Do TFP and the Relative Price of Investment Share a Common I(1) Component?*

Luca Benati
University of Bern†

Abstract

Although results from cointegration tests show that TFP and the relative price of investment are most likely not cointegrated, either cross-spectral methods, or a test in the spirit of Cochrane and Sbordone (1988), suggest that they share a common I(1) component, which induces positive long-horizon covariation between the two series.

I explore three alternative possible explanations for this finding: first, that it is genuine; second, that it results from non-linearity of the technology transforming consumption goods into investment goods, which implies than the relative price of investment is also impacted upon by neutral shocks; third, that it is the figment of not controlling for idiosyncratic breaks in the drifts of the two series. I argue that the third explanation is the most likely, so that the two series are in fact independent at the very low frequencies, and they correctly measure neutral and investment-specific technology, respectively. However, I illustrate the extreme sensitivity of these results to the exact nature and timing of the breaks characterizing the two series: conditional on a common break identified via the Bai, Lumsdaine, and Stock (1998) procedure, evidence is compatible with the notion that the two series contain a common I(1) component inducing negative-long horizon covariation between them.

These results illustrate the difficulty of making strong statements about the exact nature of the long-horizon relationship between TFP and the relative price of investment, which automatically injects an element of fragility into investigations about the role played by neutral and investment-specific shocks in macroeconomic fluctuations. For illustrative purposes, I explore via structural VAR methods the role played by the common I(1) component conditional on the common break, and I find that it is not negligible.

Keywords: Technology shocks; unit roots; cointegration; structural VARs.

---

*I wish to thank Franck Portier for kindly providing data on the relative price of investment, and Stephanie Schmitt-Grohé for comments. Usual disclaimers apply.
†Department of Economics, University of Bern, Schanzeneckstrasse 1, CH-3001 Bern, Switzerland. Email: luca.benati@vwi.unibe.ch
1 Introduction

Since Fisher (2006), several researchers have attempted to assess the relative importance of neutral and investment-specific technology shocks in macroeconomic fluctuations. The standard assumption made in this literature is that total factor productivity (henceforth, TFP) measures neutral technology, whereas the relative price of investment (henceforth, RPI) measures investment-specific technology, and that the two series evolve independently (which within DSGE models is indeed the case if the technology transforming consumption goods into investment goods is linear).

As recently pointed out by Schmitt-Grohé and Uribe (2011, p. 123; henceforth, SGU), however,

‘[...] all existing studies assume that neutral and investment-specific productivity shocks follow independent stochastic processes. This assumption, however, is not based on empirical evidence, but appears to be made in a purely ad-hoc fashion.’

Based on Johansen’s trace test, SGU (2011) detect cointegration between the two series, and proceed to estimate an RBC model featuring a common stochastic trend between TFP and the RPI. Their results suggest that shocks to the common trend play a key role in business-cycle fluctuations, explaining about 75 per cent of the variance of output and investment growth, and about one third of the variance of hours.

The importance of SGU’s (2011) analysis is that it questions the meaningfulness of the traditional distinction between the two types of technology, and it suggests that previous analyses focusing on disturbances specific to either TFP or the RPI may have produced a distorted picture of the role played by technology shocks in macroeconomic fluctuations.

In this paper I make three contributions.

First, I reconsider the time-series bivariate relationship between TFP and the RPI based on either SGU’s (2011) original series (which is essentially an updated version of the RPI series used by Fisher (2006)), or the alternative series constructed by Justiniano, Primiceri, and Tambalotti (2011, henceforth JPT) and Liu, Waggoner, and Zha (2011, henceforth LWZ). I show that, in contrast to SGU’s results, evidence from Johansen’s trace test clearly suggests that TFP and the RPI are most likely not cointegrated. I also trace the origin of SGU’s finding of cointegration between the two series, showing that it stems from their use of (i) an inconsistent criterion for lag order selection for VARs containing integrated variables in the Johansen procedure, and (ii) a sub-optimal TFP series. Eschewing either (i) or (ii)—that is: either selecting the VAR lag order for the Johansen procedure based on a consistent lag selection criterion such as Schwartz or Hannan-Quinn, or using John Fernald’s corrected TFP series, which is widely regarded as the best available measure of neutral technological progress—‘kills off’ the finding of cointegration between TFP and SGU’s RPI series.
Second, I show that either cross-spectral methods, or a simple test in the spirit of Cochrane and Sbordone (1988), suggest that the two series, although not cointegrated, share nonetheless a common I(1) component, which induces positive long-horizon covariation between them. I explore three alternative possible explanations for this finding: first, that it is genuine; second, that it results from non-linearity of the technology transforming consumption goods into investment goods, which implies that the relative price of investment is also impacted upon by neutral shocks; third, that it is the figment of not controlling for breaks in the drifts of the two series. I argue that the first two explanations can be ruled out based on a combination of econometric evidence (specifically, results from break tests) and simple logic. As for the third explanation, once controlling for the idiosyncratic breaks in the drifts of the two series identified via the Bai and Perron (1998, 2003) methodology, TFP and the RPI clearly appear to be orthogonal at all horizons/frequencies. However, I illustrate the extreme sensitivity of these results to the exact nature and timing of the breaks characterizing the two series: conditional on a common break identified via the Bai, Lumsdaine, and Stock (1998) procedure, evidence is compatible with the notion that the two series contain a common I(1) component inducing negative-long horizon covariation between them. Further, I show via Monte Carlo that if the common break identified by Bai et al.’s tests were the ‘truth’, Bai and Perron’s tests applied to this DGP would generate, with non-negligible probabilities, results in line with those they produced based on the actual data. These results illustrate the difficulty of making strong statements about the exact nature of the long-horizon relationship between TFP and the relative price of investment, which automatically injects an element of fragility into investigations about the role played by neutral and investment-specific shocks in macroeconomic fluctuations.

Third, assuming, for illustrative purposes, that TFP and the RPI share in fact a common break (and therefore may contain a common component), based on estimated VARs for TFP, the RPI, and six other standard macroeconomic time series, I identify the news and non-news sub-components of such common I(1) component based on a simple modification of the ‘maximum fraction of long-horizon forecast error variance’ procedure originally proposed by Uhlig (2003) and Uhlig (2004). Specifically, I identify the two shocks based on the restrictions that (i) the news shock induces the largest (in absolute value) negative correlation between TFP’s and the RPI’s forecast errors at the 10-year horizon, under the constraint that it does not impact contemporaneously upon either of the two series; and (ii) taking as given the previously identified news shock, the non-news shock induces the largest (in absolute value) negative correlation between the two series’ forecast errors at the 10-year horizon. Overall, the news and non-news components jointly explain non-negligible fractions of the business-cycle frequency and/or the long-horizon forecast error variance of several series, including GDP and hours per capita, and inflation.

The paper is organized as follows. The next section reconsiders the long-horizon relationship between TFP and the RPI, whereas Section 3 presents results based on
2 Investigating the Relationship Between TFP and the RPI

For reasons of robustness, all of the empirical work has been conducted based on three alternative quarterly seasonally adjusted series for the RPI, produced by SGU (2011), LWZ (2011), and JPT (2011) respectively. The sample periods for the three series are 1948Q1-2006Q4, 1959Q1-2007Q4, and 1954Q4-2009Q1, respectively.

With a single exception, the TFP series used in all of the empirical work has been computed based on the quarterly seasonally adjusted series for the log-difference of 'purified TFP' produced by John Fernald, which is widely regarded as the best available measure of neutral technology, and found at the San Francisco FED’s website. Specifically, the series for the logarithm of TFP has been computed as the cumulative sum of Fernald’s series for the log-difference of purified TFP. The only instance in which I use an alternative TFP series is when I explore the possible issue of cointegration between TFP and the RPI recently raised by SGU (2011). In that case, I will perform cointegration tests based on both Fernald’s TFP series, and the alternative series produced by SGU (2011), in order to show how their finding of cointegration crucially depends on using their alternative series. Apart from this, all of the subsequent analysis will be based on Fernald’s series.

2.1 Results from unit root tests

Table 1 reports bootstrapped \( p \)-values for augmented Dickey-Fuller (henceforth, ADF) tests for the logarithms of TFP and the RPI. For all series, the estimated models include an intercept and a linear time trend, and \( p \)-values have been computed by bootstrapping 2,000 times estimated ARIMA(\( p,1,0 \)) processes.\(^1\) In all cases, the bootstrapped processes are of length equal to the series under investigation.\(^2\) As for the lag order, since, as it is well known, results from unit root tests may be sensitive to the specific lag order which is being used, for reasons of robustness I consider three alternative lag orders, one, two, and four. For each series, I consider results for the longest sample period for which the series is available.

\(^1\)Specifically, I have estimated AR(\( p \)) processes for the log-difference of either series, selecting the lag order as the maximum between the ones chosen by the Schwartz and the Hannan-Quinn criteria. Then, I have bootstrapped the estimated AR(\( p \)) processes, and I have cumulated each bootstrapped replication in order to get bootstrapped ARIMA processes for the log-levels of TFP and the RPI.

\(^2\)To be precise, letting \( T \) be the length of the series under investigation, we bootstrap an artificial series of length \( T+100 \), and we then discard the first 100 observations in order to eliminate dependence on initial conditions.
For the logarithm of the RPI results strongly point towards the presence of a unit root based on either of the three series I consider. By the same token, evidence of a unit root in the logarithm of TFP is strong based on either the series originally used by SGU (2011), or the one produced by Fernald.

Let’s now turn to exploring the bivariate relationship between the two series.

2.2 Are TFP and the RPI cointegrated?

In this sub-section I reconsider the issue of whether TFP and the RPI are, in fact, cointegrated, showing that

*First*, even based on SGU’s (2011) TFP series, their finding of cointegration crucially depends on their use of the Akaike information criterion (henceforth, AIC) in order to select the lag order. Since the AIC is an inconsistent criterion for lag order selection for VARs containing integrated variables— in the sense that it does not choose the lag order ‘correctly in large samples’—this result should therefore not be regarded as reliable. Further, I show that when the lag order is chosen based on the Schwartz and Hannan-Quinn criteria (henceforth, SIC and HQ, respectively), the null of no cointegration cannot be rejected at the 10 per cent level. Finally, based on Fernald’s corrected TFP series, and choosing the lag order based on the SIC and HQ criteria, results are even stronger, and in no way allow to reject the null of no cointegration.

*Second*, the same result obtains based on the RPI series produced by LWZ (2011).

*Third*, based on the JPT series the bootstrapped *p*-value does not provide conclusive evidence, but the ‘candidate cointegration residual’ appears as hardly stationary based on the ‘eyeball metric’, thus casting doubts, once again, on the notion that TFP and the RPI might be cointegrated.

2.2.1 Results from Johansen’s cointegration tests

Table 2 reports bootstrapped *p*-values for Johansen’s trace test of the null of no cointegration between the logarithms of TFP and the RPI. *p*-values have been computed

---

3 See, e.g., Section 11.4.1 of Lütkepohl (1991). In particular, on page 383 he points out that ‘[...] AIC is not a consistent criterion while HQ [Hannan-Quinn] and SC [Schwartz] are both consistent. Thus if consistent estimation is the objective we may apply HQ and SC for stable and unstable processes.’


5 Since I am here bootstrapping critical and *p*-values, my results are robust to (i) the possible non-normality of the shocks, and (ii) small-sample problems. In particular, as for the former issue, under very general conditions the residuals of the VAR estimated under the null of no cointegration are consistent estimates of the true underlying shocks, in the sense that they do converge to such shocks in probability limit. As a result, the underlying shocks may have, in population, any non-degenerate distribution we can think of, because bootstrapping the residuals automatically takes care of that.
by bootstrapping the VAR estimated for the first differences of the two series.\(^6\)

Starting from the results based on SGU’s (2011) RPI and TFP series, in accordance with the results reported in their paper, the AIC selects a lag order of 7. At the 10 per cent level, the null of no cointegration is comfortably rejected, with a \(p\)-value of 0.088.\(^7\) Since, as previously pointed out, the AIC is an inconsistent criterion for lag order selection for VARs containing integrated variables, this result should however be regarded as uninformative for the issue at hand. Selecting the lag order as the maximum between those chosen by the SIC and HQ criteria, on the other hand, produces a lag order equal to 3, with the corresponding \(p\)-value equal to 0.1212. So, even based on SGU’s (2011) TFP series, the finding of cointegration between TFP and the RPI crucially depends on the use of an inconsistent criterion for lag order selection. Finally, Figure 1 provides additional evidence on this, by showing, in the first panel, the cointegration residual obtained when the lag order is selected based on the AIC. Simple ‘eyeball econometrics’ casts doubts on the notion that the two series may be cointegrated, as the residual hardly appears to be stationary.

Turning to the results based on SGU’s (2011) RPI and Fernald’s corrected TFP series, with the lag order selected as the maximum between those chosen by the SIC and HQ criteria, results are even stronger, with bootstrapped \(p\)-value equal to 0.3336, thus clearly suggesting that the null of no cointegration cannot be rejected at conventional significance levels.

Turning now to the results based on JPT’s (2011) and LWZ’s (2011) RPI series and Fernald’s corrected TFP, whereas the \(p\)-value for LWZ’s series, at 0.1531, does not allow to reject the null of no cointegration, the one based on JPT’s series, at 0.0755, points towards cointegration at the 10 per cent level. The second panel of Figure 1, plotting the cointegration residual between the two series, raises however

\(^6\)To be clear, this means that, given the vector \(Y_t = [\ln(TFP_t), \ln(RPI_t)]\), I start by selecting the lag order for the cointegration tests as the maximum between the lag orders selected based on the SIC and HQ criteria for the VAR in levels for \(Y_t\), and I perform Johansen’s trace test of the null of no cointegration. Then, I estimate the VAR for \(\Delta Y_t\) (again, to be clear: the VAR I am estimating here is not the cointegrated VAR, that is, it is equal to the VECM representation without the error-correction term); I bootstrap it 2,000 times, thus generating bootstrapped artificial series \(\tilde{\Delta Y}_t\); based on each of them I compute corresponding bootstrapped artificial series \(\tilde{Y}_t\)—that is, those for the levels of the series; and based on each of them I perform the same trace test I previously computed based on the actual data, thus building up the empirical distribution of the trace statistic under the null of no cointegration. Then, based on this distribution, I compute critical values (not reported here) and \(p\)-values. Once again, for each pair of series results are based on the longest sample period for which both are available (the exact sample periods are reported in Table 1).

\(^7\)It is to be noticed that nowhere in the paper SGU (2011) mention the issue of bootstrapping critical and \(p\)-values, and it might therefore be reasonably assumed that the \(p\)-values they show in their Table 3, ranging between 0.01 and 0.07, are asymptotic, and therefore potentially subject to small-sample size distortion. On the other hand, as I have shown via Monte Carlo in Benati (2013), even for sample sizes significantly shorter than the ones I am using here, the size distortion associated with the methodology I am using herein is remarkably small, ranging between essentially nil to about 1 per cent at most. As a result, the bootstrapped \(p\)-values here reported in Table 2 should probably be regarded as the more reliable.
doubts on the soundness of this result, as the residual, once again, does not clearly appear to be stationary.

Overall, my own reading of the evidence is that TFP and the RPI are most likely not cointegrated, which immediately raises the next set of questions: If they are not cointegrated, what is their relationship at the very low frequencies? Are the stochastic trends driving the two series orthogonal to each other, or do they contain a common I(1) component? These are key questions because, in the latter case, SGU’s (2011) criticism of the traditional approach of assuming that neutral and investment-specific technology follow independent processes would still be valid, even if the two series are not cointegrated.

2.3 Do the two series contain a common I(1) component?

There are (at least) two alternative, but conceptually related ways to search for a common I(1) component between two non-stationary, non-cointegrated series. Since the key issue here is the extent of long-horizon—that is: low-frequency—covariation between the two series, this can be explored either by

(i) searching, in the spirit of Cochrane and Sbordone (1988), for a statistically significant extent of co-variation between the two series long-horizon’s differences, or by

(ii) looking at their cross-spectrum at the very low frequencies, and testing whether it is significantly different from zero.

I now consider the two methods in turn.

2.3.1 A simple test in the spirit of Cochrane and Sbordone (1988)

A simple, but powerful reduced-form implication of lack of orthogonality between the stochastic trends driving two non cointegrated I(1) processes is that, whatever may happen at short-to-medium horizons, the two series’ long-horizon differences should exhibit a statistically significant extent of co-variation. The intuition for this is straightforward, and can be formalized as follows. Consider the two I(1) processes $y_{1,t}$ and $y_{2,t}$, defined as

$$y_{1,t} = \alpha_1 z_t + \beta_1 x_{1,t} + \gamma_1 s_{1,t}$$

$$y_{2,t} = \alpha_2 z_t + \beta_2 x_{2,t} + \gamma_2 s_{2,t}$$

where $z_t = z_{t-1} + u_t$ is the common I(1) component; $x_{i,t} = x_{i,t-1} + v_{i,t}$, $i = 1, 2$, are idiosyncratic I(1) components; and $s_{i,t} = \rho_i s_{i,t-1} + \epsilon_t$, $i = 1, 2$, with $|\rho_i| < 1$ are idiosyncratic, but correlated I(0) components: notice that the shock, $\epsilon_t$, is the same for each of the two $s_{i,t}$’s. Finally, the four shocks—$u_t$, $v_{1,t}$, $v_{2,t}$, and $\epsilon_t$—are uncorrelated both contemporaneously, and at all leads and lags, with variances $\sigma_u^2$, $\sigma_{v1}^2$, $\sigma_{v2}^2$, and $\sigma_\epsilon^2$, respectively, whereas the $\alpha_i$’s, $\beta_i$’s, and $\gamma_i$’s are parameters. This data-generation process (henceforth, DGP) captures the notion of two I(1) processes which are not cointegrated but, nonetheless, share a common random-walk component, the
Further, in order to capture the notion of short-to-medium horizon business-cycle co-movements between the two series, they are postulated to share the common shock $\epsilon_t$, which drives the I(0) processes $s_{i,t}$, $i = 1, 2$. Based on (1)-(2), there are two sources of co-variation between $y_{1,t}$ and $y_{2,t}$, originating from (i) the common random-walk component, and (ii) the two I(0) processes, which are driven by the common shock $\epsilon_t$. The intuition—and the key issue for our purposes—is that, since $z_t$ is I(1) and the $s_{i,t}$’s are I(0), co-variation between $y_{1,t}$ and $y_{2,t}$ at long horizons will be dominated by the common random-walk component, whereas the relative importance of the extent of co-variation originating from the two correlated I(0) processes will asymptotically go to zero. In turn, this implies that the presence of a statistically significant extent of long-horizon co-variation between $y_{1,t}$ and $y_{2,t}$ necessarily implies the presence of a common I(1) component.

Formally, it can be easily shown that the $k$-horizon differences for the two processes are given by

$$y_{1,t+k} - y_{1,t} = \alpha_1 \sum_{j=1}^{k} u_{t+j} + \beta_1 \sum_{j=1}^{k} v_{1,t+j} + \gamma_1 \left( \rho_1^k - 1 \right) s_{1,t} + \sum_{j=1}^{k} \rho_1^{-j} \epsilon_{t+j} \right) \right]$$

$$y_{2,t+k} - y_{2,t} = \alpha_2 \sum_{j=1}^{k} u_{t+j} + \beta_2 \sum_{j=1}^{k} v_{2,t+j} + \gamma_2 \left( \rho_2^k - 1 \right) s_{2,t} + \sum_{j=1}^{k} \rho_2^{-j} \epsilon_{t+j} \right]$$

which implies that the covariance between them is given by

$$E[(y_{1,t+k} - y_{1,t})(y_{2,t+k} - y_{2,t})] = \alpha_1 \alpha_2 k \sigma_u^2 + \gamma_1 \gamma_2 \sigma^2 \left( \rho_1^k + \rho_2^k \right)$$

so that

$$\lim_{k \to \infty} \frac{E[(y_{1,t+k} - y_{1,t})(y_{2,t+k} - y_{2,t})]}{k} = \alpha_1 \alpha_2 \sigma_u^2$$

thus implying that $(1/k)$ times the covariance between the two series’ long-horizon differences (i) comes to be dominated by the common random walk component, and, as a consequence (ii) it is different from zero if and only if the two series share a common I(1) component.

Following Cochrane and Sbordone (1988), I estimate $(1/k)$ times the covariance of long-horizon differences between $y_{1,t}$ and $y_{2,t}$ as

$$\frac{T}{k(T-k)(T-k+1)} \sum_{t=k}^{T} ((y_{1,t} - y_{1,t-k}) - \frac{k}{T} (y_{1,T} - y_{1,0})) ((y_{2,t} - y_{2,t-k}) - \frac{k}{T} (y_{2,T} - y_{2,0}))$$

where $T$ is the sample length, and $(y_{1,T} - y_{1,0})/T$ and $(y_{2,T} - y_{2,0})/T$ are the estimated drifts in the two processes.

Figure 2 reports the Monte Carlo distribution of the Cochrane-Sbordone (henceforth, CS) estimator (7) at horizons up to 30 years ahead, based on 10,000 stochastic
simulations of two independent random walks of length $T$, with $T = 200, 1000,$ and $10000$. It is to be noticed that 200 is roughly the number of observations we are here working with.\footnote{To be precise, sample lengths are equal to 235 quarters based on SGU’s RPI, and to 195 quarters based on the other two RPI series.} Two things emerge from the figure. \textit{First}—and reassuringly—as the sample length increases, the CS estimator ‘captures the truth’, in the sense that, at all horizons, the distribution becomes more and more tightly clustered around zero. \textit{Second}, for samples of the length we are here working with, uncertainty is substantial, as reflected by a pretty much spread-out distribution.

Figure 3 reports evidence based on Fernald’s TFP series, and either of the three RPI series we are here working with, showing the simple estimate of the CS estimator at horizons up to 30 years ahead, together with the median and the 90%- and 95%-coverage percentiles of the bootstrapped distribution of the estimator computed under the null hypothesis that the two processes follow orthogonal ARIMA($p,1,0$) processes.\footnote{Bootstrapping has been performed exactly as in Section 3.1. The two ARMA processes for the log-differences of TFP and the RPI have been estimated independently, bootstrapped, and each individual bootstrapped replication for the log-difference of either process has been cumulated in order to get the corresponding bootstrapped ARIMA processes.} Overall, evidence suggests that TFP and the RPI share a common I(1) component, as

(i) based on JPT’s RPI series, the CS estimator exceeds the 90%- and 95%-coverage upper percentiles of the bootstrapped distribution at all horizons beyond about 13 and 17 years, respectively;

(ii) based on SGU’s RPI series, it exceeds the 90%-coverage upper percentile at all horizons beyond about 15 years, whereas it exceeds the 95%-coverage upper percentile at horizons between about 23 and 28 years ahead; on the other hand,

(iii) based on LZW’s RPI series results are much weaker, with the estimator exceeding the 90%-coverage upper percentile marginally, and for just a few quarters.

An important point to stress is that, in all cases, evidence points towards the common I(1) component inducing a positive co-variation between the two series at long horizons. I will get back to this issue in Section 2.4.

Let’s now turn to evidence based on cross-spectral methods.

\subsection*{2.3.2 An alternative test based on the cross-spectrum of the log-differences of the two series at the very low frequencies}

A reduced-form implication of orthogonality between the stochastic trends driving two I(1) series is that, whatever the relationship between the series may be at frequencies greater than zero, the cross-spectrum between their first differences at the frequency $\omega=0$—and, more generally, at the very low frequencies—ought to be equal to zero.\footnote{To fix ideas, let’s consider two I(1) processes, $y_{1,t}$ and $y_{2,t}$, with $y_{i,t} = x_{i,t} + z_{i,t}$, $[1-\phi_1(L)]\Delta x_{i,t} = \theta_i(L)u_{i,t}$ and $z_{i,t} \sim I(0)$, $i = 1, 2$. Since the frequency-response function of the first-difference filter $\Delta \equiv (1 - L)$ at the frequency $\omega=0$ is equal to zero, it automatically follows that, at $\omega=0$, the}
This implies that a simple way to test whether TFP and the RPI’s two stochastic trends are orthogonal to each other is to test whether the cross-spectrum between their log-differences at \( \omega = 0 \) is equal to zero.

Figure 4 reports evidence on this, by showing (in black) the real and the imaginary parts of the cross-spectrum between the log-differences of TFP and the RPI at the very low frequencies, together with the medians and the 90\%-\%, 95\%-\%, 99\%-coverage percentiles of the bootstrapped distributions of the same objects computed under the assumption that the logarithms of the two series evolve according to orthogonal ARIMA\((p,1,0)\) processes.\(^{11}\) In spite of the fact that, by definition, the imaginary part of the cross-spectrum at \( \omega = 0 \) is equal to zero, I also report results for this object because the Fourier frequencies just immediately above zero provide additional evidence on the issue under investigation.

Overall, evidence points against the notion that TFP and the RPI’s two stochastic trends may be orthogonal to each other. Specifically,

- **first**, the null hypothesis that the real part of the cross-spectrum between the log-differences of the two series at \( \omega = 0 \) is equal to zero can be rejected at the 1\% level based on either JPT’s or SGU’s RPI series—strongly in the former case, and marginally in the latter—and at the 5\% level based on LWZ’s series.

- **Second**, at the frequencies just immediately above zero, the notion that the imaginary part of the cross-spectrum is equal to zero can likewise be strongly rejected, at the 1\% level based on either SGU’s or LWZ’s series, and at the 5\% level based on JPT’s series.

### 2.4 Interpreting the results pointing towards a common I(1) component

Overall, results based on either CS’ estimator or cross-spectral methods suggest that TFP and the RPI, although not cointegrated, share a common I(1) component, which induces a positive co-variation between the two series at long horizons. If correct this would be an important finding, because, conceptually in line with SGU (2011), it would imply that all previous analyses, being based on the assumption that TFP

---

\(^{11}\)The spectral density matrix of the data has been computed via the ‘lag window’ estimator as described in Hamilton (1994, chapter 6). The bootstrapping methodology used herein is exactly the same I used in Section 3.1 to compute bootstrapped \( p \)-values for ADF tests.
and the RPI evolve as independent processes, are in fact unreliable. In this section I consider three alternative possible explanations for this result.

### 2.4.1 First interpretation: the result is genuine

The first possibility is that the result is genuine, and the two series contain indeed both idiosyncratic unit root components, and a common I(1) component. As I will argue this interpretation is in fact implausible, because after controlling for econometrically identified breaks in the means of the log-differences of TFP and the RPI, evidence of positive long-horizon covariation disappears, irrespective of the specific type of break which is being controlled for (idiosyncratic, or common).

### 2.4.2 An interpretation in terms of non-linearity of the technology transforming consumption into investment

A second possible explanation is non-linearity of the technology transforming consumption goods into investment goods, which would imply that the RPI is impacted upon by both investment-specific and neutral shocks, and that—in contrast with the literature which has followed Hercowitz, Greenwood, and their co-authors\(^{12}\)—permanent disturbances to the RPI cannot therefore be interpreted as uniquely reflecting investment-specific technology. The easiest way to show this is based on the simple RBC model used by SGU (2011). Their equation for the RPI (on page 127), is given by

\[
P_t' = \frac{1}{a_tX_t^\alpha I_t^\xi - 1}
\]

where \(a_t\) and \(X_t^\alpha\) are, respectively, stationary and non-stationary investment-specific technology, \(I_t\) is gross investment measured in consumption units, and \(\xi\) is the parameter capturing the curvature of the technology transforming consumption goods into investment goods. If \(\xi = 1\), then \(I_t\) disappears from (8) and \(P_t\) uniquely reflects investment-specific shocks. If \(\xi \neq 1\), taking logs we have

\[
\ln P_t' \equiv p_t' = -\ln a_t - x_t^\alpha - \ln \xi - (\xi - 1) \ln I_t
\]

where \(x_t^\alpha = \ln X_t^\alpha\). Since \(I_t\) is expressed in consumption units, this means that \(\ln I_t = \ln I_t^\alpha + p_t'\), where \(I_t^\alpha\) is investment expressed in physical units. Since, within any DSGE model driven by I(1) neutral and investment-specific technology, GDP, consumption, and investment are all cointegrated with the two technology processes, we have that

\[
\ln I_t^\alpha = \alpha_1^\prime x_t^\alpha + \alpha_2^\prime z_t + \chi_t,
\]

where \(z_t\) is the logarithm of neutral technology, which is postulated to follow a random walk, and \(\chi_t\) is an I(0) process. This implies that,

---

ignoring $\chi_t$, which for the present purposes is irrelevant,

$$p^f_t = \frac{(1 - \xi)\alpha^f_x z_t - 1 + (\xi - 1)\alpha^f_x x_t}{\xi > 0}$$  \hspace{1cm} (10)$$

As (10) shows, assuming that Fernald’s purified TFP correctly captures neutral technology, a simple explanation for positive long-horizon covariation between TFP and the RPI is that $\xi \neq 1$, so that the RPI is also impacted upon by neutral technology.

2.4.3 An interpretation in terms of breaks in the drifts of the two series

Tables 3 and 4 report results from tests for multiple structural breaks at unknown points in the sample in the mean for the log-differences of TFP and the RPI, and the remaining series which will enter the VARs in Section 3, based on Bai and Perron (1998, 2003). The methodology I am using herein is identical to the one I used in Benati (2007). Specifically, I exactly follow the recommendations of Bai and Perron (2003), with the only difference that, instead of relying on the asymptotic critical values tabulated in Bai and Perron (1998), I bootstrap both critical and $p$-values as in Diebold and Chen (1996), setting the number of bootstrap replications to 1,000. I start by looking at the $UDmax$ and $WDmax$ double maximum test statistics. Conditional on both statistics being significant at the 10% level—thus indicating the presence of at least one break—I decide on the number of breaks by sequentially examining the $sup-F(t+1|t)$ test statistics, starting from the $sup-F(2|1)$ one. I set the maximum allowed number of structural changes to $m=4$, and the amount of ‘trimming’ at the beginning and end of the sample to Bai and Perron’s (2003) recommended value of 15 per cent. Confidence intervals for estimated break dates are computed as in Bai (1997).

Results from the double maximum tests identify breaks in the mean for the log-differences of TFP, of real consumption per capita, and of the RPI based on either LWZ’s or JPT’s series. As for SGU’s series, although no break is identified for the full sample period (1948Q2-2006Q4), it is to be noticed that the $p$-values are not far from the 10 per cent threshold. Further, the most likely explanation for the failure to identify a break in the mean based on the full sample period is that, since the early 2003, the log-difference of SGU’s RPI has visibly jumped upwards, thus pointing towards a deceleration in the rate of growth of investment-specific technology. Since this last portion of the sample represents just 6.8 per cent of the overall sample period, it automatically ends up in the 15 per cent end-of-sample ‘trimmed portion’. This means that the algorithm does not even consider (say) 2003Q1 as a possible break date, with the result that the increase in the series during the last portion of the sample has the only effect of making it more difficult to identify the break in the mean which most likely took place in the early 1980s. In order to check for this

\[13\text{See Bai and Perron (2003) section 5.5, ‘Summary and Practical Recommendations’}.

12
possibility, Table 3 also reports results for the shorter sample period 1948Q2-2002Q4. Results are completely different, with p-values equal to 0.043 and 0.045 respectively, and strongly point towards a break in the mean.

For either of the series for which the presence of at least one break has been detected, p-values for the the sup-$F(2|1)$ statistic do not point towards the presence of a second break. For the present purposes, the key results are

(i) the break in the mean of the log-difference of TFP identified in 1968Q2, with a decrease in the mean rate of growth from 2.01 to 0.80 per cent, and

(ii) the break in the mean of the log-difference of the RPI identified in 1982Q2 based on either of the three RPI series, with a decrease in the mean rate of growth from -1.27 to -2.47, from -1.25 to -3.20, and from -0.63 to -2.36 per cent, respectively.

As I will now show, these breaks in the means of the log-difference of TFP and the RPI are a logical explanation of the results reported in Figures 3 and 4, pointing towards lack of orthogonality between the two series at long horizons. Figures 5 and 6 reports results for the very same tests show in Figures 3 and 4, respectively, but with the key difference that the series under investigation have now been adjusted to control for the identified breaks in the mean. The results are completely different from those reported in Figures 3 and 4, and uniformly and strongly point towards long-horizon and low-frequency orthogonality between the TFP and the RPI. The explanation for this is straightforward, and it is conceptually in line with Fernald’s (2007) discussion of how failure to control for breaks in the mean of labor productivity can distort estimates of the impact of a technology shock on hours. Since, over the sample period, TFP and the RPI have experienced breaks in the drift in the same direction (that is: downwards), such breaks have spuriously introduced a positive extent of long-horizon (low-frequency) co-variation between the two series. As a result, failing to control for such breaks produces the illusion that they do contain a common I(1) component.

2.4.4 Are the breaks in the drifts independent?

An important objection to this interpretation is that, if the explanation discussed in section 2.4.2 were correct, the break in the mean of the log-difference of the RPI might have been caused by the break in the mean of the log-difference of TFP. This is immediately apparent from equation (10): since $\xi^{-1}(1 - \xi)\alpha_z > 0$, a decrease in $\Delta z_t$ automatically translates into a corresponding decrease in $\Delta p_r^t$. At first sight this explanation does not appear very likely, as the identified break dates for TFP and the RPI, 1968Q2 and 1982Q2, respectively, are 14 years apart, which is hardly

---

14In the case of the log-differences (Figure 6) I have imposed in Period II the same mean estimated for Period I (see the estimates reported in Table 4). As for the log-levels (Figure 5), I have performed the just-mentioned adjustment on the log-differences of the relevant series, and I have then reconstructed the log-levels controlling for the breaks in the mean of the log-differences.

15This explanation is nothing but a multivariate ‘version’ of Perron (1989)’s classic point that failure to control for breaks in the mean gives the illusion that a series contains a unit root.
compatible with the notion that the former break may have caused the latter. It is important to stress, however, that based on either the JPT or the LWZ RPI series the 90%-coverage confidence intervals for the estimated break dates for TFP and the RPI do indeed overlap, so that, technically, a common break cannot be ruled out.

In order to explore the possibility that the two series may have experienced a common break I therefore follow Bai, Lumsdaine, and Stock (1998), and I test for a common break in the intercepts of the VAR for the log-differences of TFP and the RPI. For reasons of robustness, I consider results from both sup- and exp-Wald tests, and I either set the VAR lag order to four, or select it as the maximum between the lag orders chosen by the SIC and HQ criteria. Finally, I compute critical values (not reported here) and $p$-values by bootstrapping the VAR 2,000 times under the null of no break in any of its features, and I estimate confidence intervals for the break dates based on the formulas found in Bai, Lumsdaine, and Stock (1998). Results are reported in Table 5. Evidence is remarkably robust across the different RPI series, lag orders, and test statistics, and points towards a common break in the means of the log-differences of the two series. The only difference between the three sets of results based on the alternative RPI series pertains to the specification, which is estimated as 1975Q2, 1977Q4, and 1982Q2 based on the SGU, JPT, and LWZ series, respectively.

The key question now is: Does controlling for the identified common breaks make a material difference for the extent of estimated long-horizon covariation between the two series, compared to controlling for the idiosyncratic breaks identified by Bai and Perron’s tests?

### 2.4.5 An illustration of the importance of controlling for the ‘right’ break

As an illustration of the importance of controlling for the ‘right’ break, and of how results are sensitive to the specific type of break (common, or idiosyncratic) which is being controlled for, and to its dating, Figures 7 and 8 show results from applying the CS and cross-spectral estimators to the log levels and log-differences, respectively, of TFP and the RPI after controlling for the common breaks identified by the Bai, Lumsdaine, and Stock (1998) tests. Overall, evidence is not clear-cut. Specifically,

(i) based on either cross-spectral methods or CS’ estimator, results based on JPT’s RPI series point towards the presence of non-zero long-horizon covariation. Results based on CS’ estimator, in particular, point towards negative covariation at horizons up to 20 years at the 10 per cent level (although at longer horizons the CS estimator becomes not significantly different from zero), thus suggesting that neutral and investment-specific technology may contain a common I(1) component which has a positive permanent impact on both series.

(ii) Results based on the LWZ series are the polar opposite of those based on the JPT series, as based on either methodology they do not point towards any statistically significant extent of covariation at any horizon.
Finally, results based on SGU’s series are somehow ‘in between’, and point towards negative covariation at the 10 per cent level essentially at all horizons based on CS’ estimator, whereas they are basically insignificant, with the exception of a single Fourier frequency, based on cross-spectral methods.

These results definitely reject the explanation based on the non-linearity of the technology transforming consumption goods into investment goods. If that explanation were correct, indeed, even after controlling for breaks in the means of the log-differences of the two series, still, neutral technology shocks should have a long-run positive impact on the RPI. The fact that in no way Figures 7 and 8 point towards positive long-horizon covariation represents a clear rejection of this explanation.

The question now becomes: Can we reject the notion of a common break—together with its ‘unnatural’ implication that the two series may contain a common I(1) component—in favor of the notion of idiosyncratic breaks, with the associated ‘natural’ implication that they evolve independently at long horizons? Unfortunately, as I will now show via Monte Carlo, this is not easy.

2.4.6 Can we reject the notion of a common break in favor of the alternative of independent breaks?

In this section I perform Monte Carlo simulations in order to address the following question:

‘Suppose that the true DGP is the one identified by Bai et al.’s break tests, so that the log-differences of TFP and the RPI feature a common break in the mean. Would the application of Bai and Perron’s tests to this DGP generate a set of results compatible with those we obtained based on the actual data?’

Table 6 reports results from applying Bai and Perron’s tests to the bootstrapped log-differences of TFP and the RPI, respectively, which have been generated by bootstrapping 1,000 times¹⁶ a VAR for the two actual series estimated conditional on the previously identified common break. The Bai and Perron tests I apply to the artificial series are identical to those I applied to the actual data in Section 2.4.3. In particular, critical and p-values have been computed by bootstrapping estimated AR(p) processes for the artificial series as in Diebold and Chen (1996), setting the number of bootstrap replications to 1,000. So, to be clear, the results reported in the table are exactly comparable to those we discussed in Section 2.4.3, because they have been produced by applying the very same methodology, with the only difference that there we applied it to the actual data, and here we are instead applying it to artificial data generated by bootstrapping a specific DGP. The table reports (i) the fractions of bootstrap replications for which Bai and Perron’s tests identify a break

¹⁶I am here using a smaller number of bootstrap replications due to the extreme computational intensity of the entire exercise. In particular, I am here not only generating the artificial TFP and RPI series via bootstrapping, but, crucially, I am also performing the Bai and Perron tests by bootstrapping critical and p-values as previously done in Section 2.4.3.
(as in Section 2.4.3, the criterion for identifying at least one break is that the both p-values associated with double-maximum statistics must be smaller than 0.1); and (ii) among this subset of bootstrap replications, the fraction of replications for which the 90 per cent confidence interval for the estimated break date does not contain the common break (which, within the context of the present experiment, is the ‘truth’ we wish to capture). Three things stand out.

First, Bai and Perron tests identify a break in the mean of the log-difference of the RPI between 48.7 and 97.2 per cent of the time. This is due to the fact that, for either of the three RPI series, the size of the downward jump in the mean of the log-difference of the RPI is quite sizeable compared to the standard deviation of the OLS residuals in the corresponding VAR’s equation, so that identifying the break is comparatively easy.

Second, for TFP, for which the size of the downward jump in its log-difference, relative to the standard deviation of OLS residuals, is much smaller, the tests identify a break between 11.1 and 46.2 per cent of the time.

Third, and crucially, confidence intervals for estimated break dates are quite imprecise, and do not contain the common break a non-negligible fraction of the time. Specifically, if we only consider the subset of replications for which a break has been detected, between 29.3 and 58.4 per cent of the time the confidence interval does not contain the ‘true’ common break. Out of all bootstrapped 1,000 replications—which should be regarded as the correct metric for assessing the reliability of this methodology—the fraction (which is computed by multiplying the two numbers in each column) ranges between 5.6 and 54.4 per cent. This implies that results such as those based on the JPT series in Tables 4 and 5—where the confidence interval for the break identified by Bai and Perron, [1981Q4; 1988Q1], does not contain the common break identified by Bai et al.’s methodology, should not be regarded as strong evidence against the identified common break, because conditional on the DGP featuring the identified common breaks this result should be expected a non-negligible fraction of the time.

2.5 Summing up

Although I personally find the notion that TFP and the RPI possess idiosyncratic breaks (with the associated ‘natural’ implication that they evolve independently at long-horizons) more plausible, and more appealing than the alternative of a common break, from a strictly statistical point of view, the latter notion cannot be easily rejected in favor of the former. More generally, these results illustrate the difficulty of making strong statements about the exact nature of the long-horizon relationship between TFP and the relative price of investment, which automatically injects an element of fragility into investigations about the role played by neutral and investment-specific shocks in macroeconomic fluctuations. As I will show in the next section, indeed, controlling TFP and the RPI for the common break identified by Bai et al.’s
tests, and using structural VAR methods to extract the news and non-news portions of the common I(1) component between the two series, suggests that, overall, such common component plays a non-negligible role for series such as GDP, hours, and inflation. In turn, this implies that if this were the truth, the exact nature of the role played by technology and news shocks in macroeconomic fluctuations could well be pretty much different from what the literature has produced so far.

3 Extracting a Common Component from TFP and the RPI

3.1 The reduced-form VARs

The VARs used in the empirical analysis that follows feature the logarithms of the RPI and TFP, controlled for the common breaks in the means of their log-differences identified by Bai et al.’s tests; inflation; the real ex post Federal Funds rate; and the logarithms of hours worked, real GDP, real consumption, and real stock prices, all of them expressed in per capita terms. I use the real ex post Federal Funds rate, rather than the nominal FED Funds rate, because it is a better—although inevitably imprecise—indicator of the stance of monetary policy. Since one of the objects of interest in the analysis that follows is to explore how the news and news common shocks affect the monetary policy stance, this is a natural choice. Finally, following Fisher (2006) I take the price of consumption (that is: the chain-weighted deflator for non-durables and services) to be the numeraire of the system (see the appendix for details on the construction of the dataset, which is standard).

The sample periods are 1954Q3-2006Q4, 1954Q3-2007Q4, and 1954Q3-2008Q2 based on the SGU, LWZ, and JPT RPI series, respectively. The beginning of the sample is dictated by the fact that the Federal Funds rate is available only starting from July 1954, whereas the end of the sample is due to the fact that, on the one hand, the first two RPI series are available until 2006Q4 and 2007Q4, respectively; on the other hand, as for the series from Justiniano, Primiceri, and Tambalotti (2011), I end the sample in 2008Q2 in order to exclude the period immediately following the collapse of Lehman Brothers, which was characterized by an extraordinary extent of turbulence, and might therefore have distorted the inference.

All of the VARs feature 4 lags, and are estimated in levels via OLS.

3.2 Identification

My identification strategy is conceptually in line with the ‘maximum fraction of forecast error variance’ approach to identification pioneered by Uhlig (2003) and Uhlig (2004), with the key difference that, instead of looking for shocks which explain the maximum fraction of the forecast error variance of a series at a given horizon, I look
for shocks which induce the largest (in absolute value) possible negative correlation between the two series’ forecast errors at long horizons. Different from the ‘standard’ Uhlig methodology, the modified version used herein cannot be formulated as a straightforward eigenvalue-eigenvector problem (as least, I was not able to formulate it in this way), and requires instead numerical optimization methods. Given the VAR($p$) model

$$Y_t = B_0 + B_1 Y_{t-1} + \ldots + B_p Y_{t-p} + u_t, \quad E[u_t u'_t] = \Omega$$

(11)

with the moving-average representation $Y_t = [B(1)]^{-1} B_0 + [B(L)]^{-1} u_t \equiv C(L) u_t$, where $B(L) \equiv I - B_1 L - \ldots - B_p L^p$, and $C(L) \equiv I + C_1 L + C_2 L^2 + C_3 L^3 + \ldots$, identification boils down to finding a mapping between the reduced-form forecast errors, the $u_t$, and the structural shocks, the $\epsilon_t$, such that $u_t = A_0 \epsilon_t$, with $A_0 A'_0 = \Omega$. In turn, since, given a matrix $A_0^*$ such that $A_0^* (A_0^*)' = \Omega$, and an orthonormal matrix $Q$ such that $QQ' = I$, the matrix $A_0 = Q A_0^* Q$ also satisfies $A_0 A_0' = \Omega$, the search for $A_0$ boils down—for a given starting matrix $A_0^*$ satisfying $A_0^* (A_0^*)' = \Omega$ (e.g., the Cholesky factor of $\Omega$)—to finding an appropriate orthonormal matrix $Q$. For a given starting structural impact matrix $A_0^*$ satisfying $A_0^* (A_0^*)' = \Omega$, I define the orthonormal matrix $Q$ as the product of all of the available rotation matrices $R_i(\theta_i, K)$, where $2 \leq K \leq N$ is the dimension of the square sub-matrix along the diagonal of $R_i(\theta_i, K)$ containing the $\sin(\theta_i)$ and $\cos(\theta_i)$ functions. For example, if $N$ were equal to 4—so that the VAR only contained 4 variables—there would be three rotation matrices with $K = 2$:

$$R_1(\theta_1, K) = \begin{bmatrix}
\sin(\theta_1) & \cos(\theta_1) & 0 & 0 \\
-cos(\theta_1) & \sin(\theta_1) & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1
\end{bmatrix} \quad R_2(\theta_2, K) = \begin{bmatrix}
1 & 0 & 0 & 0 \\
0 & \sin(\theta_2) & \cos(\theta_2) & 0 \\
0 & -\cos(\theta_2) & \sin(\theta_2) & 0 \\
0 & 0 & 0 & 1
\end{bmatrix}$$

$$R_3(\theta_3, K) = \begin{bmatrix}
1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 \\
0 & 0 & \sin(\theta_3) & \cos(\theta_3) \\
0 & 0 & -\cos(\theta_3) & \sin(\theta_3)
\end{bmatrix}$$

(12)

By the same token, there would be two rotation matrices with $K = 3$, and one with $K = 4$, for a total of six. In our case, with $N = 8$, the overall number of available rotation matrices is 28, and $Q$ is defined as

$$Q = \prod_{K=2}^{N} \prod_{i=1}^{N-K+1} R_i(\theta_i, K)$$

(13)

Then, I search, over the parameter space, for those specific values of the rotation angles—that is, the $\theta_i$’s—which maximize the criterion function. Although obviously not as fast as the standard Uhlig methodology—which, being based on the eigenvalue-eigenvector decomposition, is performed essentially ‘in zero time’—numerical opti-
mization is quite fast, and identifying the first two shocks which maximize the criterion function took, on average, about 25 seconds. Optimization was performed via MATLAB’s routine \texttt{fminsearch.m}, for random initial conditions.

Since identification of the shocks is here based on numerical optimization of a criterion function, rather than on a matrix decomposition, the reader may have a legitimate question: ‘How strong is identification? Is the criterion function unimodal, bimodal, or what?’. In order to address this question, I performed the following Monte Carlo experiments. For either of the three VARs estimated based on the three available RPI series, and conditional on the OLS estimate of the covariance matrix of the VAR’s innovations—that is, \( \Omega \) in (11)—I performed the numerical optimization 1,000 times with random initial conditions, each time identifying the two shocks I discuss in the next sub-section (specifically: news and non-news). For either of the three VARs, and for each of the 1,000 numerical optimizations I have performed for each VAR, the algorithm always converged to the same solution (up to machine precision). This clearly illustrates that identification here is as robust as it can possibly be.

Finally, the strongest piece of evidence in favor of the reliability of this numerical optimization-based methodology is that when I used it—just to perform a check—in order to identify the single shock which explains the maximum fraction of the horizon-\( H \) forecast error variance of either TFP or the RPI, the solution I got was, up to machine precision, identical to the one produced by the ‘standard’ Uhlig (2003, 2004) methodology.

Let’s now turn to the results.

3.3 Evidence

I order the RPI and TFP first and second, I estimate the VAR, and I proceed to identify two common shocks—that is: news and non-news, respectively—as follows.

First, I identify the news shock based on the restrictions that (i) it generates the largest (in absolute value) possible negative correlation between the two series’ forecast errors at the 10-year ahead horizon, and (ii) it has a zero impact on both series at \( t=0 \).

Second, conditional on having identified the news shock, I identify the non-news shock based on the only restriction that it generates the largest (in absolute value) possible negative correlation between the two series’ forecast errors at the 10-year ahead horizon.

Figures 9 and 10 show the results based on SGU’s series. (For reasons of space, and since this section’s analysis is purely illustrative, I do not report those based on JPT’s and LWZ’s series, but they are available upon request.) Figure 9 reports, for either series, the fractions of forecast error variance (henceforth, FEV) explained by either of the two identified common shocks, at horizons up to 10 years ahead, together with the one and two-standard deviations bootstrapped confidence bands. Figure
10 reports, for either series, bias-corrected impulse-response functions (henceforth, IRFs) to each of the two shocks, together with one and two-standard deviations bootstrapped confidence bands. Bias-correction of the IRFs and computation of the confidence bands has been implemented as in Kilian (1998). As for the fractions of forecast error variance, on the other hand, I do not perform any bias correction, so that the thick black lines shown in the two figures are just the simple estimates. The vertical bars in Figure 9, corresponding to 6 and 32 quarters, mark the boundaries of the business-cycle frequency band.

### 3.4 The fractions of forecast error variance

Starting from the FEVs, several facts clearly emerge from Figure 9. The news common shock is estimated to explain negligible fractions of the FEVs of either TFP or the RPI at any horizon. By the same token, it plays a marginal role in explaining fluctuations in either real stock prices, inflation, or the real FED Funds rate. On the other hand, it plays a non-negligible role for consumption and especially for GDP, for which, based on median estimates, it explains about half of the FEV at horizons up to one year and a half, although the fraction decreases drastically at longer horizons, reaching about 10 per cent at the 10-year horizon. The non-news common shock, on the other hand, explains non-negligible fractions of the FEVs of both TFP and, especially, the RPI. Based on median estimates, in particular, this shock explains, at the 10-year horizon, almost 30 per cent of the FEV of either series, thus implying that—if we are truly willing to believe that TFP and the RPI share a common I(1) component—the size of this component is clearly not negligible. As for the other series, the common non-news shock appears to play, once again, a non-negligible role for all of them, with the single exception of the real ex post FED Funds rate, for which the contribution to the FEV is, based on median estimates, consistently smaller than 10 per cent at all horizons. For GDP, for example, the fraction of FEV explained by this shock increases from virtually zero on impact, to about 28 per cent at the 10 year horizon, whereas the analogous figures for stock prices are about 5 per cent and 19 per cent. As for inflation, this shock explains about one fifth of its FEV at all horizons beyond three years.

---

17The reason for this is the following. In a previous version of the paper I performed indeed a bias-correction for the fractions of FEV, with the result that, in some cases, some of the bootstrapped confidence bands ended up being either above 1 or below 0 for some horizon. Because of this, I then performed the bias-corrections based on the logit transformations of the relevant objects, and then I took the inverse-logit transformations of the resulting quantities. Although this approach, by construction, delivers bias-corrected fractions of FEV, and confidence bands, which are bounded between 0 and 1, it suffers form the drawback that, in some cases, the extent of the bias-correction turned out to be extremely (and, in my view, implausibly) large, due to the extreme non-linearity of the logit transformation for values quite close to either 0 or 1. Since neither of the two approaches appeared to be problem-free, in the end I simply decided to perform no bias correction, and to just show the simple estimates.
3.5 The impulse-responses to the two common shocks

Turning to the IRFs, either of the two common shocks has, at long horizons, a negative impact on the RPI, and a positive impact on GDP, consumption, and stock prices. As for TFP, the impact is mostly positive, but strictly speaking not statistically significant. On the other hand, an unexpected result is the permanent positive impact of either shock on hours, and the permanent negative impact on inflation. These results are problematic because it is not immediately clear based on what kind of general equilibrium mechanism permanent technology shocks may have permanent effects on either hours or inflation. As for the real ex post FED Funds rate, on the other hand, the impact of either shock is estimated to be not significantly different from zero at either horizon. As for the nature of the dynamic response, two things deserve to be stressed. First, based on median estimates, the response on impact of either TFP or the RPI to the non-news shock is very close to zero, and in fact it is not significantly different from zero at any significance level. Second, the shapes of the dynamic responses of hours, GDP, and consumption to the news shock are remarkably similar, with a comparatively large jump upwards at short horizons, and sizeable decreases further out, although the long-horizon impacts remain significantly different from zero.

4 Conclusions

Results from cointegration tests show that TFP and the relative price of investment are most likely not cointegrated. However, either cross-spectral methods, or a test in the spirit of Cochrane and Sbordone (1988), suggest that they share a common I(1) component, which induces positive long-horizon covariation between them. I have explored three alternative possible explanations for this finding: first, that it is genuine; second, that it results from non-linearity of the technology transforming consumption goods into investment goods, which implies that the relative price of investment is also impacted upon by neutral shocks; third, that it is the figment of not controlling for idiosyncratic breaks in the drifts of the two series. I have argued that the third explanation is the most likely, so that the two series are in fact independent at the very low frequencies, and they correctly measure neutral and investment-specific technology, respectively. However, I have also illustrated the extreme sensitivity of these results to the exact nature and timing of the breaks characterizing the two series: conditional on a common break identified via the Bai, Lumsdaine, and Stock (1998) procedure, evidence is compatible with the notion that the two series contain a common I(1) component inducing negative-long horizon covariation between them. These results illustrate the difficulty of making strong statements about the exact nature of the long-horizon relationship between TFP and the relative price of investment, which automatically injects an element of fragility into investigations about the role played by neutral and investment-specific shocks in macroeconomic fluctuations.
References


A The Data

The three RPI series are from Schmitt-Grohé and Uribe (2011), Liu, Waggoner, and Zha (2011), and Justiniano, Primiceri, and Tambalotti (2011). The sample periods for the three series are 1948Q1-2006Q4, 1959Q1-2007Q4, and 1954Q4-2009Q1, respectively. The main TFP series is the ‘purified TFP’ produced by John Fernald, available from the San Francisco FED’s website.

The VARs used in the empirical analysis in the final section feature the logarithms of TFP and the RPI; inflation; the ex post real Federal Funds rate; and the logarithms of hours worked, real GDP, real consumption, and real stock prices, all of them expressed in per capita terms. Following a standard convention—see, e.g., Fisher (2006)—the price of consumption is taken to be the numeraire of the system, which implies that (i) all real variables are computed by deflating the corresponding nominal variables by the consumption deflator, and (ii) real GDP is rescaled by the price of consumption (as described below).

A quarterly seasonally adjusted series for the consumption deflator has been computed by chain-weighting the deflators for non-durables and services consumption based on the data found in Tables 1.1.6, 1.1.6B, 1.1.6C, and 1.1.6D of the National Income and Product Accounts. Inflation has been computed as the log-difference of the consumption deflator. By the same token, a quarterly seasonally adjusted series for real consumption of non-durables and services has been computed by chain-weighting the respective series for real chain-weighted consumption of non-durables, and of services, respectively, based on the data found in the same tables.

A monthly series for the effective Federal Funds rate (FEDFUNDS) is from the St. Louis FED’s website, and it has been converted to the quarterly frequency by taking averages within the quarter. The series is quoted at a non-annualized rate in order to make its scale exactly comparable to that of inflation. The real ex post Federal Funds Rate has been computed as the difference between the thus rescaled FED Funds rate series and inflation.

A monthly series for the civilian noninstitutional population (CNP16OV) is from the U.S. Department of Labor, Bureau of Labor Statistics, and it has been converted to the quarterly frequency by taking averages within the quarter.

A quarterly seasonally adjusted series for hours of all persons in the nonfarm business sector (HOANBS) is from the U.S. Department of Labor, Bureau of Labor Statistics.

A monthly series for the nominal Standard & Poor’s composite index is from Robert Shiller’s website, and it has been converted to the quarterly frequency by taking averages within the quarter. Real stock prices have been computed by deflating the nominal Standard & Poor’s composite index by the consumption deflator, whereas real stock prices per capita has been computed as the ratio between real stock prices

\[ r_t = \left(1 + \frac{R_t}{100}\right)^{1/4} - 1. \]

18 So, to be clear, defining the original FED Funds rate series as $R_t$—with its scale such that, e.g., a ten per cent rate is represented as 10.0—the rescaled series is computed as $r_t = \left(1 + \frac{R_t}{100}\right)^{1/4} - 1$. 

and population.

Finally, quarterly seasonally adjusted series for real GDP in chained 2005 dollars (GDPC96) and for the GDP deflator (GDPCTPI) are from the U.S. Department of Commerce, Bureau of Economic Analysis.

As previously pointed out, the price of consumption is taken to be the numeraire of the system, and real GDP is therefore rescaled as follows. Real GDP is multiplied by the GDP deflator, and then divided by the consumption deflator (the resulting series is near-numerically identical to the one obtained by simply deflating nominal GDP by the consumption deflator).

Finally, hours worked, real GDP, real consumption, and real stock prices are all expressed in per capita terms by dividing them by population.
### Table 1 Bootstrapped p-values for augmented Dickey-Fuller tests with a time trend\(^a\)

<table>
<thead>
<tr>
<th></th>
<th>(p=1)</th>
<th>(p=2)</th>
<th>(p=4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Logarithm of the RPI from</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Schmitt-Grohé and Uribe (1948Q1-2006Q4)</td>
<td>0.903</td>
<td>0.892</td>
<td>0.888</td>
</tr>
<tr>
<td>Liu, Waggoner and Zha (1959Q1-2007Q4)</td>
<td>0.029</td>
<td>0.252</td>
<td>0.303</td>
</tr>
<tr>
<td>Justiniano, Primiceri, and Tambalotti (1954Q4-2009Q1)</td>
<td>0.847</td>
<td>0.899</td>
<td>0.928</td>
</tr>
<tr>
<td>Logarithm of TFP from</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Schmitt-Grohé and Uribe (1948Q1-2006Q4)</td>
<td>0.400</td>
<td>0.221</td>
<td>0.337</td>
</tr>
<tr>
<td>Fernald (1948Q1-2012Q4)</td>
<td>0.334</td>
<td>0.564</td>
<td>0.818</td>
</tr>
</tbody>
</table>

\(^a\) Based on 2,000 bootstrap replications of estimated ARIMA processes.

### Table 2 Bootstrapped p-values\(^a\) for Johansen’s trace test of the null of no cointegration between the logarithms of TFP and the RPI

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Based on Schmitt-Grohé and Uribe’s RPI</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>based on Schmitt-Grohé and Uribe’s TFP:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>lag order selected based on AIC(^b)</td>
<td>0.088</td>
<td></td>
<td></td>
</tr>
<tr>
<td>lag order selected based on SIC-HQ(^c)</td>
<td>0.121</td>
<td></td>
<td></td>
</tr>
<tr>
<td>based on Fernald’s TFP:</td>
<td>0.334</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Based on Liu, Waggoner, and Zha’s RPI and Fernald’s TFP</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Based on Justiniano, Primiceri, and Tambalotti’s RPI and Fernald’s TFP</td>
<td></td>
<td></td>
<td>0.076</td>
</tr>
</tbody>
</table>

\(^a\) Based on 2,000 bootstrap replications. \(^b\) Lag order = 7. \(^c\) Lag order = 3.

### Table 3 Tests for multiple breaks at unknown points in the sample in the mean based on Bai and Perron: bootstrapped p-values for double maximum tests

<table>
<thead>
<tr>
<th></th>
<th>(UD_{\text{max}})</th>
<th>(WD_{\text{max}})</th>
</tr>
</thead>
<tbody>
<tr>
<td>Log-difference of the RPI from</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Schmitt-Grohé and Uribe (1948Q2-2006Q4)</td>
<td>0.109</td>
<td>0.123</td>
</tr>
<tr>
<td>Liu, Waggoner and Zha (1959Q2-2007Q4)</td>
<td>0.043</td>
<td>0.045</td>
</tr>
<tr>
<td>Justiniano, Primiceri, and Tambalotti (1955Q1-2008Q2)</td>
<td>0.007</td>
<td>0.007</td>
</tr>
<tr>
<td>Log-difference of TFP from Fernald (1948Q2-2012Q4)</td>
<td>0.052</td>
<td>0.063</td>
</tr>
<tr>
<td>Log-difference of real GDP per capita (1948Q2-2012Q4)</td>
<td>0.594</td>
<td>0.561</td>
</tr>
<tr>
<td>Log-difference of real consumption per capita (1948Q2-2012Q4)</td>
<td>0.009</td>
<td>0.009</td>
</tr>
<tr>
<td>Log-difference of real S&amp;P 500 per capita (1948Q2-2012Q4)</td>
<td>0.591</td>
<td>0.565</td>
</tr>
<tr>
<td>Logarithm of hours worked per capita (1948Q2-2012Q4)</td>
<td>0.653</td>
<td>0.630</td>
</tr>
<tr>
<td>Inflation (1948Q2-2012Q4)</td>
<td>0.113</td>
<td>0.090</td>
</tr>
<tr>
<td>Real ex post Federal Funds rate (1954Q3-2012Q4)</td>
<td>0.359</td>
<td>0.326</td>
</tr>
</tbody>
</table>

\(^a\) Based on 2,000 bootstrap replications of estimated AR\(p\) processes.
Table 4 Tests for multiple breaks at unknown points in the sample in the mean based on Bai and Perron: bootstrapped $p$-values for $sup$-$F((t+1)/t)$ test statistics, estimated break dates, and sub-sample means (quoted at an annual rate)

|                              | $sup$-$F(2|1)$ $p$-value | Estimated break date, and 90% confidence interval | Sub-sample means |
|------------------------------|--------------------------|--------------------------------------------------|-----------------|
| Log-difference of the RPI from |                          |                                                  |                 |
| Schmitt-Grohé and Uribe (1948Q2-2002Q4) | 0.732                 | 1982Q2 [1981Q4; 1987Q2]                          | -1.270           |
| Justiniano, Primiceri, and Tambalotti (1954Q4-2008Q2) | 0.925               | 1982Q2 [1981Q4; 1988Q1]                          | -0.630           |
| Log-difference of TFP from Fernald (1948Q1-2008Q2) | 0.932                | 1968Q2 [1965Q3; 1982Q1]                          | 2.007            |
| Log-difference of real consumption per capita (1948Q1-2008Q2) | 0.656               | 2000Q4 [2000Q1; 2006Q4]                          | 1.861            |

Table 5 Tests for a single joint break at an unknown point in the sample in the intercepts of the VAR for the log-differences of TFP and the RPI based on Bai, Lumsdaine, and Stock (1998): bootstrapped $p$-values and estimated break dates

<table>
<thead>
<tr>
<th></th>
<th>$sup$-Wald $p$-value</th>
<th>$exp$-Wald $p$-value</th>
<th>Estimated break date, and 90% confidence interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>Log-difference of the RPI from:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Schmitt-Grohé and Uribe (1948Q2-2006Q4)</td>
<td>0.055</td>
<td>0.050</td>
<td>1975Q2 [1967Q4; 1982Q4]</td>
</tr>
<tr>
<td>Liu, Waggoner and Zha (1959Q2-2007Q4)</td>
<td>0.003</td>
<td>0.001</td>
<td>1982Q2 [1979Q3; 1985Q1]</td>
</tr>
<tr>
<td>Justiniano, Primiceri, and Tambalotti (1955Q1-2008Q2)</td>
<td>0.004</td>
<td>0.003</td>
<td>1977Q4 [1973Q1; 1982Q3]</td>
</tr>
</tbody>
</table>

Lag order chosen based on Schwartz and Hannan-Quinn:

Lag order set to 4:
Table 6 Results from applying Bai and Perron’s tests to the bootstrapped log differences of TFP and the RPI generated based on the VAR estimated conditional on the common break identified by Bai et al.’s (1998) tests

<table>
<thead>
<tr>
<th>RPI series from:</th>
<th>TFP</th>
<th>RPI</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SGU</td>
<td>JPT</td>
</tr>
<tr>
<td>Fractions of bootstrap replications for which a break has been identified</td>
<td>0.462</td>
<td>0.410</td>
</tr>
<tr>
<td>of which, fraction of replications for which the estimated confidence interval for the break date does not contain the common break:</td>
<td>0.352</td>
<td>0.382</td>
</tr>
</tbody>
</table>
Based on Schmitt-Grohe and Uribe’s log RPI and log TFP (lag order chosen based on the AIC)

Based on Justiniano, Primiceri, and Tambalotti’s log RPI, and Fernald’s log TFP (lag order chosen based on SIC and Hannan-Quinn)

Figure 1 Cointegration residuals between the logarithms of TFP and the RPI
Figure 2 Monte Carlo distribution of \((1/k)\) times Cochrane and Sbordone’s (1988) estimator, based on 10,000 simulations of two independent random walks of length \(T\).
Figure 3  Simple estimates of $(1/k)$ times Cochrane and Sbordone’s estimator, and bootstrapped distributions of the estimator under the null hypothesis that the two series are orthogonal.
Figure 4  Estimated cross-spectrum between the log-differences of TFP and the RPI
Figure 5  Simple estimates of \((1/k)\) times Cochrane and Sbordone’s estimator, and bootstrapped distributions of the estimator under the null hypothesis that the two series are orthogonal (controlling for breaks in the mean rates of growth identified by Bai and Perron)
Figure 6  Estimated cross-spectrum between the log-differences of TFP and the RPI (controlling for breaks in the mean rates of growth identified by Bai and Perron)
Figure 7  Simple estimates of $(1/k)$ times Cochrane and Sbordone’s estimator, and bootstrapped distributions of the estimator under the null hypothesis that the two series are orthogonal (controlling for breaks in the mean rates of growth identified by Bai, Lumsdaine, and Stock)
Figure 8  Estimated cross-spectrum between the log-differences of TFP and the RPI (controlling for breaks in the mean rates of growth identified by Bai, Lumsdaine, and Stock)
Figure 9 Fractions of forecast error variance explained by individual common shocks, with one- and two-standard deviations bootstrapped confidence bands (based on Schmitt-Grohe’ and Uribe’s RPI series)
Figure 10  Impulse-response functions to individual common shocks, with one- and two-standard deviations bootstrapped confidence bands (based on Schmitt-Grohe’ and Uribe’s RPI series)