such that $\beta_1 \neq 0$, the the null hypothesis of $H_0 : \tilde{\beta}_{1(h)} = 0$ and $H_0 : \bar{\beta}_{1(h)} = 0$ is rejected in 100 % of the tests for $h = 3$. For large samples the rejection rate remains close to 1, but for small samples there is a declining rejection frequency as h increases, suggesting that using a direct projection model under the null of cointegration requires large samples for correct inference. DP has higher rejection frequencies than ADP in small samples. As well as estimating more parameters which is costly in small samples, the omitted dynamics in the DP model could be attributed to the real exchange rate, inflating its significance. Given mis-specification, the different rejection frequencies cannot be interpreted as differences in power for the hypothesis test because of the differences in controls.

For DGP#3, $\beta_1 = 0$ but the null hypothesis is rejected well above the significance level of 1% and 5%, both for DP and ADP. Table 3 records a sample of rejection frequencies for DP, comparing rejection rates using the Student-$\mathtt{t}$ distribution which assumes stationarity and the Ericsson and MacKinnon (2002) (EM) critical values under the null of no cointegration. The rejection frequency increases as h gets larger, although it is much better sized using the EM critical values. However, both lead to a conclusion of a significant cointegrating relation more often than not as h increases which is spurious. Hence, incorrect inference is quite likely under M3 and M4.

Figures 4 and 5 record the distribution of $\widehat{\tilde{\beta}}_{1(h)}; \widehat{\bar{\beta}}_{1(h)}$ and $\widehat{t}_{\tilde{\beta}_{1(h)}}; \widehat{t}_{\bar{\beta}_{1(h)}}$ for four values of $h(= 1, 3, 5, 28)$, for $T = 200$, corresponding to DP and ADP respectively. For DGPs #1 and #2, when $h = 1$, the distribution is tightly centred around the dynamic response coefficient for ADP but the bias is evident for DP due to the omitted short-run dynamics. As h increases, the distribution shifts towards more negative values and becomes less precise. Weak exogeneity (DGP#2) has less non-monotonicity in that the $h = 3$ and $h = 5$ distributions more closely overlap, but the dynamic response coefficient is more negative. Only y_1 adjusts to the equilibrium error so the response of y_1 to an initial disequilibrium

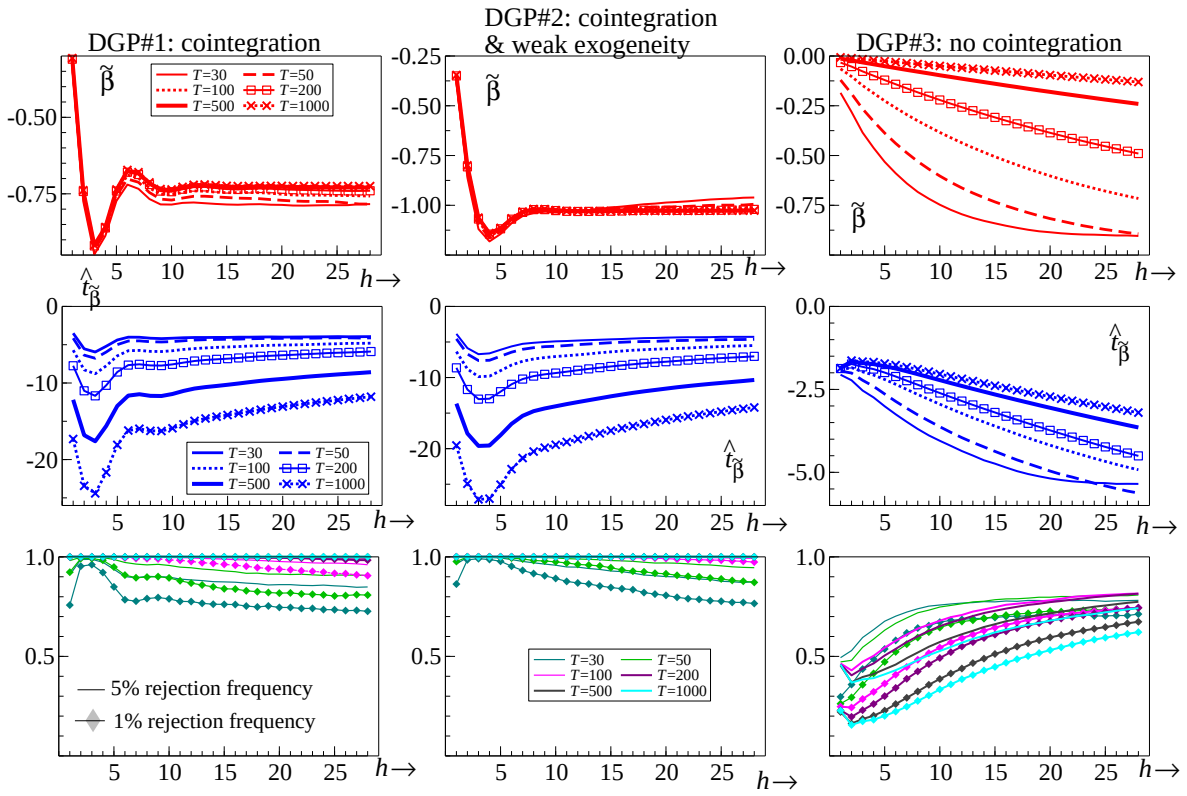


Figure 2: Mean estimates of $\widehat{\tilde{\beta}}_{1(h)}$ (top row), $\widehat{t}_{\tilde{\beta}_{1(h)}}$ (middle row) and rejection frequencies for $H_0 : \tilde{\beta}_{1(h)} = 0$ at 5% and 1% (bottom row) for the three DGP specifications (19)-(21), estimating the model given by (30). $M = 10{,}000$ replications.

	Student-t		EM	
	5%	1%	5%	1%
$h = 1$	0.442	0.218	0.082	0.028
$h = 14$	0.703	0.572	0.436	0.320
$h = 28$	0.806	0.731	0.646	0.557

Table 3: Rejection frequencies for $H_0 : \tilde{\beta}_1 = 0$ given DGP#3, for $T = 200$, $M = 10{,}000$ comparing 1% and 5% critical values from the t-distribution (which assumes stationarity: 2.6 & 1.97 respectively) and the Ericsson and MacKinnon (2002) critical values for $k = 2$ with a constant and no trend (which assumes a null hypothesis of no cointegration: -3.834 & -3.231 respectively).

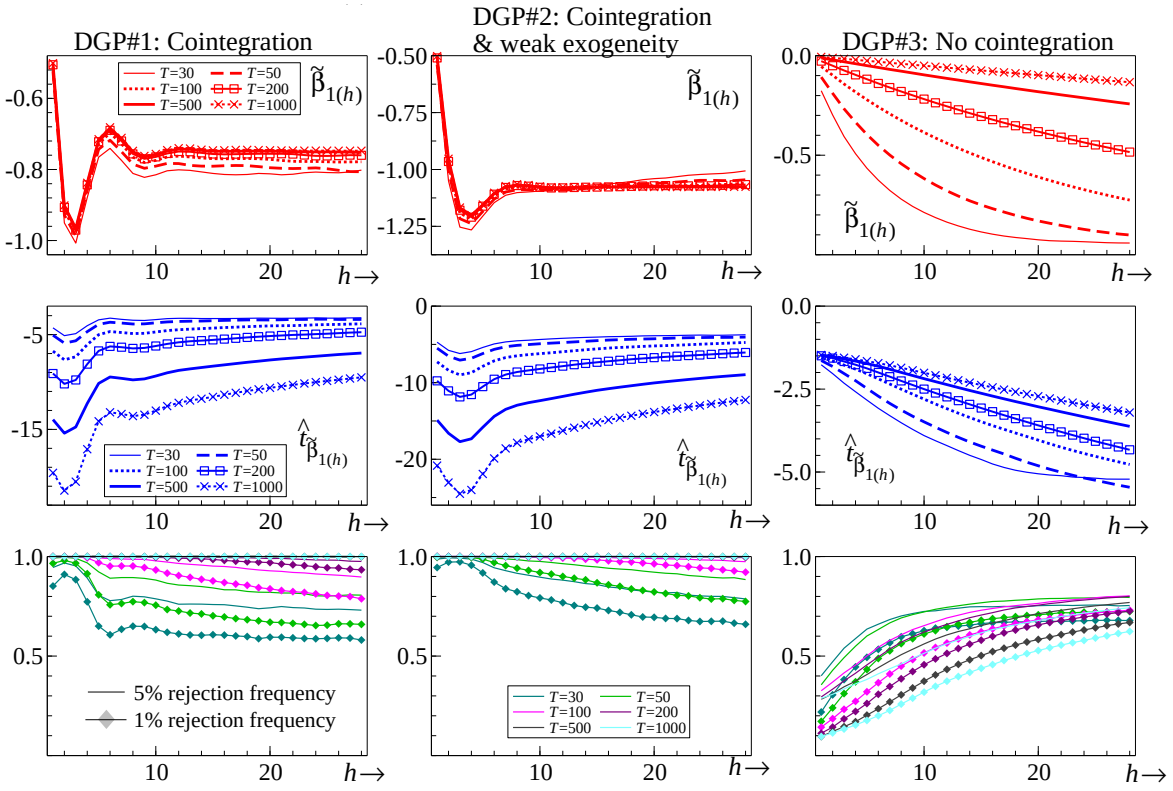


Figure 3: Mean estimates of $\widehat{\bar{\beta}}_{1(h)}$ (top row), $\widehat{t}_{\bar{\beta}_{1(h)}}$ (middle row) and rejection frequencies for $H_0 : \bar{\beta}_{1(h)} = 0$ at 5% and 1% (bottom row) for the three DGP specifications (19)-(21), estimating the model given by (36). $M = 10{,}000$ replications.

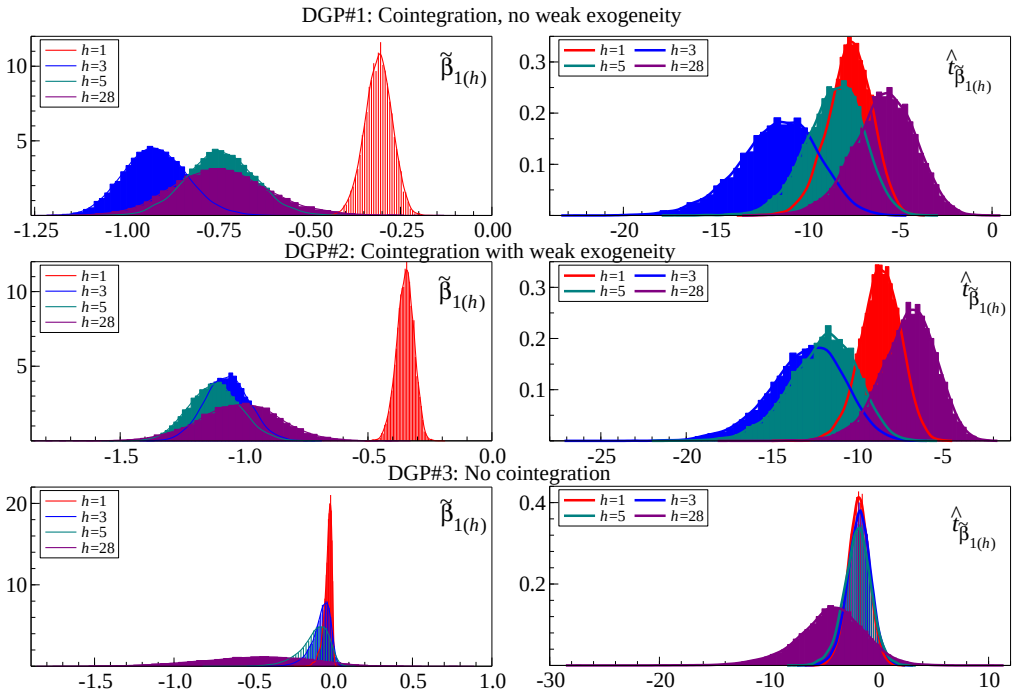


Figure 4: Distribution of $\widehat{\tilde{\beta}}_{1(h)}$ and $\widehat{t}_{\tilde{\beta}_{1(h)}}$ using model (30) for $h = 1, 3, 5, 28$. $T = 200$ and $M = 10{,}000$. DGP#1 in top panel, DGP#2 in middle panel & DGP#3 in bottom panel.

is stronger than for DGP#1. By $h = 28$ the degree of uncertainty is very wide, reflected in the t−statistic shifting closer to zero. In the case with no weak exogeneity (DGP#1) the real exchange rate determines both Δy_1 and Δy_2 such that the short-run dynamics embody more of the information contained in the initial disequilibrium. This makes the conditional contribution of q_{t-1} substantially smaller once those dynamics are controlled for, leading to a less significant $\widehat{t}_{\tilde{\beta}_{1(h)}}$ for DGP#1 compared to DGP#2.

For DGP#3, $h = 1$ is tightly concentrated around 0 but has a negative skew. This skew increases as h increases, along with the absolute value of the mean. By $h = 28$ the distribution is highly dispersed and negative for both DP and ADP.

Simulations for varying T highlight that as the sample size increases, $\widehat{\sigma}_{\tilde{\beta}_{1(h)}}$ reduces, and this increases the significance of the t−statistic. There is increasing power against the null hypothesis, and as such, large samples are needed if using DP/ADP rather than estimating the VECM. For DGP#3, there is a non-standard Dickey-Fuller type distribution. This is not a standard Dickey-Fuller distribution as the dependent variable is one variable of the system (Δe_t), while the integrated regressor is the linear combination (q_t). As such, the limiting distribution depends on the joint stochastic trends and innovations of the system. The distributions are not centred on 0 despite the null hypothesis $H_0 : \tilde{\beta}_{1(h)} = 0$ being correct, and the distributions become more skewed as h increases.

4.5 Model mis-specification

We have established that DP is mis-specified for $h \geq 1$ and ADP is mis-specified for $h > 1$ under the DGP given. This means that caution must be applied to conducting inference from the direct projection approach, even if HAC standard errors are used, e.g.,

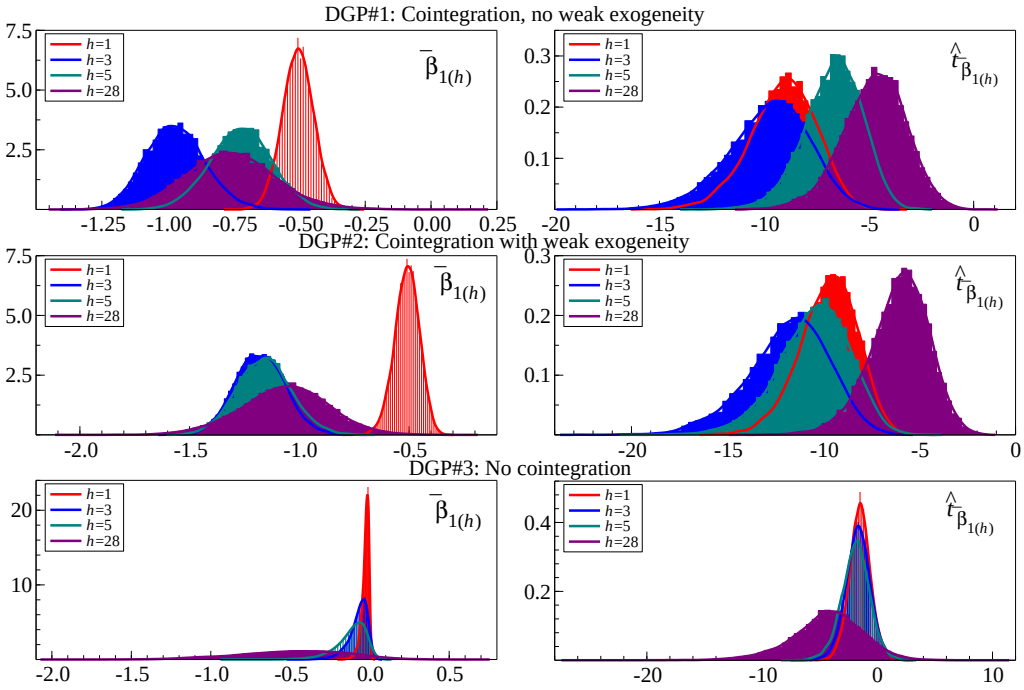


Figure 5: Distribution of $\widehat{\bar{\beta}}_{1(h)}$ and $\widehat{t}_{\bar{\beta}_{1(h)}}$ using model (36) for $h = 1, 3, 5, 28$. $T = 200$ and $M = 10{,}000$. DGP#1 in top panel, DGP#2 in middle panel & DGP#3 in bottom panel.

Hendry (1995, Ch. 10.10), Spanos (2021). Figures 6 and 7 record a range of diagnostic test rejection frequencies to highlight where mis-specification arises and how pervasive it is. Each row corresponds to a different diagnostic test, with each column specifying the DGP (19), (20) & (21). The sample sizes range from $T = 30, 50, 100, 200, 500$ to $1,000$ and $M = 10{,}000$ replications.

Row 1 corresponds to the test for residual autocorrelation, see Godfrey (1978) & Nielsen (2006), and as expected, there is strong evidence of residual autocorrelation for $h \geq 1$ for DP and $h \geq 2$ for ADP. The ADP test for AR is correctly sized at $h = 1$ for all DGPs, corresponding to the fact that the ADP nests all DGP specifications for $h = 1$. Row 2 records the rejection frequency of the test for residual normality (see Doornik and Hansen (2008)). At $h = 1$, the test is correctly sized for all DGPs but as h increases the test rejects with a higher frequency. HAC standard errors would not correct for this mis-specification. Row 3 records the rejection frequencies for the residual heteroskedasticity test (see White (1980)) which are well behaved for DGP#1 and DGP#2 but reject with increasing probability as h increases for DGP#3. The test for Autoregressive Conditional Heteroskedasticity (ARCH: Engle (1982)) also over-rejects across all DGPs for $h > 1$, with increasing rejection as T increases. The bottom row records the Ramsey RESET test for functional form (Ramsey (1969)) which correctly rejects for DGP#1 and DGP#2 but rejects too frequently for DGP#3. For DGPs #1 and #2, HACSEs will address some of the mis-specification but a failure of residual normality remains a problem and augmenting the direct projection regression for short run dynamics does not resolve the mis-specification at $h > 1$. For DGP#3, all diagnostics reject, highlighting the spurious nature of the regression.

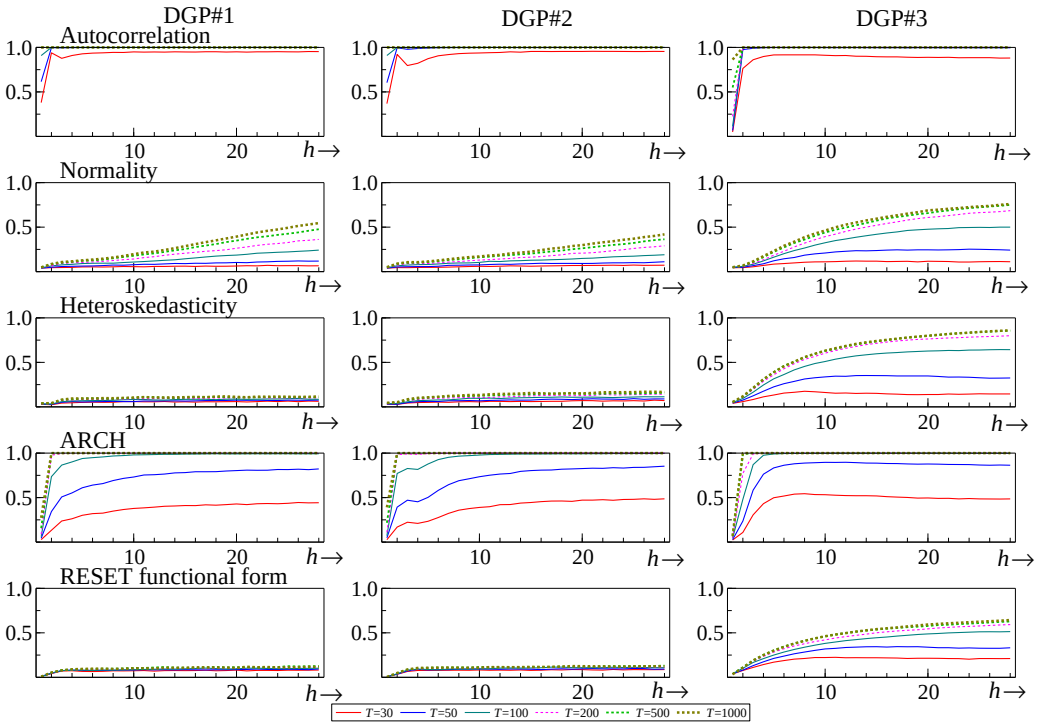


Figure 6: Diagnostic test rejection frequencies from DP (30) for residual autocorrelation, residual normality, heteroskedasticity, autoregressive conditional heteroskedasticity, and functional form specification. $M = 10{,}000$. DGP#1 (left column); DGP#2 (middle column); DGP#3 (right column).

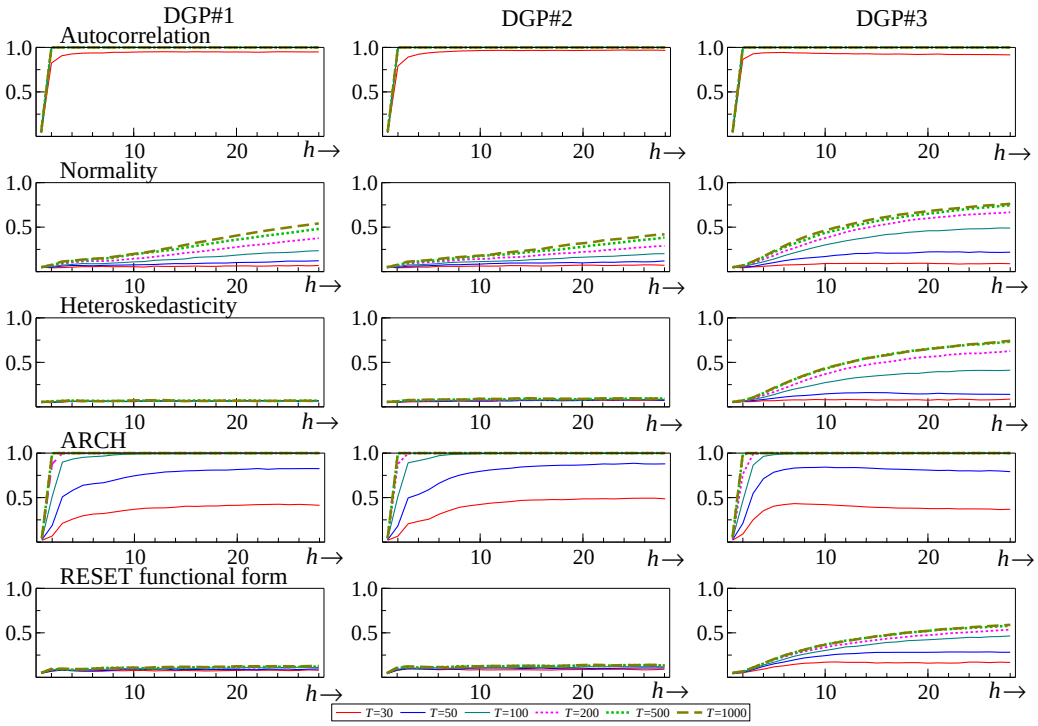


Figure 7: Diagnostic test rejection frequencies from ADP (36) for residual autocorrelation, residual normality, heteroskedasticity, autoregressive conditional heteroskedasticity, and functional form specification. $M = 10{,}000$. DGP#1 (left column); DGP#2 (middle column); DGP#3 (right column).

4.6 Forecasting the nominal exchange rate

The DP model is often used for forecasting, particularly if there is mis-specification, see Chevillon (2007). We undertake a forecast comparison exercise to compare alternative model specifications for the three DGPs (19)-(21). The models include the VECM, conditional ECM, DP and ADP. We focus on the forecast performance of $y_{1,t}$ which corresponds to the nominal exchange rate, e_t.

The VECM model is the correct specification for DGPs #1 and #2. h dynamic forecasts are produced from the system assuming the known cointegrating relation. We could compute the forecasts estimating the cointegrating relation which will add some uncertainty in small samples but does not affect the results much. The system (39) is estimated using two identities (40) and constrained full information maximum likelihood. This ensures a closed system which allows for dynamic forecasts and imposes the cointegrating relation $\begin{pmatrix} 1 & 1 \end{pmatrix}$. Given the system (for the nesting DGP#1) is:

$$\begin{aligned} \Delta y_{1,t} &= \beta_0 + \beta_1 \left(y_{1,t-1} + y_{2,t-1}\right) + \beta_2 \Delta y_{1,t-1} + \beta_3 \Delta y_{2,t-1} + \epsilon_{1,t} \\ \Delta y_{2,t} &= \gamma_0 + \gamma_1 \left(y_{1,t-1} + y_{2,t-1}\right) + \gamma_2 \Delta y_{1,t-1} + \gamma_3 \Delta y_{2,t-1} + \epsilon_{2,t} \end{aligned} \tag{39}$$

and

$$y_{1,t} = \Delta y_{1,t} + y_{1,t-1}, \qquad y_{2,t} = \Delta y_{2,t} + y_{2,t-1} \tag{40}$$

the forecasts are computed recursively as:

$$\begin{aligned} \widehat{\Delta y}^{VECM}_{1,T+h|T} &= \hat{\beta}_0 + \hat{\beta}_1 \left(\widehat{y}_{1,T+h-1|T} + \widehat{y}_{2,T+h-1|T}\right) + \hat{\beta}_2 \widehat{\Delta y}_{1,T+h-1|T} + \hat{\beta}_3 \widehat{\Delta y}_{2,T+h-1|T} \\ \widehat{y}^{VECM}_{1,T+h|T} &= \widehat{\Delta y}^{VECM}_{1,T+h|T} + \widehat{y}_{1,T+h-1|T} \end{aligned} \tag{41}$$

where $\hat{\cdot}$ denotes an estimated parameter and $\widehat{\cdot}$ denotes a forecast.

For the conditional ECM we use the forecasting model given in equation (22), again assuming the known cointegrating relation. We need to make an assumption on the path of Δy_2 as it is a conditional model. We assume y_2 follows a random walk such that $\widehat{y}_{2,T+h|T} = y_{2,T}$ and $\widehat{\Delta y}_{2,T+h|T+h-1} = 0$. The 1-step ahead forecast is given by:

$$\widehat{\Delta y}^{ECM}_{1,T+1|T} = \hat{\beta}_0 + \hat{\beta}_1 q_T + \hat{\beta}_2 \Delta y_{1,T} + \hat{\beta}_3 \Delta y_{2,T}. \tag{42}$$

For $h = 2$:

$$\widehat{\Delta y}^{ECM}_{1,T+2|T} = \hat{\beta}_0 + \hat{\beta}_1 (\widehat{y}_{1,T+1|T} + y_{2,T}) + \hat{\beta}_2 \widehat{\Delta y}^{ECM}_{1,T+1|T} \tag{43}$$

where $\widehat{y}_{1,T+1|T} = \widehat{\Delta y}_{1,T+1|T} + y_{1,T}$ and the $\Delta y_{2,t}$ component drops out. For $h-$steps ahead:

$$\widehat{\Delta y}^{ECM}_{1,T+h|T} = \hat{\beta}_0 + \hat{\beta}_1 (\widehat{y}_{1,T+h-1|T} + y_{2,T}) + \hat{\beta}_2 \widehat{\Delta y}_{1,T+h-1|T} \tag{44}$$

with $\widehat{y}_{1,T+h-1|T} = \sum_{j=1}^{h-1} \widehat{\Delta y}_{1,T+j|T} + y_{1,T}$. From the forecasts for the differenced term we obtain the levels forecast:

$$\widehat{y}^{ECM}_{1,T+h|T} = y_{1,T} + \sum_{j=1}^{h} \widehat{\Delta y}^{ECM}_{1,T+j|T} \tag{45}$$

For the DP and ADP models, forecasts for the levels are computed directly from (30) and (36):

$$\widehat{y}^{DP}_{1,T+h|T} = y_{1,T} + \hat{\tilde{\beta}}_{0(h)} + \hat{\tilde{\beta}}_{1(h)} q_T \tag{46}$$

$$\widehat{y}^{ADP}_{1,T+h|T} = y_{1,T} + \hat{\tilde{\beta}}_{0(h)} + \hat{\tilde{\beta}}_{1(h)} q_T + \hat{\tilde{\beta}}_{2(h)} \Delta y_{1,T} + \hat{\tilde{\beta}}_{3(h)} \Delta y_{2,T} \tag{47}$$

	VECM	ECM	DP	ADP	RW	DDV	DVECM
				DGP#1			
$h=1$	**0.202**	**0.202**	0.221	**0.202**	0.237	0.318	0.256
$h=3$	**0.294**	0.299	0.296	0.295	0.396	0.522	0.388
$h=12$	**0.386**	0.387	0.387	0.387	0.432	0.539	0.427
$h=28$	0.527	**0.504**	0.528	0.529	0.540	0.629	0.536
				DGP#2			
$h=1$	**0.202**	**0.202**	0.220	**0.202**	0.240	0.313	0.267
$h=3$	**0.294**	0.296	0.301	0.295	0.429	0.528	0.459
$h=12$	0.382	**0.380**	0.385	0.385	0.486	0.587	0.496
$h=28$	0.510	**0.492**	0.513	0.513	0.576	0.663	0.586
				DGP#3			
$h=1$	0.203	0.203	0.204	0.203	**0.202**	0.301	**0.202**
$h=3$	0.339	0.339	0.339	0.340	**0.330**	0.396	**0.330**
$h=12$	0.688	0.686	0.704	0.707	**0.643**	0.677	0.646
$h=28$	1.080	1.072	1.166	1.174	**0.972**	0.997	0.978

Table 4: Root mean square forecast errors (RMSFE) for models (41), (45), (46) & (47) along with benchmark devices (48), (49) & (50). $T = 200$ with $H = 28$ forecasts and $M = 10{,}000$ replications. Bold denotes smallest RMSFE.

Again, we impose the cointegrating relation to compare like with like.

Three benchmark forecasting models are computed for comparison. These include the Random Walk (RW: (48)), Double Differenced VAR (DDV: (49)) and Differenced VECM (DVECM: (50)):

$$\widehat{y}^{RW}_{1,T+h|T} = y_{1,T} \tag{48}$$

$$\widehat{y}^{DDV}_{1,T+h|T} = y_{1,T} + \Delta_1 y_{1,T} \tag{49}$$

$$\widehat{y}^{DVECM}_{1,T+h|T} = y_{1,T} + \hat{\hat{\beta}}_{1(h)} \Delta_1 y_{1,T} \tag{50}$$

We simulate the three DGP specifications $M = 10{,}000$ times and record the RMSFEs for a forecast horizon of $h = 1, 3, 12, 28$ for $T = 200$. For the cointegrated DGPs #1 and #2, the VECM and ECM models are correctly specified. The ADP is also correctly specified at $h = 1$ and the forecasts are identical, but the mis-specified DP model forecasts worse. They do outperform the benchmark forecast devices at all horizons. The ECM does the best at long horizons, suggesting there is little benefit to system estimation over the conditional model due to the diagonal error variance structure. The differences between the DP and ADP forecasts are negligible at long horizons, and are almost as accurate as the VECM.

For DGP #3 without cointegration, robust devices such as the random walk dominate at all horizons. Under this DGP specification, the nominal exchange rate is unpredictable. The DP and ADP models forecast worse than the VECM and ECM at longer horizons, but if there is a risk of spurious regression one cannot do better than a random walk forecast.

4.7 Structural Breaks in the adjustment coefficient

To investigate EJR's hypothesis that the monetary policy regime determines the adjustment mechanism, we introduce a structural break corresponding to a country's transition from a non-inflation targeting to an inflation targeting regime. Throughout, the real exchange rate is assumed to be stationary, implying cointegration between the nominal exchange rate and the relative price. Prior to the regime change, deviations from the long run equilibrium are corrected through the relative price, with the nominal exchange rate weakly exogenous. Following the adoption of inflation targeting, the adjustment mechanism switches and deviations from equilibrium are corrected through the nominal exchange rate, while the relative price becomes weakly exogenous. The DGP design is given by:

$$\Delta \mathbf{y}_t = \begin{cases} \mathbf{\Pi}^{(1)}\mathbf{y}_{t-1} + \mathbf{\Gamma}^{(1)}\Delta \mathbf{y}_{t-1} + \epsilon_t, & t \leq T_b \\ \mathbf{\Pi}^{(2)}\mathbf{y}_{t-1} + \mathbf{\Gamma}^{(2)}\Delta \mathbf{y}_{t-1} + \epsilon_t, & t > T_b \end{cases} \tag{51}$$

where

$$\begin{aligned} \mathbf{\Gamma}^{(1)} &= -A_2^{(1)} = \begin{pmatrix} -0.1 & -0.2 \\ 0.4 & 0.1 \end{pmatrix} \\ \mathbf{\Gamma}^{(2)} &= -A_2^{(2)} = \begin{pmatrix} 0.4 & 0.3 \\ -0.1 & -0.4 \end{pmatrix} \end{aligned}$$

but we keep the original $A_1 = \begin{pmatrix} 0.9 & -0.2 \\ -0.1 & 0.6 \end{pmatrix}$. The implication of this DGP design is that the cointegrating relation is still $\beta' = \begin{pmatrix} 1 & 1 \end{pmatrix}$ representing the real exchange rate but the adjustment mechanism shifts from $\alpha'^{(1)} = \begin{pmatrix} 0 & -0.5 \end{pmatrix}$ to $\alpha'^{(2)} = \begin{pmatrix} -0.5 & 0 \end{pmatrix}$, along with the short run dynamics. The VECM model is:

$$\begin{aligned} \Delta y_{1,t} &= -0.1\Delta y_{1,t-1} - 0.2\Delta y_{2,t-1} + \epsilon_{1,t}, \quad t \leq T_b \\ \Delta y_{2,t} &= -0.5q_{t-1} + 0.4\Delta y_{1,t-1} + 0.1\Delta y_{2,t-1} + \epsilon_{2,t}, \quad t \leq T_b \\ \Delta y_{1,t} &= -0.5q_{t-1} + 0.4\Delta y_{1,t-1} + 0.3\Delta y_{2,t-1} + \epsilon_{1,t}, \quad t > T_b \\ \Delta y_{2,t} &= -0.1\Delta y_{1,t-1} - 0.4\Delta y_{2,t-1} + \epsilon_{2,t}, \quad t > T_b \end{aligned} \tag{52}$$

where T_b is the break date.

Figure 8 records the distribution of $\widehat{\tilde{\beta}}_{1(h)}$, in the left panel, and the distribution of $\widehat{t}_{\tilde{\beta}_{1(h)}}$ in the right panel, for $T = 200$ where $T_b = 100$, estimating (30) over the full sample. The parameter shift in the adjustment coefficient α is very difficult to detect. As such, the distribution of $\widehat{\tilde{\beta}}_{1(h)}$ and $\widehat{t}_{\tilde{\beta}_{1(h)}}$ looks very similar to those for DGP#2 above. As the break does not involve a shift in mean it is not easily observed. While this means the inference described above still holds, the problem is that there is no power to detect the huge shift from $\begin{pmatrix} 0 & -0.5 \end{pmatrix}$ to $\begin{pmatrix} -0.5 & 0 \end{pmatrix}$. As such, policy could be based on the wrong economic interpretation. Hendry (2000) shows that substantial changes in the dynamics and adjustment speeds of cointegrated systems can remain relatively undetectable when the unconditional means of the $I(0)$ components are unchanged. Consequently, an economically important change in the policy transmission mechanism may occur without generating an obvious indication of model non-constancy. In contrast, a location shift that changes the equilibrium mean of the *RER* would generally be easier to detect, but could have more serious consequences for forecasting. Unless the shift is incorporated into the model, forecasts would continue to reflect the previous equilibrium, potentially

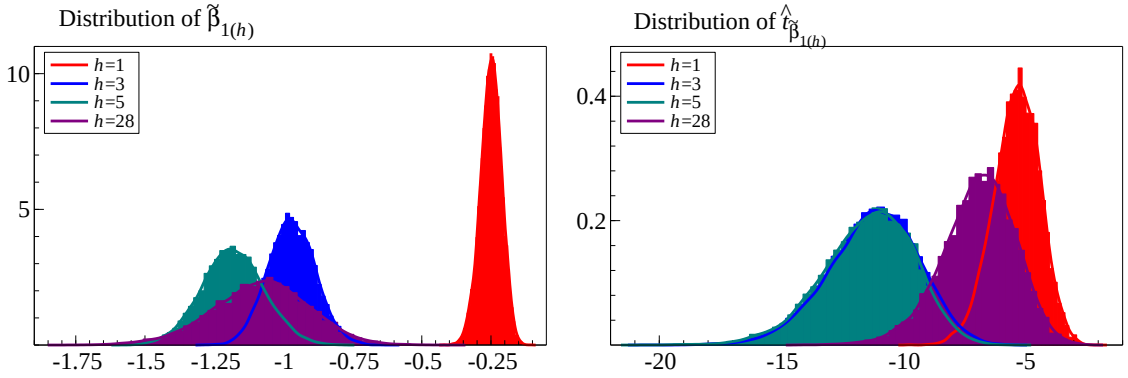


Figure 8: Distribution of $\widehat{\tilde{\beta}}_{1(h)}$ and $\widehat{t}_{\tilde{\beta}_{1(h)}}$ for $T = 200$ and $T_b = 100$ where weak exogeneity switches equation.

generating persistent systematic forecast errors even when the underlying adjustment mechanism remains unchanged.

Table 5 records the RMSFEs for the models where the break occurs at $T_b = 30, 100$ and 170 and the forecast origin is $T = 200$. The results suggest that forecast performance is relatively robust to the change in the adjustment mechanism. The break leaves the cointegrating relation and its equilibrium mean unchanged, so that the system continues to exhibit equilibrium correction, albeit through a different variable.

For 1-step forecasts the VECM, ECM and ADP produce almost identical forecasts, while the DP performs less well because it omits the short run dynamics. At longer horizons, the ECM generally performs best, with RMSFEs that are stable across the different break dates. The VECM and ADP also perform well, although their forecast accuracy deteriorates as the break occurs later in the sample. The DP RMSFEs are almost unaffected by the timing of the break.

Following the break, the relative price is weakly exogenous and the ECM forecasts the nominal exchange rate conditional on a random-walk path for the relative price. This avoids the need to estimate and recursively forecast the second equation of the system, which is potentially affected by the structural break. The VECM, in contrast, estimates and propagates the dynamics of both variables, introducing additional estimation uncertainty and some sensitivity to the timing of the break. The DP avoids recursive propagation altogether, which may explain its relatively stable forecasting performance despite omitting the short-run dynamics.

A substantial change in the economic adjustment mechanism does not generate a corresponding deterioration in forecast performance. This finding illustrates the distinction between models designed to represent economic mechanisms and models designed for forecasting. A model that provides a valid structural representation for policy analysis need not provide the most robust forecasts in the presence of structural change, see Hendry and Mizon (2000).

The DP forecasts outperform the RW so despite the structural break, the DP is still extracting useful predictive information from the current state. This is because the structural break does not change the predictive information in the cointegrating relationship (i.e., cointegration is not lost), but it does make the constant parameter dynamic representation unreliable. The structural break is a change in the adjustment

	VECM	ECM	DP	ADP	RW	DDV	DVECM
			$T_b = 30$				
$h = 1$	**0.202**	**0.202**	0.220	**0.202**	0.240	0.313	0.264
$h = 3$	**0.295**	0.296	0.301	0.296	0.429	0.528	0.456
$h = 12$	0.383	**0.380**	0.385	0.385	0.486	0.587	0.498
$h = 28$	0.510	**0.492**	0.513	0.514	0.576	0.663	0.587
			$T_b = 100$				
$h = 1$	**0.204**	**0.204**	0.222	**0.204**	0.240	0.313	0.257
$h = 3$	0.304	**0.299**	0.302	0.303	0.429	0.528	0.450
$h = 12$	0.389	**0.380**	0.386	0.387	0.486	0.587	0.500
$h = 28$	0.515	**0.492**	0.514	0.516	0.576	0.663	0.590
			$T_b = 170$				
$h = 1$	**0.208**	**0.208**	0.225	**0.208**	0.240	0.313	0.264
$h = 3$	0.321	0.309	**0.304**	0.316	0.429	0.528	0.472
$h = 12$	0.402	**0.380**	0.386	0.393	0.486	0.587	0.519
$h = 28$	0.526	**0.492**	0.514	0.520	0.576	0.663	0.607

Table 5: Root mean square forecast errors (RMSFE) for models (41), (45), (46) & (47) along with benchmark devices (48), (49) & (50) for the DGP given by a structural break in which weak exogeneity switches from Δy_1 to Δy_2. $T_b = 30, 100, 170$ and $T = 200$ with $H = 28$ forecasts and $M = 10{,}000$ replications. Bold denotes smallest RMSFE.

coefficient which does not change the equilibrium mean of the system. Therefore, failure to detect the break is not costly in terms of relative forecast performance, but it is misleading as the forecasts do not provide a signal as to the changed economic interpretation of the DGP. These results highlight that forecast performance is not a reliable criterion for determining whether an econometric model is a good representation of the DGP, see Castle and Hendry (2015) who show that a valid model can forecast badly while a misspecified model can forecast successfully.

5 Application: The *NER*-regression based evidence for Norway and the UK.

We use the bilateral exchange rates for UK and Norway with respect to USD in our application. EJR include UK in the benchmark group of countries that formally adopted inflation targeting in the 1990s.[6] Norway adopted inflation targeting in 2001.

Table 6 shows ADF-tests statistics for the differences of the two *NER*s and for the *RER*, q_t. In the same way as with the simulated data, the ADF-tests fail to reject the unit-root hypothesis for the long horizons of $h = 20$ and $h = 28$. This is relevant since EJR attribute particular importance to the *NER*-regressions that have $\Delta_{20} e_t$ and $\Delta_{28} e_t$variables as dependent variables.

Also for the *RER*, the ADF test fails to reject the null of non-stationarity. Hence care should be taken when assessing the significance of *RER*-lags in the nominal exchange rate equation.

Table 7 shows *NER*-regression results for the same horizons as in EJR's Table 1,

[6]Together with Australia, Canada, New Zealand and Sweden.

Table 6: ADF tests. Norway: Kroner/USD, 2001(1)-2025(4). UK: GBP/USD, 1997(2)-2025(4). Intercept included, no trend.

	$\Delta_1 e_t$	$\Delta_4 e_t$	$\Delta_{12} e_t$	$\Delta_{20} e_t$	$\Delta_{28} e_t$	q_t
Norway	-7.7^{**}	-5.0^{**}	-2.6	-2.0	-1.6	-1.1
UK	-6.0^{**}	-5.2^{**}	-3.0^{*}	-2.6	-1.9	-1.6

Data sources: Appendix A.1 % significance marked by ** and 5 % significance by *.Critical values based on response surfaces in MacKinnon (1991), Doornik and Hendry (2022).

and extended by $h = 1$ given its relevance for direct tests of equilibrium correction and hence non-stationarity of *RER*. We estimate three regressions for each country. The sample starts were set to when the two central banks got their mandates as independent monetary policy authorities, so the difference is the end of sample quarters: 2025(4), 2019(4) and 2008(3), which was also the end-of-sample in the EJR regressions.

For Norway, in the full sample results, we see that the *RER* is significant for the 20 and 28 quarter horizon – i.e., judged by the HAC standard errors. For the sample that ends in 2019(4), *RER* is significant also for the 12 quarter horizon. This is also seen in the line with estimates obtained for the shortest sample. Considering that there are some differences between the sample starts, and quite likely in the measurement of the price indices, there is close similarity between the 2008(3) sample-end results, and those from EJR Table 1 (as noted, the purpose is not exact replication).

We note the numerical size of all estimated β_h^e's is smaller on the full sample than on the two others. In the spirit of EJR, an interpretation of this is that the normal operation of inflation targeting policies was disrupted by the Financial crisis and COVID-19, and therefore the results about *RER*'s strength as a predictor become more muted on samples that include the two crises.[7]

However, the analysis in the previous section suggests a different explanation. If the *NER* and the relative price are not cointegrated, and therefore the *RER* is not stationary, the estimated β_h^e will tend to be reduced in absolute value as the sample length is increased, cf. Figure 2 and the discussion of it.

For the UK, the estimation results show the same pattern, with the remark that they are even more sensitive to the sample length. Both for the pre COVID-19 sample, and the full sample, the estimated β_{28}^e is less than one in absolute value.

Tests of residual mis-specification show that the equations are mis-specified. Again, this is as expected from theoretical analysis, see section 4.5. For the UK, on the sample that ends in 2008(3), already the *NER*-regression for $h = 4$ has massive residual autocorrelation as well as significant ARCH features. The same is the case for the longer horizons, where the residuals in addition are marred by non-normality. The mis-specification problem is not something that is reduced by increasing the sample. If anything, it gets worse. For the UK, the residuals of the $h = 1$ regression are approximately white-noise on the original sample, but on the full sample there is significant autocorrelation.

HCSE calculations of White (1980) assume that the heteroskedasticity is due to changes in the error variance that is reflected in the squared regressor, rather than the

[7]EJR give this type of motivation for excluding data after 2008(3) from their benchmark sample period.

Table 7: Nominal exchange rate regressions. Estimation method OLS. Robust standard errors (HACSE).

$$\Delta_h e_t = \alpha_h + \beta^e_h q_{t-h} + \epsilon_{h,t+h}$$

	β^e_h				
	Horizon in quarters				
	1	4	12	20	28
Norway:					
Sample 2001(1)-2025(4):	−0.02 (0.02)	−0.09 (0.09)	−0.22 (0.21)	−0.63 (0.29)	−1.12 (0.39)
Sample 2001(1)-2019(4):	−0.04 (0.04)	−0.24 (0.13)	−0.70 (0.22)	−1.54 (0.20)	−1.88 (0.13)
Sample 2001(1)-2008(3):	−0.06 (0.04)	−0.19 (0.20)	−0.57 (0.31)	−1.69 (0.39)	−1.93 (0.17)
EJR Table 1	n/a	−0.36 (0.14)	−0.90 (0.21)	−1.68 (0.28)	−1.85 (0.14)
UK:					
Sample 1997(2)-2025(4):	−0.03 (0.04)	−0.17 (0.10)	−0.37 (0.16)	−0.57 (0.17)	−0.42 (0.18)
Sample 1997(2)-2019(4):	−0.03 (0.05)	−0.23 (0.16)	−0.78 (0.31)	−1.32 (0.23)	−0.78 (0.29)
Sample 1997(2)-2008(3):	−0.06 (0.05)	−0.25 (0.15)	−1.13 (0.38)	−1.46 (0.22)	−1.18 (0.26)
EJR Table 1	n/a	−0.26 (0.10)	−1.09 (0.45)	−1.68 (0.15)	−1.08 (0.44)

Data sources and variable transformations: Appendix A

wrong specification of the regression model. It does not seem plausible that using, say q_{t-4} rather than q_{t-1} in a *NER*-regression will make much difference in that regard, yet the $\Delta_1 e_t$-regression residuals are empirically white-noise, while there is significant heteroskedasticity in the $\Delta_4 e_t$-equation's residuals. Hence, mis-specification of the model equation is the likely culprit, here as in much empirical research, Hendry and Nielsen (2007). However, model re-specification is warranted and needed, with the aim of establishing the approximate white-noise errors that underlie valid inference, Spanos (2021).

A feasible method is to augment the *NER*-regression by lags of Δe_t and $\Delta(p^f - p)_t$, as suggested above. For example, equation (35) is a general, fully augmented, *NER*-equation. However, as noted above, when h is large, the estimation of fully augmented *NER*-regression will consume many degrees of freedom. Instead, we analysed the specific ADP (augmented direct projection) given by equation (36).

As an example, consider the ADP-regressions for $h = 1$, $h = 4$ and $h = 28$ for UK, on the original sample, 1997(2)-2008(3):

$$\Delta_1 e_t = \underset{(0.19)}{-0.32} - \underset{(0.04)}{0.06} q_{t-1} - \underset{(0.69)}{0.35} \Delta_1 e_{t-1} + \underset{(0.66)}{0.75} \Delta_1 (p^f - p)_{t-1} + \hat{\epsilon}_{1,t} \tag{53}$$

$$\Delta_4 e_t = \underset{(0.55)}{-1.59} - \underset{(0.11)}{0.32} q_{t-4} - \underset{(1.20)}{1.43} \Delta_1 e_{t-4} + \underset{(1.31)}{2.19} \Delta_1 (p^f - p)_{t-4} + \hat{\epsilon}_{4,t} \tag{54}$$

$$\Delta_{28} e_t = \underset{(1.43)}{-5.43} - \underset{(0.29)}{1.10} q_{t-28} + \underset{(1.14)}{1.92} \Delta_1 e_{t-28} - \underset{(1.24)}{2.22} \Delta_1 (p^f - p)_{t-28} + \hat{\epsilon}_{28,t} \tag{55}$$

In fact, for $h = 1$, *NER*-regression was already well specified, which explains why the two additional regressors in (53) are insignificant. The implied t-value of q_{t-1} can therefore be compared with the critical values of the standard t-distribution, if the stationarity of

q_t is maintained, or from a non-standard, if the null hypothesis under test is $q_t \sim I(1)$. Clearly, the t-value of -1.6 is insignificant in both interpretations.

The estimated projection coefficient in equation (55), -1.1 is practically the same as the original sample results in Table 7. Note that this observation is in line with the theoretical analysis above (section 4.2), which showed that the longer the horizon h, the less the effect of the two lagged differences on the direct projection coefficients.

However, what is also true is that the residuals of (55) are significantly autocorrelated, so the mis-specification issue is not resolved by the inclusion of the two lagged first differences. An alternative approach could be to include $\Delta_h e_{t-1}$ and $\Delta_h(p^f - p)_{t-1}$ in the hope that those regressors will be sufficient to "whiten the residuals". For horizon $h = 1$, this approach gives exactly the same ADP as (53). However, for $h > 1$ it does not yield direct projections, but rather ad-hoc equilibrium correction equations. For example, for $h = 28$, we obtain:

$$\Delta_{28} e_t = \underset{(0.22)}{-1.44} - \underset{(0.18)}{0.29} q_{t-28} + \underset{(0.22)}{1.13} \Delta_{28} e_{t-1} - \underset{(0.25)}{0.35} \Delta_{28}(p^f - p)_{t-1} + \hat{\epsilon}_{28,t} \tag{56}$$

Diagnostic tests show that, unlike for (55), the residuals of (56) are not autocorrelated. Therefore, due to its relative congruency, (56) is the more appropriate model equation for testing the significance of q_{t-28} as a mere explanatory variable for $\Delta_{28} e_t$. The estimated β^e_{28} is -0.29 and the t-value of -1.6 is insignificant when compared to critical values of both standard (Student-t) and non-standard distributions.

Similar results are obtained when augmented *NER*-regressions are estimated for the other horizons in Table 7, for the UK and for Norway. Specifically, also for Norway, the lagged *RER* becomes insignificant in the $h = 28$ regression augmented with $\Delta_{28} e_{t-1}$ and $\Delta_{28}(p^f - p)_{t-1}$.

6 Conclusions

Establishing empirical regularities about exchange rates, and then formulating theoretical mechanisms that can account for them, is a commendable research strategy. Eichenbaum et al. (2021) is a prominent example: the finding that, under inflation targeting, the current real exchange rate predicts future changes in the nominal exchange rate at horizons of three years and longer has been widely cited, and has so far been challenged on its interpretation rather than on its econometric status. Because the regularity is used to discipline models that guide monetary policy, the econometrics of the fact-finding process deserves attention in its own right. In this paper we have focused on the main econometric vehicle used by EJR, the *NER*-regression, and on its key parameter β^e_h, the coefficient of the lagged real exchange rate q_{t-h}.

Our Monte Carlo analysis shows that the sampling distribution of $\hat{\beta}^e_h$, and of its HAC-based t-value, depends critically on the constellation of three factors: the horizon h; whether the *RER* is stationary, i.e., whether the *NER* and the relative price are cointegrated with unit coefficients; and the sample size T. The common thread is that the *NER*-regression for $h > 1$ does not encompass the data generating process. It omits the short-run dynamics that the DGP implies for $\Delta_h e_t$, and as h increases the overlapping innovations make the regressand behave increasingly like an $I(1)$ series in samples of realistic length. Inference about the predictive power of the *RER* is therefore inefficient, and in some constellations invalid. The residual autocorrelation, non-normality and ARCH

that follow are symptoms of mis-specification of the regression, not a nuisance that HAC standard errors can be relied upon to remove.

When the *RER* is stationary, the inefficiency grows with h, since the regression moves further away from the DGP and becomes increasingly unbalanced. The estimated coefficient no longer measures the equilibrium-correction coefficient of the DGP but a horizon-dependent projection coefficient, which in our simulations exceeds the true adjustment coefficient in absolute value. The probability of correctly rejecting the null of no predictive power nevertheless remains reasonably high, at least in large samples: OLS tends to confirm a single predictor in a simple regression also when the DGP implies a multiple regression. It is therefore not the sign of the regularity, but its magnitude and its interpretation as a measure of the adjustment mechanism, that is compromised in this case. In particular, we found that a large shift in the adjustment mechanism, from relative prices to the nominal exchange rate, is very difficult to detect from *NER*-regression estimates.

When the *RER* is non-stationary, the consequences are serious. The t-value of $\hat{\beta}_h^e$ then has a non-standard, Dickey-Fuller type distribution, and comparison with Student-t critical values leads to substantial over-rejection of $\beta_h^e = 0$ already at $h = 1$. For short horizons, the critical values of Ericsson and MacKinnon (2002) are an effective remedy. For horizons of three years and longer they no longer help: With approximately integrated variables on both sides of the equation, $\hat{\beta}_h^e$ estimates something different from an equilibrium-correction coefficient, and the *NER*-regression comes to resemble a spurious regression in the sense of Granger and Newbold (1974), or the first step of the Engle–Granger two-step procedure. Hence, the horizons that EJR regard as most informative are exactly those where the risk of finding a relationship when none exists is largest.

The empirical evidence for Norway and the UK is in line with these results. ADF tests do not reject a unit root in $\Delta_{20}e_t$, $\Delta_{28}e_t$ or in q_t. On samples that end in 2008 we obtain estimates close to EJR's, but the estimated coefficients shrink as the sample is extended, which is what the simulations predict if the *RER* is non-stationary, rather than what would be expected if the regularity were a stable feature of inflation targeting regimes. The *NER*-regressions are clearly mis-specified, and the mis-specification does not diminish as the sample is extended. When the regression is augmented by lagged changes in the *NER* and the relative price, the residuals become approximately white noise, the long-horizon coefficient is reduced substantially in magnitude, and the lagged *RER* is no longer significant, whether judged by standard or non-standard critical values.

These results suggest some practical recommendations for applied work with *NER*-regressions. First, the order of integration of the *RER* should be investigated rather than assumed, since the relevant inference theory depends on it. Second, for large h, it may be more relevant to test the residuals of the *NER*-regression for a unit root than to test $\beta_h^e = 0$ against standard critical values. Hence, it suggests simply using the Engle-Granger test. Third, when the regression is to be interpreted as an adjustment equation, it should be re-specified, for example by augmenting it with lagged differences, until it passes standard mis-specification tests.

This does not rule out that it can be efficient to regress, for example, the fourth difference of *NER* on the fourth lag of *RER*, but that will be because the DGP has that dynamic specification. Without paying attention to the congruency of the *NER*-regression, it is difficult to assess whether a given *NER*-regression is a reasonable approximation to the DGP.

Hence, many of the issues raised come down to the relationship between the estimated model and the DGP that generates the *NER* and the relative prices. Also when the

research purpose is "merely" to establish the significance and strength of the *RER* as a predictor of changes in the *NER*, the advice is to model the DGP as well as possible. Error probabilities will not go away, but they become easier to assess, and the facts that are established are more likely to be trustworthy.

Finally, the merits of the *NER* regression as a forecasting device are best assessed separately, since modelling for explanation and policy analysis and modelling for forecasting place different demands on the model. In our simulations, direct projections forecast almost as well as the correctly specified VECM when the *RER* is stationary, but are outperformed by a random walk when it is not. Forecast performance is also affected by structural change whereby an unmodelled shift in the adjustment mechanism can substantially alter the relative forecasting performance of competing models, with the consequences depending on the timing of the break. Good forecast performance is therefore a weak guide to the validity of the economic interpretation of a model and, in particular, provides limited evidence in support of the assumed stationarity of the *RER*. Forecast success cannot substitute for establishing that the econometric model used to identify an empirical regularity is congruent with the properties of the data. For policy analysis, forecast accuracy alone is therefore insufficient. Hence, also this analysis leads to the conclusion that the model must be well specified if its estimated relationships are to be given a structural economic interpretation, with policy implications.

A Data sources and variable transformations

With the exception of CPI USA, the series are not seasonally adjusted. All series were downloaded as monthly series and aggregated to quarterly series by averaging.

Table 8: Variable definitions and sources

	Definition	Table or Series-ID	Unit
NOK/GBP[1]	Kroner per GBP	EXR_M_GBP_NOK_SP	Kroner
NOK/USD[1]	Kroner per USD	EXR_M_USD_NOK_SP	Kroner
CPI Norway[2]	CPI all items	Table 14700 (all items)	Index
CPI USA[3]	CPI for all urban consumers	CPIAUCNS	Index
CPI UK[4]	CPIH all items	CPIH INDEX 00	Index

[1] www.norges-bank.no/en/topics/statistics/exchange_rates/
[2] www.ssb.no/en/priser-og-prisindekser/konsumpriser/statistikk/konsumprisindeksen
[3] www.fred.stlouisfed.org/.
[4] www.ons.gov.uk/economy/inflationandpriceindices/timeseries/l522/mm23

The GBP/USD exchange rate series was obtained as:

$$\text{GBP/USD} = (\text{NOK/USD}) \times (\text{GBP/NOK})$$

For use in the nominal exchange rate regressions, all series were log-transformed. Specifically, in Table 7: Norway: $e = log(\text{NOK/USD})$, $q = e + log(\text{CPI USA}) - log(\text{CPI Norway})$. UK: $e = log(\text{GBP/USD})$, $q = e + log(\text{CPI USA}) - log(\text{CPI UK})$.

References

Akram, Q. F. (2006). PPP in the medium run: The case of Norway. *Journal of Macroeconomics*, *28*, 700–719.

Castle, J. L. and D. F. Hendry (2015). On Not Evaluating Economic Models By Forecast Outcomes. *Istanbul University Journal of the School of Business Administration*, *40*, 1–21.

Chahrour, R., V. Cormun, P. D. Leo, P. A. Guerrón-Quintana and R. Valchev (2024). Exchange Rate Disconnect Revisited. Working Paper 32596, NBER. https://www.nber.org/papers/w32596.

Chevillon, G. (2007). Direct Multi-Step Estimation And Forecasting. *Journal of Economic Surveys*, *21*, 746–785.

Doornik, J. A. and H. Hansen (2008). An Omnibus Test for Univariate and Multivariate Normality. *Oxford Bulletin of Economics and Statistics*, *70*, 927–939.

Doornik, J. A. and D. F. Hendry (2022). *Empirical Econometric Modelling PcGive 16. Volume I*. Timberlake Consultants, London.

Eichenbaum, M. S., B. K. Johannsen and S. T. Rebello (2021). Monetary Policy and the Predictability of Nominal Exchange Rates. *Review of Economic Studies*, *88*, 192–228.

Engel, C. and S. P. Y. Wu (2023). Forecasting the U.S. Dollar in the 21st Century. *Journal of International Economics*, *141*, 103715. Https://doi.org/10.1016/j.jinteco.2023.103715.

Engle, R. F. (1982). Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation. *Econometrica*, *50*, 987–1007.

Engle, R. F. and C. W. J. Granger (1987). Co-integration and Error Correction: Representation, Estimation and Testing. *Econometrica*, *55*, 251–276.

Ericsson, N. R. (1994). Testing Exogeneity: An Introduction. In Ericsson, N. R. and J. S. Irons (eds.), *Testing Exogeneity*, chap. 1, 3–38. Oxford University Press, Oxford.

Ericsson, N. R. and J. G. MacKinnon (2002). Distributions of Error-Correction Tests for Cointegration. *Econometrics Journal*, *5*, 285–318.

Godfrey, L. G. (1978). Testing for Higher Order Serial Correlation in Regression Equations When the Regressors Include Lagged Dependent Variables. *Econometrica*, *46*, 1303–1310.

Granger, C. W. J., N. Hyung and Y. Jeon (2001). Spurious Regression with Stationary Series. *Applied Economics*, *23*(7), 899–904.

Granger, C. W. J. and P. Newbold (1974). Spurious regressions in econometrics. *Journal of Econometrics*, *2*, 111–120.

Hendry, D. F. (1995). *Dynamic Econometrics*. Oxford University Press, Oxford.

Hendry, D. F. (2000). On Detectable and Non-detectable Structural Change. *Structural Change and Economic Dynamics*, *11*, 45–65.

Hendry, D. F. and G. E. Mizon (2000). Reformulating Empirical Macroeconometric Modelling. *Oxford Review of Economic Policy*, *16*(4), 138–159.

Hendry, D. F. and B. Nielsen (2007). *Econometric Modeling*. Princeton University Press, Princeton.

Itskhoki, O. (2021). The Story of the Real Exchange Rate. *Annual Review of Economics*, *13*, 423–455.

Itskhoki, O. and D. Mukhin (2021). Exchange Rate Disconnect in General Equilibrium. *Journal of Political Economy*, *129*(8), 2183–2232. https://doi.org/10.1086/714447.

Johansen, S. (1995). *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models*. Oxford University Press, Oxford.

Jordà, O. (2005). Estimation and Inference of Impulse Responses by Local Projections. *American Economic Review*, *95*(1), 161–182.

MacKinnon, J. G. (1991). Critical Values for Cointegration Tests. In Engle, R. F. and C. W. J. Granger (eds.), *Long-Run Economic Relationships: Readings in cointegration*, chap. 13. Oxford University Press, Oxford.

Marcellino, M., J. H. Stock and M. W. Watson (2006). A comparison of direct and iterated multistep AR methods for forecasting macroeconomic time series. *Journal of Econometrics*, *135*(1-2), 499–526.

Meese, R. and K. Rogoff (1983). Empirical Exchange Rate Models of the Seventies: Do They Fit Out of Sample? *Journal of International Economics*, *14*, 3–24.

Nielsen, B. (2006). Order determination in general vector autoregressions. In Ho, H.-C., C.-K. Ing and T. Lai (eds.), *Time Series and Related Topics: In Memory of Ching-Zong Wei*, vol. 52 of *IMS Lecture Notes and Monograph Series*, 93–112. Institute of Mathematical Statistics. https://doi.org/10.1214/074921706000000978.

Phillips, P. C. B. (1986). Understanding Spurious Regressions in Econometrics. *Journal of Econometrics*, *33*, 311–340.

Ramsey, J. B. (1969). Tests for Specification Errors in Classical Linear Least-Squares regression analysis. *Journal of the Royal Statistical Society, Series B*, *31*, 350–371.

Spanos, A. (2021). Methodology of Macroeconometrics. *Oxford Research Encyclopedias*. https://doi.org/10.1093/acrefore/9780190625979.013.175.

Stavrakeva, V. and J. Tang (2026). A Fundamental Connection: Exchange Rates and Macroeconomic Expectations. *The Review of Economics and Statistics*, *108*, 1332–1348.

White, H. (1980). A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test of Heteroskedasticity. *Econometrica*, *48*, 817–838.