# Is the Fisher effect asymmetric?
# Cointegration analysis and expectations measurement

David O. Cushman[1], Glauco De Vita[2], Emmanouil Trachanas[3]

[1] Department of Economics, University of Saskatchewan, Saskatoon, Canada.

[2] Centre for Business in Society, Coventry University, Coventry, UK. Corresponding author. Email: glauco.devita@coventry.ac.uk

[3] Department of Accounting and Finance, University of Macedonia, Thessaloniki, Greece.

*International Journal of Finance & Economics*

https://doi.org/10.1002/ijfe.2616

**ABSTRACT**

Using U.S. post-war data, we investigate whether the interest rate response to inflation known as the Fisher effect could be asymmetric. The asymmetry considered is that the long-run change in the interest rate is larger when inflation rises than when it falls. The possibility follows from behavioral hypotheses about the relationship of inflation expectations to actual inflation. Using an asymmetric cointegration approach, we find asymmetric cointegration in the Fisher effect for the post-war period through 1979, but not subsequently. We then find that, starting in 1980, a breakdown developed in the relationship between inflation expectations from surveys and actual recent inflation rates, a breakdown not accounted for by asymmetry. If the survey results approximate true expectations, then econometric testing using actual recent inflation to compute expected inflation will suffer from mismeasurement, which could explain the finding of no cointegrating Fisher effect post-1979. The paper accounts for breakpoints and uses bootstrapping to conservatively estimate statistical significance.

**KEYWORDS:** Asymmetric cointegration, bootstrapping, Fisher effect, inflation expectations

**JEL CLASSIFICATION:** C22; E31; E43; E71

## 1 | INTRODUCTION

The relationship known as the Fisher effect, that a change in expected inflation will change the interest rate in the same direction by the same amount, has been a cornerstone of monetary and macroeconomic theories. Consistent empirical support for the Fisher effect has, however, remained elusive despite extensive research. A frequent conclusion is that the interest rate response to inflation in the long run is less than one for one, or even that there is no long-run relationship at all. If so, real interest rates respond permanently to monetary shocks and are nonstationary, effects that are problematic for various standard macroeconomic models (Neely & Rapach, 2008). Researchers have considered numerous reasons for the failure to resolve the Fisher effect puzzles, such as mismeasurement of expectations, use of short- versus long-term rates, breakpoints, and choice of estimator (Caporale & Pittis, 2004). Various nonlinearities have also been proposed (e.g., by Christopoulos & León-Ledesma, 2007), which if unspecified could impede accurate estimation.

This paper's purpose is to examine the Fisher effect while accounting for mismeasurement of expectations based on possible nonlinearity in the form of asymmetric expectation formation. The argument for asymmetry draws from behavioral economic hypotheses and findings related to loss aversion (Tversky & Kahneman, 1991) and a model of household inflation expectation formation with uncertain inflation news (Baqaee, 2020). People don't like inflation (Shiller, 1997), and they tend to notice or take seriously increases in its rate more than decreases. If inflation expectation formation corresponds to these responses, then the rise in the interest rate when the inflation rate increases will be larger than the fall in the interest rate when the inflation rate decreases.

Using quarterly U.S. post-war data, 1953:1-2019:1, we test whether such asymmetry has existed in the Fisher effect using the asymmetric cointegration procedure of Shin, Yu, and Greenwood-Nimmo (2014). The model assumes that the interest and inflation rates are nonstationary, I(1) variables, a typical finding in the literature and one our own tests support. Meanwhile, because the hypothesized asymmetric expectations are asymmetrically erroneous, there is likely to at some point be a noticeable divergence between perception and reality. In accordance with Hong, Stein, and Yu (2007), the divergence, although ignored for a while, would eventually bring on a regime change or break, and we therefore pay detailed attention to the specification of breaks.

We find statistically significant asymmetric cointegration between the interest rate and inflation rate for the post-war period until the end of 1979. The statistical significance is estimated via bootstrapping, which generates much more conservative results than does use of tabulated critical values. The asymmetry is in the direction predicted by our model: the interest rate responds in the long run more to increases than to decreases in inflation. Nevertheless, the estimated interest rate responses remain much smaller than the inflation rate changes.

In contrast, after 1979 we do not find cointegration, symmetric or asymmetric, at statistically significant levels with any credibility. Our post-1979 results contradict several recent papers. The contradiction could reflect these papers' specification of inaccurate breakpoints or use of tabulated critical values that are likely to give oversized results. However, if the Fisher effect did hold post-1979, our results could reflect the declining ability of actual recent inflation to be a basis for inflation expectations in econometric tests. Using survey forecasts as a measure of inflation expectations, Blanchard (2018) finds such a decline within the 1980-2016 period. Using a data set starting much earlier, we confirm the breakdown, finding a strong relationship between actual inflation and forecasts until 1980, but a poor relationship afterward, particularly in the recent post-financial crisis era. Thus, actual inflation is a plausible basis for expectation measurement for the period which gives Fisher-effect asymmetric cointegration, while the post-1980 deterioration in the actual inflation-expectations relationship can explain our finding of no cointegration in the Fisher effect from 1980 onward.

## 2 | A CRITICAL SYNTHESIS OF RELATED LITERATURE

To clarify the form of the asymmetry we investigate, and to compare our approach with previous empirical work on the Fisher effect, let us start with the basic Fisher equation:

$$i_t = \alpha_t + \beta_t E_t \pi_{t+1} + \varepsilon_t. \qquad (1)$$

The nominal interest rate $i_t$ is the rate determined in period *t* for a bond maturing in period *t* + 1, and $\pi_{t+1}$ is the inflation rate between periods *t* and *t* + 1. $E_t$ indicates expectations formed in period *t*. In the basic case, $\alpha_t$ and $\beta_t$ are constant, $\beta_t = \beta = 1$, and $\varepsilon_t$ is stationary with a mean of 0. Estimation of the equation requires some observable measure of $E_t\pi_{t+1}$, and an estimation approach that avoids econometric problems such as errors-in-variables bias from mismeasurement of $E_t\pi_{t+1}$. The value used in the literature for expected inflation is generally the actual inflation rate, $\pi_t$ or sometimes $\pi_{t+1}$ under an assumption of rational expectations. The result of central interest is the estimate of $\beta$.

Table 1 gives a list of papers examining the Fisher effect, and summarizes their key empirical approaches and findings. Fisher (1930) concluded that $\beta < 1$, and suggested the explanation was money illusion. Many other researchers listed in Table 1 have also concluded that $\beta < 1$. Although the Tobin-Mundell effect does suggest $\beta < 1$, this is probably a rather small effect, and should be swamped in data starting in the mid-20th century by a tax effect giving $\beta \approx 1.3$ (Darby, 1975; Summers, 1983). The finding that $\beta < 1$ is thus a classic puzzle, although a number of more recent papers in Table 1 have not rejected the hypothesis that $\beta = 1$. Table 1 also shows that over the last 30 years or so researchers have often concluded that $i_t$ and $\pi_t$ are well modeled as unit root (I(1)) variables. Consequently, researchers have looked for cointegration between $i$ and $\pi$. If it is found, there is a long-run relationship between the two variables and $\alpha$ and $\beta$ define the cointegration vector. The approach also avoids the bias from random mismeasurement of expected inflation (Stock, 1987; MacDonald & Murphy, 1989).

**< Table 1 here >**

In the cointegration context, the finding that $\beta < 1$ not only contradicts the standard Fisher effect, but also implies that the real interest rate is nonstationary (even if $\alpha_t$ is constant), which is another theoretical contradiction (Rose, 1988). The (expected) real interest rate $r = i_t - E_t\pi_{t+1} = \alpha + (\beta - 1)E_t\pi_{t+1} + \varepsilon_t$, and thus if $\beta < 1$ then *r* incorporates the random walk component in the inflation rate.

Papers have often allowed for the possibility that $\alpha$ and $\beta$ are not constant. One way this has been specified, seen in many papers in Table 1, follows from some major change in the economic environment, usually the change in U.S. monetary policy in 1979. Results from papers in Table 1 that do not specify such changes could be wrong. Because changes could have affected $\alpha$, $\beta$, and lag parameters for the error term in (1), various papers have therefore allowed for breaks by dividing the data into subsamples, or simply used a dummy variable, although a dummy variable only allows a change in $\alpha$. However, in most of these papers $\beta$ is constant within the subsamples, or even in the full sample in the dummy variable approach.

Christopoulos and León-Ledesma (2007) consider a more radical case for fluctuations in $\beta$ in which $\beta$ is a nonlinear function of the inflation rate: $\beta_t = f(\pi_t)$ in Equation (1). They report evidence for of this, which appears mostly for the post-1979 period. They argue that the effect could reflect Federal Reserve policies that vary with the inflation rate. The result is a form of asymmetry in which $\beta$ is smaller under low inflation than under high inflation. However, Haug, Beyer, and Dewald (2011) contradict the result, being unable to reject a linear cointegration vector. Tsong and Lee (2013) examine the Fisher effect using quantile cointegrating regression. Their finding of different quantile regression coefficients is a symptom that $\beta$ could be a possibly nonlinear and asymmetric function of shocks, i.e., $\beta_t = f(\varepsilon_t)$, an interpretation discussed by Xiao (2009). Million (2004) considers nonlinearity and asymmetry in the adjustment toward cointegrating equilibrium in the Fisher equation (1). The rate of adjustment to equilibrium depends on the size and sign of the deviation, $\varepsilon_t$, from disequilibrium.

## 3 | THE MODEL: ASYMMETRY IN THE FISHER EFFECT

To explain the asymmetry we consider, we start with the finding that households significantly dislike inflation, because, among other things, they suspect that the purchasing power of their incomes is likely to fall and that inflation creates situations where the unscrupulous can take advantage of them (Shiller, 1997). The expectation of a fall in purchasing power is not rational for inflation that is a purely monetary phenomenon, but it is reasonable for inflation related to a negative supply-side shock (Mankiw, 1997). Next, loss aversion and reference dependence (Kahneman & Tversky, 1979; Tversky & Kahneman, 1991) suggest that households view the loss from an increase in inflation as larger than the gain from a same-sized decrease in inflation, with the previous inflation rate being the reference point. The stronger impact on perceptions of bad relative to good events is also discussed in the general and economic psychology literature (Bates & Gabor 1986; Baumeister, Bratslavsky, Finkenauer, & Vohs, 2001; Bruine de Bruin, van der Klaauw, & Topa, 2011). Soroka (2006) finds that the media are more likely to report bad economic news than good, which would contribute to households being more aware of increases in inflation than decreases. Ranyard et al. (2008) and Detmeister, Lebow, and Peneva (2016) present summaries of the issue, with reference to recent surveys of inflation perceptions and expectations. Consumer perceptions are biased in the direction of sensitivity to higher inflation.

Baqaee (2020) provides a related model of asymmetric inflation expectations. Households view higher inflation as raising costs more than equal disinflation lowers them, and households have Knightian uncertainty about the quality of price change information (e.g., news of higher inflation could be false). They pursue a mini-max strategy, and thus particularly sensitive to worst-case scenarios. News of rising inflation is noisy and thus may be right or wrong, but the worst-case scenario is that it is right, and therefore it is given weight. News of falling inflation may also be right or wrong, but now the worst-case scenario is that it is wrong, and therefore it tends to be ignored. Households' expectations are, therefore, once again, biased toward sensitivity to increases in inflation. Using the Michigan Survey of Inflation Expectations over the years 1983-2015, Baqaee (2020) finds that households do indeed adjust their inflation expectations more to increases in recent inflation than to decreases. However, he finds that the Philadelphia Fed Survey's professional forecasters do not exhibit the asymmetry. The importance of asymmetry for the Fisher effect thus depends on the importance of households relative to professional forecasters in financial markets. Some financial professionals might be more like the households than the professional forecasters in their expectations.

Ball (2000) and Yellen (2007) argue on near-rationality grounds that people may simply use recent past inflation to form expectations of future inflation rather than behaving as if they use a more complete economic model. Use of simple models for forecasting is also hypothesized by Hong, Stein, and Yu (2007). Meanwhile, the behavioral arguments and survey findings we have noted suggest that it is perceptions of recent inflation rather than actual recent inflation that matter for expectations. Combining the ideas, we hypothesize that expected inflation is an asymmetric function of recent inflation (actual recent inflation is used as a proxy for expectations in every relevant paper in Table 1):

$$E_t\pi_{t+1} = \gamma + \psi P_t\pi_t. \tag{2}$$

In Equation (2), $P$ is an asymmetrical "perceptions" operator. $P_t\pi_t$ is the rate of inflation perceived (or paid attention to) in period $t$ to have occurred between periods $t-1$ and $t$. The $t$ subscript on $P$ allows for occasional changes in the perceptions operator.

Next, we assume that period $t$'s perceived inflation rate is period $(t-1)$'s perceived rate plus period $t$'s perceived change in the inflation rate. Likewise, period $(t-1)$'s perceived rate is period $(t-2)$'s perceived rate plus period $(t-1)$'s perceived change, and so on. Thus, today's perceived inflation rate is the sum of its past perceived changes (plus the initial perceived rate). We assume for now that past perceptions are not revised. Let $\Delta\pi_j = (\pi_j - \pi_{j-1})$. Then:

$$P_t\pi_t = P_0\pi_0 + \sum_{j=1}^{t}\tilde{P}_j\Delta\pi_j. \tag{3}$$

The perception asymmetry for the level of inflation in (3) comes from asymmetry in past perceived *changes* in the inflation rate: $\tilde{P}_j\Delta\pi_j = \delta_j^{+}\Delta\pi_j$ if $\Delta\pi_j > 0$, and $\tilde{P}_j\Delta\pi_j = \delta_j^{-}\Delta\pi_j$ if $\Delta\pi_j \leq 0$, with $0 <$

$\delta_j^- < \delta_j^+$ in line with our behavioral economic assumption. Finally, let us (for now) assume that the perceptions operator for changes in the inflation rate is constant. Thus, the $\delta$'s are constant, and we drop their subscript $j$. With various substitutions, we obtain:

$$E_t\pi_{t+1} = \gamma + \psi(P_0\pi_0 + \delta^+\pi_t^+ + \delta^-\pi_t^-) \quad (4)$$

where we have the partial sum processes:

$$\delta^+\pi_t^+ = \sum_{j=1}^t \delta^+\Delta\pi_j^+ = \sum_{j=1}^t \max(\delta^+\Delta\pi_j, 0), \quad (5a)$$

$$\delta^-\pi_t^- = \sum_{j=1}^t \delta^-\Delta\pi_j^- = \sum_{j=1}^t \min(\delta^-\Delta\pi_j, 0). \quad (5b)$$

And, finally, Equation (1) becomes:

$$i_t = \alpha + \beta\gamma + \beta\psi(P_0\pi_0 + \delta^+\pi_t^+ + \delta^-\pi_t^-) + \varepsilon_t. \quad (6)$$

Equation (6) presents the possible asymmetrical relationship of the interest rate with the inflation rate. Increases in the inflation rate tend to increase the nominal interest rate more than decreases in inflation reduce it. The simple Fisher effect, where the responses are symmetric, does not hold. If $i_t$ and $\pi_t$ are I(1) variables, then (6) is the basic asymmetric cointegration equation of Shin et al. (2014), which, with additions to deal with serial correlation and endogeneity (discussed below), leads to the Nonlinear Autoregressive Distributed Lag (NARDL) model. Like $\alpha$ and $\beta$, the parameters $\psi$, $\delta^+$, and $\delta^-$ could vary over time, but in Equation (6) we omit $t$ subscripts on these coefficients to lessen the clutter. The cointegration vector coefficients $\beta\psi\delta^+$ and $\beta\psi\delta^-$ in Equation (6) are equivalent to the value we call "$\beta$" in Table 1 for past papers' Fisher effect results. Past papers in effect assume that $\psi = \delta^+ = \delta^- = 1$.

Our model of inflation expectations is not rational, as eventually the perceived inflation rate could be significantly different from the actual inflation rate. At some point households will surely notice this and adjust their perceptions. This is consistent with the behavioral-finance, regime-change model of Hong, Stein, and Yu (2007). People prefer a simple model (such as one based only on recent inflation), and do not abandon it unless significant evidence against it appears. When this happens (here, when the inflation forecast errors become sufficiently large and persistent, perhaps noticed with the catalyst of a monetary policy regime change), the existing simple model is abandoned in favor of a now more plausible simple model. In our application, such a regime change would be reflected by a change in the parameters in (6), or a resetting of the calendar date associated with period 0. Either way, past perceptions are now adjusted. Thus, we should account for breaks in our empirical testing.

## 4 | THE ASYMMETRIC COINTEGRATION TESTS AND THE DATA

Following Shin et al. (2014), we convert (6) to an error correction equation, an extension of the ARDL equation of Pesaran, Shin, and Smith (2001). The result is the NARDL equation:

$$\Delta i_t = \mu + \rho i_{t-1} + \theta^+ \pi_{t-1}^+ + \theta^- \pi_{t-1}^- \\ + \sum_{j=1}^{p} \gamma_j \Delta i_{t-j} + \sum_{j=0}^{q} \left( \phi_j^+ \Delta \pi_{t-j}^+ + \phi_j^- \Delta \pi_{t-j}^- \right) + e_t, \quad (7)$$

where $\rho$ is the error correction adjustment coefficient, $\beta\psi\delta^+ = -\theta^+/\rho$, and $\beta\psi\delta^- = -\theta^-/\rho$. The values $\beta\psi\delta^+$ and $\beta\psi\delta^-$ are the cointegration vector coefficients for $\pi_{t-1}^+$ and $\pi_{t-1}^-$when the vector coefficient for $i_{t-1}$ is normalized as $-1$. Then, $\mu = \alpha + \beta\gamma + \beta\psi P_0 \pi_0$, and the remaining terms in (7) control for endogeneity and possibly asymmetric dynamics. If $\rho = 0$, then (7) is a first difference equation with no cointegration. After estimating Equation (7) with OLS, Shin et al. (2014) apply two tests for cointegration. One is the *t* test of Banerjee, Dolado, and Mestre (1998) for the null of $\rho \geq 0$ against the alternative of $\rho < 0$. The other is the *F* test of Pesaran et al. (2001) for the joint null of $\rho = \theta^+ = \theta^- = 0$. Finally, if the non-cointegration null is rejected, then a test of the null of long-run symmetry is of interest. As in Shin et al. (2014), we apply a chi square test of the null that $-\theta^+/\rho = -\theta^-/\rho$, equivalent to $\beta\psi\delta^+ = \beta\psi\delta^-$.

Our data are quarterly from 1953:1 through 2019:1 and were downloaded from FRED (dataset is available from the authors upon reasonable request). The starting date accounts for Fama's (1975) points that interest rates were controlled prior to 1953 and that important improvements in the accuracy of the consumer price index (CPI) came into effect in 1953. Thus, Fama (1975) advises that data prior to 1953 should not be used to investigate the Fisher effect. The interest rate is the average three-month treasury bill rate in the third month of a given quarter. We then use an annualized three-month inflation rate available to economic agents to form inflation expectations in the third month. Because we assume that the economic agents are using a very recent inflation rate to form their expectations, we use the most recently available annualized three-month CPI inflation rate. This is the three-month inflation lagged one and one-half months. Thus, our first usable observation is 1953:2. Since we use quarterly values at a quarterly, and not, say, a monthly frequency, we avoid the overlapping data problem possible in some papers in Table 1.[1]

[1] Use of overlapping data (e.g., quarterly inflation rates at monthly intervals) tends to generate moving average serial correlation in a regression, which is difficult to control for with standard autoregressive techniques.

## 5 | TESTING FOR LINEARITY, UNIT ROOTS, AND BREAK POINTS

We begin the empirical work by testing whether the interest and inflation rate variables can be modeled as linear autoregressive processes. If so, then doubt would be cast on the effort to search for any sort of nonlinearity in their relationship. If, however, linearity is rejected, then the asymmetric responses we have proposed could be at least a partial explanation. We apply two tests: the BDS test of Brock, Dechert, Scheinkman, and LeBaron (1996), a broad misspecification test often used to test for linearity, and the Harvey and Leybourne (2007) test (HL), which is specifically a test of the null of linearity versus nonlinearity. The HL test has the advantage that the tested variable can be either I(0) or I(1).

We apply the BDS test to the residuals of a (linear) AR(4) and AR(8) model for each variable. Acceptance of the null of i.i.d. residuals for a given AR model suggests the model does not suffer from significant omitted variable bias (among other problems) and consequently nonlinear processes including asymmetrical ones in the original series are trivial or absent. But a rejection allows for the possibility of nonlinearity in the original series.

We apply the HL test to the series themselves. The null in this case is that a linear AR model is sufficient with no nonlinear terms needed. The HL procedure examines this by testing the significance of certain nonlinear terms added to a linear AR model. The nonlinear terms are from a Taylor series expansion meant to approximate the unknown nonlinear component.

Table 2 shows the BDS test results from various specifications of the test statistic using residuals from the AR(8) model. The AR(4) results are quite similar; the AR(8) results seem preferable because of a statistically significant lag 8 coefficient in the AR model for each variable. The null of i.i.d. residuals is rejected. Table 3 shows the HL test results for three choices of the desired significance level. The null of zero coefficient values for the nonlinear terms is rejected.

Unfortunately, these results do not give much guidance on the form of any nonlinearity. In fact, the BDS test rejections could reflect many other factors, not necessarily nonlinearity. Meanwhile, Harvey and Leybourne (2007) state that their Taylor expansion in the HL test is reasonable for ESTAR and LSTAR processes. But the test probably has power against other nonlinear processes including the asymmetry that the NARDL test focuses on.

**< Tables 2 and 3 here >**

Now we examine whether our two variables are more likely I(0) or I(1). Although the NARDL procedure allows for a mixture of I(1) and I(0) variables, interpretation of the results will be aided with additional information about our two variables' orders of integration. If both are I(1) and the NARDL tests indicate a relationship, it will be a long-run cointegrating one, which could be the Fisher effect as in Equation (6). However, if one variable is I(0), then any detected relationship (Fisher or otherwise) between the two will only reflect a correlation between stationary components, but not long-run co-movement. If both variables are I(0) and correlated, the NARDL tests will likely indicate a relationship that could be a Fisher effect, but we wouldn't call it cointegration.[2] Determining the level of integration of each variable also helps us formulate plausible data generating processes for our bootstrapping.

As shown in Table 1, many authors have already concluded that both variables are I(1). We corroborate the conclusion using unit root tests that also allow for breaks as suggested by, e.g., Hong et al.'s (2007) behavioral argument. The tests are those of Carrion-i-Silvestre, Kim, and Perron (2009), using generalized least squares (GLS) detrending to gain power. They allow data-determined breaks (under both the null and alternative hypotheses) and estimate the breakpoint dates. We specify constants in first differences under the null, and deterministic trends in levels under the alternative hypothesis, with breaks in levels and trends.

We apply the tests specifying 0, 1, and 2 breaks. Lag order is from the modified Akaike Information Criterion (AIC) of Ng and Perron (2001). The maximum first-difference lag order is from Schwert's (1989) formula $maxlag = int(\ell(T/100)^{.25})$ with $\ell = 12$, the value working best with large moving average errors; *maxlag* = 15 for our full sample. Exceptions to using $\ell = 12$ are noted later in the paper when they occur. Table 4 presents the modified $Z_\alpha$ ($MZ_\alpha$) and *ADF* results. The unit root null is never rejected at anywhere near the 0.05 level.[3]

**< Table 4 here >**

The tests indicate break dates of 1972:2, 1980:1 (or 1980:2), and 2008:3. The dates give the final quarter of the pre-break subsample. The first two are approximately those used by one or more papers in Table 1, and the third coincides with the 2008 financial crisis. However, Gauss

[2] Here are two examples (without asymmetry, for clarity). Suppose $x_t$ and $u_t$ are i.i.d. with 0 means. Suppose $y_t = y_{t-1} + \beta x_t + u_t$. Then *x* is I(0) and *y* is I(1) and the equation is equivalent to the ARDL equation $\Delta y_t = \beta x_{t-1} + \beta \Delta x_t + u_t$ . The ARDL error correction term is absent and thus its sample *t* statistic would likely be small, but the ARDL *F* test could well be significant from the non-zero $\beta$ coefficient for $x_{t-1}$. Now suppose instead that $y_t = \beta x_t + u_t$. Both *x* and *y* are now I(0). The corresponding ARDL equation is $\Delta y_t = -y_{t-1} + \beta x_{t-1} + \beta \Delta x_t + u_t$. This may well lead to results that could be interpreted as cointegration, but it is only a trivial version since both variables are I(0) and any linear combination is stationary.

[3] The same testing procedure indicates the first differences are probably stationary, supporting the I(1) conclusion for the variables in levels (results not reported to conserve space but available from the authors upon request).

code from Carrion-i-Silvestre does not give their statistical significance. To estimate it, we assume the two variables are related under the null hypotheses of no cointegration and no asymmetry, and first test the significance of the $MZ_{\alpha}/ADF$-suggested breaks using a bi-variate, first-difference vector autoregression (VAR), allowing all parameters to change at the breaks. We then apply multiple breakpoint tests of Bai and Perron (1998, 2003).

We begin with Wald chi square tests of the null that all constants and lag parameters are unchanged across the three sub-periods defined by the two break dates. The lag order for the unconstrained VAR is determined using the $MZ_{\alpha}$/ADF-suggested breakpoints and the (standard) AIC; *maxlag* = 15 as before. The chosen lag order is 3. Given uncertainty about the $MZ_{\alpha}/ADF$-suggested dates, we apply the tests to the break dates one quarter before and after the $MZ_{\alpha}$ dates in addition to the $MZ_{\alpha}/ADF$ dates of 1972:2, 1980:1, and 2008:3. With three break dates and the $\pm 1$ quarter window, there are 27 tests. The three dates giving the most significant rejection of the hypothesis of no breaks are 1972:1, 1980:1 and 2008:2. The chi square statistic is 141.6 with 42 degrees of freedom with a p-value of 1.0E-12. Thus, there are surely breaks. But can they be attributed to changes in only the constants and thus be specified simply with constant dummies? The chi square statistic for the null of no differences among the constants between time periods is 5.39 with 6 degrees of freedom, giving a p-value of 0.50. The breaks are thus attributable to changes in the lag parameters. Indeed, the null of lag parameter equality is rejected with a chi square value of 129.92 with 36 degrees of freedom and a p-value of 1.5E-12. Thus, specifying a cointegration test only with constant dummies for the breaks will likely give misleading results.

Are the breaks all significant? Chi square tests for differences in the VAR parameters show that the breakpoints between sub-periods 2 and 3, and between 3 and 4 are the significant ones, giving breakpoint dates of 1980:1 and 2008:2. See Table 5.

**< Table 5 here >**

Our second approach to determining breaks employs the tests of Bai and Perron (1998, 2003). They allow the researcher to (1) estimate one or more break dates for some or all coefficients in a regression, (2) compute the break dates' statistical significance, and (3) generate confidence intervals for the break dates. We continue to assume all coefficients could change at the break dates. As with the $MZ_{\alpha}$ test, the Bai-Perron approach tests for breaks in one regression, not in two simultaneously in a VAR. However, we do use the first difference VAR specification in that we apply the Bai-Perron approach to each equation in the bivariate VAR.

The optimal lag order for the regressions is first estimated. In order to start afresh with a lag-length choice independent of past lag-length choices, and yet proceed under the nulls of no cointegration and no breaks as well as the notion that the two variables are related, we determine the lag orders using the AIC from the bivariate first difference VAR for the full sample period with no breaks. Unfortunately, the Bai-Perron Gauss code does not work with the 19 right-hand variables that result from the Schwert-formula's *maxlag* = 15. But with *maxlag* = 8, the AIC lag order is 4 and the code works (our first exception to $\ell = 12$ in the Schwert formula.)

Bai and Perron (1998, 2003) present many methods for finding breakpoints. However, their preferred approach is to use the sequential approach with the supF test, after having first checked the UD max (or WD max) test for the null of no breaks. The results are in Table 6. The UD max test clearly rejects the null of no breaks for both variables' equations. For each equation, the supF sequential approach rejects the null of 0 breaks versus 1, but not 1 versus 2, which terminates the procedure (the null of 2 versus 3 breaks is also not rejected). For the interest rate the breakpoint is 1980:3 and for the inflation rate it is 2008:2. The interest rate date is similar to the one we previously estimated (just two quarters later) and the inflation rate date is exactly the same. The previously suggested break dates of 1980:1 and 2008:2 clearly lie within the Bai-Perron confidence intervals. The various breakpoint tests thus give three subsamples.

**< Table 6 here >**

In the NARDL tests, we employ various ending and starting dates for the subsamples implied by the estimated break dates. This accounts for uncertainty about the break dates and for the possibility of gradual transitions. Including transition periods in subsamples could generate distortion. The beginning date for the first subsample is always the start of our data set, 1953:2. The ending dates range from two quarters before the $MZ_\alpha/ADF$/VAR suggested break date of 1980:1 through the Bai-Perron break date of 1980:3, that is, the end dates are 1979:3-1980:3. The 1979:3 date is the lower limit of one of the Bai-Perron confidence intervals.

For the second subsample, the breakpoint analysis suggests possible starting dates of 1980:2 from the $MZ_\alpha/ADF$/VAR-Wald analysis through 1980:4 from the Bai-Perron analysis. Both analyses suggest an ending date of 2008:2. We also examine results from later starting dates and earlier ending dates to account for transition periods and the 1981-1982 recession. The third subsample starts in 2008:4, derived from the $MZ_\alpha/ADF$ breakpoint and Bai-Perron upper confidence-interval date, and runs through the end of our data set, 2019:1. For each subsample, no data prior to that period are used for lags; thus, there is no overlap in the data used for the subsamples.[4]

## 6 | EMPIRICAL COINTEGRATION ANALYSIS

### 6.1 | Procedure

In estimating the NARDL Equation (7), we employ two methods to determine the lag order. First, the AIC is used, and $p = q$. In the second method, we start with the same *maxlag* value, but use sequential elimination of regressors (SEQ), suggested by Brüggemann and Lütkepohl (2001) and a version of general-to-specific modeling. It is the method used by Shin et al. (2014) in their empirical example. Starting with the maximum lag order, the least significant differenced variable is dropped, the equation re-estimated, and once again the least significant differenced variable is dropped. This is done until all remaining differenced variables give a p-value of at most 0.05 according to standard OLS and $t$-table calculations. Thus, $p$ might not equal $q$, and some number of shorter lag coefficients will be set to zero. With both lag-order choice methods, after the lag order is chosen but before the NARDL statistics are computed, we adjust the sample period to the maximum possible sample size for the specified data period.

Asymptotic critical values for the $t$ and $F$ cointegration tests are in Pesaran et al. (2001) and for the symmetry tests in chi square tables with one degree of freedom. However, many papers have found for other unit root and cointegration tests that p-values using the tabulated values are likely to be oversized. Thus, although we report whether test results give rejections with the Pesaran et al. (2001) and chi square critical values, we rely on bootstrapped results for our conclusions. Shin et al. (2014) also mention the problem and apply a bootstrap approach.[5]

---

[4] A popular approach for estimating a time series model with breaks is to use some form of the smooth threshold/transition autoregressive (STAR) model. STAR models can be, of course, very useful, but we do not use one for several reasons. (1) STAR models generally specify that a regime change occurs from a triggering value of a specific variable in the model. Our break-date determining approaches do not require this to be specified. (2) STAR models generally have the same specification of variables (such as their lag orders), error distributions, and cointegration restrictions (if any) across the various regimes. We do not impose these restrictions. (3) In addition to estimating parameters within regimes, STAR models must also specify transition functions. We avoid this additional specification by simply having gaps between regimes. And (4), most importantly, standard STAR models cannot easily model or test for the asymmetric responses in a cointegration framework, which is fundamental to our paper. For example, Rothman et al. (2001) are forced to use a prespecified cointegrating vector, rather than an estimated one, in their smooth-transition, money-output analysis.

[5] We do not bootstrap the already reported MZA results because they do not give any values close to rejections.

Our bootstrap approach generates data sets that have characteristics similar to the real data under the relevant null hypothesis. For the tests of no-cointegration, for each estimation period we estimate two first difference bivariate VAR models of $i$ and $\pi$. In the first VAR the lag order is determined by the AIC, and in the second VAR the lag order is determined by the Schwarz Bayesian (BIC) approach. For the chi square test of asymmetric cointegration vector coefficients, two vector error correction (VECM) models are estimated. The data generating process (DGP) based on the VAR has no cointegration and no asymmetry, while one from the VECM has cointegration but no asymmetry. To create simulated data sets with the estimated VAR or VECM parameters, we also need random residuals (shocks) to apply to the equations over time, with their distribution to be similar to that in the actual data under the null.

We wish our selected residuals to account for: (1) serial correlation; (2) homo- or heteroskedasticity; and (3) normality or non-normality. Ideally the AIC or BIC give lag orders whose VAR or VECM lag orders account for serial correlation. We can then create random residuals that possess the other two distributional characteristics of the residuals seen in the VAR or VECM model. To check for the three characteristics, we apply to every VAR/VECM the Ljung-Box Q test for serial correlation, a Lagrange Multiplier test for heteroskedasticity, and the Jarque-Bera test for non-normality. Results are in the Appendix. Serial correlation does not appear to be a problem, except sometimes the BIC is apparently too short. However, there are many rejections of homoskedasticity and normality, sometimes both. It would not be easy to account for both problems in a DGP unless we were willing to assume specific forms, but we have insufficient information to assume specific forms. However, lack of normality is the more frequent problem among the results. Therefore, we most often account for that. When heteroskedasticity seems more of a problem, we use a wild bootstrap based on Gonçalves and Kilian (2004). In either case, we make use of the estimated VAR's residual matrix, which embodies any non-normality, heteroskedasticity, and also contemporaneous correlation.

The bootstrap procedure uses the initial, actual data values as starting values for the lagged variables, and for each replication recursively builds up simulated data sets using the estimated VAR parameters and residuals. If the residuals seem to be normal and homoskedastic, or if non-normality seems to be a greater problem than heteroskedasticity, the procedure randomly selects residual pairs from rows of the estimated VAR's residual matrix. This accounts for contemporaneous correlation and any non-normality. If heteroskedasticity seems more of a problem, the wild bootstrap is employed. In recursively building up the simulated data sets, the pair of residuals for time period $t$ are randomly selected from the bivariate normal distribution defined by the covariance matrix computed from the actual VAR residuals for periods $t - 1$ through $t + 1$.[6] To minimize the influence of initial conditions, we follow Chang (2004): each simulated data set starts after additional simulated values. We use 30.

From each such simulated data set, we compute the $t$ and $F$ statistics, or the chi square statistic if testing asymmetry, having chosen the NARDL lag order with the AIC and SEQ just as with the actual data. Hence, the method accounts for the influence of lag parameters, possible non-normality and contemporaneous correlation (or heteroskedasticity if using the wild bootstrap), and the effect of choosing the lag order. Accounting for lag-order choosing gives what Murray and Nelson (2000) call exact p-values. For each DGP there are 4,000 replications when the AIC determines the NARDL lag order, and 2,000 replications when the SEQ does (fewer replications with the SEQ because the SEQ estimations take much longer). The resulting distributions of the simulated statistics are compared with the actual values to get the p-values.

### 6.2 | Results

Table 7 gives the results for the first-period samples ending in 1979 and 1980. Using the tabulated critical values from Pesaran et al. (2001), there is considerable support for cointegration.[7] The bootstrapped p-values, however, give more conservative results. Using them, rejection of the non-cointegration null at (or very near) the 0.05 level occurs for the data sets ending in 1979:3 and 1979:4, but not for those ending later. The significant decline in statistical significance with bootstrapping compared with using the tabulated critical values suggests significant size distortion if the tabulated critical values are used. Therefore, we focus on the more conservative, bootstrapped results. They suggest a cointegrated relationship that ended just before 1980:1. The failures to reject when the next few quarters are included are consistent with our previously documented change in structure, which could introduce the appearance of non-cointegration in the longer data sets, even if cointegration exists in the shorter data sets.

---

[6] For observation 1 of a simulated data set, we use the covariance matrix computed for periods 1-3, and for the last observation ($T$), we use the covariance matrix computed for periods $T - 2$ through $T$. For the initial 30 discarded simulated values, we use the covariance matrix from the estimated VECM

[7] The tabulated critical values depend on the number of right-hand variables, unclear for NARDL ($k$ = 1 or 2). Our significance asterisks use the number giving the more conservative values for $t$ and F (as in Shin et al., 2014).

**< Table 7 here >**

We conclude there is cointegration from 1953:2 through 1979:4. But is it asymmetric? The various estimated cointegration vector coefficients for inflation increases are larger than for decreases, as earlier proposed. The chi square tests indicate, with the bootstrapped p-values, that the differences are statistically significant at the 0.05 level (0.01 in one case). Once again, bootstrapping gives much more conservative conclusions than using tabulated critical values – the computed chi square statistics are all very much larger than the 0.01 critical value of 6.64. Indirect support for asymmetry comes from the ARDL test of Pesaran et al. (2001), which is the same test but without asymmetry. The ARDL test gives no rejections of non-cointegration for any estimation period (specific results available upon request). That the (N)ARDL procedure finds cointegration with asymmetry, but not without, supports the significance of asymmetry.

The magnitude of the difference in the relationship of the interest rate to increases versus decreases in the inflation rate is about 15% according to the AIC models (Panel (a) of Table 5), but close to 50% according to the SEQ models (Panel (b) of Table 5). The 50% figure is dramatic; the 15% figure is more plausible. But even 15% is likely an overestimate. Ioannidis (2008) points out that if a test is underpowered, a point estimate that is in the rejection region is likely to be larger than the true effect. Finally, the Table 5 cointegrating coefficient estimates have, in common with a lot of past work, the puzzling trait of having values far less than 1.0.

Now let us turn to analysis of the 1980-2008 period. Table 8 gives the results for 1980:4-2008:2 (the NARDL results for five similar periods with starting dates of 1980:2-1980:4 and ending dates of 2008:1:2008:2 are almost the same and omitted to save space. Just as before, the bootstrapped p-values give much more conservative results than do tabulated critical values. There is some support for rejection, but it is not robust, existing at the 0.05 level from only the SEQ lag approach with BIC-based DGPs. If there is cointegration, the estimated interest-inflation relationship is nearly absent or significantly negative, which is not in accordance with the Fisher effect. A cointegrating Fisher effect in this time period, therefore, seems doubtful.

**< Table 8 here >**

The 1980:4-2008:2 data period just analyzed includes the 1981:3-1982:4 recession, a period of highly fluctuating interest and inflation rates. Perhaps this period should be excluded. Mishkin (1992) does include a 1982 breakpoint. Also, the ending date of 2008:2 could conservatively be moved a quarter earlier. In fact, the most conservative period 2, based on the relevant boundaries of the Bai-Perron 95% confidence intervals, is 1984:4-2006:1. We therefore apply the NARDL tests to the subsamples 1983:1-2008:1 and 1984:4-2006:1. Again see Table 6. Results are similar to those for 1980:4-2008:2, with little statistical significance for cointegration and, usually, negative vector coefficients that are sometimes substantially negative. With no cointegration, the asymmetry test results are irrelevant.[8]

Finally, we turn to the 2008:4-2019:1 period. There are not enough observations to estimate the NARDL model with *maxlag* = 9 from using $\ell = 12$ in the Schwert formula. So, we set $\ell = 6$ instead (the second exception to $\ell = 12$ in the Schwert formula). The results, not reported but available upon request, show that the NARDL cointegration *t* test statistics are utterly insignificant (p-values of 0.99), reflecting a wrong-signed error correction coefficient. The NARDL *F* statistics are significant, but the wrong-signed error correction coefficient that drives the significance renders it irrelevant. Moreover, the estimated cointegration vector coefficients are nearly zero. The results surely reflect, at least in part, that the Fed held short-term interest rates at nearly zero for much of this estimation period.

## 7 | WHY NO EVIDENCE SUPPORTING COINTEGRATION POST-1979?

Our finding of no cointegration after 1979 contradicts several results in Table 1 that concern mostly the post-1979 period: Christopoulos and León-Ledesma (2007), Beyer and Farmer (2007), Westerlund (2008), and Haug et al. (2011). However, none of these papers employ bootstrapping, and they may thus be too optimistic that cointegration was present. Moreover, except for Westerlund (2008), the starting dates for the non-cointegration rejecting sample periods are earlier than ours. The biggest discrepancy occurs in Haug et al. (2011), who employ a breakpoint date for the U.S. of 1977:1, three years before ours. If a breakpoint in 1980 is better, then results from sample periods starting several years earlier could be misleading.

Blanchard (2018) concludes that in the second half of a 1981-2016 data set forecasters' and consumers' inflation expectations became "largely nonresponsive to actual inflation" (Blanchard, 2018, p. 116). Analysis of the post-1980-break Fisher effect using actual inflation as a proxy for expected inflation, as done in every paper in Table 1 as well as the present one, would

[8] In the 1980-2008 period, the interest and inflation rates both appear to have downward trends (just as they both appear to have upward trends prior to 1980). The 1980-2008 lack of cointegration means either the variables did not have the same downward trend, or they did not share long-run random movements, or both.

then be distorted. The distortion, at least for the professional forecasters, would not be related to asymmetric response, because they did not exhibit it (Baqaee, 2020).

If growing lack of expectation response to actual recent inflation explains finding no cointegration after 1980 while finding cointegration before, the implication is that there *was* responsiveness before 1980. We now present evidence to support this. Unfortunately, the Michigan consumer and Philadelphia Fed professional forecaster data used by Blanchard (2018) do not extend much (or at all) prior to 1981. Therefore, we use the inflation forecasts of Livingston survey of economists, which go back to 1946.[9] The data come from surveys conducted each June and December. We compare the six-month inflation forecasts with the most recently available six-month inflation rates, also in the survey data set. There is long-running discussion of whether the Livingston inflation forecasts accurately reflect people's unobservable expectations (Croushore, 1997). One claim is that the forecasts are not rational, but Ball (2000), Hong and Stein (2007), and Yellen (2007) suggest people's expectations do not necessarily conform to theoretical rational expectations.

Given the I(1) status of inflation already documented, and the corresponding I(1) status for the Livingston forecasts (see Table 9), we test for cointegration between the forecasts and recent inflation in our three sub-periods. Cointegration in a sub-period would mean the economists' expectations were correlated with actual inflation in the long run. The closer to 1.0 is the vector coefficient, the more accurate the relationship. Forecasts adjusting to correct past forecast errors (lack of weak exogeneity) would also be expected. However, lack of cointegration (or a small cointegration coefficient), weak exogeneity, or unstable vector coefficients would imply an absent or weak relationship over the long run. In the absence of linear cointegration there could, however, be a nonlinear relationship, perhaps asymmetric, mis-specified using the linear approach. But if the Livingston economists have shared the professional forecasters' lack of asymmetric response (Baqaee, 2020), and if the Livingston economists have had no other nonlinear response, then not finding a linear cointegration would clearly suggest a minimal long-run relationship between forecasts and actual inflation. We do, however, test for asymmetric cointegration between forecasted and actual inflation.

[9] The Philadelphia Fed forecasters' data begin in 1981 (https://www.philadelphiafed.org/research-and-data/real-time-center/survey-of-professional-forecasters/), the University of Michigan household survey expectations data begin in 1978 (http://www.sca.isr.umich.edu/tables.html), and the Livingston inflation forecast data begin in 1946 (https://www.philadelphiafed.org/research-and-data/real-time-center/livingston-survey).

We apply several cointegration and stability tests based on regressions of the Livingston forecasts on the most recently available actual inflation rate. The cointegration tests are the Engle-Granger OLS detrended $ADF\ t$ test, the GLS detrended $ADF\ t$ and $MZ_\rho$ tests (Perron and Rodríguez, 2016), and the NARDL $t$ and $F$ tests. The stability tests are the *Lc*, *MeanF*, and *SupF* tests of Hansen (1992). The *Lc* test is also a test of the null of cointegration. The OLS and Hansen (1992) fully modified (FM) cointegration vector coefficients are reported, along with the error correction coefficient (estimated using the OLS error) and its statistical significance. In the final exception to $\ell = 12$ in the Schwert formula, $\ell$ is set to 6 (and 5 for the final data period). This adjusts for the smaller number of semi-annual observations relative to quarterly observations over a given time period (and if the errors follow an ARMA($p$, 1) process, a size-distorting moving average parameter will be smaller, reducing the need for a long AR lag order).

The results are in Table 10. The bootstrap DGPs are first difference VARs for the tests of no-cointegration nulls. The DGPs are VECMs for the remaining tests (for the error correction coefficient test, the forecast variable's error correction coefficient is constrained to zero). DGP lag orders are from the AIC. The wild bootstrap is used except for the first- and third-period tests of no cointegration. There are 2,000 replications for each DGP and time period. The results are consistent with an economically significant and stable relationship between the forecasts and actual inflation before the 1980 breakpoint, but not after. For 1954:6-1980:6, the *EG-ADF t* test soundly rejects non-cointegration in favor of linear cointegration, and the stability tests suggest the vector is stable. The GLS-detrended tests do not reject non-cointegration, a seeming puzzle because Perron and Rodríguez (2016) point out that the GLS-detrended tests are generally more powerful than the *EG-ADF-t* test. However, they also point out that the OLS detrended *EG-ADF-t* delivers more power in the presence of a large initial condition, and the residuals from the OLS regression for the *EG* test do have a very sizeable initial value. When this is eliminated by starting the sample period one quarter later, the two GLS-detrended tests show strong rejections while the results from all the remaining tests for the time period remain largely unchanged. Finally, the NARDL tests do not support asymmetric cointegration.

**< Tables 9 and 10 here >**

In the 1980:12-2008:6 period, neither non-cointegration nor stability are rejected. Because the *Lc* stability test is also a test of the null of cointegration, cointegration's presence is ambiguous. If there *is* cointegration, the vector coefficient for actual inflation is much smaller than in the first period (using either the OLS or FM estimate) and the error correction coefficient is insignificant with the wrong sign. Meanwhile, the FM procedure without the default pre-whitening does generate rejections for all three tests, with bootstrapped p-values between 0.03 and 0.07 (not shown in the table). The FM test results for the other time periods are virtually unaffected by not pre-whitening. Finally, the results for the 2008:12-2018:12 period are certainly consistent with the deterioration in the relationship after 1980. Non-cointegration is not rejected by any test, and cointegration is rejected by the *Lc* test. The non-rejections by the NARDL tests for the two later periods are consistent with our conjecture that the breakdown in the Livingston economists' inflation/forecast relationship would not be explained by asymmetric expectations.

Overall, recent inflation seems to be a reasonable proxy for professional inflation expectations before 1980, but not after. There is no corresponding household evidence. But if households were the primary source of the asymmetric Fisher effect prior to 1980, and if their expectations subsequently became a more tenuous function of inflation, then an asymmetric cointegration relationship estimated from actual inflation would have broken down, as we find.

## 8 | CONCLUSION

We propose that asymmetric expectations could be relevant to the Fisher effect and use the NARDL approach of Shin et al. (2014) to test for cointegration under such a possibility and to assess the significance of any asymmetric relationship. Since asymmetric expectations will likely generate at some point a noticeable discrepancy between expectation and outcome, we expect breaks in the asymmetric Fisher effect when the discrepancies are (temporarily) corrected. We, therefore, expend considerable effort to identify and account for breaks in our testing.

The subsequent NARDL analysis rejects non-cointegration in favor of asymmetric cointegration for U.S. data in the sample period 1953:2-1979:4. As in many past papers, the interest rate response to inflation is less than one-to-one, but our central finding is that it is asymmetric. Among several estimates, the most plausible is that the interest rate response to inflation rate increases is about 15% larger than to decreases.

In contrast, we find no credible evidence for cointegration, asymmetric or otherwise, in post-1979 subsamples. This may reflect that inflation expectations, assuming they are reasonably measured by surveys of inflation forecasts, have become increasingly disconnected from corresponding actual recent inflation, as found by Blanchard (2018) in a post-1980 data set.

Extending Blanchard's (2018) work, we show that actual and expected inflation were cointegrated before 1980, but not afterward. Thus, use of actual inflation rates to study the Fisher effect is reasonable with the older data, but encounters problems with newer time periods. Consequently, it could be fruitful to examine the Fisher effect using survey data.

**REFERENCES**

Atkins, F.J., & Coe, P.J. (2002). An ARDL bounds test of the long-run Fisher effect in the United States and Canada. *Journal of Macroeconomics, 24*(2), 255–266.

Bai, J., & Perron, P. (1998). Estimating and testing linear models with multiple structural changes. *Econometrica*, *66*(1), 47–78.

Bai, J., & Perron. P. (2003). Computation and analysis of multiple structural change models. *Journal of Applied Econometrics, 18*(1), 1–22.

Ball, L. (2000). Near-rationality and inflation in two monetary regimes. Proceedings, Federal Reserve Bank of San Francisco.

Banerjee, A., Dolado, J., & Mestre, R. (1998). Error-correction mechanism tests for cointegration in a single-equation framework. *Journal of Time Series Analysis, 19*(3), 267–283.

Baqaee, D.R. (2020). Asymmetric inflation expectations, downward rigidity of wages, and asymmetric business cycles. *Journal of Monetary Economics, 114*(C), 174–193.

Bates, J.M., & Gabor, A. (1986). Price perception in creeping inflation: Report on an enquiry. *Journal of Economic Psychology, 7*(3), 291–314.

Baumeister, R.F., Bratslavsky, E., Finkenauer, C., & Vohs, K.D. (2001). Bad is stronger than good. *Review of General Psychology, 5*(4), 323–370.

Beyer, A., & Farmer, R.E.A. (2007). Natural rate doubts. *Journal of Economic Dynamics and Control, 31*(3), 797–825.

Blanchard, O. (2018). Should we reject the natural rate hypothesis? *Journal of Economic Perspectives, 32*(1), 97–120.

Brock, W.A., Dechert, W.D., Scheinkman, J.A., & LeBaron, B. (1996). A test for independence based on the correlation dimension. *Econometric Reviews*, *15*(3), 197–235.

Brüggemann, R., & Lütkepohl, H. (2001). Lag selection in subset VAR models with an application to a U.S. monetary system. In Friedmann, R., Knüppel, L., & Lütkepohl, H. (Eds.), *Econometric Studies: A Festschrift in Honour of Joachim Frohn* (pp. 107–28). Münster: LIT Verlag.

Bruine de Bruin, W., van der Klaauw, W., & Topa, G. (2011). Expectations of inflation: The biasing effect of thoughts about specific prices. *Journal of Economic Psychology, 32*(4), 834–845.

Caporale, G.M., & Pittis, N. (2004). Estimator choice and Fisher's paradox: A Monte Carlo Study. *Econometric Reviews, 23*(1), 25–52.

Carrion-i-Silvestre, J.L., Kim, D., & Perron, P. (2009). GLS-based unit root tests with multiple structural breaks under both the null and the alternative hypotheses. *Econometric Theory, 25*(6), 1754–1792.

Chang, Y. (2004). Bootstrap unit root tests in panels with cross-sectional dependency. *Journal of Econometrics, 120*(2), 263–293.

Christopoulos, D.K., & León-Ledesma, M.A. (2007). A long-run non-linear approach to the Fisher effect. *Journal of Money, Credit and Banking, 39*(2-3), 543–559.

Croushore, D. (1997). The Livingston survey: Still useful after all these years. *Business Review, Federal Reserve Bank of Philadelphia,* (March/April), 1–12.

Crowder, W.J., & Hoffman, D.L. (1996). The long-run relationship between nominal interest rates and inflation: The Fisher equation revisited. *Journal of Money, Credit and Banking, 28*(1), 102–118.

Darby, M.R. (1975). The financial and tax effects of monetary policy on interest rates. *Economic Inquiry, 13*(2), 266–276.

Detmeister, A., Lebow, D., & Peneva, E. (2016). Inflation perceptions and inflation expectations. *FEDS Notes.* Board of Governors of the Federal Reserve System (December 5).

Evans, M.D.D., & Lewis, K.K. (1995). Do expected shifts in inflation affect estimates of the long-run Fisher relation? *The Journal of Finance, 50*(1), 225–253.

Fama, E.F. (1975). Short-term interest rates as predictors of inflation. *American Economic Review, 65*(3), 269–282.

Fisher, I. (1930). *The Theory of Interest, as Determined by Impatience to Spend Income and Opportunity to Invest It*. New York: Macmillan.

Gonçalves, S., & Kilian, L. (2004). Bootstrapping autoregressions with conditional heteroskedasticity of unknown form. *Journal of Econometrics, 123*(1), 89–120.

Harvey, D.I., & Leybourne, S.J. (2007). Testing for time series linearity. *The Econometrics Journal*, *10*(1), 149–165.

Haug, A.A., Beyer, A., & Dewald, W. (2011). Structural breaks and the Fisher effect. *The B.E. Journal of Macroeconomics: Topics, 11*(1), 1–29.

Hansen, B.E. (1992). Tests for parameter instability in regressions with I(1) processes. *Journal of Business & Economic Statistics, 10*(3), 321–335.

Hong, H., Stein, J.C., & Yu, J. (2007). Simple forecasts and paradigm shifts. *The Journal of Finance, 62*(3), 1207–1242.

Ioannidis, J.P.A. (2008). Why most discovered true associations are inflated. *Epidemiology, 19*(5), 640–648.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica, 47*(2), 263–291.

Koustas, Z., & Serletis, A. (1999). On the Fisher effect. *Journal of Monetary Economics, 44*(1), 105–130.

MacDonald, R., & Murphy, P.D. (1989). Testing for the long run relationship between nominal interest rates and inflation using cointegration techniques. *Applied Economics, 21*(4), 439–447.

Mankiw, N.G. (1997). Comment on Shiller, Robert J., Why do people dislike inflation? In Romer, C.D., & Romer, D.H. (Eds.), *Reducing Inflation: Motivation and Strategy* (pp. 65–69). University of Chicago Press.

Million, N. (2004). Central bank's interventions and the Fisher hypothesis: A threshold cointegration investigation. *Economic Modelling, 21*(6), 1051–1064.

Mishkin, F.S. (1992). Is the Fisher effect for real? A reexamination of the relationship between inflation and interest Rates. *Journal of Monetary Economics, 30*(2), 195–215.

Murray, C.J., & Nelson, C.R. (2000). The uncertain trend in U.S. GDP. *Journal of Monetary Economics, 46*(1), 79–95.

Neely, C.J., & Rapach, D.E. (2008). Real interest rate persistence: Evidence and implications. *Federal Reserve Bank of St. Louis Review, 90*(6), 609–641.

Ng, S., & Perron, P. (2001). Lag length selection and the construction of unit root tests with good size and power. *Econometrica, 69*(6), 1519–1554.

Panopoulou, E., & Pantelidis, T. (2016). The Fisher effect in the presence of time-varying coefficients. *Computational Statistics and Data Analysis, 100*, 495–511.

Perron, P., & Rodríguez, G. (2016). Residuals-based tests for cointegration with generalized least-squares detrended data. *The Econometrics Journal, 19*, 84–111.

Pesaran, M.H., Shin, Y., & Smith, R.J. (2001). Bounds testing approaches to the analysis of level relationships. *Journal of Applied Econometrics, 16*(3), 289–326.

Ranyard, R., Del Missier, F., Bonini, N., Duxbury, D., & Summers, B. (2008). Perceptions and expectations of price changes and inflation: A review and conceptual framework. *Journal of Economic Psychology, 29*(4), 378–400.

Rapach, D.E. (2003). International evidence on the long-run impact of inflation. *Journal of Money, Credit and Banking, 35*(1), 23–48.

Rose, A.K. (1988). Is the real interest rate stable? *Journal of Finance, 43*(5), 1095–112.

Rothman, P., van Dijk, D., & Franses, P.H. (2001). Multivariate STAR analysis of money-output relationship. *Macroeconomic Dynamics, 5*(4), 506-532.

Schwert, G.W. (1989). Tests for unit roots: A Monte Carlo investigation. *Journal of Business and Economic Statistics, 7*(2), 147–159.

Shiller, R.J. (1997). Why do people dislike inflation? In Romer, C.D., & Romer, D.H. (Eds.), *Reducing Inflation: Motivation and Strategy* (pp. 13–65). University of Chicago Press.

Shin, Y.,Yu, B., & Greenwood-Nimmo, M. (2014). Modelling asymmetric cointegration and dynamic multipliers in a nonlinear ARDL framework. In Sickles, R.C., & Horrace, W.C. (Eds.), *Festschrift in Honor of Peter Schmidt: Econometric Methods and Applications* (pp. 281–314). New York: Springer.

Soroka, S.N. (2006). Good news and bad news: Asymmetric responses to economic information. *Journal of Politics, 68*(2), 372–385.

Stock, J.H. (1987). Asymptotic properties of least squares estimators of cointegrating vectors. *Econometrica, 55*(5), 1035–1056.

Summers, L.H. (1983). The non-adjustment of nominal interest rates: A study of the Fisher effect. In Tobin, J. (Ed.), *A Symposium in Honor of Arthur Okun* (pp. 201–41). Washington, D.C.: Brookings Institution.

Tsong, C.-C., & Cheng-Feng Lee, C.-F. (2013). Quantile cointegration analysis of the Fisher hypothesis. *Journal of Macroeconomics, 35*(1), 186–198.

Tversky, A., & Kahneman, D. (1991). Loss aversion in riskless choice: A reference-dependent model. *Quarterly Journal of Economics, 106*(4), 1039–1061.

Westerlund, J. (2008). Panel cointegration tests of the fisher effect. *Journal of Applied Econometrics, 23*(2), 193–233.

Xiao, Z. (2009). Quantile cointegrating regression. *Journal of Econometrics, 150*(2), 248–260.

Yellen, J.L. (2007). Implications of behavioral economics for monetary policy. Federal Reserve Bank of Boston Conference. http://www.frbsf.org/our-district/files/0928.pdf

**TABLE 1** Literature review

| Paper | Countries | Data period | Techniques | Breakpoints? | General findings | Paper concludes $\beta$ = ? | Bootstrap test stats? | Overlapping data? |
|---|---|---|---|---|---|---|---|---|
| Fisher (1930) | 6 incl. USA | various: 1800s to early 1900s (latest 1927) | Tables, graphs, correlations | Various sub-periods | Pos. corr($i,\pi$) for full period, but they break down with sub-periods and lags/leads | $0 < \beta < 1$ | No | No |
| Fama (1975) | USA | 1953-1971 (1959 start sometimes) | Regression | Various sub-periods for some results | Constant (stationary) $r$, fluctuations in $i$ reflect $\pi$ predicted by past $\pi$ | $\beta = 1$ | No | No |
| Summers (1983) | USA | 1860-1979 | Tables, band spectral regression | WW II, and many sub-periods | No relationship prior to 1940; some response after 1940 | After 1940*:* $0 < \beta < 1$ (so also < 1.3, the hypothesized value with tax effects) | No | No |
| Rose (1988) | 18 incl. USA | USA: annual 1892-1970, quarterly 1957-?, monthly 1948-86 | Unit root tests, EG cointegration | Various, using sub-samples, early 70s, or 1979 for US monthly | $r$ is nonstationary, $i$ and $\pi$ are not cointegrated (and $\pi$ is probably stationary) | Not estimated | No | Usually no |
| MacDonald & Murphy (1989) | 4 incl. USA | 1955-86 | Cointegration tests | 1972 or 1973, using subsamples | Non-cointegration is weakly rejected for 1st sub-period, not for 2nd | In 1st sub-period, usually $0 < \beta < 1$; $\beta < 0$ for US | No | No |
| Mishkin (1992) | USA | 1953-1990 | Regression, unit root and EG cointegration tests | 1979 and 1982, using subsamples | $i$ and $\pi$ are I(1); little sh. run forecast ability in regressions, but $i$ and $\pi$ are cointegrated until 1979, but not later | $(1/\beta) > 1$ (so $0 < \beta < 1$) for 1953-1979 period (no cointegration in later periods) | Yes | Some-times, but account-ed for |
| Evans & Lewis (1995) | USA | 1947-1987 | Markov switching, JOH coint., but $\beta$ estimates from DOLS | $\pi$ process shifts are sometimes specified with Markov switching | See next column | $0 < \beta < 1$ with standard coint. approach, but cannot reject $\beta \geq 1$ with inflation process shifts included | No | Yes, with 3-mo. maturi-ties |
| Crowder & Hoffman (1996) | USA | 1952-1991 | JOH cointegration | No | Support for "tax-adjusted" Fisher effect | $\beta > 1$ (without tax adjusted data, which is the usual approach); $\beta = 1$ not rejected with tax adjusted $i$ | No | No |
| Koustas & Serletis (1999) | 11 incl. USA | 1957-1995 | EG cointegration, then 1st difference VAR | No | $i$ and $\pi$ are I(1) but not cointegrated; 1st difference VARs estimate $\beta$ | Depends on contemporaneous identifying restrictions, but generally conclude $0 < \beta < 1$ | No | Unclear, probably not |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Atkins & Coe (2002) | USA and Canada | 1953-1999 | ARDL cointegration | No | i and $\pi$ are probably I(1) and cointegrated | $\beta$ = 1 not rejected. But higher values for tax-adjusted $\beta$ not generally supported. | No | Yes |
| Rapach (2003) | 14 incl. USA | varies by country, US: 1961-1996. | Structural 1st difference VAR to estimate effect of permanent $\pi$ shock on *i*; real GDP included in model | No | *i* and $\pi$ are I(1) but not cointegrated; one exception, permanent $\pi$ rise significantly reduces real *i* with US quarterly data | Result implies $\beta$ < 1, so l.r. full Fisher effect is rejected, except using US quarterly data | Yes, for the VAR results | Yes (long term *i* is used) |
| Million (2004) | USA | 1951-1999 | Unit root tests; EG coint. with break in constant; tests EG residuals for asym-metric and nonlin adjust to long-run equilibrium | Not for unit root tests on *i*, $\pi$ or *r*; but paper does specify breaks in mean for cointegration test: 1979 | *i*, $\pi$, and *r* are I(1); but *i* and $\pi$ are coint. when brk in mean included; residuals from assuming $\beta$ = 1 show asymmetric and nonlinear adjustment | Estimates are not reported; $\beta$ = 1 is assumed in residual analysis | No | Yes |
| Caporale & Pittis (2004) | USA | 1960-1999 | 28 estimators to estimate $\beta$ and test $\beta$ = 1 | Breaks only in cointegration vector variance in DGP for Monte Carlo analysis | *i* and $\pi$ are I(1) and cointegrated and $\beta$ = 1 is usually not rejected | Estimators with best properties in Monte Carlo analysis do not reject $\beta$ = 1 | Data-based Monte Carlo to analyze estimator properties | No |
| Christopou-los & León-Ledesma (2007) | USA | 1960-2004 | JOH cointegration initially, then specify $\beta$ as a nonlinear function of $\pi$ (not constant) | 1978-79, using subsamples | *i* and $\pi$ are I(1), but not cointegrated using JOH, but are cointegrated when $\beta$ is nonlinearly related to $\pi$ | 0 < $\beta$ < 1 for linear model and nonlinear model in 1st sub-sample; in 2nd sub-sample $\beta$ < 1 for low $\pi$ but $\beta$ > 1 for high $\pi$ | No, except for most tests of non-linearity | No |
| Beyer & Farmer (2007) | USA | 1970-1999 | JOH cointegration; test for break in vector and short-run parameters with DOLS.; model includes unemployment | 1979, using subsamples | *i* and $\pi$ are I(1) and cointegrated in an identified vector that excludes unemployment | $\beta$ < 1 in 1st subsample, $\beta$ > 1 in 2nd sub-sample, but neither individually statitically significant so, but jointly $\beta$ = 1 rejected. | No | No |
| Westerlund (2008) | 20 incl. USA | 1980-2004 | Panel cointegration test, several versions of OLS to est. $\beta$ | No, but data do not include typical breakpoint dates | Non-cointegration null rejected for the panel | Across the 20 coiuntries $\beta$ = 1 is usually not rejected (and not rejected for US) | Not for coint. tests | Unclear, probably not |
| Haug et al. (2011) | 14 incl. USA | varies by country, US: 1957-2007. | Unit root tests without breaks; cointegration break test with estimated break dates; JOH | Varies by country, single variable, and multivariate tests; cointegration tests applied to | *i* and $\pi$ are I(1) with no brks specified. JOH shows coint. in 5 of 14 no-break cases, but, with breaks, does in all | For 5 countries with no breaks, $\beta$=1 not rejected, but in remaining cases with subsamples $\beta$=1 is usually rejected (US: $\beta$=1 | No | No |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | | cointegration test but DOLS to estimate coefficients. Tests for nonlinear cointegration vectors. | subsamples; break in 1977 for the US. | pre- and post-break periods | in 1st period, but $\beta$>1 in 2nd period). No evidence of nonlinear cointegration vector when breaks included | | |
| Tsong & Lee (2013) | 6 incl. USA | 1957-2010 | Quantile cointegration, and EG cointegration for comparison | In unit root tests but not in cointegration tests | $i$ and $\pi$ are I(1); quantile cointegration is reported | $\beta$ varies across quantiles, < 1 for lower quantiles, not significantly different from 1 for higher quantiles; results argued to be consistent with shocks affecting $\beta$ | Yes for quantile cointegra-tion tests | Unclear |
| Panopoulou & Pantelidis (2016) | 19 incl. USA | varies by country, US: 1881-2009. | 9 cointegration tests to estimate $\beta$; time-varying serial correlation specified in DGP errors when testing $\beta$ = 1 using simulations | No | $i$ and $\pi$ are I(1), and cointegrated in JOH tests for 15 countries including the US; but 9 tests of $\beta$ = 1 for each country are the main focus (see next column) | Null of $\beta$ = 1 usually cannot be rejected, but $\beta$ estimates mostly differ substantially from 1; US: $\beta$ = 1.0 rejected for only 1 of 9 tests, but $\beta$ estimates differ substantially from 1.0 in 6 of 9 tests | Yes for tests of $\beta$ = 1 with VAR(1) time-varying error processes | No |

*Notes*: DOLS denotes the Stock-Watson dynamic ordinary least squares, JOH denotes the Johansen cointegration approach; EG denotes the Engle-Granger cointegration approach.

**TABLE 2** BDS tests (1953:2-2019:1)

| **Interest rate** | | | | **Inflation rate** | | | |
|---|---|---|---|---|---|---|---|
| | $\varepsilon$ | *em* | BDS statistic | | $\varepsilon$ | *em* | BDS statistic |
| 0.5 × SD | 0.43 | 2 | 10.129 (0.000) | 0.5 × SD | 0.97 | 2 | 4.468 (0.000) |
| 0.5 × SD | 0.43 | 3 | 15.069 (0.000) | 0.5 × SD | 0.97 | 3 | 4.615 (0.000) |
| 0.5 × SD | 0.43 | 4 | 20.103 (0.000) | 0.5 × SD | 0.97 | 4 | 5.233 (0.000) |
| 1.0 × SD | 0.87 | 2 | 8.252 (0.000) | 1.0 × SD | 1.93 | 2 | 4.528 (0.000) |
| 1.0 × SD | 0.87 | 3 | 9.859 (0.000) | 1.0 × SD | 1.93 | 3 | 5.511 (0.000) |
| 1.0 × SD | 0.87 | 4 | 10.730 (0.000) | 1.0 × SD | 1.93 | 4 | 6.209 (0.000) |
| 1.5 × SD | 1.30 | 2 | 6.453 (0.000) | 1.5 × SD | 2.90 | 2 | 3.960 (0.000) |
| 1.5 × SD | 1.30 | 3 | 7.594 (0.000) | 1.5 × SD | 2.90 | 3 | 4.853 (0.000) |
| 1.5 × SD | 1.30 | 4 | 8.124 (0.000) | 1.5 × SD | 2.90 | 4 | 5.588 (0.000) |
| 2.0 × SD | 1.74 | 2 | 7.243 (0.000) | 2.0 × SD | 3.86 | 2 | 2.898 (0.003) |
| 2.0 × SD | 1.74 | 3 | 8.508 (0.000) | 2.0 × SD | 3.86 | 3 | 3.468 (0.000) |
| 2.0 × SD | 1.74 | 4 | 8.843 (0.000) | 2.0 × SD | 3.86 | 4 | 4.212 (0.000) |
| 2.5 × SD | 2.17 | 2 | 10.356 (0.000) | 2.5 × SD | 4.83 | 2 | 1.889 (0.058) |
| 2.5 × SD | 2.17 | 3 | 10.658 (0.000) | 2.5 × SD | 4.83 | 3 | 1.831 (0.067) |
| 2.5 × SD | 2.17 | 4 | 10.267 (0.000) | 2.5 × SD | 4.83 | 4 | 2.618 (0.008) |

*Notes:* SD denotes the standard deviation of the data, $\varepsilon$ (the given multiple of SD) denotes the maximum distance between data points used for computing the correlation integral, and *em* denotes the embedded dimension. The BDS statistic is asymptotically standard normal, and corresponding two-sided *p*-values are in parentheses.

**TABLE 3** Harvey and Leybourne (2007) linearity tests (1953:2-2019:1)

| | $W^*_{10\%}$ | $W^*_{5\%}$ | $W^*_{1\%}$ |
|---|---|---|---|
| **Interest rate** | 64.06 (0.000) | 64.46 (0.000) | 65.17 (0.000) |
| **Inflation rate** | 10.04 (0.040) | 10.09 (0.039) | 10.16 (0.038) |

*Notes:* The subscript on $W^*$ indicates the desired significance level of the test, which determines *b* in the formula at the bottom of p. 152 of Harvey and Leybourne (2007). The $W^*$ statistic asymptotically follows the $\chi^2_{(4)}$ distribution. Resulting *p*-values are in parentheses.

**TABLE 4** $MZ_\alpha$ and *ADF* unit root test statistics, critical values, and break dates (1953:2 - 2019:1)

**Interest rate**

| number of breaks | $MZ_\alpha$ | $MZ_\alpha$ 0.05 | *ADF* | *ADF* 0.05 | break dates |
|---|---|---|---|---|---|
| 0 | -6.40 | -17.33 | -1.70 | -2.90 | |
| 1 | -12.67 | -23.94 | -2.49 | -3.44 | 1980:1 |
| 2 | -23.85 | -29.16 | -3.37 | -3.79 | 1972:2; 1980:1 |

**Inflation rate**

| number of breaks | $MZ_\alpha$ | $MZ_\alpha$ 0.05 | *ADF* | *ADF* 0.05 | break dates |
|---|---|---|---|---|---|
| 0 | -6.19 | -17.33 | -1.82 | -2.90 | |
| 1 | -11.91 | -20.47 | -2.26 | -3.19 | 2008:3 |
| 2 | -14.92 | -28.00 | -2.85 | -3.73 | 1980:2; 2008:3 |

*Notes*: The unit root null is rejected for sufficiently negative values of the test statistic. The breakpoint date designates the final quarter of the earlier regime. The new regime starts with the next quarter.

**TABLE 5** Tests for breaks between various periods in the first-difference VAR

| Periods compared | Chi-square | *p*-value |
|---|---|---|
| 1, 2 | 15.3 | 0.350 |
| 1, 3 | 14.8 | 0.390 |
| 1, 4 | 70.2 | **0.000** |
| 2, 3 | 44.1 | **0.000** |
| 2, 4 | 47.2 | **0.000** |
| 3, 4 | 84.5 | **0.000** |

*Notes*: Period 1 is 1953:2-1972:1, period 2 is 1972:2-1980:1, period 3 is 1980:2-2008:2, and period 4 is 2008:3-2019:1. The chi-square statistics have 14 degrees of freedom.

**TABLE 6** Bai-Perron breakpoint test results

| **Interest rate equation** | | | **Inflation rate equation** | | |
|---|---|---|---|---|---|
| **UD max test** | | | **UD max test** | | |
| null; alt. | statistic | 0.05 crit. | null; alt. | statistic | 0.05 crit. |
| 0; >0 breaks | 50.19 | 26.48 | 0; >0 breaks | 50.43 | 26.48 |
| **supF test** | | | **supF test** | | |
| null; alt. | statistic | 0.05 crit. | null; alt. | statistic | 0.05 crit. |
| 0; 1 break | 50.19 | 26.20 | 0; 1 break | 50.43 | 26.20 |
| 0; 2 breaks | 31.38 | 23.36 | 0; 2 breaks | 33.29 | 23.36 |
| 0; 3 breaks | 27.84 | 21.63 | 0; 3 breaks | 26.41 | 21.63 |
| 1; 2 breaks | 26.02 | 28.23 | 1; 2 breaks | 19.53 | 28.23 |
| 2; 3 breaks | 16.36 | 29.44 | 2; 3 breaks | 10.78 | 29.44 |
| **conf. int. sequential procedure break** | | | **conf. int. sequential procedure break** | | |
| 0.025 val. | 0.500 val. | 0.975 val. | 0.025 val. | 0.500 val. | 0.975 val. |
| 1979:3 | 1980:3 | 1984:3 | 2006:4 | 2008:2 | 2008:3 |

**TABLE 7** First period NARDL results, 1953:2 through end date given in the table

**Panel A: NARDL regression lag order chosen by AIC**

| End date | 1979:3 | 1979:4 | 1980:1 | 1980:2 | 1980:3 |
|---|---|---|---|---|---|
| *T* | -4.13*** | -4.13*** | -3.37* | -3.45* | -3.43* |
| *t* p-val (AIC) | **0.033** | **0.016** | 0.066 | 0.084 | 0.066 |
| *t* p-val (BIC) | 0.056 | 0.054 | 0.15 | 0.12 | 0.19 |
| *F* | 6.43** | 6.64* | 4.78* | 4.29 | 4.26 |
| *F* p-val (AIC) | 0.079 | **0.048** | 0.18 | 0.20 | 0.17 |
| *F* p-val (BIC) | 0.066 | 0.060 | 0.17 | 0.19 | 0.25 |
| $\hat{\rho}$ | -0.57 | -0.58 | -0.84 | -1.15 | -1.16 |
| $\widehat{\beta\psi\delta^{+}}$ | 0.35 | 0.37 | 0.46 | 0.38 | 0.39 |
| $\widehat{\beta\psi\delta^{-}}$ | 0.30 | 0.32 | 0.41 | 0.33 | 0.34 |
| chi sq. sym. test | 12.46*** | 12.28*** | 10.50*** | 11.72*** | 11.54*** |
| chi sq. p-val (AIC) | **0.027** | **0.028** | **0.021** | **0.040** | **0.018** |
| chi sq. p-val (BIC) | 0.053 | 0.051 | **0.045** | **0.035** | **0.028** |

**Panel B: NARDL regression lag order chosen by SEQ**

| | | | | | |
|---|---|---|---|---|---|
| *T* | -5.72*** | -5.49*** | -3.06 | -3.69** | -4.45*** |
| *t* p-val (AIC) | 0.058 | **0.047** | 0.34 | 0.50 | 0.33 |
| *t* p-val (BIC) | **0.039** | 0.061 | 0.43 | 0.35 | 0.12 |
| *F* | 13.17*** | 11.88*** | 4.26 | 6.38** | 7.65** |
| *F* p-val (AIC) | 0.053 | 0.065 | 0.56 | 0.47 | 0.37 |
| *F* p-val (BIC) | **0.024** | 0.059 | 0.48 | 0.27 | 0.13 |
| $\hat{\rho}$ | -0.40 | -0.40 | -0.24 | -0.33 | -0.39 |
| $\widehat{\beta\psi\delta^{+}}$ | 0.20 | 0.23 | 0.28 | 0.14 | 0.22 |
| $\widehat{\beta\psi\delta^{-}}$ | 0.13 | 0.16 | 0.21 | 0.07 | 0.15 |
| chi sq. sym. test | 32.55*** | 29.39*** | 8.67*** | 8.58*** | 10.08*** |
| chi sq. p-val (AIC) | **0.008** | **0.012** | 0.094 | 0.11 | 0.19 |
| chi sq. p-val (BIC) | **0.014** | **0.013** | 0.093 | 0.11 | 0.088 |

*Notes*: $t$ = $t$ statistic for $\rho$; p-val (AIC or BIC) = bootstrapped p-values for $t$ and $F$ statistics using VAR with AIC or BIC lag choice for the DGP; $\rho$ = error correction coefficient; $\beta\psi\delta^{+}$ and $\beta\psi\delta^{-}$ = estimated cointegration vector coefficients for $\pi_{t-1}^{+}$ and $\pi_{t-1}^{-}$when the vector coefficient for $i_{t-1}$ is normalized as $-1$; chi sq. sym. test = chi square statistic with 1 degree of freedom for the null that $\beta\psi\delta^{+} = \beta\psi\delta^{-}$; chi sq. p-val (AIC or BIC) = bootstrapped p-values for the chi square statistic using VECM with AIC or BIC lag choice. The statistical significances of the test statistics from tables are indicated by *** for the 0.01 level, ** for the 0.05 level, and * for the 0.10 level. The tables for $t$ and $F$ are from the I(1) columns in Tables CII(iii) and CI(iii) in Pesaran et al. (2001) and for chi square are from its standard distribution. Bootstrapped p-values significant at the 0.10 level are reported to three decimal places instead of two, and those additionally significant at the 0.05 level are highlighted in boldface italics.

**TABLE 8** Second period NARDL results

**Panel A: NARDL with AIC lag choice**

| **Estimation period** | **1980:4-2008:2** | **1983:1-2008:1** | **1984:4-2006:1** |
|---|---|---|---|
| $T$ | -3.90** | -2.56 | -3.58** |
| $t$ p-val (AIC) | 0.13 | 0.34 | 0.11 |
| $t$ p-val (BIC) | 0.062 | 0.35 | 0.11 |
| $F$ | 5.62* | 2.32 | 5.43* |
| $F$ p-val (AIC) | 0.16 | 0.56 | 0.13 |
| $F$ p-val (BIC) | 0.090 | 0.64 | 0.13 |
| $\hat{\rho}$ | -0.31 | -0.19 | -0.25 |
| $\widehat{\beta\psi\delta}^{+}$ | -0.32 | -0.38 | -0.80 |
| $\widehat{\beta\psi\delta}^{-}$ | -0.21 | -0.25 | -0.65 |
| chi sq. sym. test | 64.04*** | 35.22*** | 25.50*** |
| chi sq. p-val (AIC) | **0.010** | 0.059 | **0.034** |
| chi sq. p-val (BIC) | **0.005** | **0.009** | 0.054 |

**Panel B: NARDL with SEQ lag choice**

| **Estimation period** | **1980:4-2008:2** | **1983:1-2008:1** | **1984:4-2006:1** |
|---|---|---|---|
| $T$ | -5.47*** | 2.86 | -4.69*** |
| $t$ p-val (AIC) | 0.13 | 0.43 | 0.12 |
| $t$ p-val (BIC) | **0.015** | 0.38 | 0.087 |
| $F$ | 10.05*** | 3.03 | 7.54** |
| $F$ p-val (AIC) | 0.18 | 0.62 | 0.25 |
| $F$ p-val (BIC) | **0.032** | 0.62 | 0.17 |
| $\hat{\rho}$ | -0.21 | -0.12 | -0.22 |
| $\widehat{\beta\psi\delta}^{+}$ | 0.02 | -0.07 | -0.04 |
| $\widehat{\beta\psi\delta}^{-}$ | 0.11 | 0.04 | -0.16 |
| chi sq. sym. test | 43.85*** | 28.37*** | 47.40*** |
| chi sq. p-val (AIC) | **0.020** | 0.11 | **0.021** |
| chi sq. p-val (BIC) | **0.012** | **0.015** | **0.028** |

*Notes*: See notes to Table 5.

**TABLE 9** MZA unit root test statistics for Livingston forecast data

| Estimation period | *forc* | *Infl* | 0.05 crit. |
|---|---|---|---|
| 1953:12-1980:6 | -8.52 | -13.42 | -17.33 |
| 1980:12-2008:6 | -4.76 | -0.81 | -17.33 |
| 2008:12-2018:12 | -8.51 | -2.54 | -17.33 |

*Notes*: The MZA unit root test without breaks is applied to the three sub-periods identified in the main text. The $\ell$ coefficient in the Schwert formula for max lag order is reduced from 12 to 6 to adjust for the fewer observations for the given span of time in the sample of semi-annual frequency compared with our usually available quarterly frequency. *forc* is the Livingston forecasted annualized six-month CPI inflation rate for June to December and December to June (the survey months). *infl* is the Base Period rate, the rate from the previous April to October and October to April. The number after the colon in the estimation period is the month of the survey. There are no rejections of the unit root at the 0.05 level.

**TABLE 10** Cointegration and stability tests for the Livingston forecast and actual inflation data

| **Estimation period** | **1953:12-1980:6** | **1954:6-1980:6** | **1980:12-2008:6** | **2008:12-2018:12** |
|---|---|---|---|---|
| *EG-ADF t* | -5.97*** (**0.001**) | -4.40*** (**0.036**) | -1.51 (0.49) | -1.27 (0.79) |
| EG coint. vec. coef. | 0.69 | 0.68 | 0.59 | 0.05 |
| EG e.c. coef. | -0.33** (**0.017**) | -0.37*** (**0.011**) | 0.20 (0.85) | 0.41 (0.23) |
| *ADF-GLS t* | -2.44 (0.25) | -4.36*** (**0.020**) | -0.43 (0.77) | -1.26 (0.77) |
| $MZ_\rho$ | -9.95 (0.28) | -19.03** (0.088) | 0.34 (0.82) | -2.30 (0.74) |
| FM coint. coef. (s.e.) | 1.25 (0.09) | 1.22 (0.09) | 0.60 (0.09) | 0.09 (0.05) |
| *Lc* | 0.11 (0.78) | 0.13 (0.77) | 0.53* (0.40) | 2.30*** (**0.040**) |
| *MeanF* | 1.15 (0.84) | 1.35 (0.79) | 4.13* (0.36) | 27.54*** (0.13) |
| *SupF* | 4.26 (0.76) | 4.22 (0.79) | 5.61 (0.67) | 76.49*** (0.19) |
| NARDL *t* | -3.60** (0.12) | -3.43** (0.22) | -3.59** (0.28) | 0.67 (0.94) |
| NARDL *F* | 4.86* (0.18) | 5.18* (0.20) | 12.16*** (0.057) | 0.65 (0.94) |

*Notes*: The dates give the month of the Livingston survey. Asterisks on the test statistics indicate rejection using tabulated critical values as in previous tables. Except for the FM cointegration coefficient, numbers in parentheses give test statistic p-values bootstrapped from VARs or VECMs as described in the main text (for the EG error correction coefficient, the test statistic is its *t* statistic). For the FM cointegration coefficient, the numbers in parentheses give the standard errors.

**APPENDIX**

The following tables give the p-values from diagnostic tests applied to various DGP equations. The tests are for three possible problems in the residuals: heteroskedasticity, non-normality, and serial correlation. The heteroskedasticity test, LM, is the Lagrange multiplier test from the regression of the squared residuals on the right-hand variables. The non-normality test, JB, is the Jarque-Bera test that compares the skewness and kurtosis of the sample residuals with the skewness and kurtosis of the normal distribution. The serial correlation test is the Ljung-Box Q test, where the tables report the most significant p-value for the 10 tests for autocorrelation orders 1–10. In interpreting the normality results, one should keep in mind that if the residuals are heteroskedastic, the JB test will tend to suggest they are non-normal even if they are normal for a given variance.

**TABLE A1** P-values of diagnostic tests for DGP equations, 1953:2 through 1979 or 1980

**Panel A: VAR DGP with AIC lag order**

| | Interest rate ($i$) equation | | | Inflation rate ($\pi$) equation | | |
|---|---|---|---|---|---|---|
| End date | LM | JB | Q | LM | JB | Q |
| 1979:3 | 0.98 | 0.15 | 0.41 | 0.63 | 0.39 | 0.58 |
| 1979:4 | 0.37 | 0.15 | 0.37 | 0.76 | 0.69 | 0.36 |
| 1980:1 | 0.21 | **0.000** | 0.88 | 0.72 | 0.68 | 0.32 |
| 1980:2 | **0.000** | **0.000** | 0.80 | 0.63 | 0.32 | 0.22 |
| 1980:3 | **0.000** | **0.000** | 0.81 | 0.87 | 0.49 | 0.38 |

**Panel B: VAR DGP with BIC lag order**

| | Interest rate ($i$) equation | | | Inflation rate ($\pi$) equation | | |
|---|---|---|---|---|---|---|
| End date | LM | JB | Q | LM | JB | Q |
| 1979:3 | 0.73 | 0.15 | 0.056 | 0.23 | 0.86 | 0.12 |
| 1979:4 | 0.37 | 0.19 | 0.11 | 0.23 | 0.89 | 0.12 |
| 1980:1 | 0.22 | **0.000** | 0.41 | 0.11 | 0.88 | 0.12 |
| 1980:2 | **0.000** | **0.000** | 0.51 | 0.47 | 0.70 | **0.047** |
| 1980:3 | 0.16 | **0.000** | 0.38 | 0.76 | 0.57 | **0.010** |

**Panel C: VECM DGP with AIC lag order**

| | Interest rate ($i$) equation | | | Inflation rate ($\pi$) equation | | |
|---|---|---|---|---|---|---|
| End date | LM | JB | Q | LM | JB | Q |
| 1979:3 | **0.002** | 0.20 | 0.055 | 0.12 | 1.00 | 0.31 |
| 1979:4 | **0.001** | 0.23 | 0.10 | 0.19 | 0.98 | 0.35 |
| 1980:1 | **0.000** | **0.000** | 0.40 | 0.38 | 0.92 | 0.35 |
| 1980:2 | **0.000** | **0.000** | 0.69 | **0.002** | 0.66 | 0.12 |
| 1980:3 | **0.000** | **0.000** | 0.81 | **0.001** | 0.54 | 0.47 |

**Panel D: VECM DGP with BIC lag order**

| | Interest rate (i) equation | | | Inflation rate ($\pi$) equation | | |
|---|---|---|---|---|---|---|
| End date | LM | JB | Q | LM | JB | Q |
| 1979:3 | **0.007** | 0.081 | **0.000** | 0.30 | 0.81 | **0.043** |
| 1979:4 | **0.000** | 0.17 | **0.001** | 0.29 | 0.77 | **0.050** |
| 1980:1 | **0.000** | **0.000** | **0.004** | 0.51 | 0.79 | **0.038** |
| 1980:2 | **0.000** | **0.000** | 0.63 | **0.000** | 0.84 | **0.042** |
| 1980:3 | **0.000** | **0.000** | 0.15 | **0.000** | 0.58 | 0.12 |

**TABLE A2** P-values of diagnostic tests for DGP equations, 1980-2008

**Panel A: VAR DGP with AIC lag order ending 2008:2, with start date given in table**

| | Interest rate ($i$) equation | | | Inflation rate ($\pi$) equation | | |
|---|---|---|---|---|---|---|
| Start date | LM | JB | Q | LM | JB | Q |
| 1980:2 | **0.000** | **0.014** | 0.13 | 0.25 | **0.000** | 0.39 |
| 1980:3 | **0.001** | 0.11 | 0.32 | **0.021** | **0.020** | 0.45 |
| 1980:4 | **0.016** | **0.001** | 0.66 | **0.014** | **0.003** | 0.74 |
| 1983:1 | 0.78 | 0.16 | 0.69 | 0.27 | **0.035** | 0.47 |

**Panel B: VAR DGP with BIC lag order ending 2008:2, with start date given in table**

| | Interest rate ($i$) equation | | | Inflation rate ($\pi$) equation | | |
|---|---|---|---|---|---|---|
| Start date | LM | JB | Q | LM | JB | Q |
| 1980:2 | **0.058** | **0.000** | 0.072 | 0.089 | **0.000** | 0.35 |
| 1980:3 | **0.056** | **0.000** | 0.15 | 0.37 | **0.000** | 0.19 |
| 1980:4 | **0.000** | **0.000** | 0.22 | 0.51 | **0.000** | 0.100 |
| 1983:1 | 0.42 | **0.000** | 0.24 | 0.34 | **0.000** | **0.029** |

**Panel C: VAR DGP with AIC lag order ending 2008:1, with start date given in table**

| | Interest rate ($i$) equation | | | Inflation rate ($\pi$) equation | | |
|---|---|---|---|---|---|---|
| Start date | LM | JB | Q | LM | JB | Q |
| 1980:2 | **0.000** | **0.004** | 0.11 | 0.39 | **0.000** | 0.41 |
| 1980:3 | **0.003** | **0.000** | **0.021** | 0.34 | **0.000** | 0.46 |
| 1980:4 | 0.075 | **0.000** | 0.68 | 0.021 | **0.003** | 0.76 |
| 1983:1 | 0.95 | 0.57 | 0.46 | 0.29 | **0.018** | 0.47 |

**Panel D: VAR DGP with BIC lag order ending 2008:1, with start date given in table**

| | Interest rate ($i$) equation | | | Inflation rate ($\pi$) equation | | |
|---|---|---|---|---|---|---|
| Start date | LM | JB | Q | LM | JB | Q |
| 1980:2 | 0.053 | **0.000** | 0.057 | 0.084 | **0.000** | 0.32 |
| 1980:3 | 0.077 | **0.000** | 0.12 | 0.40 | **0.000** | 0.12 |
| 1980:4 | **0.000** | **0.000** | 0.18 | 0.55 | **0.000** | 0.056 |
| 1983:1 | 0.58 | **0.000** | 0.16 | 0.32 | **0.000** | **0.015** |

**Panel E: VECM DGP with AIC lag order ending 2008:2, with start date given in table**

| | Interest rate ($i$) equation | | | Inflation rate ($\pi$) equation | | |
|---|---|---|---|---|---|---|
| Start date | LM | JB | Q | LM | JB | Q |
| 1980:2 | 0.34 | 0.40 | 0.22 | 0.39 | 0.14 | 0.42 |
| 1980:3 | 0.39 | 0.34 | 0.34 | 0.37 | **0.015** | 0.47 |
| 1980:4 | 0.23 | **0.019** | 0.63 | 0.21 | **0.002** | 0.76 |
| 1983:1 | 0.78 | 0.16 | 0.69 | 0.27 | **0.035** | 0.47 |

**Panel F: VECM DGP with BIC lag order ending 2008:2, with start date given in table**

| | Interest rate ($i$) equation | | | Inflation rate ($\pi$) equation | | |
|---|---|---|---|---|---|---|
| Start date | LM | JB | Q | LM | JB | Q |
| 1980:2 | **0.000** | **0.000** | 0.078 | 0.085 | **0.000** | 0.31 |
| 1980:3 | **0.000** | **0.000** | 0.13 | 0.20 | **0.000** | 0.16 |
| 1980:4 | **0.000** | **0.000** | 0.19 | 0.18 | **0.000** | 0.087 |
| 1983:1 | 0.42 | **0.000** | 0.24 | 0.34 | **0.000** | **0.029** |

**Panel G: VECM DGP with AIC lag order ending 2008:1, with start date given in table**

| | Interest rate ($i$) equation | | | Inflation rate ($\pi$) equation | | |
|---|---|---|---|---|---|---|
| Start date | LM | JB | Q | LM | JB | Q |
| 1980:2 | **0.003** | **0.044** | 0.082 | 0.28 | **0.000** | 0.41 |
| 1980:3 | **0.023** | **0.004** | **0.019** | 0.32 | **0.000** | 0.47 |
| 1980:4 | **0.017** | **0.000** | 0.28 | 0.34 | **0.000** | 0.35 |
| 1983:1 | 0.95 | 0.57 | 0.46 | 0.29 | **0.018** | 0.47 |

**Panel H: VECM DGP with BIC lag order ending 2008:1, with start date given in table**

| | Interest rate ($i$) equation | | | Inflation rate ($\pi$) equation | | |
|---|---|---|---|---|---|---|
| Start date | LM | JB | Q | LM | JB | Q |
| 1980:1 | **0.000** | **0.000** | 0.19 | 0.158 | **0.000** | 0.054 |
| 1980:2 | 0.58 | **0.000** | 0.16 | 0.317 | **0.000** | **0.015** |
| 1980:3 | **0.000** | **0.000** | **0.000** | **0.000** | **0.000** | **0.000** |
| 1980:4 | **0.000** | **0.000** | **0.006** | 0.081 | **0.000** | 0.30 |
| 1983:1 | **0.000** | **0.000** | **0.039** | 0.056 | **0.000** | 0.26 |