A Panel Analysis of the Fisher Effect with an Unobserved $I(1)$ World Real Interest Rate

Gerdie Everaert∗†

SHERPPA, Ghent University

April 22, 2014

Abstract

The Fisher effect states that inflation expectations should be reflected in nominal interest rates in a one-for-one manner to compensate for changes in the purchasing power of money. Despite its wide acceptance in theory, much of the empirical work fails to find favorable evidence. This paper examines the Fisher effect in a panel of 21 OECD countries over the period 1983-2010. Using the Panel Analysis of Non-stationarity in Idiosyncratic and Common Components (PANIC), a non-stationary common factor is detected in the real interest rate. This may reflect permanent common shifts in e.g. time preferences, risk aversion and the steady-state growth rate of technological change. We therefore control for an unobserved non-stationary common factor in estimating the Fisher equation using both the Common Correlated Effects Pooled (CCEP) and the Continuously Updated (Cup) estimation approach. The impact of inflation on the nominal interest rate is found to be insignificantly different from 1, providing support of the Fisher effect.

JEL Classification: C23, E31, E43

Keywords: Fisher effect, panel cointegration, cross-sectional dependence, unobserved common factors

∗I thank an anonymous referee, Koen Inghelbrecht, Lorenzo Pozzi and the participants of the Amsterdam Econometric Seminar (Tinbergen Institute, November 2011), the 5th CSDA International Conference on Computational and Financial Econometrics (University of London, December 2011) and of the 20th Annual Symposium of the Society for Nonlinear Dynamics and Econometrics (Istanbul, April 2012) for helpful suggestions and constructive comments. I further acknowledge financial support from the Interuniversity Attraction Poles Program - Belgian Science Policy, contract no. P5/21.

†Sint-Pietersplein 6, B-9000 Ghent, Belgium. Tel.: +32 92647878; fax: +32 92648996. E-mail address: gerdie.everaert@ugent.be.
1 Introduction

The Fisher effect states that inflation expectations should be reflected in nominal interest rates in a one-for-one manner to compensate for changes in the purchasing power of money (Fisher, 1930). This implies that the ex ante real interest rate, defined as the difference between the nominal interest rate and expected inflation, is not affected by changes in inflation expectations. While probably not being valid in the short run, the Fisher effect is expected to hold as a long-run equilibrium concept. Insofar as permanent changes in expected inflation originate from permanent shocks in the rate of money growth, this is in accordance with the so-called long-run superneutrality of money. Despite its wide acceptance in theory, most of the empirical work fails to find convincing evidence in favor of the Fisher effect. As nominal interest rates and inflation are typically found to be non-stationary, the long-run Fisher effect implies that these two variables should cointegrate with a unit slope coefficient such that the real interest rate is stationary and therefore not affected by permanent shocks to inflation. A survey of this literature shows that unit root tests find real interest rates to be non-stationary (see e.g. Rose, 1988; Rapach and Weber, 2004; Lai, 2008) while cointegration analysis either finds no cointegration between nominal interest rates and inflation (see e.g. MacDonald and Murphy, 1989; Koustas and Serletis, 1999) or when cointegration is found the estimated slope is significantly less (see e.g. Evans and Lewis, 1995) or significantly greater (see e.g. Crowder and Hoffman, 1996) than one.

A number of theoretical explanations for the empirical failure of the Fisher effect have been put forward. First, inflation expectations are not observed and are therefore replaced by ex post observed inflation to calculate ex post real interest rates. Evans and Lewis (1995) argue that the alleged permanent component in these ex post real interest rates may be due to people incorporating anticipated shifts in the inflation process into their expectations implying a persistent deviation of observed inflation from expected inflation over the period these shifts do not materialize. Second, Darby (1975) argues that the presence of taxes on interest income implies that nominal interest rates have to rise by more than one-for-one in response to a change in inflation expectations in order to keep the after-tax real interest rate constant. These tax effects may thus explain why nominal interest rates and inflation cointegrate with a slope coefficient greater than one. Third, in the seminal papers of Mundell (1963) and Tobin (1965) higher inflation causes a substitution out of money balances into bonds and real assets, putting downward pressure on real interest rates. This may explain why nominal interest rates and inflation cointegrate with a slope coefficient less than one.

A plausible econometric explanation is that the existing empirical evidence on the Fisher effect is flawed as it is based on a country-by-country analysis often using at most 50 annual observations. Using such relatively small data span results in low power of conventional unit root and cointegration tests, especially when there is high persistence under the alternative hypothesis of stationarity. Westerlund (2008) therefore suggests to test the Fisher effect in a panel of quarterly data covering 20 OECD countries between 1980 and 2004. Taking into account error cross-sectional dependence when testing for cointegration, he shows that the null hypothesis of no panel cointegration between interest rates and inflation can be rejected while the hypothesis of a unit slope coefficient on inflation cannot be rejected.
An alternative, but yet unexplored, explanation is that the real factors behind the real interest rate are not stable over time. Standard neoclassical growth models with household intertemporal utility maximization imply that the real interest rate is a function of time preference, risk aversion and the steady-state growth rate of technological change. While time preference and risk aversion are generally believed to be fairly stable, or at least changing only slowly over extended periods of time, shifts in steady-state growth, such as the ‘productivity slowdown’ of the early 1970s and the ‘New Economy’ resurgence of growth in the late 1990s, have been widely documented in the literature (see e.g. Oliner and Sichel, 2000; Roberts, 2001). Additional determinants of real interest rates suggested in the literature are demographic changes, changes in the stance of fiscal policy and the evolution of public debt, changes in the taxation of profits, (de)regulation of financial markets, ... (see e.g. Blanchard and Summers, 1984; Chadha andDimsdale, 1999; Ardlagna, 2009). Permanent shifts in any of these factors induce a unit root in the real interest rate and by extension in the residuals of a regression of the nominal interest rate on inflation. However, this does not automatically invalidate the Fisher hypothesis of a one-for-one relation between nominal interest rates and inflation. Basically, the sources of the non-stationary behaviour of real interest rates are omitted non-stationary variables that should be added to a regression of nominal interest rates on inflation for this to be a cointegrating relation. Note that this argument not only explains the failure to find cointegration but also the large variety of estimated slope coefficients over empirical studies that do find cointegration as Everaert (2011) shows that omitting relevant non-stationary variables yields spurious estimation results with standard cointegration tests indicating these results to be a cointegration regression in far too many cases.

Ideally, the non-stationary determinants of real interest rates should be included in a regression of nominal interest rates on inflation. However, there is a large variety of possible determinants which are, moreover, not directly observable or at least hard to measure. A promising way out of this problem is to identify these determinants by exploiting the strong cross-section correlation between interest rates observed over countries. Increasing economic integration leads to a substantial degree of linkage between real interest rates of different countries and has led a number of authors to construct and analyze a world real interest rate (Barro and Sala-i Martin, 1990; Koedijk et al., 1994; Lee, 2002) or to relate national real interest rates to international rather than to domestic events (Blanchard and Summers, 1984).

This paper uses recent advances in panel data econometrics to identify and account for unobserved common factors in a panel of quarterly data for nominal interest rates and inflation covering 21 OECD countries between 1983 and 2010. The analysis consists of two steps. In the first step, we investigate the integration properties of the data using the Panel Analysis of Non-stationarity in Idiosyncratic and Common Components (PANIC) of Bai and Ng (2004). The most important conclusion from this analysis is that real interest rates and the error terms of a fixed effects (FE) regression of nominal interest rates on inflation can both be decomposed in a single non-stationary common factor and a stationary idiosyncratic component. The latter finding is consistent with Westerlund (2008) who also shows that real interest rates and Fisher equation regression errors are stationary after filtering out common factors. This leads him to conclude that nominal interest rates and inflation cointegrate. However, Westerlund does not analyze the integration properties of the common factors but simply assumes these to be stationary.
Our finding of a non-stationary common factor invalidates his conclusion and implies that regression results ignoring this common factor are spurious (see Urbain and Westerlund, 2011). In the second step we therefore estimate the relation between nominal interest rates and inflation taking into account a common non-stationary component in real interest rates. In particular, we use the common correlated effects pooled (CCEP) estimation approach proposed by Pesaran (2006) and Kapetanios et al. (2011) and the continuously-updated (Cup) estimation approach proposed by Bai et al. (2009). The advantage of both approaches is that they can consistently estimate the relationship between nominal interest rates and inflation under very general integration properties of the data without the need to identify and measure the determinants of real interest rates as long as these determinants are common to all countries. Thus, rather than treating the cross-section correlation as a nuisance, which requires adjustment of standard unit root and cointegration tests, we exploit the comovement of interest rates to identify unobserved common determinants of real interest rates. This allows us to test the Fisher effect in the presence of a non-stationary world real interest rate. Endogeneity of observed inflation induced by a rational expectations forecasting error is taken into account using CupBC, a bias-corrected version of the Cup estimator, and CCEP_GMM, a GMM version of the CCEP estimator. We also propose how to test for cointegration from the error terms of the CCEP and CUP estimators. A small-scaled Monte Carlo simulation shows that the CupBC and CCEP_GMM estimators and cointegration tests perform reasonably well for the modest sample size $T = 112, N = 21$ that is available for our empirical analysis. From the estimation results, the hypothesis of a one-for-one relation between the nominal interest rate and inflation cannot be rejected using either the CupBC or the CCEP_GMM estimator.

The paper is organized as follows. Section 2 outlines the standard Fisher equation. Section 3 analyses the time series properties of the data. Section 4 augments the standard Fisher equation with a non-stationary factor and discusses how this factor-augmented equation can be estimated. Section 5 presents the estimation results and analyses the small sample properties of the proposed estimators using a Monte Carlo simulation. Section 6 concludes.

2 The standard Fisher equation

Fisher (1930) hypothesized that inflation expectations should be reflected in the nominal interest rate in a one-for-one manner to compensate for changes in the purchasing power of money. This implies that the real interest rate should be invariant to changes in expected inflation. Formally, the Fisher hypothesis can be stated as $\beta = 1$ in

$$i_{it} = r_{it}^e + \beta \pi_{it}^e,$$

where $i_{it}$ is the nominal interest rate observed in country $i$ at time $t$, $r_{it}^e$ is the ex ante real interest rate and $\pi_{it}^e$ the expected rate of inflation.

The validity of the Fisher effect cannot be directly analyzed using (1) as $r_{it}^e$ and $\pi_{it}^e$ are unobserved ex ante variables. The Fisher equation can be written in terms of ex post observed variables after making
two assumptions. First, the ex ante real interest rate is driven by real factors which are typically assumed to be more or less stable over time such that \( r_{it}^e \) can be written as
\[
r_{it}^e = \alpha_i + \nu_{it},
\]
(2)
where \( \alpha_i \) is a country-specific constant and \( \nu_{it} \) is a stationary error term which captures temporary fluctuations in \( r_{it}^e \). Second, assuming rational expectations
\[
\pi_{it} = \pi_{it}^e + \zeta_{it},
\]
(3)
where \( \zeta_{it} \) is a mean zero stationary forecast error orthogonal to any information known at time \( t \). Inserting (2) and (3) in (1) yields
\[
i_{it} = \alpha_i + \beta \pi_{it} + \epsilon_{it},
\]
(4)
where \( \epsilon_{it} \) is a composite error term, comprised of the forecast error \(-\beta \zeta_{it}\) and the term \( \nu_{it} \).

Equation (4) forms the basis for testing the Fisher hypothesis. Given that \( i_{it} \) and \( \pi_{it} \) are typically found to be \( I(1) \) series, this is nowadays done using unit root testing and cointegration analysis (see e.g. MacDonald and Murphy, 1989). In fact, this alleged non-stationarity significantly simplifies testing the Fisher hypothesis. First, when \( i_{it} \) and \( \pi_{it} \) are cointegrated, superconsistency of the LS estimator implies that (4) can be estimated ignoring the correlation between \( \pi_{it} \) and \( \epsilon_{it} \) and any dynamics in \( \epsilon_{it} \). Note that cointegration between \( i_{it} \) and \( \pi_{it} \) requires \( \epsilon_{it} \) to be stationary, but does not depend on the specific value of \( \beta \) with \( \beta = 1 \) then being denoted as the full Fisher effect and \( \beta \neq 1 \) as the partial Fisher effect. Popular theoretical explanations for \( \beta \neq 1 \) are (i) taxes on interest income which imply that the nominal interest rate has to raise by more than one-for-one \( (\beta > 1) \) in response to a change in inflation expectations to keep the after-tax real interest rate constant and (ii) portfolio shifts out of money balances into interest bearing assets in response to an increase in inflation expectations which puts downward pressure on real interest rates \( (\beta < 1) \).

Second, defining the ex post observed real interest rate \( r_{it} \) as
\[
r_{it} \equiv i_{it} - \pi_{it} = \alpha_i - (1 - \beta) \pi_{it} + \epsilon_{it},
\]
(5)
the Fisher hypothesis boils down to a simple unit root test on \( r_{it} \). This is a test for the full Fisher effect as stationarity of \( r_{it} \) requires \( \beta = 1 \) when \( \pi_{it} \) is found to be non-stationary.
3 Time series properties of the data

As a first step in the empirical analysis, in this section we look at the time series properties of the data. We start with unit root and cointegration tests on the country-specific level. We then look at the cross-sectional correlation in the data. As strong evidence of cross-sectional dependence is found, we next use second generation panel tests and decompose all series in a common factor and an idiosyncratic component and analyze the time series properties of these components separately using PANIC.

3.1 Data

We use quarterly data taken from the International Financial Statistics database of the International Monetary Fund. The sample includes 21 OECD countries (see Table 1 for the full list of countries) covering the period from 1983Q1 to 2010Q4. The selection of the countries and the sample period is determined by data availability and the aim to have as many time periods as possible for a reasonably large set of countries. The nominal interest rate is either the three months treasury bill rate, if available, or the three months money market rate. Expected inflation is proxied by the ex post observed inflation rate calculated as the year-on-year percent change in the consumer price index (CPI). We use year-on-year percent changes as this attenuates the strong noise in annualized quarter-on-quarter percent changes (also see Bekaert and Wang, 2010). Year-on-year changes are also the most prominent way inflation is reported and is also the subject of most professional inflation forecasts (central bank forecasts, survey forecasts, ...). Studies examining the forecasting power of alternative methods (see e.g. Stock and Watson, 1999; Ang et al., 2007) typically also focus on a one-year inflation horizon. One disadvantage of using annual inflation at a quarterly frequency is that we will have to take into account that the forecast errors $\zeta_{it}$ in equation (3) follow a MA(3) process due to overlapping observations. This will be taken into account in the econometric analysis.

3.2 Country-by-country analysis

As outlined in Section 2, a first way to test the Fisher effect is to analyze the time series properties of $i_{it}$, $\pi_{it}$ and $r_{it}$. Table 1 presents country-by-country ADF-GLS unit root tests for a model with a constant and no trend. This ADF-GLS test is the modified augmented Dickey and Fuller (1979) (ADF) test based on generalized least squares (GLS) demeaning of the data as suggested by Elliott et al. (1996). Compared to the standard ADF test, removal of the constant term by means of GLS demeaning yields substantial power improvements, especially in small samples. More details on the exact implementation of the test are provided as a note to Table 1. First looking at $i_{it}$ and $\pi_{it}$, the unit root hypothesis cannot be rejected at the 5% level of significance for any of the individual countries. Given this finding, the full Fisher effect requires $r_{it}$ to be stationary. However, at the 5% level of significance, the null hypothesis of a unit root in $r_{it}$ can only be rejected for Norway and Portugal. Thus, these results provide no clear support for the full Fisher.

An alternative way to test the Fisher effect is to infer whether there is a one-for-one cointegration relation between $i_{it}$ and $\pi_{it}$. Table 2 reports OLS (together with DOLS and FMOLS) coefficient estimates
for $\beta$ in equation (4). ADF cointegration test results on the OLS residuals can be found in Table 1. The results are clearly not in support of the Fisher effect. The null hypothesis of a unit coefficient is clearly rejected in most countries. More importantly, the ADF cointegration test results show that the null hypothesis of no cointegration cannot be rejected in 19 out of 21 countries. This implies that the country-specific regression results should be considered spurious.

As the individual county data span a relatively short period of 28 years, the failure to find evidence in favor of the Fisher effect may be due to inaccurate estimation and/or a lack of power to reject the null hypothesis of a unit root in either $r_{it}$ or $\bar{\epsilon}_{it}$. Exploiting the panel dimension of the data using the Maddala and Wu (1999) (MW) panel unit root/cointegration test, the null hypothesis of a unit root can indeed be rejected for both $r_{it}$ or $\bar{\epsilon}_{it}$. Moreover, the FE estimate for $\beta$ in the Fisher equation (4) is 1.09 but and found to be insignificantly different from 1 using bootstrap inference (see Table 7 below for highly similar results (not reported) are obtained for a cointegration test using the DOLS and FMOLS estimation results.

### Table 1: Country-by-country unit root and cointegration tests

<table>
<thead>
<tr>
<th>Country</th>
<th>$i_{it}$</th>
<th>$\pi_{it}$</th>
<th>$r_{it}$</th>
<th>$\hat{e}_{it}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Australia</td>
<td>9</td>
<td>-0.31</td>
<td>0.56</td>
<td>4</td>
</tr>
<tr>
<td>Austria</td>
<td>1</td>
<td>-1.09</td>
<td>0.25</td>
<td>4</td>
</tr>
<tr>
<td>Belgium</td>
<td>1</td>
<td>0.23</td>
<td>0.75</td>
<td>8</td>
</tr>
<tr>
<td>Canada</td>
<td>7</td>
<td>-0.44</td>
<td>0.51</td>
<td>4</td>
</tr>
<tr>
<td>Denmark</td>
<td>2</td>
<td>0.21</td>
<td>0.74</td>
<td>8</td>
</tr>
<tr>
<td>Finland</td>
<td>0</td>
<td>-0.16</td>
<td>0.62</td>
<td>4</td>
</tr>
<tr>
<td>France</td>
<td>1</td>
<td>0.27</td>
<td>0.76</td>
<td>9</td>
</tr>
<tr>
<td>Germany</td>
<td>1</td>
<td>-1.34</td>
<td>0.17</td>
<td>0</td>
</tr>
<tr>
<td>Greece</td>
<td>0</td>
<td>0.03</td>
<td>0.68</td>
<td>5</td>
</tr>
<tr>
<td>Ireland</td>
<td>2</td>
<td>-0.26</td>
<td>0.58</td>
<td>5</td>
</tr>
<tr>
<td>Italy</td>
<td>1</td>
<td>0.27</td>
<td>0.76</td>
<td>5</td>
</tr>
<tr>
<td>Japan</td>
<td>3</td>
<td>-0.58</td>
<td>0.46</td>
<td>4</td>
</tr>
<tr>
<td>Netherlands</td>
<td>1</td>
<td>-1.56</td>
<td>0.11</td>
<td>4</td>
</tr>
<tr>
<td>New Zealand</td>
<td>12</td>
<td>-1.03</td>
<td>0.27</td>
<td>12</td>
</tr>
<tr>
<td>Norway</td>
<td>0</td>
<td>-1.03</td>
<td>0.28</td>
<td>4</td>
</tr>
<tr>
<td>Portugal</td>
<td>2</td>
<td>-0.61</td>
<td>0.44</td>
<td>1</td>
</tr>
<tr>
<td>Spain</td>
<td>3</td>
<td>-0.21</td>
<td>0.60</td>
<td>4</td>
</tr>
<tr>
<td>Sweden</td>
<td>8</td>
<td>0.01</td>
<td>0.68</td>
<td>1</td>
</tr>
<tr>
<td>Switzerland</td>
<td>6</td>
<td>-1.82</td>
<td>0.06</td>
<td>6</td>
</tr>
<tr>
<td>UK</td>
<td>1</td>
<td>-0.27</td>
<td>0.58</td>
<td>8</td>
</tr>
<tr>
<td>US</td>
<td>1</td>
<td>-0.88</td>
<td>0.33</td>
<td>5</td>
</tr>
</tbody>
</table>

Notes: The unit root test statistic $\tau_\mu$ is the augmented Dickey-Fuller (ADF) generalised least squares (ADF-GLS) test of Elliott et al. (1996) for a specification including a constant. The lag length $k$ is selected using the modified Akaike information criterion (MAIC) suggested by Ng and Perron (2001) with the maximum lag length $k_{\max}$ set according to the Schwert (1989) rule: $k_{\max} = \text{int}(12(T/100)^{0.25}) = 12$. The cointegration test statistic $\tau$ is a standard ADF test for a specification with no deterministic terms on the estimated residuals of the country-specific OLS regressions ($\hat{\sigma}_{it}^2$) reported in Table 2. The $p$-values for the country-specific unit root and cointegration tests are obtained by simulating their finite-sample distributions (based on 20,000 Monte Carlo iterations) taking into account, as in Cook and Manning (2004), that the augmentation of the test equation is optimized using the MAIC criterion. The 1% and 5% critical values taken from these distributions are -2.59 and -2.02 for the ADF-GLS unit root test and -3.88 and -3.30 for the ADF cointegration test on the OLS residuals.

### 3.3 Panel analysis

As the individual county data span a relatively short period of 28 years, the failure to find evidence in favor of the Fisher effect may be due to inaccurate estimation and/or a lack of power to reject the null hypothesis of a unit root in either $r_{it}$ or $\bar{\epsilon}_{it}$. Exploiting the panel dimension of the data using the Maddala and Wu (1999) (MW) panel unit root/cointegration test, the null hypothesis of a unit root can indeed be rejected for both $r_{it}$ or $\bar{\epsilon}_{it}$. Moreover, the FE estimate for $\beta$ in the Fisher equation (4) is 1.09 but and found to be insignificantly different from 1 using bootstrap inference (see Table 7 below for

\[1\] Highly similar results (not reported) are obtained for a cointegration test using the DOLS and FMOLS estimation results.

\[2\] Results are available on request.
Next, we compute the cross-sectional dependence (CD) test of Pesaran (2004). This shows that the null

full results in a slightly different specification including EMS dummies). However, there are at least two
reasons for why these results may not be trustworthy. First, the MW panel unit root/cointegration test is
only valid for combining p-values from cross-sectionally independent tests. O’Connell (1998) documents
that the alleged power gain of panel unit root tests developed under cross-sectional independence may in
practice very well be the consequence of nontrivial size distortions induced by cross-sectional dependence,
raising the real size of tests with a nominal size of 5% to as much as 50%. A similar conclusion can be
found in Banerjee et al. (2004, 2005). Second, insofar as the cross-sectional dependence is induced by
non-stationary omitted common factors that are relatively small compared to the stationary component
in the data, unit root tests are biased towards rejection of the null hypothesis of a unit root (Bai and
Ng, 2004). Below we therefore test for cross-sectional dependence and the presence of (non-stationary)
common factors in the data and in the residuals of the FE Fisher regression.

### 3.3.1 Testing for cross-sectional dependence

Table 3 presents information on the extent of the cross-sectional dependence in the original data, in the
residuals of the ADF-GLS regressions and in the FE residuals $\tilde{z}_{it}^{FE}$. For those series that are potentially
non-stationary, we also report results for the first differences to avoid spurious non-zero correlations. We
first compute the average cross-correlation coefficient $\tilde{p}$ which is the average of the country-by-country
cross-correlation coefficients $\tilde{p}_{ij}$ (for $i \neq j$). The original data and the residuals from both the ADF-
GLS regressions and the FE Fisher regression all exhibit considerable positive cross-sectional correlation.

Next, we compute the cross-sectional dependence (CD) test of Pesaran (2004). This shows that the null

<table>
<thead>
<tr>
<th>Country</th>
<th>$\hat{\beta}$</th>
<th>se</th>
<th>t-stat</th>
<th>p-val</th>
<th>$\hat{\beta}$</th>
<th>se</th>
<th>t-stat</th>
<th>p-val</th>
<th>$\hat{\beta}$</th>
<th>se</th>
<th>t-stat</th>
<th>p-val</th>
</tr>
</thead>
<tbody>
<tr>
<td>Australia</td>
<td>1.26</td>
<td>0.09</td>
<td>2.98</td>
<td>0.00</td>
<td>1.60</td>
<td>0.18</td>
<td>3.30</td>
<td>0.00</td>
<td>1.45</td>
<td>0.18</td>
<td>2.53</td>
<td>0.01</td>
</tr>
<tr>
<td>Austria</td>
<td>1.34</td>
<td>0.13</td>
<td>2.54</td>
<td>0.01</td>
<td>1.79</td>
<td>0.32</td>
<td>2.49</td>
<td>0.01</td>
<td>1.57</td>
<td>0.27</td>
<td>2.10</td>
<td>0.04</td>
</tr>
<tr>
<td>Belgium</td>
<td>1.18</td>
<td>0.14</td>
<td>1.30</td>
<td>0.20</td>
<td>1.67</td>
<td>0.42</td>
<td>1.61</td>
<td>0.11</td>
<td>1.35</td>
<td>0.31</td>
<td>1.13</td>
<td>0.26</td>
</tr>
<tr>
<td>Canada</td>
<td>1.56</td>
<td>0.14</td>
<td>4.07</td>
<td>0.00</td>
<td>2.10</td>
<td>0.29</td>
<td>3.80</td>
<td>0.00</td>
<td>1.97</td>
<td>0.27</td>
<td>3.57</td>
<td>0.00</td>
</tr>
<tr>
<td>Denmark</td>
<td>1.54</td>
<td>0.19</td>
<td>2.86</td>
<td>0.01</td>
<td>1.62</td>
<td>0.48</td>
<td>1.28</td>
<td>0.20</td>
<td>1.51</td>
<td>0.40</td>
<td>1.28</td>
<td>0.20</td>
</tr>
<tr>
<td>Finland</td>
<td>1.80</td>
<td>0.11</td>
<td>7.58</td>
<td>0.00</td>
<td>2.07</td>
<td>0.21</td>
<td>4.97</td>
<td>0.00</td>
<td>1.95</td>
<td>0.20</td>
<td>4.84</td>
<td>0.00</td>
</tr>
<tr>
<td>France</td>
<td>1.32</td>
<td>0.11</td>
<td>2.96</td>
<td>0.00</td>
<td>1.61</td>
<td>0.33</td>
<td>1.83</td>
<td>0.07</td>
<td>1.30</td>
<td>0.24</td>
<td>1.62</td>
<td>0.11</td>
</tr>
<tr>
<td>Germany</td>
<td>1.33</td>
<td>0.11</td>
<td>2.95</td>
<td>0.00</td>
<td>1.60</td>
<td>0.25</td>
<td>2.42</td>
<td>0.02</td>
<td>1.51</td>
<td>0.22</td>
<td>2.29</td>
<td>0.02</td>
</tr>
<tr>
<td>Greece</td>
<td>0.96</td>
<td>0.05</td>
<td>-0.74</td>
<td>0.46</td>
<td>1.03</td>
<td>0.10</td>
<td>0.32</td>
<td>0.75</td>
<td>1.01</td>
<td>0.10</td>
<td>0.04</td>
<td>0.97</td>
</tr>
<tr>
<td>Ireland</td>
<td>0.80</td>
<td>0.14</td>
<td>-1.42</td>
<td>0.16</td>
<td>0.83</td>
<td>0.41</td>
<td>-0.41</td>
<td>0.68</td>
<td>0.83</td>
<td>0.29</td>
<td>-0.58</td>
<td>0.56</td>
</tr>
<tr>
<td>Italy</td>
<td>1.48</td>
<td>0.07</td>
<td>6.73</td>
<td>0.00</td>
<td>1.83</td>
<td>0.18</td>
<td>4.69</td>
<td>0.00</td>
<td>1.58</td>
<td>0.15</td>
<td>3.84</td>
<td>0.00</td>
</tr>
<tr>
<td>Japan</td>
<td>1.68</td>
<td>0.13</td>
<td>5.35</td>
<td>0.00</td>
<td>2.10</td>
<td>0.25</td>
<td>4.37</td>
<td>0.00</td>
<td>2.00</td>
<td>0.24</td>
<td>4.07</td>
<td>0.00</td>
</tr>
<tr>
<td>Netherlands</td>
<td>0.57</td>
<td>0.19</td>
<td>-2.31</td>
<td>0.02</td>
<td>0.56</td>
<td>0.50</td>
<td>-0.89</td>
<td>0.37</td>
<td>0.50</td>
<td>0.41</td>
<td>-1.02</td>
<td>0.31</td>
</tr>
<tr>
<td>New Zealand</td>
<td>1.10</td>
<td>0.05</td>
<td>1.88</td>
<td>0.06</td>
<td>1.21</td>
<td>0.10</td>
<td>2.17</td>
<td>0.03</td>
<td>1.20</td>
<td>0.10</td>
<td>2.11</td>
<td>0.04</td>
</tr>
<tr>
<td>Norway</td>
<td>1.37</td>
<td>0.11</td>
<td>3.39</td>
<td>0.00</td>
<td>1.65</td>
<td>0.22</td>
<td>2.91</td>
<td>0.00</td>
<td>1.57</td>
<td>0.21</td>
<td>2.71</td>
<td>0.01</td>
</tr>
<tr>
<td>Portugal</td>
<td>0.91</td>
<td>0.04</td>
<td>-2.48</td>
<td>0.01</td>
<td>1.00</td>
<td>0.10</td>
<td>-0.03</td>
<td>0.97</td>
<td>0.95</td>
<td>0.07</td>
<td>-0.87</td>
<td>0.38</td>
</tr>
<tr>
<td>Spain</td>
<td>1.29</td>
<td>0.08</td>
<td>3.60</td>
<td>0.00</td>
<td>1.50</td>
<td>0.18</td>
<td>2.69</td>
<td>0.01</td>
<td>1.41</td>
<td>0.16</td>
<td>2.52</td>
<td>0.01</td>
</tr>
<tr>
<td>Sweden</td>
<td>1.14</td>
<td>0.07</td>
<td>1.89</td>
<td>0.06</td>
<td>1.33</td>
<td>0.13</td>
<td>2.45</td>
<td>0.02</td>
<td>1.28</td>
<td>0.13</td>
<td>2.06</td>
<td>0.04</td>
</tr>
<tr>
<td>Switzerland</td>
<td>1.23</td>
<td>0.10</td>
<td>2.41</td>
<td>0.02</td>
<td>1.56</td>
<td>0.16</td>
<td>3.46</td>
<td>0.00</td>
<td>1.48</td>
<td>0.17</td>
<td>2.79</td>
<td>0.01</td>
</tr>
<tr>
<td>UK</td>
<td>1.29</td>
<td>0.10</td>
<td>2.90</td>
<td>0.00</td>
<td>1.68</td>
<td>0.18</td>
<td>3.89</td>
<td>0.00</td>
<td>1.54</td>
<td>0.18</td>
<td>3.00</td>
<td>0.00</td>
</tr>
<tr>
<td>US</td>
<td>1.19</td>
<td>0.15</td>
<td>1.24</td>
<td>0.22</td>
<td>1.72</td>
<td>0.37</td>
<td>1.95</td>
<td>0.05</td>
<td>1.58</td>
<td>0.30</td>
<td>1.91</td>
<td>0.06</td>
</tr>
</tbody>
</table>

Notes: Reported are country-specific coefficient estimates for $\beta$ in equation (4). The t-statistics and p-values are calculated under the null hypothesis that $\beta = 1$. The DOLS estimator is based on two leads and lags. The FMOLS estimator is based on the Bartlett kernel with bandwidth set to five.
hypothesis of no cross-sectional dependence is strongly rejected for all variables and residuals. The finding of significant cross-sectional dependence implies that first generation panel unit root and cointegration tests, which ignore cross-sectional dependence, are inappropriate.

Table 3: Cross-sectional dependence test

<table>
<thead>
<tr>
<th></th>
<th>Sample period: 1983:Q1-2010:Q4, 21 countries</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Levels First-differences</td>
</tr>
<tr>
<td></td>
<td>( \hat{\rho} ) CD p-val</td>
</tr>
<tr>
<td>Original data i_{it}</td>
<td>0.79 120.99 0.00</td>
</tr>
<tr>
<td>( \pi_{it} )</td>
<td>0.54 83.50 0.00</td>
</tr>
<tr>
<td>( r_{it} )</td>
<td>0.58 89.19 0.00</td>
</tr>
<tr>
<td>Residuals ADF-GLS regression i_{it}</td>
<td>0.22 32.20 0.00</td>
</tr>
<tr>
<td>( \pi_{it} )</td>
<td>0.25 36.65 0.00</td>
</tr>
<tr>
<td>( r_{it} )</td>
<td>0.15 22.63 0.00</td>
</tr>
<tr>
<td>Residuals Fisher FE regression ( \hat{\epsilon}^{FE}_{it} )</td>
<td>0.53 81.49 0.00</td>
</tr>
</tbody>
</table>

Notes: the average cross-correlation coefficient \( \bar{\hat{\rho}} = (2/N(N-1)) \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \hat{\rho}_{ij} \) is the average of the country-by-country cross-correlation coefficients \( \hat{\rho}_{ij} \) (for \( i \neq j \)). CD is the Pesaran (2004) test defined as \( \sqrt{2T/N(N-1)} \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \hat{\rho}_{ij} \), which is asymptotically standard normal under the null of cross-sectional independence.

3.3.2 Common factor structure

In the recent panel literature, cross-sectional dependence is typically assumed to stem from omitted common variables or global shocks that affect each country differently and is therefore modelled using a common factor structure with country-specific factor loadings (see e.g. Bai and Ng, 2004; Coakley et al., 2006; Pesaran, 2006). More precisely, assume that the data generating process (DGP) of a series \( X_{it} \) is given by the following prototypical common factor model

\[
X_{it} = \lambda_i^iF_t + e_{it},
\]

where \( F_t \) is an \( r \times 1 \) vector of common factors with country-specific factor loadings \( \lambda_i \) and \( e_{it} \) is an idiosyncratic error term. Cross-sectional dependence stems from the common component \( \lambda_i^iF_t \) which is correlated over countries. The series \( X_{it} \) is non-stationary if at least one of the common factors in \( F_t \) is non-stationary, or the idiosyncratic error \( e_{it} \) is non-stationary, or both.

Table 4 reports results for estimating the total number of relevant common factors \( r \) in the data for \( i_{it}, \pi_{it} \) and \( r_{it} \) and in the residuals of the FE Fisher regression using the panel information criteria suggested by Bai and Ng (2002). As consistency of these criteria requires stationary data, we take first-differences of all series (also see Bai and Ng, 2004, p. 1144). The top panel of Table 4 reports the \( IC_{1,2,3}, PC_{1,2,3} \), \( AIC_3 \) and \( BIC_3 \) criteria with the maximum number of factors (\( r_{max} \)) ranging from 2 to 6. Using the \( PC_{1,2,3} \) and \( AIC_3 \) criteria, the optimal number of factors is found to increase with \( r_{max} \) for all series. The results of the \( IC_{1,2,3} \) and \( BIC_3 \) criteria are more stable over alternative choices of \( r_{max} \) and point
to a single common factor in all series, except when using the \( IC_1 \) and \( IC_3 \) criteria on \( i_t \) for which the number of common factors increases with \( r_{\text{max}} \) and when using the \( BIC_3 \) criterion on \( r_{\text{it}} \) and \( \hat{\epsilon}_{\text{FE}}^\text{it} \) for which no factors are found for lower values of \( r_{\text{max}} \). Note that the contradictory results over the various information criteria are in line with the Monte Carlo simulations in Bai and Ng (2002) which show that in samples of moderate size, i.e. \( \min\{N,T\} < 40 \), the \( IC \) criteria tend to underparameterize (especially for larger values of \( r \)) while the \( PC \) criteria tend to overparameterize (estimated number of components is found to increase with \( r_{\text{max}} \)), with the problem being even more severe for the \( AIC \) and \( BIC \) criteria.

Taking this into account, the information criteria suggest that there is at least 1 common factor in \( i_{\text{it}} \), \( \pi_{\text{it}} \) and \( r_{\text{it}} \) and in the residuals of the FE Fisher regression.

**Table 4: Estimating the number of common factors**

<table>
<thead>
<tr>
<th>Sample period: 1983:Q1-2010:Q4, 21 countries</th>
</tr>
</thead>
<tbody>
<tr>
<td>( r_{\text{max}} = 2 )</td>
</tr>
<tr>
<td>( i_{\text{it}} )</td>
</tr>
<tr>
<td>Data in first-differences: estimating the total number of factors</td>
</tr>
<tr>
<td>( IC_1 )</td>
</tr>
<tr>
<td>( IC_2 )</td>
</tr>
<tr>
<td>( IC_3 )</td>
</tr>
<tr>
<td>( PC_1 )</td>
</tr>
<tr>
<td>( PC_2 )</td>
</tr>
<tr>
<td>( PC_3 )</td>
</tr>
<tr>
<td>( AIC_3 )</td>
</tr>
<tr>
<td>( BIC_3 )</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Data in levels: estimating the number of non-stationary factors</th>
</tr>
</thead>
<tbody>
<tr>
<td>( IPC_1 )</td>
</tr>
<tr>
<td>( IPC_2 )</td>
</tr>
<tr>
<td>( IPC_3 )</td>
</tr>
</tbody>
</table>

Notes: Prior to computation of the eigenvectors, each first-differenced series is demeaned and standardized to have unit variance (see Bai and Ng, 2002, p. 203).

Bai (2004) has proposed a set of information criteria that are closely related to those of Bai and Ng (2002) but that can be applied to the levels of the series to determine the number of non-stationary factors. The results for the \( IPC_{1,2,3} \) criteria are reported in the bottom panel of Table 4, again with the maximum number of factors \( r_{\text{max}} \) ranging from 2 to 6. The results suggest a single non-stationary common factor in \( r_{\text{it}} \) and \( \hat{\epsilon}_{\text{FE}}^\text{it} \) and at least one non-stationary common factor in \( i_{\text{it}} \) and \( \pi_{\text{it}} \).

To visualize the importance of the common factors, Figure 1 plots the data for \( i_{\text{it}} \), \( \pi_{\text{it}} \) and \( r_{\text{it}} \) together with the first 3 factors estimated using the differencing and recumulating approach outlined in Bai and Ng (2004). Because the true factors can only be identified up to scale, the factors are rotated such that (i) the average of the factor loadings on each factor equals 1 and (ii) the average of each factor coincides with the panel wide average of the plotted data. First, the graph for \( i_{\text{it}} \) in panel (a) of Figure 1 shows that the first two factors are important, while the third is clearly unimportant. Looking more closely at the first two factors shows that they are virtually the same apart from a short period in 1992-1993. This is the period of the EMS crisis during which a lot of European countries sharply raised their short-term interest rates to defend their currencies. To visualize more

---

\(^3\)Note that consistency of the information criteria in Bai (2004) requires the idiosyncratic component to be stationary. Evidence that this is indeed the case is presented in Section 3.3.3 below.
clearly how this is picked up by the common factors, panel (b) plots an alternative representation by combining the first two factors leaving the full effect $\lambda_i'F_t$ unchanged for each country.\textsuperscript{4} The first factor seems to be non-stationary, decreasing from about 11% in the early 1980s to just below 2% in the late 2000s. The second factor now shows up as a stationary component capturing the temporary increase in many European nominal interest rates during the EMS crisis. Second, from the graph for $\pi_{it}$ in panel (c) of Figure 1 it is clear that only the first factor is an important global driver of inflation. It exhibits non-stationary behavior, starting around 8% in the early 1980s to stabilize around 2% in the late 1990s and 2000s. This factor captures the disinflation process all OECD countries went through in the 1980s and the early 1990s and relative stable inflation around 2% from the mid 1990s onwards. The second and the third factor are of no overall importance at all. The graphs for $r_{it}$ and $\hat{\epsilon}_{FE_{it}}$ in panels (d) and (e) of Figure 1 are highly similar. Only the first and to a lesser extent also the second factor seem important.

In line with the results for $i_{it}$ the EMS crisis shows up as a clear spike. However, the EMS crisis is a very specific event, common to only a part of the countries in the sample over a limited period of time. Moreover, it implied higher nominal interest rates mainly for reasons other than inflation expectations. Therefore, instead of trying to capture it using the common factor structure, in the remainder we will control for the EMS crisis using dummy variables\textsuperscript{5} when estimating the Fisher equation. Using the FE estimator including the EMS dummies the point estimate of $\beta$ in (4) is virtually unchanged (estimation results are reported in Table 7 below), but the residuals $\hat{\epsilon}_{FE_{it}}$ and the common factors in panel (f) of Figure 1 are now purged of the EMS crisis. Only the first factor is important now. It seems to be non-stationary, increasing from around 0% in the early 1980s to over 2% in the mid 1980s and then decreasing slowly to around -2% at the end of the sample.

As a final check, Table 5 reports the cross-sectional correlation in the idiosyncratic part of the data, i.e. $\epsilon_{it}$ in (6), after taking out the contribution of $r$ common factors with $r$ ranging from 0 to 3. In line with the picture emerging from Figure 1, one factor seems to be sufficient to remove the cross-sectional dependence from $\pi_{it}$. For $i_{it}$, $r_{it}$ and $\hat{\epsilon}_{FE_{it}}$, at least two factors seem to be necessary. However, after including the EMS dummies one factor seems to be sufficient to remove the cross-sectional dependence from $\hat{\epsilon}_{FE_{it}}$.

The tentative conclusion from Tables 4 and 5 and Figure 1 is that, after correcting for the EMS crisis, the cross-sectional correlation observed in the data and in the residuals of the FE estimator is due to a single non-stationary common factor in each of these series. In the next section, we more formally test the time series properties of the data using unit root tests that allow for cross-sectional dependence induced by unobserved common factors.

\textsuperscript{4}We first set $F_{1t}^* = F_{1t} - \lambda F_{2t}$ and $F_{2t}^* = F_{2t}$ and recalculate the factor loadings as $\lambda_{11}' = \lambda_{11}$ and $\lambda_{22}' = \lambda_{22} + \lambda \lambda_{11}$, with $\lambda$ set equal to 1.3. Next, $F_{1t}^*$ and $F_{2t}^*$ and their factor loadings are again rotated such that the average of the factor loadings on each factor equals 1, the average of the first factor coincides with the panel wide average of the plotted data and the average of the second factor is zero.

\textsuperscript{5}After careful studying the evolution of short-term interest rates during the EMS crisis, country- and time-specific intervention dummies were constructed for the following quarters: Belgium 1993Q3-1993Q4; Denmark 1992Q2-1993Q1 and 1993Q3; Finland 1992Q3; France 1993Q1; Greece 1994Q2; Ireland 1992Q3-1993Q1; Italy 1992Q3-1992Q4; Norway 1992Q3-1992Q4; Sweden 1992Q3-1992Q4.
Figure 1: Time plots of the data and the FE residuals together with the first 3 estimated factors

Notes: Country-specific data: thin solid gray lines
Factor 1: bold solid line, Factor 2: bold dashed line, Factor 3: bold dotted line

Table 5: Cross-sectional correlation $\bar{\rho}$ after taking out $r$ common factors

<table>
<thead>
<tr>
<th>Sample period: 1983:Q1-2010:Q4, 21 countries</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>$r$</th>
<th>Levels</th>
<th>First-differences</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$i_{it}$</td>
<td>$\pi_{it}$</td>
</tr>
<tr>
<td>$r = 0$</td>
<td>0.79</td>
<td>0.54</td>
</tr>
<tr>
<td>$r = 1$</td>
<td>0.38</td>
<td>0.00</td>
</tr>
<tr>
<td>$r = 2$</td>
<td>0.02</td>
<td>-0.01</td>
</tr>
<tr>
<td>$r = 3$</td>
<td>-0.01</td>
<td>-0.01</td>
</tr>
</tbody>
</table>

Note: see Table 3 for definition of $\bar{\rho}$. 
### 3.3.3 Second generation panel unit root tests

Unit root tests allowing for cross-sectional dependence have been proposed by, most notably, Pesaran (2007), Moon and Perron (2004) and Bai and Ng (2004). These tests are similar in that they assume an observed data series to be, in the spirit of the representation in equation (6), the sum of an unobserved idiosyncratic component and a number of unobserved common factors to which each individual can react differently. The tests differ in the allowed number and order of integration of the unobserved common factors and in the way these factors are eliminated. The most general approach is the PANIC of Bai and Ng (2004), which allows for non-stationarity in either the common factors, or in the idiosyncratic errors or in both. Rather than testing the order of integration of the observed data, these are first decomposed in unobserved common factors and idiosyncratic errors which are then tested separately. The key to this is a ‘differencing and recumulating’ procedure that permits consistent estimation of the unobserved components when it is not known a priori whether they are $I(0)$ or $I(1)$.

**Table 6**: Bai and Ng (2004) PANIC unit root tests

<table>
<thead>
<tr>
<th>Sample period: 1983:Q1-2010:Q4, 21 countries</th>
<th>$r=1$</th>
<th>$r=2$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ADF-GLS $\hat{F}_t$</td>
<td>MW $\hat{c}_{it}$</td>
</tr>
<tr>
<td>$i_{it}$</td>
<td>0.32 (0.77)</td>
<td>45.89 (0.31)</td>
</tr>
<tr>
<td>$\pi_{it}$</td>
<td>0.43 (0.81)</td>
<td>66.21 (0.01)</td>
</tr>
<tr>
<td>$\rho_{it}$</td>
<td>-0.84 (0.35)</td>
<td>81.85 (0.00)</td>
</tr>
<tr>
<td>$\hat{\epsilon}_{it}$</td>
<td>-0.99 (0.29)</td>
<td>86.61 (0.00)</td>
</tr>
<tr>
<td>$\hat{\epsilon}^{FE}_{it}$</td>
<td>-0.64 (0.44)</td>
<td>96.63 (0.00)</td>
</tr>
<tr>
<td>$\hat{\epsilon}^{FEd}_{it}$</td>
<td>-0.64 (0.44)</td>
<td>96.63 (0.00)</td>
</tr>
</tbody>
</table>

For $r=1$ the unit root test on the single common factor $\hat{F}_t$ is a ADF-GLS test for a model with constant. The corresponding (simulated) $p$-values are reported in parentheses. For $r>1$, the $MQ_c$ statistic tests the number of independent non-stationary factors ($m$) in the vector $\hat{F}_t$. The critical values at the 1%, 5% and 10% level of significance are -20.151, -13.730 and -11.022 for $m=1$ and -31.621, -23.535 and -19.923 for $m=2$. *** indicates that the $MQ_c$ test is significant at the 1% level, ** at the 5% and * at the 10% level.

The panel unit root test on the estimated idiosyncratic errors $\hat{c}_{it}$ for different number of common factors $r=1, 2$ is the Maddala and Wu (1999) (MW) test defined as $-2\sum_{t=1}^{N}\ln(p_t)$ where $p_t$ is the $p$-value corresponding to the unit root test of the $t$th country. The $p$-value of the MW test, reported in parentheses, is obtained from the $\chi^2$ distribution with $2N$ degrees of freedom.

The results of the Bai and Ng (2004) PANIC unit root tests reported in Table 6 imply that each of the three variables is non-stationary, with this non-stationarity being induced by the common factor(s) leaving the idiosyncratic error terms stationary. First consider $i_{it}$. The analysis in Section 3.3.2 suggests that 2 common factors are necessary to capture the cross-sectional dependence in the data. Setting $r=2$, the idiosyncratic errors $\hat{c}_{it}$ are found to be stationary using the MW test. The $MQ_c$ statistic shows that the space spanned by the two common factors is non-stationary but there is only 1 independent non-stationary common factor. This is consistent with the interpretation above that the second factor captures the EMS crisis. Second, for both $\pi_{it}$ and $\rho_{it}$ the analysis in Section 3.3.2 suggests 1 common factor which is found to be non-stationary using the ADF-GLS test. Setting $r=1$, the idiosyncratic errors $\hat{c}_{it}$ are found to be stationary using the MW test. Finally, also the residuals from the FE regressions (with or without EMS dummies) are found to be non-stationary, with a single non-stationary common factor and
stationary idiosyncratic errors. Urbain and Westerlund (2011) show that the standard result in Phillips and Moon (1999) that panel regressions yield consistent results even if there is no cointegration does not longer hold when the non-stationary in the error term is induced by a common factor. This implies that the results from the FE estimator should be considered spurious.

4 The Fisher equation in the presence of an unobserved $I(1)$ common factor

In this section, we augment the standard Fisher specification (4) by allowing for an $I(1)$ unobserved common component which we interpret as representing permanent fluctuations in the world real interest rate. We discuss how this common factor-augmented specification can be estimated and how to test whether this is a cointegrating relation.

4.1 An $I(1)$ world real interest rate

The main conclusion from the PANIC in Section 3.3.3 is that there is an $I(1)$ common factor in both the real interest rate $r_{it}$ and the residuals $\epsilon_{it}$ of the Fisher equation (4). This has two important implications for modelling the Fisher effect.

First, the finding that $\epsilon_{it}$ is $I(1)$ implies that $i_{it}$ and $\pi_{it}$ are not cointegrated, but does not automatically invalidate the Fisher effect. It does signal, though, that equation (4) is miss-specified, i.e. the assumption that the composite error term $\epsilon_{it} = \nu_{it} - \beta\zeta_{it}$ is stationary is wrong. As non-stationarity of the forecast error $\zeta_{it}$ would be at odds with rational expectations, the observed non-stationarity in $\epsilon_{it}$ is most probably due to $\nu_{it}$ which represents time variation in the real factors driving the ex ante real interest rate. Standard neoclassical growth models with household intertemporal utility maximization imply that the real interest rate is a function of time preference, risk aversion and the steady-state growth rate of technological change. While time preference and risk aversion are generally believed to be fairly stable, or at least changing only slowly over extended periods of time, shifts in steady-state growth, such as the ‘productivity slowdown’ of the early 1970s and the ‘New Economy’ resurgence of growth in the late 1990s, have been widely documented in the literature (see e.g. Oliner and Sichel, 2000; Roberts, 2001). Moreover, in the Diamond overlapping-generations model, a permanent increase in government spending leads to a permanently higher real interest rate. Additional determinants of real interest rates suggested in the literature are demographic changes, changes in the stance of fiscal policy and the evolution of public debt, changes in the taxation of profits, (de)regulation of financial markets, ... (see e.g. Blanchard and Summers, 1984; Chadha and Dimsdale, 1999; Ardagna, 2009). Permanent shifts in any of these factors induce a unit root in the ex ante real interest rate $r_{eit}$ which implies a unit root in the ex post real interest rate $r_{it}$ and in the residuals $\epsilon_{it}$ of the Fisher equation (4). Ideally, the non-stationary determinants of real interest rates should be included as covariates in the Fisher equation. Unfortunately, there is a large variety of possible determinants which are, moreover, not directly observable or at least hard to measure.

Second, the finding that only the common factor in $\epsilon_{it}$ is $I(1)$ while the idiosyncratic part is $I(0)$
suggests that the permanent shifts in the real interest rate are common to all countries in the sample. This is in line with the results in e.g. Gagnon and Unferth (1995), Pain and Thomas (1997) and Lee (2002) who show that country-specific deviations from an $I(1)$ world real interest rate are stationary. Note that Blanchard and Summers (1984) already argued that increasing economic integration leads to a substantial degree of linkage between real interest rates of different countries such that national real interest rates should be related to international rather than to domestic events. The main advantage of this $I(1)$ world real interest rate is that it can be identified by exploiting the strong cross-section correlation observed over countries.

### 4.2 Common factor-augmented Fisher equation

To allow for an $I(1)$ world real interest rate, the DGP of ex ante real interest rates in equation (2) is rewritten to

$$r_{it}^e = \alpha_i + \gamma_i r_{it}^w + \mu_{it}, \quad (7)$$

where $r_{it}^w$ is a single non-stationary common factor with idiosyncratic factor loadings $\gamma_i$ and $\mu_{it}$ a stationary idiosyncratic component. Inserting (7) and (3) in (1) yields

$$i_{it} = \alpha_i + \beta \pi_{it} + \epsilon_{it}, \quad (8)$$
$$\epsilon_{it} = \gamma_i r_{it}^w + \varepsilon_{it}, \quad (9)$$

with $\varepsilon_{it} = \mu_{it} - \beta \zeta_{it}$. The Fisher equation in (8) is the basic specification in (4) augmented with a unobserved non-stationary common factor in the residuals $\epsilon_{it}$ modelled in equation (9).

The model in equations (8)-(9) in vector notation is

$$i_i = \alpha_i + \beta \pi_i + \gamma_i r_{iw}^w + \varepsilon_i, \quad (10)$$

where $i_i = (i_{i1}, \ldots, i_{iT})'$, $\pi_i = (\pi_{i1}, \ldots, \pi_{iT})'$, $r_{iw}^w = (r_{w1}^w, \ldots, r_{wT}^w)'$, $\varepsilon_i = (\varepsilon_{i1}, \ldots, \varepsilon_{iT})'$.

### 4.3 Estimation in the presence of unobserved $I(1)$ common factors

#### 4.3.1 Principal Component Estimators

Bai et al. (2009) suggest a ‘continuously-updated’ (Cup) procedure that jointly estimates the slope coefficient $\beta$ and the unobserved common factor $r_{iw}^w$ in equation (10). More specifically, the solution $(\hat{\beta}_{cup}, \hat{r}_{cup}^w)$ is obtained by iteratively estimating (i) $\hat{\beta}$ as the FE estimator for $\beta$ in equation (10) conditional on $\hat{r}_{cup}^w$

$$\hat{\beta} = \left( \sum_{i=1}^{N} \pi_i' M_{Fw} \pi_i \right)^{-1} \sum_{i=1}^{N} \pi_i' M_{Fw} i_i, \quad (11)$$

where $M_{Fw} = I_T - \hat{r}_{cup}^w \left( \hat{r}_{cup}^w, \hat{r}_{cup}^w \right)^{-1} \hat{r}_{cup}^w$ and (ii) $\hat{r}_{cup}^w$ as the first $r$ eigenvectors (multiplied by $T$) of the matrix
\[ \frac{1}{NT} \sum_{i=1}^{N} \left( i - \hat{\beta} \pi_i \right) \left( i - \hat{\beta} \pi_i \right)' \] conditional on \( \hat{\beta} \). Bai et al. (2009) show that \( \hat{\beta}_{\text{Cup}} \) is \( T \) consistent for \( \beta \) but has an asymptotic bias (for \( N \to \infty \)) arising from endogeneity of \( \pi_{it} \) and serial correlation in \( \varepsilon_{it} \). They therefore suggest a bias-corrected (CupBC) and a fully modified (CupFM) version of the Cup estimator. The first estimates the asymptotic bias directly while the second modifies the data so that the limiting distribution does not depend on nuisance parameters. Both are \( \sqrt{NT} \) consistent for the common slope coefficient \( \beta \) and are robust to mixed \( I(1)/I(0) \) factors and regressors. Moreover, the estimators enable the use of standard test statistics for inference. This approach requires specifying the number of common factors \( r \).

### 4.3.2 CCEP estimators

Pesaran (2006) proposes to eliminate the cross-sectional dependence in \( \varepsilon_{it} \) by projecting out the common factor \( r_w \) using the cross-sectional averages of \( i_{it} \) and \( \pi_{it} \). For a model with a single factor, inserting (9) in (8) and taking cross-sectional averages yields

\[ \tilde{i}_t = \bar{i} + \beta \bar{\pi}_t + \gamma r_w^t + \tilde{\varepsilon}_t, \quad (12) \]

where \( \tilde{i}_t = N^{-1} \sum_{i=1}^{N} i_{it} \) and similarly for the other variables. Solving (12) for \( r_w^t \)

\[ r_w^t = \frac{1}{\gamma} \left( \tilde{i}_t - \bar{i} - \beta \bar{\pi}_t - \tilde{\varepsilon}_t \right), \quad (13) \]

and inserting (13) in (8)-(9) yields

\[ i_{it} = \alpha_i + \beta \pi_{it} + \frac{\gamma_i}{\gamma} \left( \tilde{i}_t - \bar{i} - \beta \bar{\pi}_t - \tilde{\varepsilon}_t \right) + \varepsilon_{it}, \]

\[ = \tilde{\alpha}_i + \beta \pi_{it} + c_{1i} \tilde{i}_t + c_{2i} \bar{\pi}_t + \tilde{\varepsilon}_{it}, \quad (14) \]

with \( \tilde{\alpha}_i = \alpha_i - \left( \gamma_i / \gamma \right) \bar{\pi}, c_{1i} = \left( \gamma_i / \gamma \right), c_{2i} = -\beta \left( \gamma_i / \gamma \right) \) and \( \tilde{\varepsilon}_{it} = \varepsilon_{it} - \left( \gamma_i / \gamma \right) \tilde{\varepsilon}_t \).

The CCEP estimator proposed by Pesaran (2006) is the FE estimator applied to the augmented regression in (14), ignoring the non-linear coefficient restrictions, given by

\[ \hat{\beta}_{\text{CCEP}} = \left( \sum_{i=1}^{N} \pi_i' M_H \pi_i \right)^{-1} \sum_{i=1}^{N} \pi_i' M_H \tilde{i}_i, \quad (15) \]

where \( M_H = I_T - H (H'H)^{-1} H' \) with \( H = (\tilde{i}, \bar{\pi}), \tilde{i} = (\tilde{i}_1, \ldots, \tilde{i}_T)' \) and \( \pi = (\pi_1, \ldots, \pi_T)' \).

As the assumption that \( \varepsilon_{it} \) is cross-sectionally independent implies that \( \lim_{N \to \infty} \tau_t = 0 \), the error made when approximating \( r_w^t \) by \( \tilde{i}_t \) and \( \pi_i \) in (13) becomes negligibly small for \( N \to \infty \) such that \( \tilde{\varepsilon}_{it} \overset{p}{\to} \varepsilon_{it} \) in (14). This is the basic result in Pesaran (2006) that the inclusion of cross-sectional averages asymptotically eliminates the error cross-sectional dependence induced by the unobserved common factors such that the CCEP estimator is \( \sqrt{N} \) consistent regardless of whether \( T \) is fixed or \( T \to \infty \). These results hold for

\[ \text{Multiple factors can be treated in the same way (see Phillips and Sul, 2007), and yield the same (unrestricted) model as the one presented in (14), but are not presented here for notational convenience.} \]
any fixed number of unobserved factors \( r \), which implies that there is no need to estimate or specify \( r \). Kapetanios et al. (2011) further shows that these results continue to hold regardless of whether the common factors are stationary or non-stationary.

An important restriction is that consistency of the CCEP estimator requires that the idiosyncratic errors \( \varepsilon_{it} \) are distributed independently of the explanatory variable \( \pi_{it} \). To see why, note that the CCEP estimator in equation (15) is equivalent to the least squares estimator for \( \beta \) after projecting out the individual effects and the cross-sectional means from the model in equation (14)

\[
\tilde{i}_{it} = \beta \tilde{\pi}_{it} + \tilde{\varepsilon}_{it},
\]

where \( \tilde{i}_{it} = (\tilde{i}_{i1}, \ldots, \tilde{i}_{iT})' = M_Hi_i \) and \( \tilde{\pi}_{i} = (\tilde{\pi}_{i1}, \ldots, \tilde{\pi}_{iT})' = M_H\pi_i \) are the residuals from country-by-country regressions of \( i_{it} \) and \( \pi_{it} \) on a constant, \( i_t \) and \( \pi_t \). Pesaran (2006) and Kapetanios et al. (2011) show that this orthogonalisation on the cross-sectional averages makes \( \tilde{i}_{it} \) and \( \tilde{\pi}_{it} \) estimates of the idiosyncratic part in \( i_{it} \) and \( \pi_{it} \) respectively. As these idiosyncratic parts are found to be stationary by the PANIC in Section 3.3.3, which is also the working assumption in Kapetanios et al. (2011), equation (16) is a regression model including stationary variables. This implies that, in contrast to the Cup estimator, the CCEP estimator is not super consistent such that endogeneity cannot be ignored asymptotically.

As the forecast error \( \zeta_{it} \) implies that \( \pi_{it} \) and \( \varepsilon_{it} \) are correlated, equation (16) is estimated using GMM. Valid moment conditions are

\[
E(\tilde{\pi}_{i,t-l}\tilde{\varepsilon}_{it}) = 0 \quad \text{for each} \quad t = l + 1, \ldots, T \quad \text{and} \quad l \geq q + 1, \tag{17}
\]

with \( q \) being the order of the MA process in \( \zeta_{it} \). Equation (17) defines a relatively large set of moment conditions. Using more instruments from deeper lags of \( \tilde{\pi}_{it} \) improves the efficiency of the GMM estimator. However, it also reduces the sample size as observations for which lagged observations are unavailable are dropped. To avoid this trade-off between instrument lag depth and sample depth, we construct instruments by zeroing out missing observations of lags as in Holtz-Eakin et al. (1988). Furthermore, in order to avoid problems related to using too many instruments, we truncate the set of available instruments at the first \( L \) available lags. This results in the following reduced set of moment conditions

\[
E(\tilde{\pi}_{i,t-l}\tilde{\varepsilon}_{it}) = 0 \quad \text{for each} \quad q + 1 \leq l \leq L + q, \tag{18}
\]

The CCEP-GMM estimator for \( \beta \) is obtained by minimizing the empirical moments \( \sum_{i} \sum_{t} \tilde{\pi}_{i,t-l}\tilde{\varepsilon}_{it} \) using a Newey-West type optimal weighting matrix.

### 4.4 Common factor-augmented panel cointegration

Cointegration in panels with unobserved non-stationary common factors has been considered by Gegenbach et al. (2006) and Banerjee and Carrion-i Silvestre (2006) and applied by e.g. Costantini and Destefanis (2009) and Auteri and Costantini (2010). Both studies would define panel cointegration in our
case as a situation where the interest rate \( i_{it} \) and the inflation rate \( \pi_{it} \) cointegrate with vector \((1, -\beta)\). Equations (8)-(9) show that this concept of panel cointegration requires both the common factor \( r^w_t \) and the idiosyncratic error term \( \varepsilon_{it} \) to be \( I(0) \). Especially the former is highly restrictive as it requires that any non-stationary common factors in \( i_{it} \) and \( \pi_{it} \) should cointegrate leaving the common factor in the error term \( \varepsilon_{it} \) stationary. However, when our interest is in estimation and inference on \( \beta \), the estimation procedures of Kapetanios et al. (2011) and Bai et al. (2009) outlined above only require \( \varepsilon_{it} = (i_{it} - \beta \pi_{it} - \gamma r^w_t) \) to be \( I(0) \). Intuitively, \( r^w_t \) is a vector of \( I(1) \) variables that should be included in the model for this to be cointegrating regression. We label this common factor-augmented panel cointegration.

The most obvious approach to test whether \( i_{it}, \pi_{it} \) and \( r^w_t \) are cointegrated or not would be to first estimate the model in (8)-(9), using either the CCEP or the Cup estimation approach, and then test for the null hypothesis of no cointegration using e.g. a MW panel cointegration test on the estimated idiosyncratic error terms \( \hat{\varepsilon}_{it} \). This direct approach is problematic for two reasons, though. First, the country-specific orthogonalisation, either on \( \hat{r}^w_t \) in (11) or on the cross-sectional averages \( \bar{i}_t \) and \( \bar{\pi}_t \) in (15), implies that the distribution of a country-by-country cointegration test on \( \hat{\varepsilon}_{it} \) depends on the number of \( I(1) \) factors in \( r^w_t \). This is problematic as the CCEP estimator does not require specifying the number of factors while the Cup estimator does only require a decision on the number of factors but not on the number of \( I(1) \) factors. Second, the fact that the orthogonalisation is on the same variable(s) in each country implies that the country-by-country cointegration tests are not independent and therefore the MW panel cointegration test does not have the standard \( \chi^2 \) distribution.

A natural alternative approach is to use \( \hat{\varepsilon}_{it} = \left( i_{it} - \hat{\beta} \pi_{it} \right) \) instead of \( \hat{\varepsilon}_{it} \) and apply a principal component analysis as in Bai and Ng (2004) to split \( \hat{\varepsilon}_{it} \) in a number of common factors and an idiosyncratic error term and then test whether the idiosyncratic error is stationary or not. The advantage of this approach is that, as shown by Bai and Ng (2004), the test whether the idiosyncratic errors are stationary does not depend on the presence or absence of common stochastic trends and/or their integration properties and thus can be tested using standard panel unit root tests. It only requires specifying the number of common factors.

5 Empirical results common factor-augmented Fisher equation

In this section we present the estimation results of the Fisher equation augmented with an unobserved common factor to capture shifts in the real interest rate. We also conduct a small-scaled Monte Carlo experiment to assess the finite sample properties in terms of estimation and inference of the CCEP and Cup estimators outlined in Section 4.3 and the size and power of the cointegration tests on \( \hat{\varepsilon}_{it} \) and \( \hat{\varepsilon}_{it} \) outlined in Section 4.4.

5.1 Estimation results

The estimation results for the common factor-augmented Fisher equation are reported in Table 7. All estimators are obtained by including the EMS dummies as outlined in Section 3.3.2. Consistent with the results reported in Section 3.3.3, the PANIC shows that there is an \( I(1) \) common factor and an \( I(0) \)
idiosyncratic component in the estimated composite residuals $\hat{\epsilon}_{it}$ of the FE, CCEP(.GMM) and Cup(BC) regressions. This implies that there is cointegration between $i_{it}$, $\pi_{it}$ and an unobserved common factor. Note that the non-stationarity of the composite error term $\hat{\epsilon}_{it}$ is not detected by the MW panel unit root test on the FE composite residuals but is picked up when using CCEP(.GMM) and Cup(BC) composite residuals.

### Table 7: Estimation results

<table>
<thead>
<tr>
<th></th>
<th>Fisher regression</th>
<th>MW</th>
<th>PANIC $\hat{\epsilon}_{it}$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\hat{\beta}$</td>
<td>se</td>
<td>$\tau (\beta = 1)$</td>
</tr>
<tr>
<td>FE</td>
<td>1.10</td>
<td>0.04</td>
<td>2.32</td>
</tr>
<tr>
<td>CCEP</td>
<td>0.60</td>
<td>0.04</td>
<td>-10.91</td>
</tr>
<tr>
<td>CCEP,.GMM</td>
<td>1.00</td>
<td>0.10</td>
<td>0.03</td>
</tr>
<tr>
<td>Cup</td>
<td>0.58</td>
<td>0.02</td>
<td>-24.63</td>
</tr>
<tr>
<td>CupBC</td>
<td>0.83</td>
<td>0.04</td>
<td>-4.52</td>
</tr>
</tbody>
</table>

Notes: All estimators are obtained by including the EMS dummies as outlined in Section 3.3.2. The Cup(BC) estimators are obtained setting the number of common factors $r = 1$. The CupBC estimator is calculated from a long-run covariance matrix estimated using the Bartlett kernel with bandwidth set to 5. The CCEP,.GMM estimator is obtained by setting $q = 3$ and $L = 8$. Reported are two-step GMM results with optimal weighting matrix constructed from a Newey-West type of estimator with lag truncation set to 3.

The $t$-statistic $\tau (\beta = 1)$ is calculated under the null hypothesis that $\beta = 1$ with $p$-value $p(\tau)$ calculated from a standard normal distribution. To obtain the bootstrapped $p$-value $p^b(\tau)$, we generate 5000 bootstrap samples by resampling cross-sections as in Kapetanios (2008), compute a bootstrap $t$-statistic $t^b(\beta = \hat{\beta})$ using each of them, and calculate $p^b(\tau)$ as the proportion of the $|t^b(\beta = \hat{\beta})|$ that are larger than $|\tau (\beta = 1)|$.

For the MW unit root tests on $\hat{\epsilon}_{it}$, $\hat{\epsilon}_{it}$ and $\hat{\epsilon}_{it}$, we report $p$-values. See notes to Table 6 for technical details on the implementation of the MW test. The PANIC on $\hat{\epsilon}_{it}$ is conducted setting $r = 1$. For $\hat{F}_t$ we report the $p$-value of an ADF-GLS test for a model with constant.

Looking at the coefficient estimates from the various estimators, these range from low values of 0.60 and 0.58 for the CCEP and Cup estimators, over 0.83 for the CupBC estimator to 1.00 for the CCEP,.GMM estimator. From the asymptotic properties of the estimators outlined in Section 4.3, CCEP,.GMM and CupBC are the preferred estimators. However, despite being consistent and asymptotically normally distributed, these estimators and especially $t$-statistic based on them are biased in small samples (see Monte Carlo simulation below). The main reason for the latter is that autocorrelation in the error terms invalidates the standard asymptotic distribution. Therefore, we calculate bootstrap $p$-values. The bootstrap data generating process (also see notes to Table 7 for the exact implementation) involves resampling whole cross-sectional units with replacement as suggested by Kapetanios (2008). The advantage of this resampling scheme is that it preserves (i) the autocorrelation structure in the data and the errors, (ii) the endogeneity of $\pi_{it}$ and (iii) the cross-sectional dependence. The bootstrap inference shows that the hypothesis that $\beta = 1$ is not rejected for both the CCEP,.GMM estimator and the CupBC estimator.

The overall conclusion is that after taking into account a non-stationary common factor, the full

---

7 This result is robust over alternative choices of $L$.

8 Note that the cross-sectional resampling scheme is not valid in the case of local cross-sectional dependence but is appropriate in the presence of the assumed factor structure which introduces global cross-sectional dependence which is symmetric across all panel units.
Fisher hypothesis is not rejected by the data. To shed some more light on the small sample behavior of the estimators and PANIC cointegration test, we conduct a Monte-Carlo simulation in the next section.

5.2 Monte Carlo simulation

To make sure that our simulation results are relevant for putting the estimation results in perspective, data are simulated for exactly the same sample size ($T=112$, $N=21$) that is available to us while the data-generating process and population parameters, under the hypothesis that the full Fisher effect holds ($\beta = 1$), are chosen such that the properties of the simulated data match with those of the observed data for $i_{it}$ and $\pi_{it}$ as much as possible.

Design

In line with the model in Section 4.2 and the results of the PANIC in Section 3.3.3, data are generated based on the following design

\[
\begin{align*}
    i_{it} &= r^c_{it} + \beta \pi^c_{it}, \\
    \pi_{it} &= \pi^c_{it} + \zeta_{it}, \\
    r^c_{it} &= \alpha_i + \gamma_i r^w_{it} + \mu_{it}, \quad \alpha_i \sim i.i.d. \mathcal{N}(\alpha, \sigma^2_\alpha), \quad \gamma_i \sim i.i.d. \mathcal{N}(1, \sigma^2_\gamma), \\
    \pi^c_{it} &= \tau_i + \lambda_i \pi^w_{it} + \eta_{it}, \quad \tau_i \sim i.i.d. \mathcal{N}(\tau, \sigma^2_\tau), \quad \lambda_i \sim i.i.d. \mathcal{N}(1, \sigma^2_\lambda), \\
\end{align*}
\]

with the common factors $r^w_{it}$ and $\pi^w_{it}$ being generated as random walks

\[
\begin{align*}
    r^w_{it} &= r^w_{t-1} + \psi_{it}, \quad \psi_{it} \sim i.i.d. \mathcal{N}(0, \sigma^2_\psi), \\
    \pi^w_{it} &= \pi^w_{t-1} + \xi_{it}, \quad \xi_{it} \sim i.i.d. \mathcal{N}(0, \sigma^2_\xi), \\
\end{align*}
\]

while the idiosyncratic components $\mu_{it}$ and $\eta_{it}$ are generated as AR(1) processes

\[
\begin{align*}
    \mu_{it} &= \theta \mu_{i,t-1} + \chi_{it}, \quad \chi_{it} \sim i.i.d. \mathcal{N}(0, \sigma^2_\chi) \\
    \eta_{it} &= \phi \eta_{i,t-1} + \omega_{it}, \quad \omega_{it} \sim i.i.d. \mathcal{N}(0, \sigma^2_\omega). \\
\end{align*}
\]

In order to obtain realistic parameter values, we calibrate the DGP outlined above to our observed sample of OECD data. As $\pi^c_{it}$ and $r^c_{it}$ are not observed we start by making the strong assumption of perfect foresight, i.e. $\zeta_{it} = 0$, such that $\pi^c_{it}$ equals ex post observed inflation $\pi_{it}$ and $r^c_{it}$ equals the ex post observed real interest rate $r_{it}$. The observed data for both $\pi_{it}$ and $r_{it}$ are then split up into a fixed effect, a common component and an idiosyncratic component using the PANIC of Bai and Ng (2004)\(^9\). Parameter values are estimated from the various estimated components. This is the experiment 1:

- Experiment 1: $\sigma_\zeta = 0$, $\beta = 1$, $\alpha = 3.03$, $\sigma_\alpha = 1.05$, $\tau = 3.44$, $\sigma_\tau = 1.93$, $\sigma_\gamma = 1.09$, $\sigma_\lambda = 0.36$, $\sigma_\psi = 0.41$, $\sigma_\xi = 0.37$, $\phi = 0.77$, $\sigma_\omega = 1.21$, $\theta = 0.67$ and $\sigma_\chi = 1.54$.

\(^9\gamma_i\) and $\lambda_i$ are normalized to have mean 1 and $r^w_{it}$ and $\pi^w_{it}$ to have mean 0
Next, we consider non-zero forecasting errors. From rational expectations, the forecast error $\zeta_{it}$ should be white noise. However, since in our dataset we measure inflation as the year-on-year percent change in the consumer price index, the white noise forecast error builds into an MA(3) process. Therefore, $\zeta_{it}$ is assumed to be generated as

$$
\zeta_{it} = \frac{\sigma_\zeta}{\sqrt{4}} \sum_{j=0}^{3} e_{i,t-j},
$$

where $e_{it} \sim i.i.d. N(0,1)$. The unconditional standard deviation $\sigma_\zeta$ of $\zeta_{it}$ is set to 1.25 which implies that 95% of the quarterly forecasting errors lies between $-2.5$ and $+2.5$ %points. Note that simply adding $\zeta_{it}$ to $\pi_{it}$ from experiment 1 would increase the variance of the simulated $\pi_{it}$. In order to ensure comparability of the simulation results over the experiments, $\sigma^2_\omega$ is therefore lowered such that the variance of the idiosyncratic component in $\pi_{it}$, i.e. $\eta_{it} + \zeta_{it}$, is constant when varying $\sigma^2_\zeta$. As a result, also $\sigma^2_\chi$ is adjusted to ensure that the variance of the idiosyncratic component in $i_{it}$, i.e. $\mu_{it} + \beta \eta_{it}$, remains constant over the experiments. Parameter values that differ compared to experiment 1 are given by

- Experiment 2: $\sigma_\zeta = 1.25$, $\sigma_\omega = 0.91$, $\sigma_\chi = 1.80$.

For each experiment we compute the FE, CCEP, CCEP-GMM, Cup and CupBC estimator. The CCEP-GMM estimator uses the first $L = 8$ available lags with $q$ being adjusted according to the MA structure in $\zeta_{it}$. Reported are two-step GMM results with optimal weighting matrix constructed from a Newey-West type of estimator with lag truncation set to 3. The CupBC estimator is calculated from a long-run covariance matrix estimated using the Bartlett kernel with bandwidth set to 5. For each estimator we report the mean bias ($\text{bias}$) of $\hat{\beta}$, the standard deviation ($\text{sd}$) of the Monte Carlo distribution of $\hat{\beta}$, the root mean squared error ($\text{rmse}$), the mean of the estimated standard error ($\text{se}$) and the size ($\text{size}$) of a $t$-test for the null hypothesis that $\beta = 1$. The standard errors ($\text{se}$) are robust to heteroscedasticity and serial correlation in the error terms. As the inference only has asymptotic validity, we also report the size from using bootstrapped $p$-values (see Section 5.1 and Table 7 for details on the implementation). Each experiment was replicated 5000 times with bootstrapped $p$-values calculated from 1000 bootstrap replications.

Note that in all of the above experiments the setting $\theta = 0.67$ implies that there is cointegration between $i_{it}$, $\pi_{it}$ and $r^w_t$. As next to the power we also want to analyze the size of the cointegration tests discussed in Section 4.4, we will first simulate data for an experiment where $\mu_{it}$ is non-stationary such that there is no cointegration between $i_{it}$, $\pi_{it}$ and $r^w_t$. Experiment 0 therefore differs from experiment 1 in the following parameter value:

- Experiment 0: $\theta = 1.00$.

We preform 3 different cointegration tests: (i) a naive MW unit root test on $\hat{\epsilon}_{it}$ which is non-stationary in all experiments, (ii) a MW cointegration test on the defactored residuals $\hat{\epsilon}_{it}$ (using either the CCEP or the Cup approach) and (iii) a PANIC which first decomposing $\hat{\epsilon}_{it}$ in a single common factor $\hat{F}_t$ and an idiosyncratic component $\hat{\epsilon}_{it}$ and next performing an ADF-GLS unit root test on $\hat{F}_t$ and a MW unit root
test on $\hat{e}_{it}$. For each of these tests, p-values are calculated from simulated finite-sample distributions (for details, see the notes to Tables 1 and 6).

**Simulation results**

First look at the cointegration tests in Table 8. In Experiment 0 $\epsilon_{it}$ and $\hat{e}_{it}$ are non-stationary as the common factor $r_t^w$ and the idiosyncratic error $\mu_{it}$ are both $I(1)$. However, the MW test on the composite error term $\hat{e}_{it}$ and on the defactored error term $\tilde{e}_{it}$ are strongly oversized. This implies that these tests should not be trusted as the null of no cointegration is wrongly rejected in far too many cases. In contrast, the PANIC has the correct size both for a unit root test on $\hat{F}_t$ and on $\hat{e}_{it}$. In Experiments 1-3 there is cointegration between $i_{it}, \pi_{it}$ and $r_t^w$. The MW test on the composite error term $\hat{e}_{it}$ rejects the null of no cointegration between $i_{it}$ and $\pi_{it}$ in almost all cases though. This shows that the $I(1)$ common factor $r_t^w$ is not detected by a standard panel unit root test ignoring the factor structure. The PANIC has good size for the unit root test on the non-stationary factor $\hat{F}_t$ while having power close to 1 for the unit root test on the stationary idiosyncratic errors $\hat{e}_{it}$ in all cases. This shows that a PANIC on the composite error term $\hat{e}_{it}$ is an appropriate approach to test for common factor-augmented panel cointegration.

Looking at the estimation results, first note that the FE estimator is spurious in all experiments which

---

**Table 8:** Monte Carlo simulation results: estimation and inference for $N = 21$ and $T = 112$

<table>
<thead>
<tr>
<th>Estimation precision</th>
<th>Inference $H_0: \beta = 1$</th>
<th>Rejection frequency cointegration tests</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>bias</td>
<td>sd</td>
</tr>
<tr>
<td>FE</td>
<td>-0.00</td>
<td>0.36</td>
</tr>
<tr>
<td>CCEP</td>
<td>-0.00</td>
<td>0.18</td>
</tr>
<tr>
<td>CCEP,GMM</td>
<td>-0.00</td>
<td>0.19</td>
</tr>
<tr>
<td>Cup</td>
<td>-0.00</td>
<td>0.19</td>
</tr>
<tr>
<td>CupBC</td>
<td>-0.01</td>
<td>1.70</td>
</tr>
</tbody>
</table>

**Experiment 0:** $\theta = 1, \zeta_{it} = 0.00$

<table>
<thead>
<tr>
<th></th>
<th>bias</th>
<th>sd</th>
<th>rmse</th>
<th>se</th>
<th>size</th>
<th>size $^b$</th>
<th>MW</th>
<th>PANIC $\hat{e}_{it}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>FE</td>
<td>0.00</td>
<td>0.20</td>
<td>0.20</td>
<td>0.02</td>
<td>0.82</td>
<td>0.60</td>
<td>0.94</td>
<td>-</td>
</tr>
<tr>
<td>CCEP</td>
<td>0.00</td>
<td>0.08</td>
<td>0.08</td>
<td>0.04</td>
<td>0.38</td>
<td>0.06</td>
<td>0.91</td>
<td>1.00</td>
</tr>
<tr>
<td>CCEP,GMM</td>
<td>0.00</td>
<td>0.09</td>
<td>0.09</td>
<td>0.05</td>
<td>0.29</td>
<td>0.07</td>
<td>0.90</td>
<td>1.00</td>
</tr>
<tr>
<td>Cup</td>
<td>-0.00</td>
<td>0.05</td>
<td>0.05</td>
<td>0.01</td>
<td>0.59</td>
<td>0.09</td>
<td>0.91</td>
<td>1.00</td>
</tr>
<tr>
<td>CupBC</td>
<td>-0.00</td>
<td>0.07</td>
<td>0.07</td>
<td>0.04</td>
<td>0.30</td>
<td>0.06</td>
<td>0.91</td>
<td>1.00</td>
</tr>
</tbody>
</table>

**Experiment 1:** $\theta = 0.67, \zeta_{it} \text{ is MA(3) with } \sigma_{\zeta} = 1.25$

<table>
<thead>
<tr>
<th></th>
<th>bias</th>
<th>sd</th>
<th>rmse</th>
<th>se</th>
<th>size</th>
<th>size $^b$</th>
<th>MW</th>
<th>PANIC $\hat{e}_{it}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>FE</td>
<td>-0.08</td>
<td>0.19</td>
<td>0.21</td>
<td>0.02</td>
<td>0.87</td>
<td>0.68</td>
<td>0.98</td>
<td>-</td>
</tr>
<tr>
<td>CCEP</td>
<td>-0.61</td>
<td>0.07</td>
<td>0.62</td>
<td>0.04</td>
<td>1.00</td>
<td>1.00</td>
<td>0.67</td>
<td>1.00</td>
</tr>
<tr>
<td>CCEP,GMM</td>
<td>0.16</td>
<td>0.42</td>
<td>0.45</td>
<td>0.42</td>
<td>0.02</td>
<td>0.05</td>
<td>0.89</td>
<td>1.00</td>
</tr>
<tr>
<td>Cup</td>
<td>-0.20</td>
<td>0.15</td>
<td>0.25</td>
<td>0.02</td>
<td>0.98</td>
<td>0.98</td>
<td>0.91</td>
<td>1.00</td>
</tr>
<tr>
<td>CupBC</td>
<td>-0.08</td>
<td>0.14</td>
<td>0.16</td>
<td>0.04</td>
<td>0.44</td>
<td>0.13</td>
<td>0.94</td>
<td>1.00</td>
</tr>
</tbody>
</table>

Notes: Results based on 5000 Monte Carlo replications. ‘Bias’ is the mean bias, ‘sd’ is the standard deviation of the Monte Carlo distribution of $\beta$ and ‘rmse’ is its root mean squared error. The standard error is the mean of the appropriate analytical estimate ‘se’ for the standard deviation of $\beta$. The reported sizes ‘size’ and ‘size $^b$’ are computed at the 5% nominal level for a double-sided $t$-test for the null hypothesis that $\beta = 1$ using ‘se’ and p-values calculated from the standard normal and bootstrap distribution (see Table 7 for details) respectively.
results in an unacceptably high size using either the analytic or the bootstrap inference. This is in line with Urbain and Westerlund (2011) who show that neglecting I(1) common factors in the residuals of a panel regression implies spurious results. Second, the CCEP(GMM) and Cup(BC) estimators yield unbiased estimates for $\beta$ in experiments 0 and 1. Although the estimated standard errors underestimate the true standard deviation of $\hat{\beta}$ resulting in oversized inference, the bootstrap inference is more or less correctly sized. Especially for experiment 0 these are remarkable results as non-stationarity of the idiosyncratic error $\mu_{it}$ implies that there is no cointegration between $i_{it}$, $\pi_{it}$ and $r_{it}^w$. So taking into account the I(1) common factor seems to re-establish the result in Phillips and Moon (1999) that in a panel consistent estimation and valid inference is possible regardless of whether there is cointegration or not (as long the non-stationary of the error terms is not induced by a common factor). Third, introducing endogeneity in experiment 2 results in a downward bias for both the CCEP and the Cup estimator. Especially for the CCEP estimator this bias is very strong. This is in line with the argument in Section 4.3.2 that the CCEP estimator is inconsistent in this case. The bias of the Cup estimator is smaller, although also sizable. The CCEP_GMM and CupBC estimators significantly improve on the performance of the CCEP and Cup estimators. Using the bootstrap inference, the size is acceptable for both estimators.

Overall, the simulation results show that the CCEP_GMM and CupBC perform reasonably well for the modest sample size $T = 112$, $N = 21$ that is available for our empirical analysis, with the former still being somewhat upward biased and the latter being somewhat downward biased. Note that this variation is quantitatively very much in line with the estimation results in Section 5.1, which further supports the validity of the Fisher hypothesis.

6 Conclusion

The Fisher effect states that inflation expectations should be reflected in nominal interest rates in a one-for-one manner to compensate for changes in the purchasing power of money. Despite its wide acceptance in theory, much of the empirical work fails to find favorable evidence. This paper examines the Fisher effect in a panel of quarterly data for 21 OECD countries over the period 1983-2010. Using a FE regression of nominal interest rates and inflation we find a slope coefficient which is not significantly different from 1 while a MW panel cointegration test finds the error terms to be stationary. These results support the full Fisher hypothesis. However, a non-stationary common factor in the error terms of this alleged cointegrating relation is detected using PANIC. This implies that the FE regression results are spurious. Our simulation results confirm that a non-stationary common factor in the error terms of the Fisher equation leads to a substantial size bias for the standard MW panel test ignoring cross-sectional dependence and to deceptive inference for the FE estimator. A possible interpretation for the non-stationary common factor is that it reflects permanent common shifts in the real interest rate induced by e.g. shifts in time preferences, risk aversion and the steady-state growth rate of technological change. We next control for an unobserved non-stationary common factor in estimating the Fisher equation using both the CCEP and the Cup estimation approach. Endogeneity of observed inflation induced by a rational expectations forecasting error is taken into account using a bias-corrected version of the Cup estimator.
and a GMM version of the CCEP estimator. A small-scale Monte Carlo simulation shows that these two estimators perform reasonably well for the modest sample size $T = 112, N = 21$ that is available for our empirical analysis. From the estimation results, the hypothesis of a one-for-one relation between the nominal interest rate and inflation cannot be rejected using either the CupBC or the CCEP\_GMM estimator.
References


Bai, J., Ng, S., 2002. Determining the number of factors in approximate factor models. Econometrica 70 (1), 191–221.


Contributions to Macroeconomics 1 (1), Article 3.